Métaheuristiques pour l'optimisation TP3

Deniz Sungurtekin

October 2020

1 Introduction

Le but de ce travail pratique est de se familiariser avec l'algorithme "Simulated Annealing" (SA), afin de pouvoir résoudre différentes configurations données du problème de "Traveling Salesman". Cette méthode est inspirée de la physique, plus précisément de la métallurgie, elle consiste a reproduire notamment deux processus de traitement de la température, le recuit (Annealing) et la trempe (Quenching). Le premier consiste à baisser de manière constante la température afin d'obtenir un métal dont les atomes seront organisés en une structure très régulière. De son côté le "Quenching" consiste à baisser très rapidement la température pour obtenir une structure dont les atomes seront dispersés dans des énergies minimums local.

Dans un premier temps, l'objectif est de simuler le phénomène de recuit, afin de résoudre le problème de "Traveling Salesman", ce problème consiste à trouver, à partir des positions de n villes, le chemin le plus court. L'intention est donc de comparé les résultats de cette algorithme à celui de Greedy afin d'observer les éventuels différences, puis d'analyser l'impact de la diminution des températures et de la taille du problème sur notre système.

Ensuite, nous discuterons du "Parallel Tempering" qui peut être considéré comme une version en parallèle de la SA où les températures ne sont pas réduites de manière constante mais de façon abrupte. Cela permet la simulation du processus de la trempe (Quenching), pour que chaque répliquât du système travaillant en parallèle obtienne un minimum local du problème. Malheureusement, l'implémentation de cette méthode n'a pas pu être faite mais nous discuterons de son fonctionnement et des difficultés qu'il en résulte.

2 Méthodologie

Premièrement, expliquons en détails le fonctionnement du "Simulated Annealing", pour cela analysons la structure d'une configuration pour le problème donné. Elle est constituée d'une liste de ville possédant un id, une position X et une position Y. Un voisin d'une configuration est défini comme étant cette

même configuration qui a subit un swap sur deux de ses villes, par conséquent on remarque que chaque configuration possède $n^*(n-1)/2$ voisins. L'ordre de la liste indique l'ordre de passage du marchand à travers les villes, il est donc important de ne pas oublier d'ajouter à cette liste la première ville afin de revenir à la position de départ sans pour autant qu'il soit possible d'effectuer une permutation avec ce dernier élément.

Dans un premier temps, l'idée est de générer un parcours complètement aléatoire qui sera la configuration initial de notre solution. Ensuite, il est nécessaire de définir une bonne température initial, pour cela on effectue 100 fois un swap à partir de la configuration initial et on calcule la différence d'énergie entre les deux configurations à chaque itération puis on mesure la moyenne obtenu sur ces 100 résultats. Dans notre cas, l'énergie est donnée par la longueur du chemin défini par une configuration, la différence d'énergie revient donc à calculer la longueur du chemin de la configuration voisine à laquelle on soustrait la longueur du chemin de la configuration courante. Finalement, on peut utiliser cette valeur pour obtenir une bonne température en utilisant l'équation ci-desous:

$$\exp(\frac{-\Delta E}{T_0}) = 0.5$$

En isolant T0, on obtient:

$$T_0 = \frac{-\Delta E}{\ln(0.5)}$$

Cependant, on remarque que ce résultat peut être négatif ce qui est un problème puisqu'on veut qu'il soit au maximum 1 et au minimum 0. Par conséquent, il suffit de prendre la valeur absolue de ce résultat afin d'obtenir une bonne valeur initial de notre température.

On peut donc désormais lancer notre algorithme qui, à chaque itération, calculera un voisin aléatoire et évaluera sa probabilité d'être choisis comme nouvel état courant. Pour cela on utilise l' "Acceptance/Rejection rule" qui consiste à utiliser la différence d'énergie pour définir à quel point ce voisin est un bon candidat. Cette estimation est établit par la fonction suivante où P désigne la probabilité de la configuration voisine:

$$P = \begin{cases} 1 & \text{if } \Delta E < 0\\ \exp(\frac{-\Delta}{T}) & \text{otherwise} \end{cases}$$

Il suffit ensuite de générer un nombre compris entre 0 et 1 et de vérifier si la probabilité obtenu est plus grande ou égale à ce nombre afin de simuler le choix du prochain voisin.

Pour savoir quand la température doit être réduite, il faut que le système atteigne un état d'équilibre. Cet état est obtenue lorsqu'on a accepté 12*n voisins ou si il y a eu 100*n tentatives d'affecter un nouveau voisin, avec n le nombre de ville. C'est donc à ce moment qu'on affecte une nouvelle valeur de T, dans

notre cas on a choisi:

$$T_{k+1} = 0.9T_k$$

Finalement, on arrête le système lorsqu'un état de "Freezing" est obtenu, celui-ci est atteint lorsque l'état d'équilibre est atteint et que le système n'a pas améliorer la valeur minimum d'énergie durant les 3 dernières itérations de modification de la température.

