

Algorithmes Génétiques - EISTI - ING 2

Yannick Le Nir
yannick.lenir@eisti.fr

Ecole Internationale des Sciences du Traitement de l'Information

Présentation
Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application :
coloration de
graphes

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

IA

- ▶ Création d'une IA dès les années 50
- ▶ Programmation auto-adaptative
- ▶ Echec initial puis reprise dans les années 80 :
 - ▶ Réseaux de neurones
 - ▶ Apprentissage automatique
 - ▶ Evolution artificielle

Pourquoi l'évolution

Mécanismes bien adaptés :

- ▶ aux problèmes nécessitant des ressources de calcul importantes
- ▶ aux programmes qui doivent être adaptatif (trajectoire robot)
- ▶ aux programmes trop complexes pour être programmés à la main

Principes de l'évolution

- ▶ Recherche parmi un nombre de possibilité gigantesque
- ▶ Conception de solutions innovantes à des problèmes complexes
- ▶ Méthodes de recherche massivement parallèles

Concept

Les individus soumis à l'évolution sont des solutions plus ou moins performantes à un problème donné.

A chaque génération, une succession d'opérations est appliquée pour engendrer une nouvelle population.

Les individus résultants sont les enfants. Les individus utilisés par un opérateur sont les parents (transmission du patrimoine génétique).

Opérateurs de sélection

A chaque génération, les individus se reproduisent, survivent ou disparaissent sous l'action des opérateurs de sélection :

- ▶ La sélection pour la reproduction (combien de fois un individu sera reproduit en une génération)
- ▶ La sélection pour le remplacement (disparition d'individus)

Sélection

Darwin : selection des meilleurs individus :

- ▶ Valeur de performance associée à chaque individu (fitness function) associée à l'objectif.
- ▶ Transmission de ses meilleurs caractéristiques à ses descendants.

Opérateurs de variation

Améliorer la qualité des solutions :

- ▶ par mutation : transformation d'un individu en un autre
- ▶ par croisement (crossover) : engendre un ou plusieurs enfants à partir de deux parents (reproduction sexuée)

Algorithme

1. Pour reproduction : sélection des parents parmi une population de μ individus pour engendrer λ enfants
2. Croisement et mutation à partir des individus sélectionnés pour engendrer λ enfants
3. Evaluation et performance des enfants
4. Sélection pour la survie de μ individus parmi les λ enfants et μ parents ou uniquement parmi les λ enfants selon les paramètres choisis, afin de générer la population suivante.

Définitions

- ▶ Individu/chromosome/séquence : une solution potentielle au problème
- ▶ Population : un ensemble de chromosomes ou de points de l'espace de recherche
- ▶ Environnement : l'espace de recherche
- ▶ Fonction de fitness : fonction (positive) que l'on cherche à maximiser

Représentation des individus

Définition

Nous appelons séquence (chromosome, individu) A de longueur $l(A)$ une suite $A = \{a_1, a_2, \dots, a_l\}$ avec $\forall i \in [1, l], a_i \in V$ où V est un codage.

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Codage des données d'un algorithme génétique

Codage des solutions (chromosomes)

Structure de données codant les gènes :

- ▶ codage binaire : chaîne de bits :
codage historique parfois peu naturel pour un problème donné.
- ▶ codage à caractères multiples :
codage linéaire constitués de différents types de caractères
- ▶ codage arborescent :
codage autorisant des solutions de taille infinie

Exemple de codage binaire

Codage classique

Sur 4 bits,

- ▶ on code la valeur 1 par 0001
- ▶ on code la valeur 2 par 0010
- ▶ ces deux valeurs sont proches mais leur codage les éloigne (distance de Huffman de 2)

Codage de Gray

On garde toujours entre 2 valeurs consécutives une distance de 1 :

- ▶ valeur 1 : 0001
- ▶ valeur 2 : 0011
- ▶ valeur 3 : 0010
- ▶ valeur 4 : 0110
- ▶ etc ...

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Fonction de fitness

Nous appelons fitness d'une séquence A toute valeur positive que nous noterons $f(A)$ (fonction de fitness).

En général, il s'agit d'une valeur réelle caractérisant l'adaptation de la solution à son environnement. Son calcul dépend fortement du codage choisi pour représenter les solutions.

