

Cours: Algorithme génétique

Introduction

- L'algorithme génétique est un algorithme d'optimisation appartient à la classe des algorithmes évolutionnaires à base de population.
- A été introduit et développé par John Holland en 1960.
- Il se base sur des techniques inspirées de la génétique et des mécanismes d'évolution naturelle:
 - Sélections,
 - Croisements,
 - Mutations.
- C'est dans le but d'améliorer une population de solutions.
- Cet algorithme fournit des solutions aux problèmes n'ayant pas de solutions calculables en temps raisonnable de façons analytique ou algorithmique.

Introduction

Définition

- Quelques notions de la génétique utilisables par l'algorithme génétique:
 - Un gène est une unité d'information génétique transmise par un individu à son descendant. Il code généralement une des variables du problème à optimiser.
 - Un chromosome est une structure contenant une séquence finie de gènes.
 - Un individu est une solution potentielle qui peut être présenté par un ou plusieurs chromosomes.
 - Une population est un ensemble d'individus.
 - Une génération est un ensemble d'opérations effectuées afin de passer d'une population à une autre.

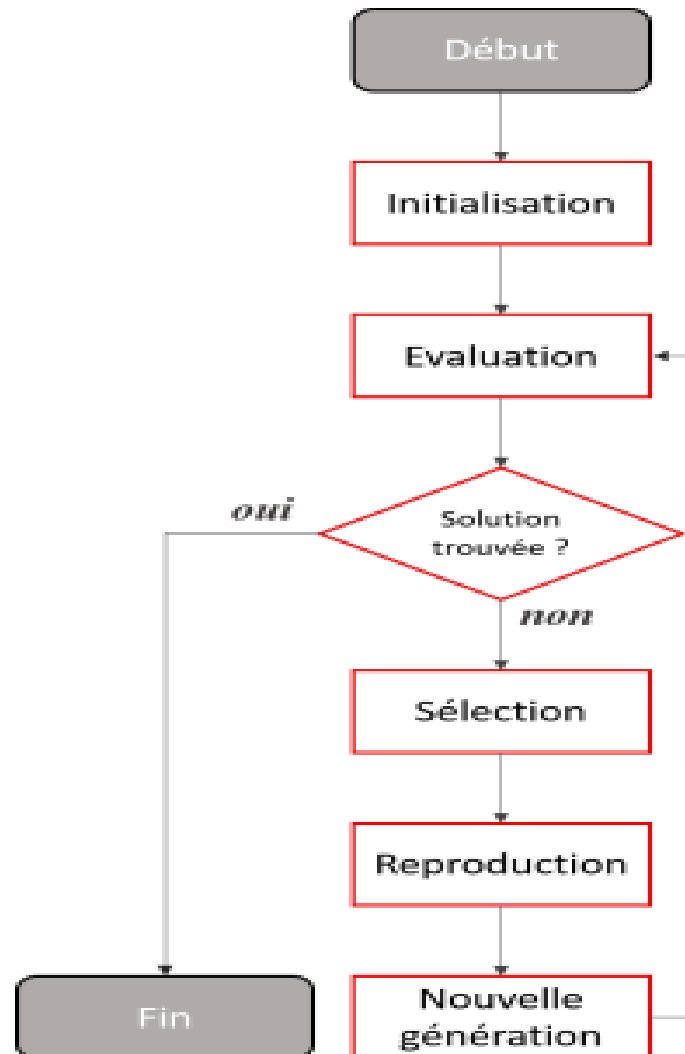
Introduction

Principe

- Le principe de fonctionnement consiste à :
 - Produire une population des individus.
 - L'adaptation de cette population à son environnement est mesurée à l'aide d'une fonction d'évaluation (fitness) habituellement liée à la fonction objectif à optimiser.
 - L'évolution de cette population c'est grâce aux opérateurs évolutionnaires : sélection, croisement et mutation.
 - Jusqu'à la satisfaction d'un test d'arrêt.

Introduction

Principe



Codage des individus

Codage des individus

- Le codage d'une solution a un rôle très important dans l'algorithme génétique.
- Il définit une représentation adéquate d'une solution d'un problème à optimiser en un ou plusieurs chromosomes.
- Ce choix du codage ne peut pas être effectué d'une façon évidente.
- Le codage du chromosome est fait généralement par trois types de codages:
 - Binaire,
 - Réel,
 - Codage de gray.
- Remarque : C'est facile de passer d'un codage à l'autre (voir Michalewicz (1992)).

