

Manuel Utilisation de GenAlg

Sébastien Liou, Antony Perzo, Tanguy Colleville, Timothée Grandchamp

29 janvier 2021



CentraleSupélec



GENALG

Table des matières

1	Introduction	4
2	Interface utilisateur	4
3	VS-Code	8
4	Conclusion	10

Table des figures

1	Application au lancement	4
2	Application avec un fichier sélectionné	5
3	Application avec un calcul en cours	6
4	Premier résultat de "Lancer les calculs : Convergence de la solution"	7
5	Second résultat de "Lancer les calculs : Molécule 3D"	7
6	Résultat de "Tracer de la trajectoire 3D de la molécule initiale" .	8
7	Exemple 1 Utilisation du Terminal	10
8	Exemple 2 Utilisation du terminal	10

1 Introduction

Nous allons dans le présent manuel d'utilisation détailler l'utilisation de notre outil GenAlg. Ce manuel se décomposera en deux parties. En effet, nous nous intéresserons aux différentes façons de lancer l'algorithme. Notez que vous trouverez le projet sur ce gitlab : https://gitlab-cw2.centralesupelec.fr/tanguy.colleville/jeu_evolutionnaire

2 Interface utilisateur

Nous avons décidé, par soucis de clarté et d'ergonomie pour l'utilisateur, de créer une interface userfriendly. Nous avons ainsi utilisé Tkinter. Placez vous dans le programme interface utilisateur.py. Lancez-le. La fenêtre suivante va s'ouvrir :

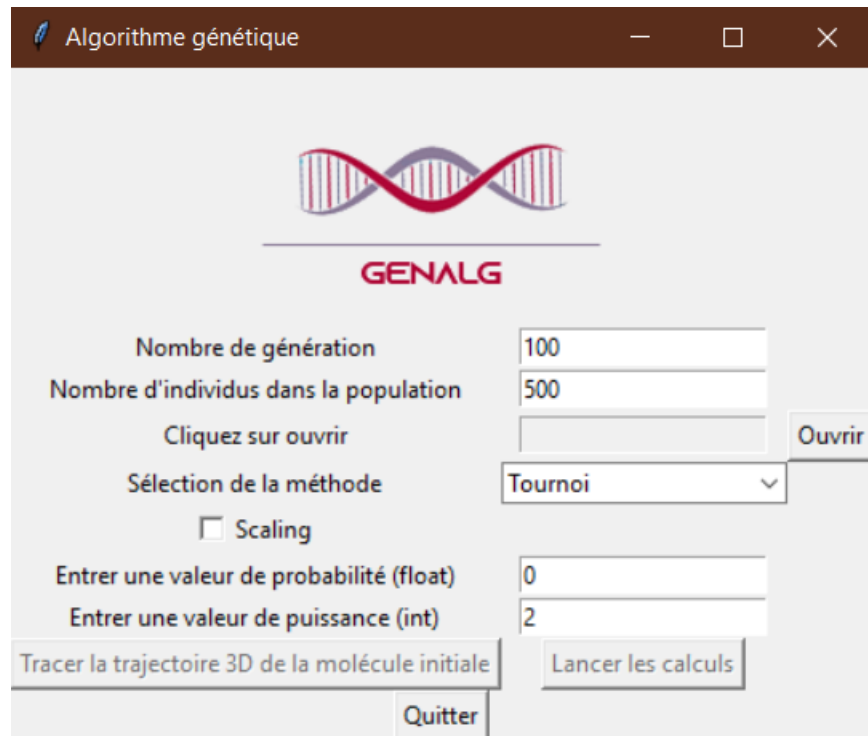


FIGURE 1 – Application au lancement

Les valeurs de nombre de génération et d'individus sont remplis par défaut, vous pouvez modifier la valeur de ces paramètres à votre guise. Ces derniers doivent être des entiers. Cliquez ensuite sur le bouton ouvrir. Un explorateur de fichier va s'ouvrir. Il vous permettra de sélectionner la chaîne de nucléotide

que vous souhaitez étudier. Vous ne pourrez sélectionner que des fichiers .fasta. Une fois votre fichier sélectionné, vous devriez obtenir une interface de la sorte :