L'algorithme greedy correspondant est beaucoup plus simple à comprendre puisqu'il consiste uniquement à choisir initialement une ville au hasard et prendre son voisin le plus proche jusqu'à qu'elles soient toute visitées par le marchand.

En ce qui conserne le Parallel Tempering, l'idée est d'effectuer l'algorithme SA avec plusieurs répliquâts du système possédant chacun une valeur T. On parle de Quenching parce qu'il est possible d'inter-changer deux configurations parmi les N systèmes. Par conséquent, si on effectue cela sur deux systèmes qui possède une température différente, on assiste à un changement de température brusque. Chaque système possède une température T tel que:

$$T_1 < T_2 < ... T_N$$

Ils résolvent indépendamment le problème tout comme SA mais la difficulté réside dans le choix de ces températures puisqu'elle a une importance cruciale sur le résultat obtenu. Avec SA, on baisse de manière constante la température, car on veut trouver le minimum global du problème. C'est pour cela qu'on passe d'une phase d'exploration avec une haute température à une phase d'exploitation avec une température basse empêchant la configuration d'effectuer des trop grand sauts de valeur. Avec une résolution en parallèle, il est plus avantageux de faire des changements de température brusque, car cela augmente la probabilité que chaque système atteigne un minimum local du problème. Par conséquent, on favorise les chances d'obtenir une valeur optimal en prenant la meilleur fitness parmi tout les systèmes.

