

Projet

Métaheuristiques pour l'optimisation d'un clavier

Préambule : ce projet peut être réalisé en binôme. Vous pouvez aussi le réaliser en trinôme si les trois méthodes sont implémentées.

Le but de ce projet est de comparer deux méthodes d'optimisation au choix parmi le recuit simulé, la recherche tabou et les algorithmes génétiques sur le problème du 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`. 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.

On supposera que le texte est tapé à l'aide d'un seul doigt.

2 Travail à réaliser

2.1 Réalisation

Vous devez implémenter efficacement le recuit simulé, la recherche tabou et/ou les algorithmes génétiques afin de comparer leurs performances sur le problème du positionnement des touches du clavier informatique français. Le choix du langage de programmation

1. Pour plus de détails, voir <http://www.nymphomath.ch/crypto/stat/francais.html>

est libre.

Paramètres modifiables dans l'interface :

- pour le recuit simulé : la valeur de départ de la température ;
- pour la recherche tabou : la taille de la liste tabou ;
- pour les algorithmes génétiques : la taille de la population et la probabilité de mutation ;
- tout autre paramètre jugé utile.

Informations à afficher dans l'interface :

- le nombre d'itérations ;
- la meilleure configuration courante ;
- l'évolution de la valeur de la fonction objectif associée à la meilleure configuration ;
- toute autre information jugée utile.

Afin de privilégier les temps de calcul, vous pouvez prévoir un paramètre supplémentaire dans l'interface qui définit le pas de rafraichissement des informations afin de réduire les temps d'affichage si besoin.

2.2 Rapport

Un court rapport doit présenter :

- la modélisation choisie pour chaque méthode (représentation d'une configuration, taille de l'espace de recherche, fonction objectif, mouvements possibles ou opérateurs de sélection/croisement/mutation, critère d'arrêt...);
- une critique de chaque méthode à travers différents exemples ;
- une comparaison des méthodes comprenant les meilleures solutions trouvées et les paramètres associés ;
- une conclusion présentant la meilleure configuration trouvée et la valeur de la fonction objectif associée ;
- une bibliographie/webographie ;
- un diagramme de classes (comportant le nom des champs et des fonctions) en annexe.

Le rapport et les sources doivent être déposés sur la page arche dédiée à l'EC optimisation au plus tard le vendredi 29 mai 2015 à 18h.