

ALGORITHMES GÉNÉTIQUES

Laurent Simon

`simon@lri.fr`

`im:lsimon@jabber.fr`

LRI - CNRS
Université Orsay Paris 11,
91405 Orsay Cedex

1^{er} semestre 2006

SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTATS

❶ PRINCIPES DES ALGORITHMES GÉNÉTIQUES

❷ RAPPELS (VACCINS?) DE BIOLOGIE

❸ ALGORITHMES GÉNÉTIQUES

❹ SÉLECTION

❺ ENCODAGE DES CHROMOSOMES

❻ REPRODUCTION DES INDIVIDUS

SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

1 PRINCIPES DES ALGORITHMES GÉNÉTIQUES

2 Rappels (vaccins?) de biologie

3 Algorithmes Génétiques

4 Sélection

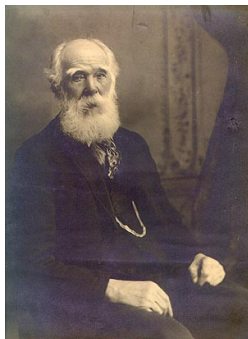
5 Encodage des chromosomes

6 Reproduction des individus

PRINCIPES DES ALGOS GÉNÉTIQUES

FAMILLE DES ALGORITHMES ÉVOLUTIONNAIRES

On reprend l'idée de l'évolution selon Darwin.



IDÉE DE DARWIN

Seuls les plus forts subsistent dans une population.

DARWIN EN I.A.

On prend une *population* (les solutions potentielles au problème). On *sélectionne* les meilleurs pour qu'ils se *reproduisent*. On élimine les plus faibles. On recommence.

PETITE HISTOIRE DES ALGOS GÉNÉTIQUES

RETOUR AUX ANNÉES 1960

L'idée de calculs *évolutionnaires* remonte à 1960 (I. Rechenberg). En 1975, le livre « *Adaption in Natural and Artificial Systems* » de John Holland sort. Les *Algorithmes génétiques* sont nés.

DES PROGRAMMES GÉNÉTIQUES ?

En 1992, John Koza utilise les algorithmes génétiques pour faire évoluer des programmes permettant de résoudre certaines tâches. La *Programmation Génétique* était née.

SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

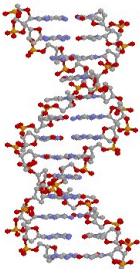
- 1 Principes des Algorithmes Génétiques
- 2 RAPPELS (VACCINS?) DE BIOLOGIE**
- 3 Algorithmes Génétiques
- 4 Sélection
- 5 Encodage des chromosomes
- 6 Reproduction des individus

RAPPELS DE BIOLOGIE

CELLULES ET CHROMOSOMES

QU'EST CE QUE LE VIVANT ?

Tous les organismes vivants sont constitués de *Cellules*. Toutes les cellules d'un organisme contiennent le même ensemble de *Chromosomes*. Les Chromosomes sont constitués de *Gènes*, des parties d'ADN. Un gène encode une protéine particulière.



LE RÔLE DES GÈNES

On dit généralement qu'un gène encode un *trait* de l'individu (par exemple la couleur des yeux). Les différents traits possibles sont appelés *allèles*. Un gène a sa propre position sur le chromosome, appelée *locus*.

GÉNOME : ENSEMBLE DU MATÉRIEL GÉNÉTIQUE

Il représente le substrat pour le développement de l'individu (couleur des yeux, intelligence, ...).

RAPPELS DE BIOLOGIE

REPRODUCTION



REPRODUCTION

Pendant la *reproduction*, les gènes des deux parents se combinent pour former de nouveaux chromosomes. L'ADN du fils peut cependant connaître des *mutations génétiques* (erreurs lors de la copie des gènes des parents). Ces mutations vont permettre de diversifier la population (création de nouvelles espèces...).

L'adaptation d'un organisme est mesuré d'après son succès dans la vie (ou sa survie).

SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

- 1 Principes des Algorithmes Génétiques
- 2 Rappels (vaccins?) de biologie
- 3 ALGORITHMES GÉNÉTIQUES**
- 4 Sélection
- 5 Encodage des chromosomes
- 6 Reproduction des individus

DE LA BIOLOGIE VERS L'I.A.

