Techniki optymalizacji

Symulowane wyżarzanie

Maciej Hapke
maciej.hapke at put.poznan.pl

Wyżarzanie

- wzrost temperatury gorącej kąpieli do takiej wartości, w której ciało stałe topnieje
- powolne zmniejszanie temperatury do chwili, w której cząsteczki ułożą się wzajemnie i osiągną (ang. ground state) temperaturę zerową

Algorytm Metropolis

- Metropolis i in. (1953) - algorytm statystycznego symulowania (Monte Carlo) zmian ciała stałego w gorącej kąpieli aż do stanu termicznej równowagi
- Losowa generacja sekwencji stanów ciała stałego:
  - stan i ciała stałego i jego energia Ei,
  - perturbacja (małe zniekształcenie) $\rightarrow$ następny stan. Energia następnego stanu wynosi Ej.
  - jeśli Ej - Ei $\leq$ 0, stan j jest akceptowany jako stan bieżący
  - w przeciwnym wypadku, stan j jest akceptowany z pewnym prawdopodobieństwem:

$$\exp\left(\frac{E_i - E_j}{k_B T}\right)$$

T – temperatura kąpieli

k_B – stała Boltzmanna

Analogie do optymalizacji

System fizyczny	Problem optymalizacji
stan	rozwiązanie
energia	koszt
stan stały	optimum
temperatura	parametr kontrolny *T*
szybkie schładzanie	lokalna optymalizacja
powolne schładzanie	symulowanie wyżarzanie

Symulowane wyżarzanie

Simulated annealing, Monte Carlo annealing, probabilistic hill climbing, stochastic relaxation

- Zastosowanie algorytmu Metropolis do optymalizacji kombinatorycznej.

Kryterium akceptacji

- i, j – rozwiązania
- f(i) i f(j) - koszty
- kryterium akceptacji określa czy j uzyskane z i jest akceptowane

$$\mathbf{P}_c\{\text{accept } j\} = \begin{cases} 1 & \text{jesli } f(j) \leq f(i) \\ \exp\left(\dfrac{f(i) - f(j)}{c}\right) & \text{jesli } f(j) > f(i) \end{cases}$$

Procedura SW

```
procedure SYMULOWANE_WYŻARZANIE;
begin
   INICJALIZUJ(x_start,C_0,L);
   k:=0;
   x:=x_start;
   repeat
      for l:=1 to L_k do
      begin
         GENERUJ(x' z Sx);
         if f(x')≤f(x) then
            x:=x'
         else
            if exp(-(f(x')-f(x))/C_k)>random[0,1) then
               x:=x';
      end;
      k:=k+1;
      OBLICZ(C_k);
   until WARUNEK_STOPU;
end;
```

Zbieżność algorytmu SW

- Można uzyskać takie L_k i c_k, które zapewniają zbieżność SW do opt.
- Dobre przybliżenie SW
    - generowanie homogenicznych łańcuchów Markowa skończonej długości dla skończonej sekwencji malejących wartości parametru kontrolnego

Łańcuch Markowa

- Łańcuch Markowa jest sekwencją prób (rozwiązań), w której prawdopodobieństwo wyniku danej próby zależy od wyniku poprzedniej próby.
- Łańcuch Markowa jest niehomogeniczny jeśli prawdopodobieństwo przejścia zależy od numeru próby k. Jeśli nie zależy łańcuch Markowa jest homogeniczny.

Sposób chłodzenia określa

- skończona sekwencja wartości parametru kontrolnego, tj.
  - początkowa wartość parametru kontrolnego c_0
  - funkcja dekrementacji parametru kontrolnego
  - końcowa wartość parametru kontrolnego
- skończona liczba przejść dla każdej wartości parametru kontrolnego, tj.
  - skończona długość każdego homogenicznego łańcucha Markowa

