

Algorithmes génétiques - Introduction 2

Historique

- A l'origine: modélisation de populations se reproduisant selon un mode asexué.

Fogel L.J., Owens A. J., Walsh M.J. (1966), Artificial Intelligence through Simulated Evolution, John Wiley.
- Holland, de l'Université du Michigan, a étendu le modèle à des populations où des individus échangent du matériel génétique:

Bagley J.D. (1967), "The Behavior of Adaptive Systems which employ Genetic and Correlation Algorithms", Doctoral Dissertation, University of Michigan, *Dissertation Abstracts International* 28(12), 5106B.

DeJong K.A. (1975), "An Analysis of the Behavior of a Class of Genetic Adaptive Systems", Doctoral Dissertation, University of Michigan, *Dissertation Abstracts International* 36(10), 5140B.

John Holland (1975), "Adaptation in Natural and Artificial Systems", The University of Michigan Press, Ann Arbor, réimpression par MIT Press (1992).
- Tâches d'apprentissage
- Optimisation de fonctions réelles complexes (multimodalité, discontinuité, etc..)
- Optimisation combinatoire

Algorithmes génétiques

Idée de base

Soit la population de 8 chromosomes suivante:

Chromosome 1: 110001	90	[0 , 90]
Chromosome 2: 010101	10]90 , 100]
Chromosome 3: 111001	100]100, 200]
Chromosome 4: 100101	5]200, 205]
Chromosome 5: 000011	95]205, 300]
Chromosome 6: 010111	90]300, 390]
Chromosome 7: 001100	5]390, 395]
Chromosome 8: 101010	5]395, 400]

Observation: les chromosomes avec des "1" occupant les deux premières positions, ou des "1" occupant les deux dernières positions sont plus performants.

Hypothèse: un chromosome plus performant peut être obtenu en juxtaposant deux "1" dans les deux premières positions, et deux "1" dans les deux dernières positions.
Exemple: le chromosome 110011.

- Il est donc important de découvrir les caractéristiques communes des chromosomes performants (schémas), de façon à propager ces caractéristiques au sein de la population.
- C'est ce qu'un algorithme génétique tente de faire.

N.B. Dans la suite, on suppose toujours que les mesures de performance des chromosomes sont non négatives.

Caractéristiques principales des algorithmes génétiques

- Technique de recherche heuristique robuste (générale)
- Stochastique
- S'applique à une population de chromosomes (représentation, encodage d'une valeur, d'une solution)
- Mécanique s'applique à la représentation chromosomique, sans aucune interprétation ou décodage de ce chromosome.
- N'utilise pas de connaissance du domaine d'application, sauf la mesure de performance.
- Algorithme génétique "pur et dur": chaînes de bits, opérateurs de croisement à un point, mutation.
- Equivalence Algorithmes génétiques-Organismes biologiques

Algorithmes génétiques

Chaîne de bits
Bit
Position
Valeur du bit
Mesure de performance (fitness)
Solution, Structure décodée
Epistatique

Organismes Biologiques

Chromosome (génotype)
Gène
Locus
Allèle
Adaptation à l'environnement
Phénotype
Non Linéaire

Exemple d'encodage sous forme de chromosome
(pour l'optimisation d'une fonction réelle)

- Soit la fonction:

$$f(x,y) = \frac{0.5 - (\sin(x^2 + y^2))^{0.5}}{1.0 + 0.001(x^2 + y^2)}$$

où $x, y \in [-100, 100]$

- Soit un chromosome de longueur 44.
- On le partitionne en deux parties égales: les 22 premiers bits encodent la valeur de x , et les 22 derniers bits encodent la valeur de y .
- Les deux chaînes de bits sont converties de la base 2 à la base 10.
- On obtient ainsi des valeurs entières pour x et y entre 0 et $2^{22}-1$.
- les valeurs entières sont projetées dans l'intervalle $[0, 200]$, en les multipliant par $200/(2^{22}-1)$.
- On soustrait 100 à la valeur obtenue de façon à obtenir un nombre réel dans l'intervalle $[-100, 100]$.
- On évalue f au point (x,y) ainsi obtenu. C'est la mesure de performance du chromosome.
- Plus le nombre de bits est élevé, plus la précision du nombre réel est grande.