3 Résultats obtenues

Voici les 10 résultats obtenu lorsqu'on lance 10 fois l'algorithme SA avec le fichier cities.dat:

```
My Configuration is: [['p', 6.0, 3.0], ['g', 4.0, 4.0], ['o', 2.0, 6.0], ['k', 1.0, 6.0], ['r', 0.0, 6.0], ['e', 0.5, 3.0], ['n', 0.5, 1.3], ['b', 0.0, 1.0], ['a', 0.0, 0.0], ['c', 1.0, 0.0], ['h', 3.0, 0.0], ['f', 3.0, 0.5], ['d', 2.0, 2.0], ['j', 1.3, 2.0], ['l', 2.0, 3.0], ['g', 3.0, 2.0], ['i', 4.0, 1.0], ['m', 7.5, 3.0], ['p', 6.0, 3.0]]
My path length is: 28.971958755080447
 My Configuration is:
(['a', 0.0, 0.0], ['b', 0.0, 1.0], ['n', 0.5, 1.3], ['j', 1.3, 2.0], ['d', 2.0, 2.0], ['f', 3.0, 0.5], ['h', 3.0, 0.0], ['i', 4.0, 1.0], ['q', 3.0, 2.0], ['l', 2.0, 3.0], ['e', 0.5, 3.0], ['r', 0.0, 6.0], ['k', 1.0, 6.0], ['o', 2.0, 6.0], ['g', 4.0, 4.0], ['p', 6.0, 3.0], ['m', 7.5, 3.0], ['c', 1.0, 0.0], ['a', 0.0, 0.0]]

My path length is: 31.156312994642544
 My Configuration is:
 ny contagua etaun as.
[['m', 7.5, 3.0], ['i', 4.0, 1.0], ['f', 3.0, 0.5], ['h', 3.0, 0.0], ['c', 1.0, 0.0], ['a', 0.0, 0.0], ['b', 0.0, 1.0], ['n', 0.5, 1.3], ['j', 1.3, 2.0], ['d', 2.0, 2.0], ['q', 3.0, 2.0], ['l', 2.0, 3.0], ['e', 0.5, 3.0], ['r', 0.0, 6.0], ['k', 1.0, 6.0], ['k', 1.0, 6.0], ['k', 1.0, 6.0], ['m', 7.5, 3.0]]
My path length is: 27.515362563425352
 ny contagua estant 25.
[['n', 0.8, 1.3], ['e', 0.8, 3.0], ['n', 0.0, 6.0], ['k', 1.0, 6.0], ['o', 2.0, 6.0], ['g', 4.0, 4.0], ['p', 6.0, 3.0], ['m', 7.5, 3.0], ['i', 4.0, 1.0], ['h', 3.0, 0.0], ['f', 3.0, 0.5], ['q', 3.0, 2.0], ['l', 2.0, 3.0], ['d', 2.0, 2.0], ['j', 1.3, 2.0], ['c', 1.0, 0.0], ['a', 0.0, 0.0], ['b', 0.0, 1.0], ['n', 0.5, 1.3]]
My path length is: 28.470902397390756
My Configuration is: [['q', 3.0, 2.0], ['i', 4.0, 1.0], ['m', 7.5, 3.0], ['p', 6.0, 3.0], ['g', 4.0, 4.0], ['o', 2.0, 6.0], ['k', 1.0, 6.0], ['r', 0.0, 6.0], ['e', 0.5, 3.0], ['l', 2.0, 3.0], ['f', 3.0, 0.5], ['h', 3.0, 0.0], ['c', 1.0, 0.0], ['a', 0.0, 0.0], ['b', 0.0, 1.0], ['n', 0.5, 1.3], ['j', 1.3, 2.0], ['d', 2.0, 2.0], ['q', 3.0, 2.0]]
My path length is: 29.089910978242706
 My Configuration is:
 [['q', 3.0, 2.0], ['d', 2.0, 2.0], ['l', 2.0, 3.0], ['o', 2.0, 6.0], ['k', 1.0, 6.0], ['r', 0.0, 6.0], ['e', 0.5, 3.0], ['j', 1.3, 2.0], ['n', 0.5, 1.3], ['b', 0.0, 1.0], ['a', 0.0, 0.0], ['c', 1.0, 0.0], ['h', 3.0, 0.0], ['f', 3.0, 0.5], ['i', 4.0, 1.0], ['m', 7.5, 3.0], ['p', 6.0, 3.0], ['g', 4.0, 4.0], ['q', 3.0, 2.0]]

My path length is: 28.58941470129242
 My Configuration is:
 THE COLOR OF THE C
 My Configuration is:
 (if', 3.0, 0.5], ['g', 4.0, 4.0], ['o', 2.0, 6.0], ['k', 1.0, 6.0], ['r', 0.0, 6.0], ['e', 0.5, 3.0], ['l', 2.0, 3.0], ['j', 1.3, 2.0], ['n', 0.5, 1.3], ['b', 0.0, 1.0], ['a', 0.0, 0.0], ['c', 1.0, 0.0], ['d', 2.0, 2.0], ['q', 3.0, 2.0], ['p', 6.0, 3.0], ['m', 7.5, 3.0], ['i', 4.0, 1.0], ['h', 3.0, 0.0], ['f', 3.0, 0.5]]
My path length is: 31.720316741057463
 My Configuration is:
 my contagua attor 15:
[['o', 2.0, 6.0], ['k', 1.0, 6.0], ['r', 0.0, 6.0], ['e', 0.5, 3.0], ['l', 2.0, 3.0], ['d', 2.0, 2.0], ['j', 1.3, 2.0], ['n', 0.5, 1.3], ['b', 0.0, 1.0], ['a', 0.0, 0.0], ['c', 1.0, 0.0], ['h', 3.0, 0.0], ['f', 3.0, 0.5], ['q', 3.0, 2.0], ['i', 4.0, 1.0], ['m', 7.5, 3.0], ['p', 6.0, 3.0], ['g', 4.0, 4.0], ['o', 2.0, 6.0]]
My path length is: 27.897228574675456
 My Configuration is:
```

Et voici les résultats avec Greedy:

```
Ny Configuration is:
[['h', 3.0, 0.0], ['r', 1.0, 0.0], ['r', 0.0, 1.0], ['n', 3.0, 2.0], ['n', 2.0, 3.0], ['n', 1.3, 2.0], ['n', 0.5, 3.0], ['n', 0.5, 3.0], ['n', 0.0, 3.0], ['n', 0.0, 0.0]]

Ny path length is: 33.86500051759808

Ny Configuration is:
[['s', 0.0, 0.0], ['r', 1.0, 0.0], ['n', 0.5, 1.3], ['b', 0.0, 0.0]]

Ny path length is: 33.86500051759808

Ny Configuration is:
[['s', 0.0, 0.0], ['n', 0.0, 0.0], ['n', 0.5, 1.3], ['b', 0.0, 0.0]]

Ny path length is: 34.351254934867

Ny Configuration is:
[['k', 1.0, 0.0], ['n', 0.0, 0.0], ['n', 7.5, 3.0], ['a', 0.0, 0.0]]

Ny path length is: 34.35125493487

Ny Configuration is:
[['k', 1.0, 0.0], ['n', 0.0, 0.0], ['n', 2.0, 0.0], ['a', 0.0, 0.0]]

Ny path length is: 34.35125493487

Ny Configuration is:
[['k', 1.0, 0.0], ['n', 0.0, 0.0], ['n', 2.0, 0.0], ['a', 1.0, 0.0]]

Ny configuration is:
[['k', 1.0, 0.0], ['n', 3.0, 0.0], ['n', 3.0, 2.0], ['a', 3.0, 2.0], ['y', 1.3, 2.0], ['n', 0.5, 1.3], ['b', 6.0, 1.0], ['a', 0.0, 0.0]]

Ny configuration is:
[['s', 4.0, 1.0], ['r', 3.0, 0.0], ['n', 3.0, 0.0], ['k', 1.0, 0.0]]

Ny configuration is:
[['s', 4.0, 1.0], ['r', 3.0, 0.0], ['n', 3.0, 0.0], ['k', 1.0, 0.0]]

Ny configuration is:
[['s', 4.0, 1.0], ['r', 3.0, 0.0], ['n', 3.0, 0.0], ['n', 3.0, 3.0], ['n', 0.5, 1.3], ['b', 6.0, 1.0], ['a', 0.0, 0.0], ['r', 1.0, 0.0], ['r', 4.0, 1.0], ['r', 4.0, 4.0], ['r', 6.0, 0.0], ['r', 1.0, 0.0], ['r', 0.0, 0.0], ['r
```