Codage des individus : binaire

- C'est le codage le plus fréquemment utilisé.
- L'alphabet binaire ne comporte que deux symboles "0" et "1".
- On les nomme couramment bit.
- En faisant le parallèle avec la biologie, un bit correspond à un nucléotide et un allèle correspond à une séquence de bits.
- On parle même parfois de génotype quand on évoque la représentation binaire d'un individu et de phénotype pour désigner sa valeur réelle.F

Codage des individus : binaire

- Exemple :
 - On veut coder en binaire le nombre 29 sur 5 bits : $29_{10} = 11101_2$
 - Dans le jargon biologiste, la séquence 11101 désigne le génotype et 29 est le phénotype.
- Formule générale :
 - Décoder un chromosome x de longueur n présenté sous forme d'une séquence de bits a_i (prends soit la valeur 0 soit 1) est la suivante :

$$x = \sum_{i=1}^{n-1} a_i 2^i$$

Codage des individus : binaire

- Dans l'exemple précédent, nous avons codé un entier, si bien que chaque bit correspondait à une puissance naturelle de 2.
- Mais on pourrait coder des nombres décimaux ou négatifs.
- Par exemple, si on voulait explorer l'espace de recherche caractérisé par l'intervalle $[-3,4]$.
- En représentant les individus par une séquence de 8 bits.

Codage des individus : binaire

- Le décodage des chaînes de bits devrait se faire en plusieurs étapes :
 - Tout d'abord, il faut envisager qu'une chaîne de 8 bits permet de coder $2^8=256$ valeurs différentes.
 - L'échelle dans la représentation binaire n'est donc pas la même que dans la représentation réelle puisque notre intervalle initial $[-3, 4]$ a une longueur 7 et que le nouvel intervalle de codage est de longueur 255 et s'étend de 0 à 255.
 - Le rapport d'échelle est le suivant : $7/255=0.0274$
 - Cela signifie qu'on se déplace dans l'intervalle $[-3,4]$ par un pas de 0.0274.
 - Ensuite, l'origine de l'intervalle réel est -3.
 - En effectuant une opération de translation.

Codage des individus : binaire

- Exemple 1
 - On souhaite décoder la séquence (10011010).
 - On procède d'abord à la conversion dans le système décimal :
$$10011010_2 = 154_{10}$$
 - On tient compte de l'échelle : $154 \times 7/255 = 4.227$
 - On effectue la translation pour obtenir la valeur finale :
$$4.227 + (-3) = 1.227$$
 - La séquence (10011010) désigne le chromosome approché par la valeur 1.227 dans notre espace de recherche.

Codage des individus : binaire

- Exemple 2
 - d le nombre de décimales désiré (la précision).
 - x_{\max} et x_{\min} les bornes supérieures et inférieures de l'intervalle de variation pour x.
 - n la taille minimale de la chaîne binaire, qui est alors le plus petit entier vérifie : $|x_{\max} - x_{\min}| 10^d \leq 2^n$
 - Pour $x \in [-1.28, 1.28]$, si l'on veut une précision de 2 décimales, il faudra prendre $n=8$ pour le codage.
 - En effet, comme : $(1.28+1.28)10^2 = 256$ et que $2^8=256$,
 - On en déduit que $n=8$ est le plus petit entier qui satisfait la relation précédente.

Codage des individus : binaire

- Exemple 2
 - La chaîne binaire 00000000 correspondra alors à $x_{\min} = -1.28$ alors que 11111111 représentera $x_{\max} = 1.28$
 - Les valeurs intermédiaires sont linéairement réparties entre les deux bornes.
 - De ce fait, on peut associer une valeur de x à n'importe quelle chaîne de longueur $n = 8$, grâce à la formule suivante :

$$2^{n-1}(x - x_{\min}) / (x_{\max} - x_{\min})$$

Génération de la population initiale

Génération de la population initiale

- La résolution d'un problème par algorithme génétique commence par la génération de la population initiale.
- Les individus de cette population initiale doivent être de la même espèce que la solution potentielle : ils doivent respecter les contraintes du problème.
- Lorsqu'on n'a aucune idée de la position de la solution optimale, on peut choisir aléatoirement cette population initiale non homogène afin d'explorer une grande partie de l'espace des solutions.