Sposób chłodzenia Kirkpatricka, Gelatta i Vecchi'ego

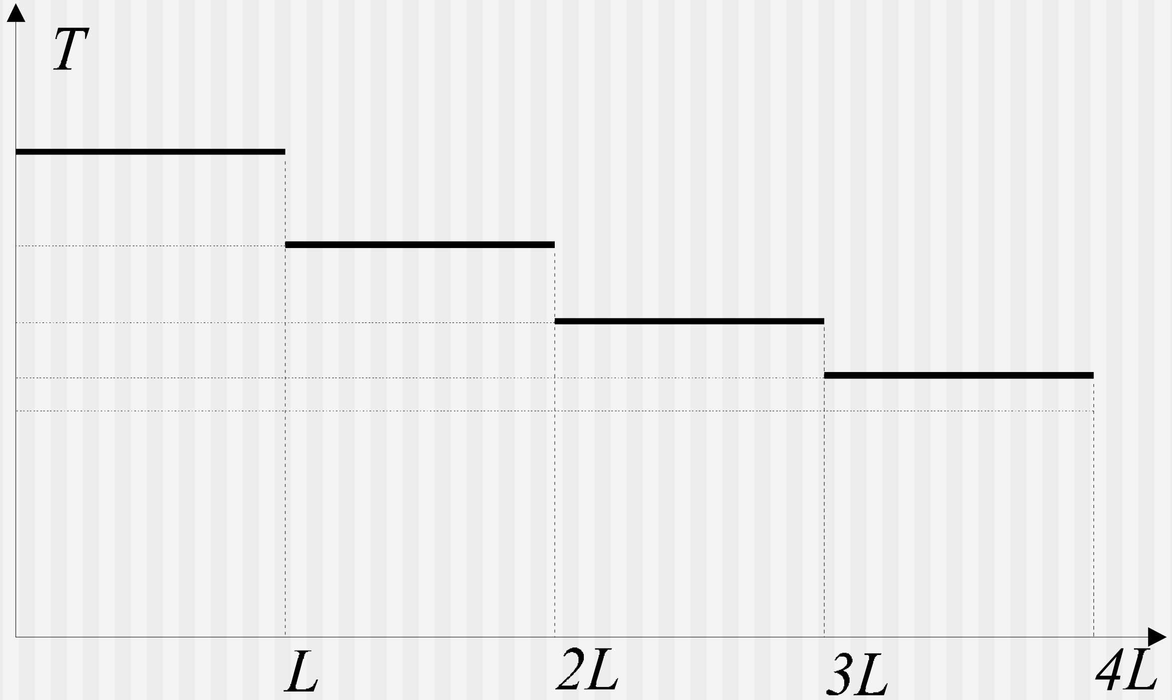
- Wartość początkowa parametru kontrolnego
  - Wartość c_0 powinna być odpowiednio wysoka by zapewnić akceptację wszystkich przejść (początkowy współczynnik akceptacji jest bliski 1).
  - Zależy od problemu (patrz formuła: Δf, c_k)
  - Np. $\Delta f = 1000$, $p \approx 0{,}98$ dla $c_k = 250$
  - Podgrzewanie
- Zmniejszanie wartości parametru kontrolnego

$$c_{k+1} = \alpha\, c_k, \quad k = 1, 2, \ldots$$

$$c_{k+1} = \alpha^k\, c_0$$

α jest stałą mniejszą ale bliską 1. (0.8 – 0.99)

Sposób chłodzenia $L_k = L$



Długość łańcucha Markowa

- Powinna być odpowiednia, tak by algorytm mógł „odwiedzić” wszystkich swoich sąsiadów przynajmniej raz (osiągnąć stan równowagi termicznej na każdym poziomie temperatury). Związek ze średnim rozmiarem sąsiedztwa.
- Ponieważ prawdopodobieństwo akceptacji zmniejsza się w czasie, można by się spodziewać, że $L_k \to \infty$ dla $c_k \to 0$.

  Dlatego istnieje pewne ograniczenie na długość łańcucha Markowa dla małych c_k.