On remarque que globalement SA produit un meilleur résultat, de plus on voit que la meilleur fitness obtenu par SA est de 28.97 et que pour Greedy, on a une longueur de 29.48. Cela signifie que la moyenne des résultats et nettement meilleur pour SA mais que la différence entre les meilleurs résultats de chaque méthode n'est pas si grande. Par conséquent, on déduit qu'il est tout de même possible d'obtenir un bon résultat avec Greedy selon la configuration initiale.

Maintenant observons les résultats pour le fichier cities2.dat avec l'algorithme SA :

```
| Wy Configuration is: | ['ca9'], ['ca9'], ['ca4'], ['c41], ['c41], ['c62'], ['c68'], ['c68']
```

puis Greedy:

```
## (case) | [case] |
```

Tout comme la comparaison précédente, on observe des meilleurs résultats pour SA. Cependant cette fois-ci, le meilleur résultat de Greedy est assez éloigné du meilleur résultat pour SA. Ceci a du sens, contrairement à SA, le chemin obtenu par Greedy dépend fortement de la configuration initiale, par conséquent plus le nombre de ville est grand, moins Greedy a des chances d'obtenir une configuration optimale.

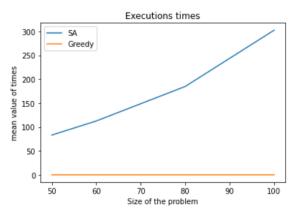
Analysons maintenant l'impact de la taille du problème sur le temps d'exécution des algorithmes. Pour cela, nous allons lancer nos algorithmes 10 fois avec 4 tailles différentes: 50,60,80 et 100.

Voici la moyenne et l'écart-type des temps en secondes obtenus:

```
mean times for SA
size 50:
          83.14425339698792
size 60:
          112.61741814613342
size 80:
          184.7649436235428
size 100:
          302.5018871068954
mean times for Greedy
size 50:
          0.0017951726913452148
          0.0025930881500244142
size 60:
          0.004487943649291992
size 80:
size 100:
           0.006781959533691406
std for SA
size 50:
          21.22596110152987
size 60:
          20.32868843752929
size 80:
          69.52204853168803
          141.19881919857642
size 100:
std for Greedy
size 50:
          0.0003923519502726057
size 60:
          0.0004875888539297336
size 80:
          0.0004989183075403973
size 100:
           0.000400263685940915
```

On observe clairement une immense différence de temps de calcule entre les deux algorithmes, Greedy trouve une solution de 50'000 à 100'000 fois plus vite que SA pour les tailles observés. Plus la taille du problème est grande plus la différence est notable. On déduit donc que SA est une méthode plus efficace mais bien plus coûteuse que Greedy. En ce qui concerne les valeurs d'écart-type, celle de Greedy sont presque nul puisque la résolution s'effectue toujours presque instantanément et donc ne varie que très peu. Contrairement à SA où l'écart-type croît très rapidement selon la taille du problème.

Voici maintenant le graphe permettant de mieux réaliser la différence de temps d'éxecution entre les deux méthodes.



Pour conclure, on peut affirmer que SA est une méthode bien plus efficace que

Greedy mais qui est extrêmement plus coûteuse en temps (Selon mon implementation de SA qui n'est sûrement pas optimale). On en déduit que son utilisation demande une très grande force de calcule pour résoudre des problèmes conséquents. De son côté Greedy donne une solution acceptable presque instantanément mais sa précision dépend fortement de la taille du problème puisque cela influe sur la probabilité d'obtenir une configuration initial optimal.