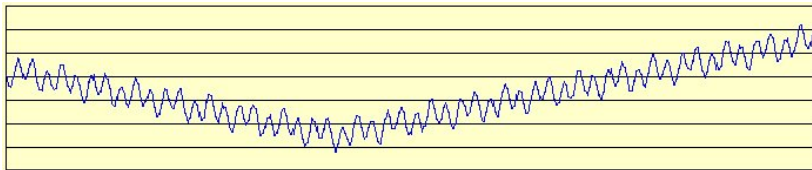
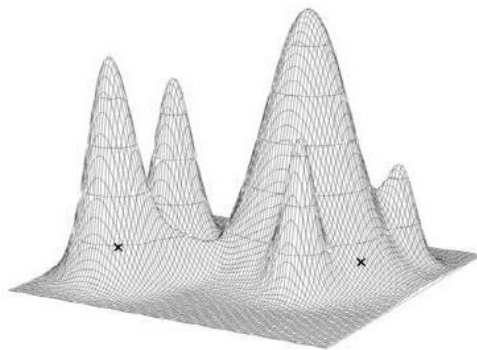
On peut faire un rapide parallèle entre les éléments de la biologie et les éléments utiles aux algorithmes génétiques :

Biologie	Algos Génétiques
Chromosome	Solution potentielle
Adaptation	<i>Fitness</i> (valeur heuristique)
Reproduction	Combinaison de deux solutions
Mutation	Remplacement par un voisin

SCHÉMA GÉNÉRAL D'UN ALGORITHME GÉNÉTIQUE

- 1: Générer une population aléatoire de n chromosomes
- 2: Tant que Continue Faire
- 3: Calculer la fonction *fitness* $f(x)$, pour tout chromosome x
- 4: Tant que NouvellePopulationARemplir Faire
- 5: Sélectionner deux chromosomes parents
- 6: En faire un nouveau chromosome
- 7: Lui imposer des mutations éventuelles
- 8: Ajoute le nouveau chromosome à la nouvelle population
- 9: Fin Tant que
- 10: Remplace l'ancienne population par la nouvelle
- 11: Fin Tant que

REPRÉSENTATION DE L'ESPACE DE RECHERCHE



SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

- 1 Principes des Algorithmes Génétiques
- 2 Rappels (vaccins?) de biologie
- 3 Algorithmes Génétiques
- 4 SÉLECTION**
- 5 Encodage des chromosomes
- 6 Reproduction des individus

SÉLECTIONNER LES PLUS FORTS

TIRAGE SUR UNE ROUE *biaisée*



LA ROUE DE LA SÉLECTION

Principes : tout individu peut se reproduire, mais les plus forts doivent avoir plus de chance.

FONCTIONNEMENT DE LA ROUE

- 1 Calculer la somme des fitness de tous les chromosomes $\sum f(x)$
- 2 Générer un nombre aléatoire a dans $[0, \sum f(x)]$
- 3 Parcourir la population ordonnée en sommant leur fitness. Dès que la somme est plus grande que a , le chromosome considéré est retourné.

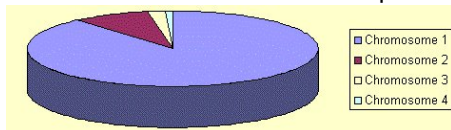
VARIANTE

Si la fonction fitness est *très grande* pour les bons individus et *très petite* pour les mauvais, cela ne leur laissera que très peu de chance d'être sélectionné.

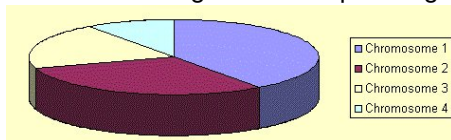
Idée : On utilise le rang des individus plutôt que la fonction fitness pour biaiser la roue.

ILLUSTRATION DE LA ROUE DES CHOIX

Avec une fonction fitness très disparate :



Avec une stratégie de choix par rang :



UNE AUTRE MÉTHODE DE CHOIX : LE TOURNOI

On choisit deux individus et celui qui a la plus grande fonction de fitness remporte le combat (ou a de plus grandes chances de remporter le combat). On effectue un tournoi entre tous les individus. Ceux qui remportent le tournois peuvent se reproduire.



SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

- 1 Principes des Algorithmes Génétiques
- 2 Rappels (vaccins?) de biologie
- 3 Algorithmes Génétiques
- 4 Sélection
- 5 ENCODAGE DES CHROMOSOMES**
- 6 Reproduction des individus

ENCODAGE DES CHROMOSOMES

Suivant la nature du problème, on devra utiliser différents types de représentations pour les chromosomes.

REPRÉSENTATION EN VECTEUR DE BITS

Un chromosome est codé sur n bits.

A : 10111000110110101000100010100011001

B : 01010001001110010101001110101010000

Exemple : Problème du sac à dos. On désire placer des objets dans un sac à dos. Chaque objet a un poids et une valeur. On désire maximiser la valeur totale du sac à dos. Le bit i représente la présence (ou non) de l'objet dans le sac à dos. C'est la représentation la plus souvent utilisée. Elle est simple, intuitive, et efficace.