Génération de la population initiale

- Si de plus on ne connaît pas le comportement de problème à résoudre, il est conseillé de répartir cette population sur tout le domaine de recherche.
- Quand on dispose d'une certaine connaissance du problème, on peut aussi générer les bons individus, plutôt que de laisser le hasard opérer.
- En ce qui concerne la taille de la population, elle n'est pas prédéfinie : c'est un paramètre laissé au libre choix de l'utilisateur.
- Mais il est bon de savoir qu'une petite population réduit le temps de calcul et peut donc évoluer plus rapidement qu'une grande population.

Génération de la population initiale

- Par conséquent, un caractère favorable présent dans une petite population pourra être transmis plus rapidement que dans une grande.
- Par contre, une population de plus grande taille augmente la diversité génétique et permet donc une plus grande possibilité d'adaptation à divers environnements.
- En tout premier lieu, on génère aléatoirement ou non quelques individus afin de composer la population initiale.
- La population va ensuite évoluer de génération en génération.

Fonction d'évaluation (d'adaptation)

Fonction d'évaluation (d'adaptation)

- La fonction d'évaluation consiste à associer à chaque individu de la population générée une valeur numérique, appelée aussi **Force ou Fitness** de l'individu.
- Cette fonction détermine la capacité de chaque individu à survivre et à se reproduire.
- Plus le fitness d'un individu est élevé, plus l'individu est "fort" et donc, plus il a de chances d'être sélectionné pour engendrer une nouvelle population.
- La fonction d'adaptation est en général une fonction à valeur réelle.

Fonction d'évaluation (d'adaptation)

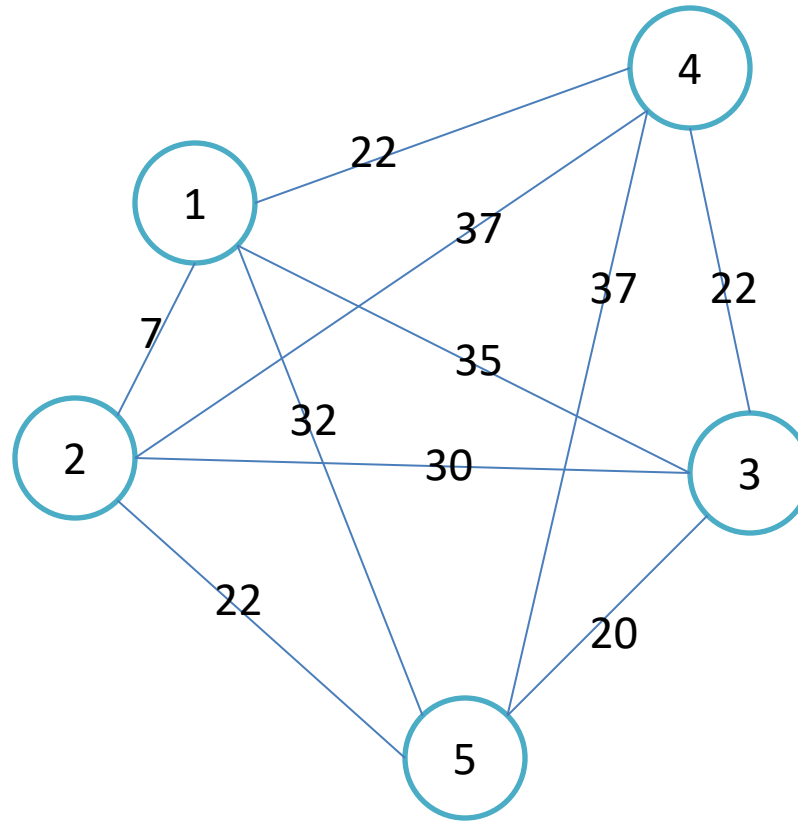
- Le processus de l'algorithme génétique cherche alors à optimiser cette fonction d'adaptation afin d'obtenir le meilleur individu, c.-à-d. la solution au problème.
- Le choix de la fonction doit donc rendre compte de la qualité réelle de l'individu.
- Si le problème consiste en la recherche d'un maximum local d'une fonction, la fonction d'adaptation peut être la fonction elle-même.