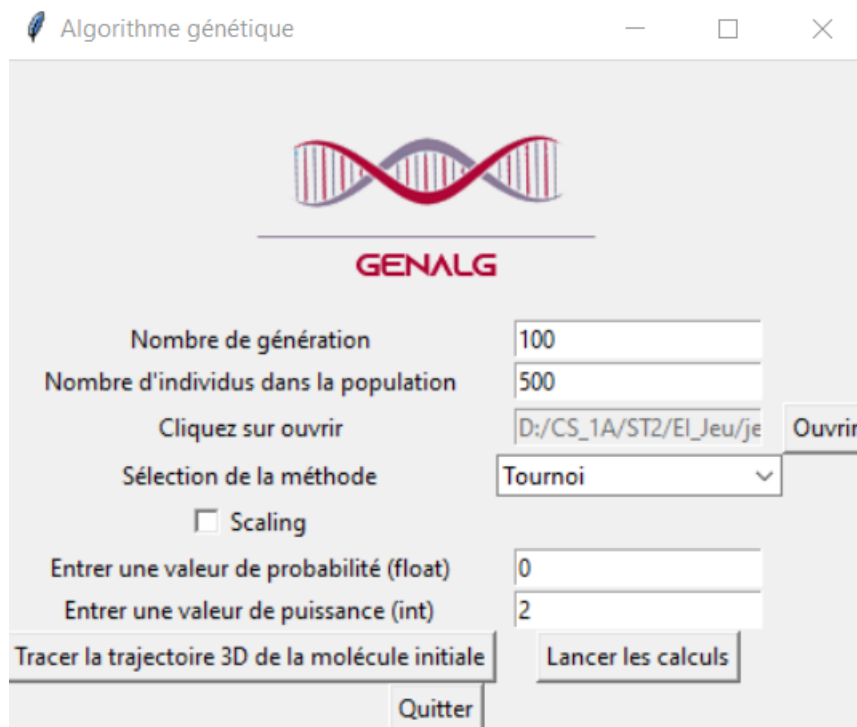


FIGURE 2 – Application avec un fichier sélectionné

Vous remarquerez qu'à présent, les boutons "Lancer les calculs" et "Tracer la trajectoire 3D de la molécule initiale" sont disponibles.

Vous pouvez dès lors sélectionner une méthode de sélection de la population parmi les choix offerts dans la liste déroulante. Vous noterez que la valeur par défaut est Tournoi. Vous pouvez également choisir d'utiliser du scaling en cochant le checkbox dédié à cet effet. Vous noterez que par défaut vous n'utilisez pas de Scaling.

Vous pouvez entrer une valeur probabilité. Cette dernière correspond à la probabilité du gain d'un individu plus faible par rapport à un autre individu dans la sélection par tournoi. Vous noterez que, par défaut, la valeur de la probabilité est de 0.

Enfin vous pouvez entrer une puissance. Cette dernière permet de donner plus ou moins d'importance à la probabilité entrée précédemment. Vous noterez que, par défaut, la valeur de la puissance est de 2.

A présent deux choix s'offre à vous :

- Lancer les calculs : ceci va vous permettre de tracer la convergence du score de la solution à mesure que les générations de la population évolue,

- et de tracer la molécule une fois les optimisations opérées sur ses angles.
- Tracer la trajectoire 3D de la molécule initiale : ceci va vous permettre de voir, en l'état actuel de la table d'angle, (i.e. les angles de base) la représentation 3D de votre molécule.
- Tant que des calculs sont en cours, vous verrez que, le bouton sur lequel vous avez appuyé reste sélectionné :

