

TD noté

Métaheuristiques pour l'optimisation d'un clavier

Préambule : ce travail est à réaliser en binôme.

Le but de ce TD est d'implémenter et d'étudier le recuit simulé, la recherche tabou ou les algorithmes génétiques sur un problème de positionnement optimal des lettres sur un clavier informatique afin d'écrire le plus rapidement possible un texte en français.

1 Présentation du problème

Soit les 26 lettres de l'alphabet à placer sur un clavier 4x10 (Fig. 1). Chaque emplacement fait un centimètre de côté.

W		P	D		I	Z	O	N	
V	J	U		R	M		A		H
	X	Q	F		E	B	S	T	
K		C		L		Y	G		

FIGURE 1 – Représentation d'une configuration possible du clavier.

Une estimation des fréquences des occurrences de bigrammes, c'est-à-dire des successions de deux lettres, dans la langue française peut être obtenue à partir du fichier freqBigrammes.txt (disponible sur arche). Ce tableau a été construit en comptant les occurrences dans un texte français de 100'000 lettres composé de textes de Gustave Flaubert (20'600 lettres), de Jules Verne (19'438) et de trois articles de l'Encyclopedia Universalis, le premier consacré à Bruges (8'182), le deuxième à l'artillerie (25'078) et le dernier à la population (26'702)¹.

2 Travail à réaliser

Vous devez choisir une métaheuristique parmi le recuit simulé, la recherche tabou et les algorithmes génétiques.

1. les occurrences étaient disponibles sur un site maintenant inexistant <http://www.nymphomath.ch/crypto/stat/francais.html>

2.1 Conception

2.1.1 Etat/Individu

Q. 1 : Quelle représentation d'un état proposez-vous en fonction de votre métaheuristique sur le problème du positionnement des touches du clavier informatique français ?

Q. 2 : Quelle est la taille de l'espace de recherche avec cette représentation ?

2.1.2 Fonction énergie/objectif/d'adaptation

Q. 3 : Quelle fonction énergie/objectif/d'adaptation proposez-vous pour le recuit simulé, la recherche tabou ou les algorithmes génétiques ?

Q. 4 : Ecrivez la formule :

- quand on tape avec un doigt ;
- quand on tape avec un doigt de chaque main (chaque doigt s'occupe de la partie droite ou gauche du clavier).

Pour la suite du travail, on supposera à minima que le texte est tapé à l'aide d'un seul doigt.

2.1.3 Précisez et justifiez

Q. 5.a : Pour le recuit simulé :

1. L'initialisation de l'état de départ ;
2. Les mouvements possibles ;
3. la valeur de départ de la température ;
4. la fonction d'évolution de la température ;
5. le critère d'arrêt ;
6. toute autre information utile.

Q. 5.b : Pour la recherche tabou :

1. L'initialisation de l'état de départ ;
2. Les mouvements possibles ;
3. la taille de la liste tabou ;
4. le critère d'arrêt ;
5. toute autre information utile.

Q. 5.c : Pour les algorithmes génétiques

1. L'initialisation des états de départ ;
2. la taille de la population ;
3. l'opérateur de sélection des parents ;
4. l'opérateur de croisement ;
5. l'opérateur de mutation et la probabilité de mutation ;

6. l'opérateur de remplacement ;
7. le critère d'arrêt ;
8. toute autre information utile.

2.2 Implémentation

Le choix du langage est libre.

2.3 Rapport

Pour expliquer vos solutions algorithmiques et vos programmes, rédigez un court rapport de quelques pages, qui devra au moins contenir des éléments de réponse (par exemple des figures) aux points mentionnés dans la section 2.1. N'hésitez pas à décrire en détails vos solutions, vos choix de paramètres, vos difficultés et n'importe quelle information ou astuce supplémentaire qui vous semblerait pertinente à la résolution du problème.

3 Travail à rendre

Une archive contenant le rapport, les sources et un fichier `readme.txt` qui décrit les options de lancement du programme doit être déposée sur la page arche dédiée à l'EC optimisation au plus tard le vendredi 14 avril 2023.