Fonction d'évaluation (d'adaptation)

- Au contraire, si on cherche un minimum local d'une fonction, la fonction d'adaptation est une modification de la fonction à minimiser de sorte que l'algorithme puisse malgré tout chercher à maximiser le fitness pour trouver le minimum.
- On voit donc que le choix de la fonction d'adaptation est crucial et dépend à la fois du problème à résoudre et de l'espace de solutions.
- De plus, la fonction d'adaptation est coûteuse en temps de calcul car elle est appelée plusieurs fois durant la procédure itérative : avant la sélection pour évaluer les individus parents potentiels, et après le croisement et la mutation, pour évaluer les enfants.

Fonction d'évaluation (d'adaptation)

- Exemple TSP.



Opérateurs de sélection

Opérateurs de Sélection

- L'opérateur de sélection est un opérateur permettant la sélection des individus d'une population (reproduction) pour la nouvelle population.
- Cet opérateur utilise le fitness comme critère de sélection et décide si les individus doivent survivre, se reproduire ou mourir.
- Il existe différentes méthodes de sélection.
- Nous en présentons ici quelques unes.
 - Sélection par roulette.
 - Sélection par tournoi.
 - Sélection par rang.
 - Sélection par élitiste.

Opérateurs de Sélection

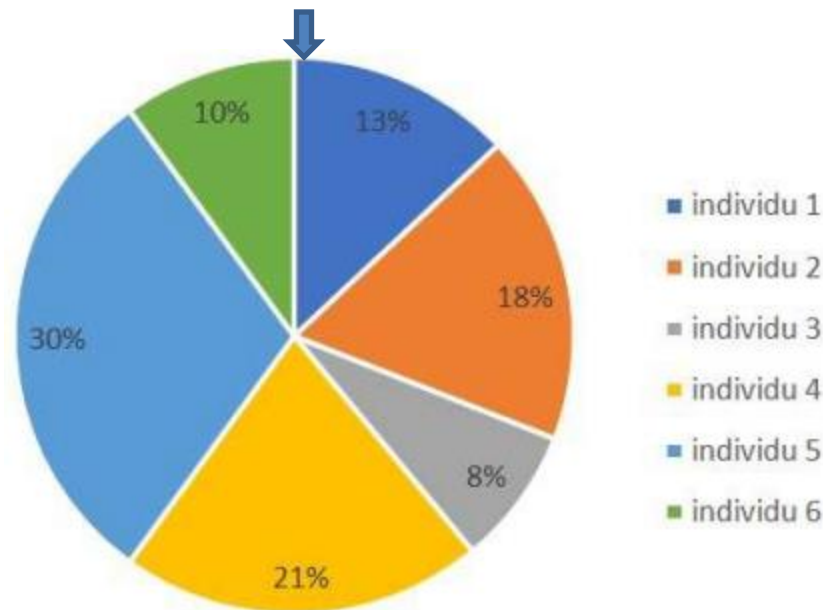
Sélection par roulette

- Il s'agit de la méthode la plus fréquemment utilisée, dont le principe d'une roue de loterie biaisée est utilisé.
- A chaque individu correspond un secteur d'une roue (la roue de loterie).
- La superficie de ce secteur est proportionnelle à la capacité (fitness) de l'individu à s'adapter.
- Plus un individu a une fonction fitness élevée plus il a de chance d'être sélectionné.
- A chaque secteur correspond une probabilité cumulée.

Opérateurs de Sélection

Sélection par roulette

- Lorsque chaque individu s'est vu attribuer un secteur, on "fait tourner" la roue jusqu'à l'arrêt.
- L'individu sélectionné est celui dont le secteur est pointé, c.-à-d. dont la plage de probabilité cumulée contient le nombre tiré au sort.



Opérateurs de Sélection

Sélection par roulette

- La population est constituée de 6 individus auxquels un secteur de roue a été attribué.
- L'individu 3 a moins de chance d'être sélectionné que l'individu 1.
- Si le hasard tire la valeur 0.5, c'est l'individu 4 qui est désigné puisque la plage de probabilité cumulée lui correspondant se situe entre 0.39 et 0.60
- On reproduit alors cette opération autant de fois que nécessaire pour obtenir tous les individus qui font désormais partie des individus parents.
- Il faut noter qu'un individu peut être sélectionné plusieurs fois.