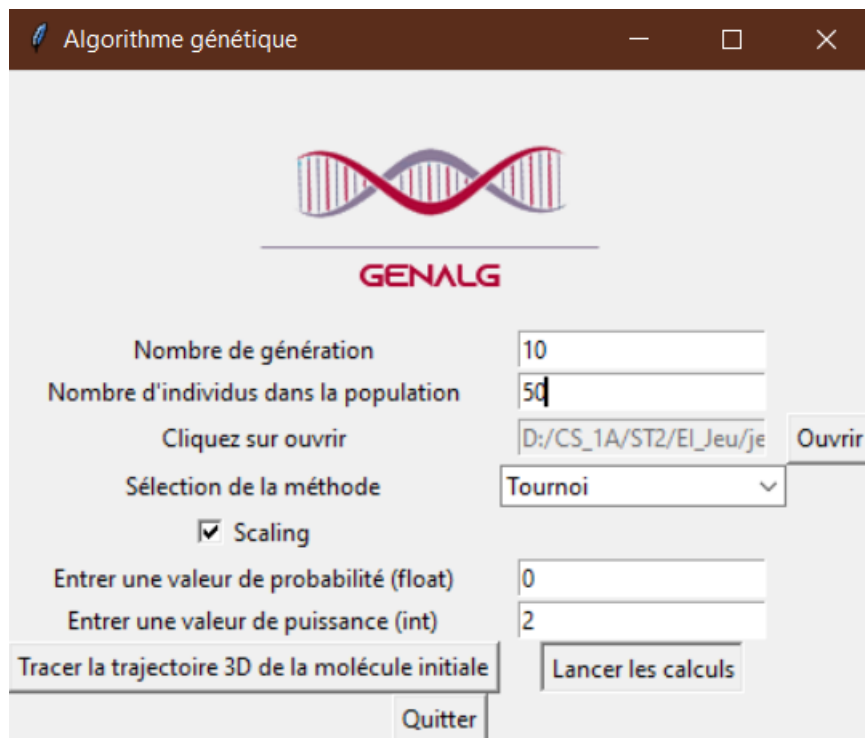


FIGURE 3 – Application avec un calcul en cours

Ainsi en fonction du choix que vous avez fait un ou des graphiques s'ouvriront. Pour relancer des calculs vous devrez les fermer. Vous trouverez les résultats dans le dossier output de votre projet.

- Si vous avez cliqué sur "Lancer les calculs" : vous obtiendrez ce genre de résultats :

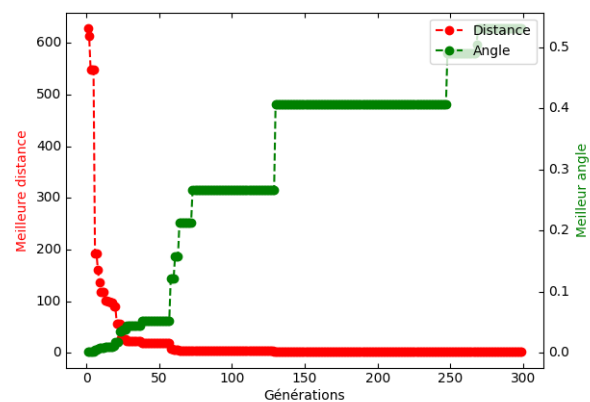


FIGURE 4 – Premier résultat de "Lancer les calculs : Convergence de la solution"

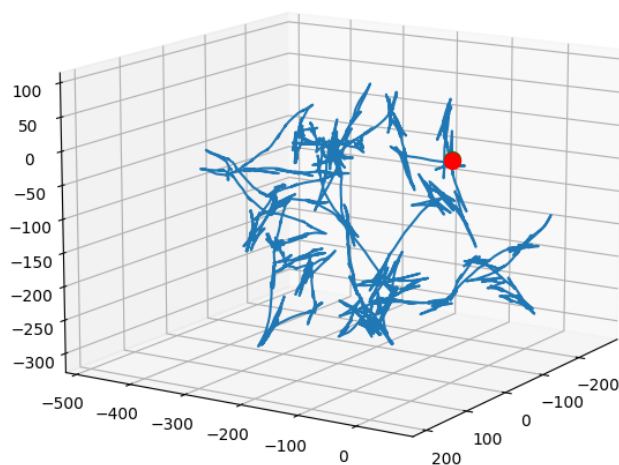


FIGURE 5 – Second résultat de "Lancer les calculs : Molécule 3D"

En vert le début de chaine, en rouge la fin de chaine et ce lorsque la table des angles a été optimisée .

- Si vous avez cliqué sur "Tracer la trajectoire 3D de la molécule initiale", vous obtiendrez ce genre de résultats :

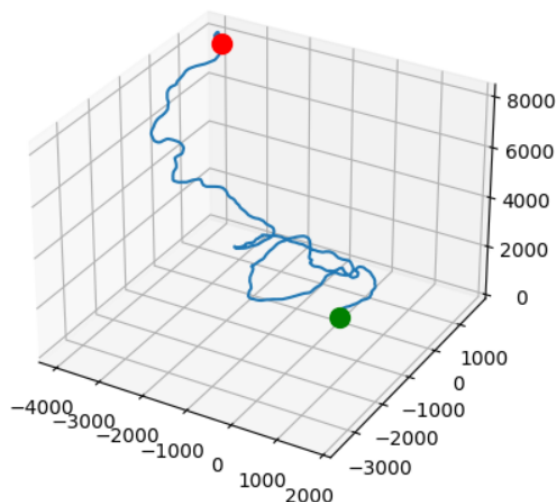


FIGURE 6 – Résultat de "Tracer de la trajectoire 3D de la molécule initiale"

En vert le début de chaîne initiale, en rouge la fin de chaîne initiale.