Wartość końcowa parametru kontrolnego

- Algorytm kończy działanie, gdy wartość funkcji celu otrzymanego rozwiązania w ostatniej próbie łańcucha Markowa nie uległa zmianie dla kilku kolejnych łańcuchów

SA for TSP
Pierwsze implementacje

- Pierwsze implementacje
  - Kirkpatrick [1983]
    - 6000 miast, brak dokładnych informacji o jakości z wyjątkiem określenia „dobry”
    - Ruchy 2-opt (wymiana 2 łuków)
    - Liczba ruchów na każdym poziomie temperatury ma być proporcjonalna do rozmiaru sąsiedztwa – rozmiar N^2 (poważna wada, inne algorytmy są lepsze)
  - Cerny [1985]
    - 2-opt oraz ruchy z wymianą miast (gorsze rezultaty)

SA for TSP
Johnson, Aragon, McGeoch, Schevon [1996]

- Parametry
  - T_0 – 3% ruchów odrzucane
  - wsp. zmiany T, $\alpha = 0.95$
  - L = N(N-1)
  - Stop, 5 zmian temperatur bez poprawy najlepszego rozwiązania LUB akceptacja niższa niż 2%
- Obserwacje
  - Mniej więcej podobne wyniki dla wielu uruchomień, niezależnie od N
  - Jeśli L=N, niewielka poprawa
  - Lepsze od 2-opt, ale gorsze od 3-opt
  - Wydłużenie wyżarzania do 10N(N-1) pomaga i daje lepszy wynik od 3-opt, ale czas obliczeń mocno rośnie a jakość wzrasta słabo
    - Gdy N=100, SA daje lepsze wyniki od LK, ale w czasie 20000 razy dłuższym
    - N=1000, SA liczy 3.5 dnia, LK – 0,77 s

Porównanie wyników

Random Euclidean Instances			
	10^2	$10^{2.5}$	10^3
SA_1 Running Time in Seconds	12.4	188	3170
2-Opt Running Time in Seconds	0.03	0.09	0.34
3-Opt Running Time in Seconds	0.04	0.11	0.41
LK Running Time in Seconds	0.06	0.20	0.77
Number of Temperatures	117	143	171
Average SA_1 Percentage Excess	5.2	4.1	4.1
Average SA_1 Excess after 2-Opting	3.4	3.7	4.0
Average 2-Opt Excess	4.5	4.8	4.9
Average 3-Opt Excess	2.5	2.5	3.1
Average LK Excess	1.5	1.7	2.0

Jak przyspieszyć SA?

- Obcinanie sąsiedztwa (neighborhood pruning)
  - Wykluczanie długich odcinków
  - Może zmniejszyć rozmiar do N
- Niższe temperatury startowe
  - Strata czasu na bezowocne poszukiwania

Porównanie

		Random Euclidean Instances					
		Average Percent Excess			Running Time in Seconds		
Variant		10^2	$10^{2.5}$	10^3	10^2	$10^{2.5}$	10^3
SA_1 (Baseline Annealing)	$\alpha = 1$	3.4	3.7	4.0	12.40	188.00	3170.00
SA_1 + Pruning	$\alpha = 1$	2.7	3.2	3.8	3.20	18.00	81.00
SA_1 + Pruning	$\alpha = 10$	1.7	1.9	2.2	32.00	155.00	758.00
SA_2 (Pruning + Low Temp)	$\alpha = 10$	1.6	1.8	2.0	14.30	50.30	229.00
SA_2	$\alpha = 40$	1.3	1.5	1.7	58.00	204.00	805.00
SA_2	$\alpha = 100$	1.1	1.3	1.6	141.00	655.00	1910.00
2-Opt		4.5	4.8	4.9	0.03	0.09	0.34
Best of 1000 2-Opts		1.9	2.8	3.6	6.60	16.20	52.00
Best of 10000 2-Opts		1.7	2.6	3.4	66.00	161.00	517.00
3-Opt		2.5	2.5	3.1	0.04	0.11	0.41
Best of 1000 3-Opts		1.0	1.3	2.1	11.30	33.00	104.00
Best of 10000 3-Opts		0.9	1.2	1.9	113.00	326.00	1040.00
Lin-Kernighan		1.5	1.7	2.0	0.06	0.20	0.77
Best of 100 LK's		0.9	1.0	1.4	4.10	14.50	48.00