ENCODAGE DES CHROMOSOMES, II

ENCODAGE PAR PERMUTATION

Chaque chromosome est une suite de nombre qui représente une position dans une séquence.

A : 1 5 3 2 6 4 7 9 8

B : 8 5 6 7 2 3 1 4 9

Exemple : Problème du voyageur de commerce. La séquence représente l'ordre de parcours des villes.

ENCODAGE PAR LISTE DE VALEURS

Chaque chromosome est une suite de valeurs possibles.

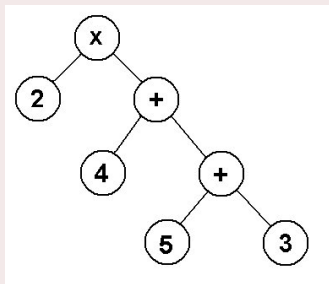
(gauche) (droite) (droite) (haut) (gauche) (gauche)
1.543232 1.6764734 5.434343 3.323232 8.473843

Exemple : Instructions de directions pour un chemin dans un labyrinthe, Coefficients d'un polynôme, ...

ENCODAGE DES CHROMOSOMES, III

ENCODAGE PAR ARBRE

Chaque chromosome est un objet représenté par un arbre.



Exemple : Programme en LISP, Instructions d'un robot avec observation, Expressions mathématiques.

SECOND COURS SUR LA RECHERCHE DANS LES GRAPHS D'ÉTAT

- 1 Principes des Algorithmes Génétiques
- 2 Rappels (vaccins?) de biologie
- 3 Algorithmes Génétiques
- 4 Sélection
- 5 Encodage des chromosomes
- 6 REPRODUCTION DES INDIVIDUS**

COURS DE REPRODUCTION

SUR DES VECTEURS DE BITS

On définit généralement la notion de point de croisement sur le chromosome (*crossover*), choisi aléatoirement.

EXEMPLE AVEC UN SEUL *crossover*

A	110010101010	0101010
B	010101001011	1100101
fil 1	110010101010	1100101
fil 2	010101001011	0101010

VARIANTES

- On peut définir plusieurs points de croisements
- On peut construire les fils en sélectionnant aléatoirement chaque bit chez l'un ou l'autre des deux parents
- On peut faire une opération logique bit à bit entre les deux parents

COURS DE REPRODUCTION

SUR DES PERMUTATIONS ET DES VALEURS

ENCODAGE PAR PERMUTATION

On choisit un point de crossover. On fait naître un seul fils tel que : tous les nombres à gauche du crossover sont ceux du premier père, et tous les nombres à droite sont ceux du second père qui ne sont pas déjà à gauche (dans le même ordre).

A	1	2	3	4	5	6	7	8	9
B	4	5	3	6	8	9	7	2	1
fils	1	2	3	4	5	6	8	9	7

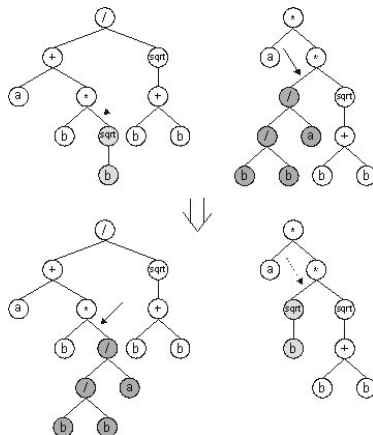
ENCODAGE PAR LISTE DE VALEURS

On peut utiliser toutes les techniques des chaînes de bits.

COURS DE REPRODUCTION

SUR DES ARBRES

On choisit chez les deux parents un sous-arbre. On les échange.



DES OGM

Les mutations génétiques sont importantes car elles permettent de s'échapper de minima locaux et d'explorer potentiellement tout l'espace de recherche. Suivant les problèmes, on effectuera une ou plusieurs mutations à chaque nouvelle génération.

- Représentation en vecteur de bit : changer la valeur d'un bit ;
- Permutation : échanger la position de deux entiers ;
- Liste de valeur : une valeur est changée mais doucement ;
- Arbre : une feuille ou un noeud est changé

QUELQUES PARAMÈTRES

TAUX DE MUTATION

Le taux doit être particulièrement bas. Entre 0.5% et 1% du chromosome.

TAILLE DE LA POPULATION

De manière très surprenante, de petites populations sont suffisantes. Généralement, entre 20 et 30. Si la taille du chromosome est importante, il faut augmenter la population (100 individus est généralement un maximum dans tous les cas).

REMPACEMENT DE LA POPULATION

Il est généralement préférable de garder les parents dans la nouvelle population (différence avec la vraie biologie), sans quoi les meilleurs éléments peuvent être rapidement perdus. On peut aussi garder les n meilleurs chromosomes d'une génération à l'autre.