Opérateurs de Sélection

Sélection par roulette

- Le calcul de la probabilité p_i associée à l'individu i , utilise la formule suivante :

$$p_i = \frac{f_i}{\sum_{j=1}^{N_p} f_j}$$

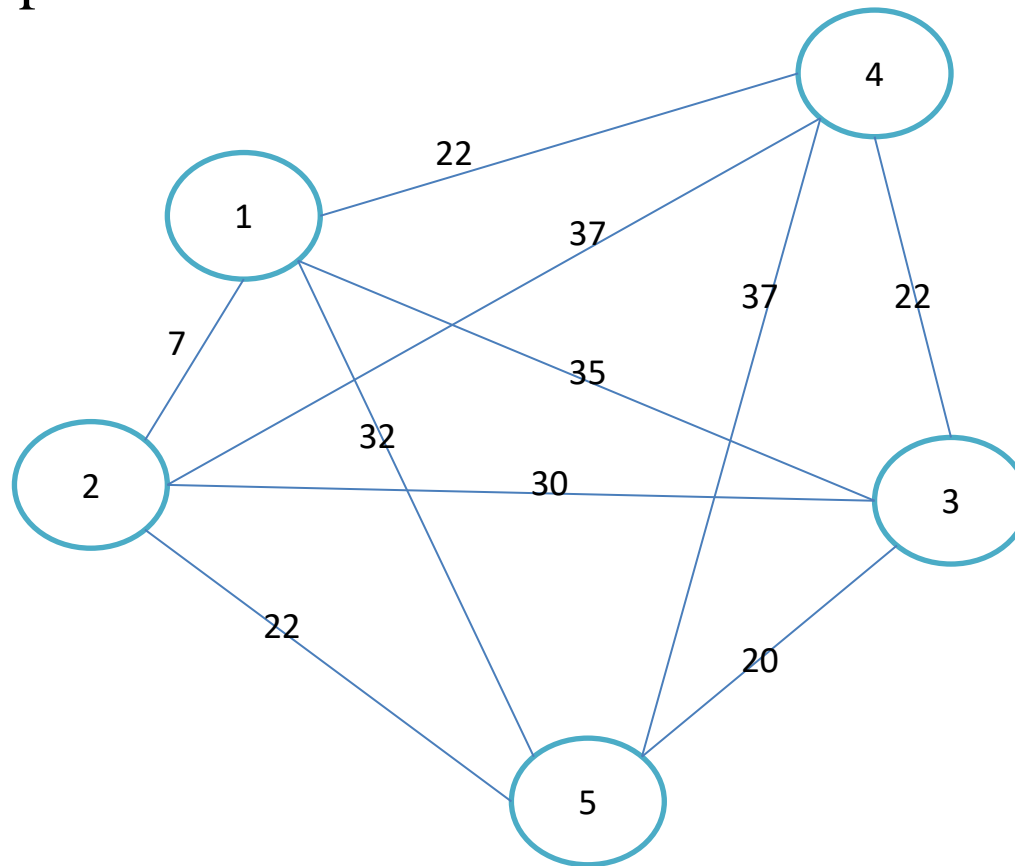
- Avec f_i représente le fitness de l'individu i et N_p est la taille de la population.
- Le nombre n_i de fois que l'individu i peut espérer être sélectionné, si la roulette tourne n fois, est donné par :

$$n_i = n \times p_i$$

Opérateurs de Sélection

Sélection par roulette

- Exemple TSP.



Opérateurs de Sélection

Sélection par roulette

- Même si la méthode de sélection la plus utilisée, elle a des inconvénients :
 - La méthode présente une forte variance : les individus avec une fonction fitness moins élevée que d'autres peuvent être sélectionnés.
 - Ce qui va totalement à l'inverse du principe des AGs
 - Il faut rappeler que le but des AGs est de sélectionner les meilleurs individus afin d'arriver au final à trouver un individu unique (le "meilleur").