Présentation
Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application :
coloration de
graphes

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Sélection proportionnelle

Utilisée pour la reproduction : Nombre espéré de sélections λ_i d'un individu i proportionnel à sa performance f_i :

$$\lambda_i = \frac{\lambda}{\sum_{j=1}^{\mu} f_j} f_i$$

Inconvénient :

- ▶ un individu qui n'est pas le meilleur peut dominer la sélection
- ▶ peut être efficace à la fin lorsque les individus se ressemblent

Roulette

Choix aléatoire : individu a une probabilité d'être choisi proportionnelle à sa qualité

Sélection par tournoi

Tournoi déterministes

Sélection aléatoire de k individus, sélection du meilleur

- ▶ autant de tournois que d'individus sélectionné
- ▶ tirage avec ou sans remise
- ▶ pression de sélection ajustée par le paramètre k

Tournoi binaire stochastique

Sur 2 individus en compétition, le meilleur gagne avec une probabilité comprise entre 0,5 et 1.

Pression de sélection

- ▶ Algorithme dans un optimum local.
- ▶ Pression sélective quasi-inexistante : dérive génétique.
- ▶ Ajustements dynamiques de la fonction de performance

Sélections pour le remplacement

Remplacement générationnel

Génération $g + 1$ constituée uniquement des enfants $\mu = \lambda$.

En général, sélection des meilleurs, $\lambda > \mu$

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Exemple avec un codage binaire

Création d'un nouvel individu à partir de 2 individus par échange de chromosomes entre individus

- ▶ Sélection des deux individus parents
- ▶ Détermination de 1 ou plusieurs points de croisement (entre 1 et $\text{longueur}(\text{individu}) - 1$)
- ▶ Selon une valeur probabiliste, échange des segments finaux des 2 parents (dans le cas d'un seul point de croisement) afin de générer 2 nouveaux individus (enfants)

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Opérateurs de mutation

Exemple avec un codage binaire

Tirage aléatoire d'un (ou plusieurs) gène du chromosome et remplacement par son inverse selon une certaine probabilité.

Rôle dans l'algorithme

- ▶ Permet d'effectuer une recherche locale
- ▶ Permet de sortir d'un optimum local
- ▶ Permet d'atteindre potentiellement tous les points de l'espace de recherche

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Paramétrisation de l'algorithme génétique

- ▶ Taille de la population N : si N est trop grand le temps de calcul peut être trop long, s'il est trop petit il y a un risque de convergence trop rapide vers un mauvais chromosome.
- ▶ Probabilité de croisement p_c . Elle dépend de la forme de la fonction de fitness. Plus elle est élevée, plus la population subit des changements importants (en général entre 0,5 et 0,9)
- ▶ Probabilité de mutation p_m . Taux généralement faible afin d'éviter de converger vers une solution sous-optimale.

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Algorithme

- ▶ Initialisation : une population initiale de N chromosomes est tirée aléatoirement.
- ▶ Evaluation : chaque chromosome est décodé, puis évalué.
- ▶ Sélection : création d'une nouvelle population de N chromosomes par la méthode de sélection appropriée.
- ▶ Reproduction : possibilité de croisement et mutation au sein de la nouvelle population.
- ▶ Retour à la phase d'évaluation jusqu'à l'arrêt de l'algorithme.

Présentation
Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application :
coloration de
graphes

Démonstration

Algorithmes
Génétiques - EISTI
- ING 2

Yannick Le Nir
yannick.lenir@eisti.fr

Présentation
Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application :
coloration de
graphes

Sommaire

Présentation Générale

Codage

Fitness

Sélection

Croisement

Mutation

Autres paramètres

Cycle de Base

Démonstration

Application : coloration de graphes

Problème de coloration de graphe

Énoncé

A partir d'un graphe non orienté non valué, il s'agit de déterminer le nombre minimal de couleurs nécessaire afin de colorier l'ensemble des sommets de telle sorte que 2 sommets adjacents n'ai jamais la même couleur.

Problématique

A partir d'un graphe nécessitant plus de 3 couleurs, ce problème n'est pas solvable efficacement de manière exacte, i.e via un algorithme polynomial. La seule solution exacte consisterait à essayer toutes les possibilités, ce qui mène à un algorithme exponentiel.

Solution

Utiliser une méthode approchée (heuristique), par exemple un algorithme génétique dont chaque chromosome va coder une solution.