Algorytmy zrandomizowane (2)

www.qed.pl/ai/nai2003

PRZESZUKIWANIE LOSOWE

Najprostsza metoda: z przestrzeni stanów *S* losujemy wiele obiektów *x* i wybieramy ten, dla którego *f(x)* ma wartość największą.

Metoda bardziej złożona: po wylosowaniu np. 1000 obiektów zapamiętujemy 100 najlepszych i kolejnych losowań dokonujemy z "sąsiedztwa" tych 100.

(Przykład - problem komiwojażera: losujemy 1000 dróg, wybieramy 100 najkrótszych, następnie losowo je modyfikujemy, otrzymując kolejne 1000 przykładów itd.)

Metoda hybrydowa: wybieramy 100 najlepszych z 1000 losowych, następnie dla każdego z nich uruchamiamy algorytm wspinaczki.

Metody (często) oparte na sąsiedztwie.

MONTE CARLO - PRZYKŁAD

Minimalne pokrycie kolumnowe macierzy zerojedynkowej A.

Algorytm:
- Losujemy permutację (kolejność) kolumn.
- Niech B będzie zbiorem wszystkich kolumn - załóżmy, że pokrywają one A.
- Usuwamy z B pierwszą kolumnę (według ustalonej kolejności).
- Sprawdzamy, czy wszystkie wiersze nadal są pokryte. Jeśli nie, przywracamy usuniętą ostatnio kolumnę.
- Powtarzamy operacje dla kolejnych kolumn.

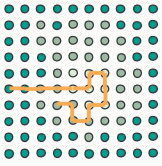
Algorytm nie musi dać globalnie optymalnego wyniku - błąd może być dowolnie duży.
Istnieje jednak taka permutacja kolumn, że algorytm znajdzie rozwiązanie optymalne - a więc, czekając odpowiednio długo, zawsze je otrzymamy.

```
1 1 0 0 0
1 0 1 0 0
1 0 0 1 0
1 0 0 0 1
```

PRZESZUKIWANIE TABU

Schemat działania: przeglądamy przestrzeń stanów (rozwiązań), przechodząc zawsze do tego sąsiada, dla którego wartość funkcji oceny jest największa (jak w algorytmie wspinaczki), o ile nie znajduje się on na liście punktów zabronionych. Na liście tej przechowujemy k ostatnio odwiedzonych punktów (np. k=1000).

Przykład: maksymalizacja funkcji dwuwymiarowej (na siatce o ustalonej dokładności).



Wersja zrandomizowana:
Z prawdopodobieństwem 1/2 wybieramy najlepszego (dozwolonego) sąsiada, z prawd. 1/4 - drugiego z kolei, z prawd. 1/8 - trzeciego itd.

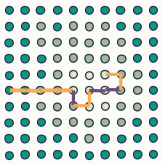
Metoda oparta na sąsiedztwie.

METODA WYCHŁADZANIA

(Simulated annealing)

Schemat działania: startujemy z losowego punktu, wybieramy ***losowo*** *nowy punkt z sąsiedztwa aktualnego, jeśli ma większą wartość - idziemy tam, jeśli mniejszą - z pewnym prawdopodobieństwem albo tam idziemy, albo testujemy następnego sąsiada.*

```
x0 = Random(X)
do
  x=Random(N(x0))
  if(f(x)>f(x0)) x0=x
  else
    if(Random()<P(x0,x))
      x0=x
while(time_out)
```



Metoda oparta na sąsiedztwie.

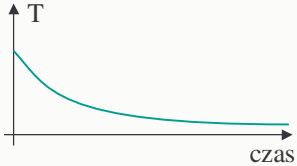
TEMPERATURA W ALG. WYCHŁADZANIA

Prawdopodobieństwo akceptacji gorszego stanu:

$$P(x_0, x) = e^{-\frac{|f(x_0)-f(x)|}{T}}$$

gdzie T - pewien parametr (“temperatura”)

Prawdopodobieństwo zależy od temperatury (im niższa, tym mniejsze), oraz od spadku wartości funkcji f.



Temperatura zwykle zmniejsza się w czasie.

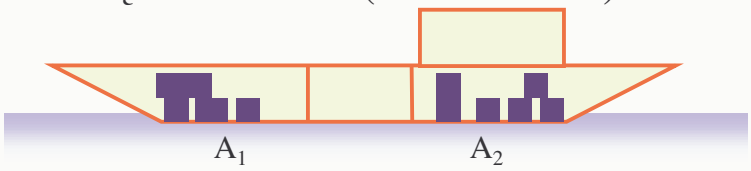
- Zalety: odporność na maksima lokalne.
- Wady: wolniejsza zbieżność.