Opérateurs de Sélection

Sélection par rang

- La méthode de la roulette existe également dans une forme appelée sélection par le rang.
- Elle consiste à faire une sélection en utilisant une roulette dont les secteurs sont proportionnels, non pas au fitness, mais au rang des individus.
- Pour une taille de population égale N , l'individu avec la meilleure performance reçoit le rang N , le deuxième meilleur individu à pour rang $N-1$, jusqu'au dernier individu, le moins bon, qui reçoit 1.

Opérateurs de Sélection

Sélection par roulette

- Même si la méthode de sélection la plus utilisée, elle a des inconvénients :
 - Il est possible qu'un individu avec une fonction fitness très élevée soit sélectionné plusieurs fois,
 - Ce qui diminuerait fortement la diversité de la seconde population.
 - Cette population ne saurait donc plus évoluer par la suite : on arriverait au phénomène de "convergence prématurée".
 - Le processus serait bloqué autour d'une solution finale, trouvée prématurément, qui n'est pas la solution optimale.

Opérateurs de Sélection

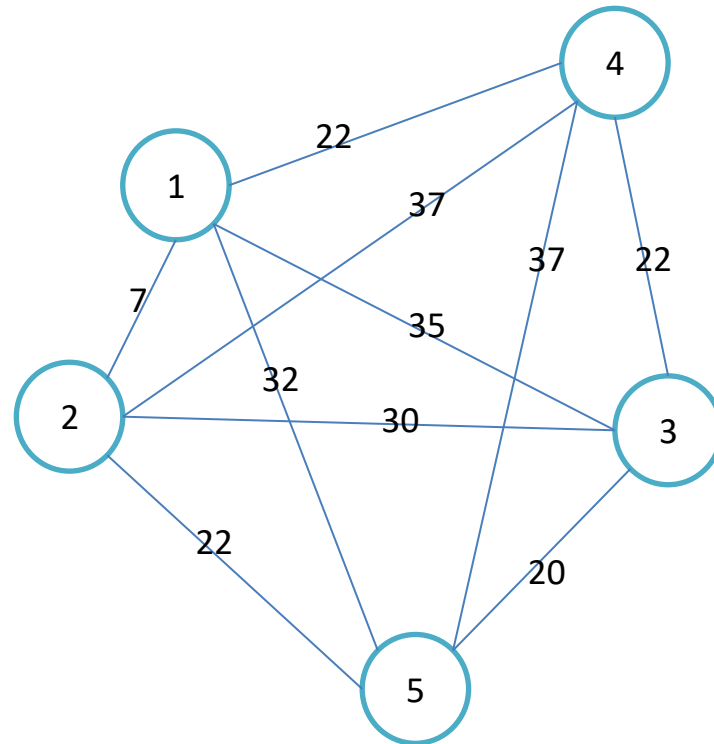
Sélection par tournoi

- C'est la méthode donnant en général les meilleurs résultats.
- Elle possède un paramètre T , la taille du tournoi.
- Le principe est le suivant : on effectue un tirage de T individus dans la population et chaque tirage donne lieu à un combat.
- L'individu dont la fonction fitness est la plus élevée, donc celui qui a remporté "le combat" est désigné comme parent.
- On reproduit ce processus autant de fois que nécessite l'obtention des N nouveaux individus de la 2^e génération.

Opérateurs de Sélection

Sélection par tournoi

- Exemple



Opérateurs de Sélection

Sélection par élitiste

- C'est la seule méthode de sélection qui soit déterministe.
- Son principe est de classer par ordre croissant les individus en fonction de leur fitness.
- Ensuite, on sélectionne un ou plusieurs individus parmi les meilleurs de ce classement qui constitueront la population de parents.
- On génère ensuite par croisement les individus enfants nécessaires à la constitution de la génération suivante.