Une fois les calculs terminés, l'interface est à nouveau disponible. Vous pouvez relancer un calcul avec des paramètres différents ou bien quitter en appuyant sur le bouton dédié.

3 VS-Code

Nous allons ici détailler l'utilisation du programme avec VS-Code en passant par un lancement de programme via un terminal python.

Ce mode d'utilisation est certainement plus pertinent dans la mesure où il s'agit de calculs importants, ils risquent d'être lancés sur des super-calculateurs. C'est pourquoi, passer par des lignes de commandes dans le terminal peut s'avérer pertinent.

Afin de l'utiliser, placez vous dans le répertoire du programme. La syntaxe est la suivante :

```
python genalg.py mode DNA_sequence [-args]*
```

- mode ne peut prendre que 2 valeurs :
 - evolution
 - display

- DNA_sequence est le chemin vers le fichier Fasta contenant la séquence de nucléotides à cycliser

En mode "evolution", le programme fait évoluer une population selon notre algorithme génétique. Il est alors possible de spécifier les arguments suivants (sinon leur valeur par défaut est prise en compte) :

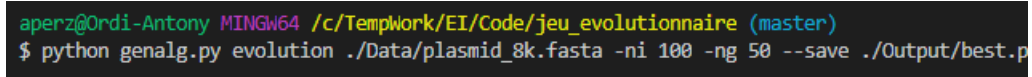
- -number_of_individuals [-ni] (le nombre d'individus dans la population initialement), par défaut : 10
- -number_of_generation [-ng] (le nombre de générations à parcourir), par défaut : 5
- -selection_method [-sel] ('Elitisme' ou 'Tournoi', la méthode de sélection des individus à utiliser par l'algorithme), par défaut : Tournoi
- -scaling (spécifier cet argument entraîne l'utilisation du Scaling lors de l'étape d'évaluation), par défaut : non spécifié
- -alpha [-a] (la probabilité de mutation d'un individu à chaque génération), par défaut : 0,59
- -luck [-l] (probabilité qui pondère la probabilité de gain du plus faible lors de la sélection par tournoi, cf details*), par défaut : 0
- -power [-p] (exposant qui modifie l'influence de luck sur la probabilité de gain du plus faible lors de la sélection par tournoi, cf details*), par défaut : 2
- -save [-s] (chemin du fichier dans lequel enregistrer le meilleur individu à l'issue de l'évolution), par défaut : pas d'enregistrement

En mode "display", le programme affiche la séquence d'ADN en 3D. Il est alors possible de spécifier les arguments suivants (sinon leur valeur par défaut est prise en compte) :

- -load (chemin vers un fichier pickle contenant un dictionnaire python de type rotation table), par défaut : le dictionnaire original est utilisé

* La probabilité de gain du plus faible lors de la sélection par tournoi est calculée selon la formule suivante : $\text{luck_prob} = \text{luck} \times (1 - (\frac{|score1 - score2|}{\max(score1, score2)})^{power})$
 où score1 et score2 sont les scores des deux individus en compétition (tournoi)
 et luck et power sont les paramètres détaillés ci-dessus.

Ci-dessous des exemples d'utilisation :

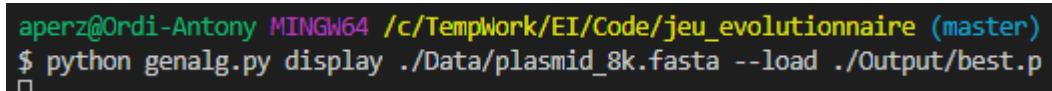


```

aperz@Ordi-Antony MINGW64 /c/Tempwork/EI/Code/jeu_evolutionnaire (master)
$ python genalg.py evolution ./Data/plasmid_8k.fasta -ni 100 -ng 50 --save ./Output/best.p

```

FIGURE 7 – Exemple 1 Utilisation du Terminal



```

aperz@Ordi-Antony MINGW64 /c/Tempwork/EI/Code/jeu_evolutionnaire (master)
$ python genalg.py display ./Data/plasmid_8k.fasta --load ./Output/best.p

```

FIGURE 8 – Exemple 2 Utilisation du terminal

4 Conclusion

Pour conclure, on remarque que l'interface permet une utilisation et expérience utilisateur plus agréable. Par ailleurs, on pourra noter que l'utilisation via un terminal est plus adapté afin de lancer ce genre de calcul sur un super-calculateur. Aussi, c'est cet usage qui semble être le plus fréquent et usuel. Si malgré ce manuel d'utilisation des questions persistes, n'hésitez pas à regarder le code qui est commenté mais aussi de nous contacter par mail.