SĄSIEDZTWO - ZASADY OGÓLNE

Metody typu wychładzanie, przeszukiwanie tabu, algorytm wspinaczki, bazują na pojęciu **sąsiedztwa** rozwiązań (punktów przestrzeni stanów). Może być ono zdefiniowane przez nas w zasadzie dowolnie, jednak należy trzymać się kilku zasad ogólnych:

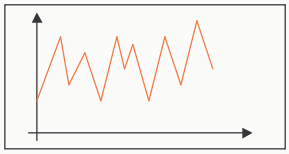
- **Sąsiedztwo nie powinno być zbyt liczne.**
  W praktyce: najwyżej kwadratowo liczne. Częścią wielu algorytmów jest pętla przeglądająca wszystkich sąsiadów, lub funkcja losująca jednego z nich.
- **Relacja sąsiedztwa powinna być spójna.**
  Tzn. do każdego rozwiązania powinniśmy dojść idąc od sąsiada do sąsiada. Inaczej część z nich będzie nieosiągalna.
- **Przejście do sąsiada powinno niewiele zmieniać funkcję celu.**
  Milczące założenie: w pobliżu dobrych rozwiązań warto szukać jeszcze lepszych - bez tego heurystyki są bezradne.

SĄSIEDZTWO (PRZYKŁAD)



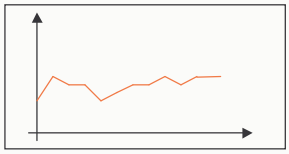
Mamy dany zbiór *n* wartości rzeczywistych $\{a_1, \ldots, a_n\}$. Podzielić zbiór na dwa rozłączne podzbiory A_1 i A_2 takie, żeby suma wartości z A_1 była bliska sumie wartości z A_2.

Funkcja celu: moduł różnicy sum zbiorów (minimalizujemy).



Sąsiednie rozwiązanie: przekładamy jeden pakunek.

Trudne do optymalizacji.



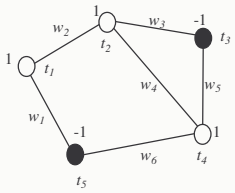
Sąsiednie rozwiązanie: zamieniamy parę pakunków.

Relacja niespójna.
Znacznie więcej sąsiadów.

Najlepiej połączyć te dwie definicje.

Maszyny Boltzmanna

Sieć (graf) złożona z **jednostek** wyposażonych w wartości progowe *t* i **krawędzi** opisanych wagami *w*. Jednostka może się znajdować w jednym z dwóch stanów: 1 lub -1.



Sieć zmienia swój stan przez równoległe wyliczenie energii każdej *i*-tej jednostki:

$$E_i = -1/2\, m_i\, (\, m_1 w_{i1} + \ldots + m_k w_{ik} + t_i\,)$$

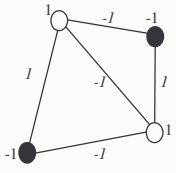
gdzie m_j - stan *j*-tej jednostki, w_{ij} - waga krawędzi łączącej jednostki *i* i *j*.

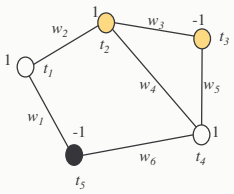
Im energia większa, tym większe prawdopodobieństwo, że jednostka zmieni stan na przeciwny (według **algorytmu wychładzania**).

Metoda oparta na sąsiedztwie (lokalna).

Maszyna Boltzmanna - przykład

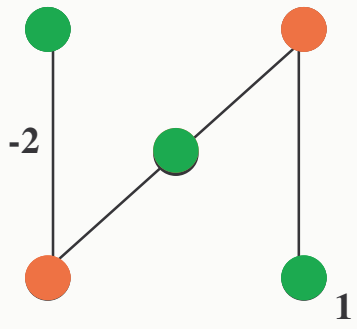
Ta sieć ustabilizuje się w takim stanie, w którym po lewej stronie będą wartości 1, a po prawej -1 (lub odwrotnie). Krawędzie o ujemnych wagach wspierają stan, w którym na obu krańcach jednostki mają przeciwne wartości.





W praktycznych zastosowaniach część jednostek może być traktowanych jako dane wejściowe (nie zmieniające swoich wartości). Zastosowania: np. rozpoznawanie wzorców.

Maszyna Boltzmanna - przykład



Przykładowe działanie dla problemu znajdowania maksymalnego zbioru wierzchołków niezależnych