Opérateurs de Sélection

Sélection par élitiste

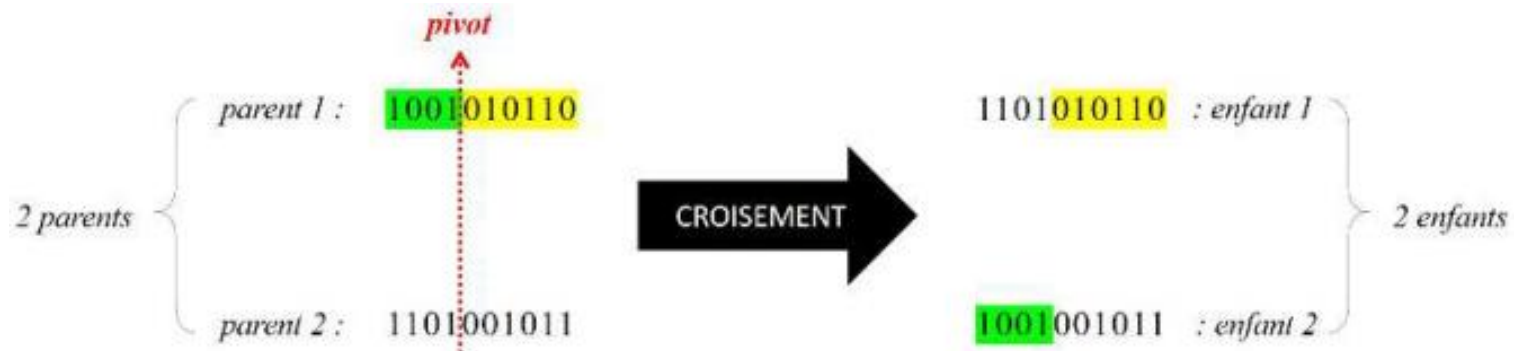
- Les individus les moins performants sont totalement éliminés de la population, et le meilleur individu est toujours sélectionné.
- On dit que cette sélection est élitiste.
- Cette méthode présente une convergence fortement prématurée car on a une variance et une diversité presque nulles : les individus les moins bons n'ont aucune chance de survivre.
- Elle permet cependant de ne pas perdre des individus avec une fonction fitness élevée donc, solution possible du problème.

Opérateurs de croisement

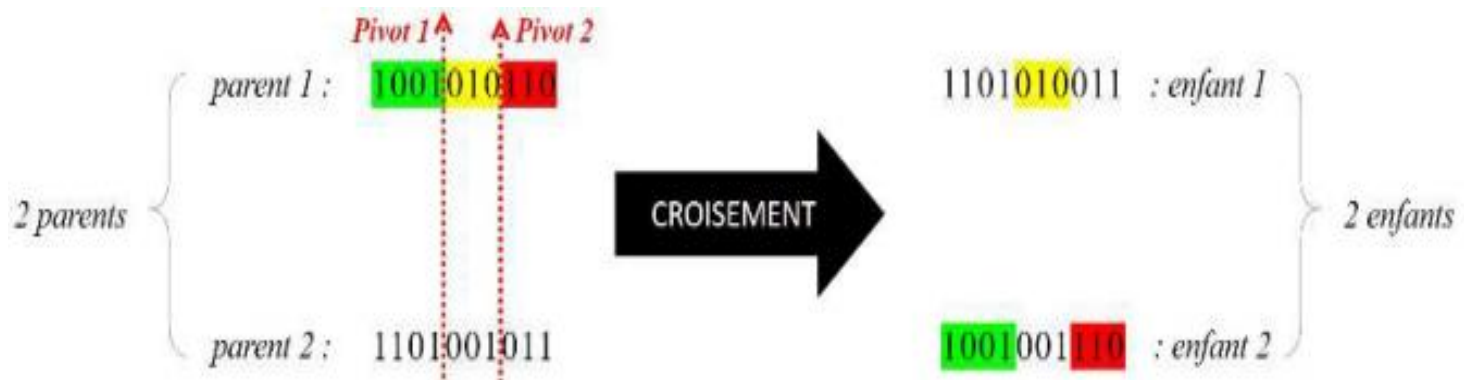
Opérateurs de croisement

- Le but de croisement est d'explorer le mieux l'espace des solutions.
- A partir de deux individus, on obtient deux nouveaux individus (enfants) qui héritent de certaines caractéristiques de leurs parents.
- Comment?
 - Un point de croisement est choisi aléatoirement sur les deux chromosomes.
 - Les parties des parents de part et d'autre du point de croisement sont intervertit pour former deux nouveaux individus.

Opérateurs de croisement



- Ils existes d'autres types de croisement:
 - Le croisement à 2 points,
 - Le croisement uniforme,
 -



Opérateurs de croisement

Taux d'opérateur de croisement

- Le taux de croisement P_c détermine la proportion des individus qui vont participer à l'opérateur de croisement.
- Plus cette valeur est élevée, plus de nouveaux individus sont introduites dans la nouvelle génération.
- Par contre, si ce taux est trop faible, la population n'évolue pas assez vite.
- Généralement, P_c varie entre 0.20 et 0.80.

Opérateurs de croisement

Taux d'opérateur de croisement

- Exemple
- Soit f une fonction à minimiser définie sur $[-4,1]$ par :
$$f(x)=x^4+5x^3+4x^2-4x+1$$
- On choisi d'encoder x en 4 bits.
- On sélectionne deux individus de la population courante.
- Effectuer l'opérateur de croisement sur les parents.

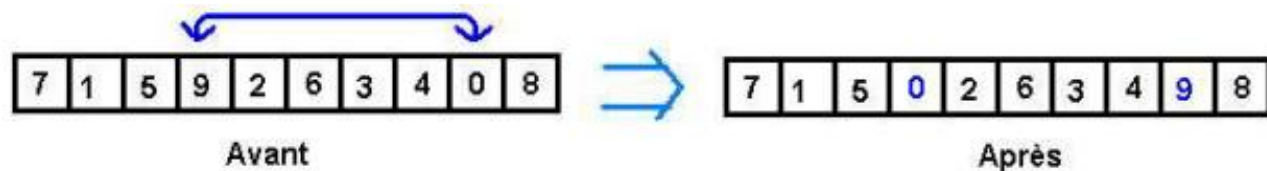
Opérateur de mutation

Opérateur de mutation

- Cet opérateur consiste à modifier un ou plusieurs allèles d'un gène avec une probabilité souvent assez faible (de l'ordre de 0.01 à 0.1).
- Il consiste donc à modifier aléatoirement les caractéristiques d'une solution.
- Il joue le rôle élément perturbateur ou diversificateur.
- Il permet d'éviter le maximum possible une convergence de l'algorithme vers un optimum local.

Opérateur de mutation

- Si l'on veut muter un individu qui est codé en binaire, il suffit de modifier la valeur d'un des bits comme le montre la figure suivante.



- L'opérateur de mutation possède plusieurs qualités :
 - Il permet une plus grande diversité de solutions.
 - Il permet d'éviter une dérive génétique, c.-à-d. que l'un ou plusieurs des gènes favorisés par le hasard ne se répandent tandis que les autres s'effacent de plus en plus.
 - Dans ce cas défavorable, les allèles favorisés sont alors présents dans la plupart des solutions au même endroit

Opérateur de mutation

Qualités

- Il permet d'éviter les risques d'une convergence prématurée.
- Lorsqu'une convergence se fait, tous les individus deviennent pratiquement identiques mais ne sont pas de réelles solutions de l'algorithme génétique. Le croisement n'a donc plus aucun effet sur la génération suivante. C'est à ce moment-là que l'opérateur de mutation devient primordial car il va permettre de modifier un individu aléatoirement, ce qui va décoincer la situation.

Opérateur de mutation

Qualités

- Il permet d'atteindre une propriété s'appelant l'ergodicité
Cette propriété garantit que chaque valeur de l'espace de recherche peut être atteinte.
- En effet, étant donné que l'opérateur de mutation intervient aléatoirement, on a la certitude que toutes les valeurs de l'espace de recherche peuvent apparaître dans les diverses générations. On est donc certain de pouvoir trouver une réelle solution au problème.

Opérateur de mutation

Taux d'opérateur de mutation

- L'opérateur de mutation est appliqué avec une probabilité P_m .
- probabilité P_m varie en général entre 0.02 et 0.2.
- Si le taux de probabilité est grand alors la recherche devient purement aléatoire, la population est diversifiée et l'algorithme génétique perd de son efficacité.
- Si au contraire ce taux est faible, la population est moins diversifiée et en plus il y a risque de stagnation.

Critères d'arrêt

Critères d'arrêt

- Plusieurs test d'arrêts pour l'algorithme génétique sont possibles :
 - Nombre de générations fixé,
 - Convergence de la population,
 - Population n'évoluant plus suffisamment,
 - Etc.