Un algorithme génétique simple

1. Créer une population initiale de N chromosomes (Génération 0).
2. Évaluer la mesure de performance de chaque chromosome.
3. Reproduction. Sélectionner N parents dans la population courante (e.g., sélection proportionnelle: la probabilité de sélection est proportionnelle à la mesure de performance).
4. Croisement. Choisir aléatoirement une paire de parents pour le croisement. Échanger des sous-chaînes de bits à l'aide d'un croisement à un point (one-point crossover) de façon à créer deux enfants.
5. Mutation. Traiter les deux enfants avec l'opérateur de mutation et ajouter ceux-ci dans la nouvelle population.
6. Répéter les étapes 4 et 5 jusqu'à ce que tous les parents aient été traités (N enfants sont créés).
7. Remplacer l'ancienne population de chromosomes par la nouvelle (nouvelle génération).
8. Évaluer la mesure de performance de chaque chromosome dans la nouvelle population.
9. Retourner en (3) si le nombre de générations actuel est moindre que le nombre maximal de générations. Autrement, le résultat final est le meilleur chromosome généré durant la recherche.

Un algorithme génétique simple

Reproduction

Sélection proportionnelle ("roulette-wheel selection")

1. Additionner les mesures de performance de tous les chromosomes dans la population.
2. Générer un nombre aléatoire entre 0 et cette somme.
3. Sélectionner le premier chromosome dont la mesure de performance additionnée à la somme partielle des mesures de performance des chromosomes précédents est plus grand ou égal au nombre aléatoire.

Remarques

- Voir l'exemple plus haut pour une illustration du principe.
- Plus la mesure de performance est élevée, plus la portion de la "roulette" couverte par le chromosome est grande, et plus la probabilité de sélection est élevée.
- De fait, la probabilité de sélection du chromosome i dans une population de n chromosomes est:

$$p_i = f_i / (\sum_{j=1..n} f_j).$$

- Il s'agit donc d'une loi de Bernouilli avec probabilité de succès p_i .
- Puisqu'on répète le processus n fois de façon indépendante, on obtient une loi binomiale $B(n, p_i)$, dont l'espérance est np_i . Le nombre espéré de sélections pour le chromosome i est donc:

$$E_i = n p_i = n f_i / (\sum_{j=1..n} f_j) = f_i / f_{\text{moy}}$$

- L'espérance du nombre de sélections correspond à l'espérance du nombre d'enfants. Donc, les chromosomes les plus performants se reproduisent davantage.
- Implémentation naive: $O(n^2)$
- Implémentation avec un "B-tree": $O(n \log n)$

Reproduction

Désavantages

- (1) La variance dans le nombre de sélections pour un chromosome donné est élevée.
- (2) En fait, tout chromosome peut être sélectionné entre 0 et n fois avec une probabilité non nulle.
- (3) Un super-chromosome peut rapidement dominer une population ce qui mène à une convergence prématurée.
- (4) A l'opposé, on obtient une recherche quasi-aléatoire, dans une population où tous les individus ont de très légères différences dans leur mesure de performance.

Améliorations suggérées pour (3) et (4)

Ajustement des mesures de performance

- (a) Ajustement linéaire (linear scaling)

$$f' = a f + b$$

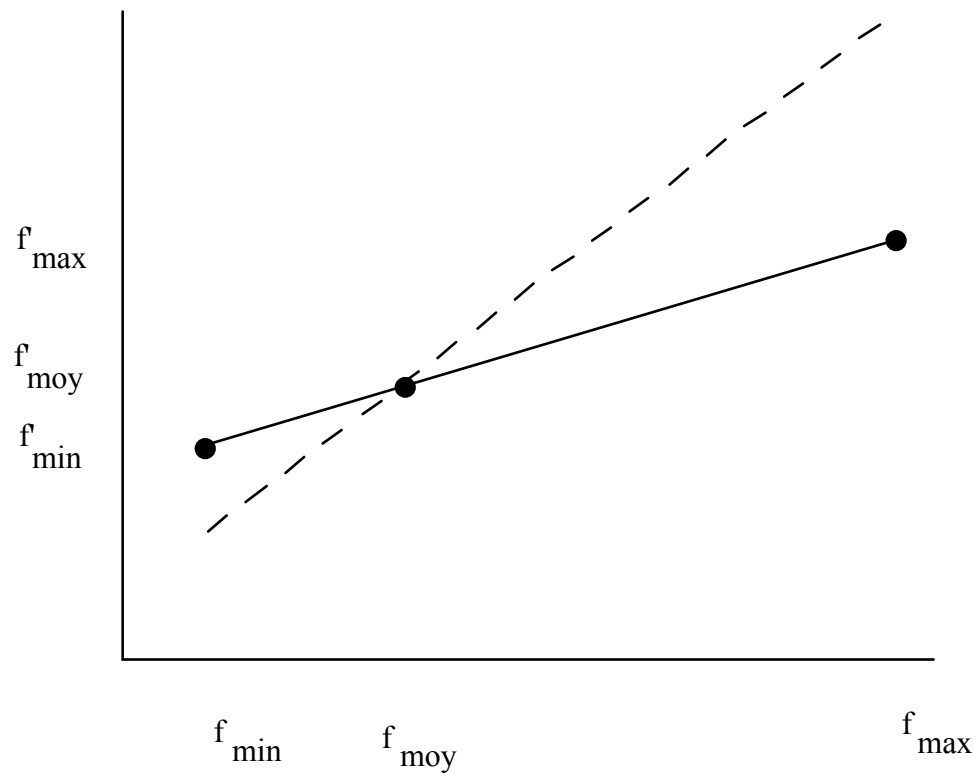
sujet à:

$$f'_{\text{moy}} = f_{\text{moy}}$$

$$f'_{\text{max}} = \text{constante} * f_{\text{moy}}$$

- Les contraintes assurent qu'un chromosome moyen génère toujours un enfant en moyenne, tandis que le meilleur chromosome génère un nombre d'enfants correspondant à la valeur de la constante.
- La constante ("selection pressure") est habituellement fixée à une valeur entre 1,2 et 2,0.
- Effet de la constante: plus sa valeur est élevée, plus l'écart entre les chromosomes les plus performants et les moins performants augmente.
- Mais, attention aux valeurs négatives!

Reproduction



Reproduction

Ajustement des mesures de performance

(b) Ajustement exponentiel

$$f' = f^k$$

(c) Ranking: on utilise le rang d'un chromosome dans la population (selon sa mesure de performance brute), pour calculer une mesure de performance transformée.

Formule de Baker pour le chromosome occupant le i^e rang, en utilisant des valeurs MAX et MIN définies a priori:

$$f_i = \text{MAX} - [(\text{MAX}-\text{MIN}) (i-1) / (n-1)]$$

et

$$\begin{aligned}\sum_i f_i &= \sum_i \text{MAX} - [(\text{MAX}-\text{MIN}) (i-1) / (n-1)] \\ &= n \text{MAX} - \sum_i (\text{MAX}-\text{MIN}) (i-1) / (n-1) \\ &= n \text{MAX} - [(\text{MAX}-\text{MIN}) / (n-1)] \sum_i (i-1) \\ &= n \text{MAX} - n (\text{MAX}-\text{MIN}) / 2 \\ &= n [(\text{MAX}+\text{MIN})/2] \\ &= n\end{aligned}$$

en supposant que $\text{MAX} + \text{MIN} = 2$.

Ainsi, $f_{\text{moy}} = 1$ et $E_i = np_i = f_i / f_{\text{moy}} = f_i$

Désavantage: la méthode n'est pas sensible à l'amplitude de la différence entre les performances de deux chromosomes.

Reproduction

Améliorations suggérées pour (1) et (2)

Modification du modèle de sélection (de façon à réduire la variance)

(a) Expected value model, Stochastic Sampling without Replacement

1. Évaluer E_i selon le modèle proportionnel, en l'occurrence:

$$E_i = f_i / f_{\text{moy}}$$

2. Chaque fois qu'un chromosome est sélectionné à l'aide de la roulette, soustraire 1 de E_i . Si la valeur devient négative, f_i est mis à 0 et le chromosome ne peut plus être choisi.

Cette approche introduit une borne supérieure (plafond de E_i) sur le nombre de sélections, mais aucune borne inférieure.

(b) Remainder Stochastic Sampling without Replacement

1. Évaluer E_i et considérer la partie entière $e(E_i)$ et la partie fractionnaire $\text{frac}(E_i)$.
2. Sélectionner chaque chromosome $e(E_i)$ fois.
3. Pour compléter la sélection, procéder à une sélection proportionnelle sur les parties fractionnaires $\text{frac}(E_i)$. Toutefois, dès qu'un chromosome i est sélectionné une fois, $\text{frac}(E_i)$ est mis à 0 (no replacement).

Cette approche introduit une borne supérieure (plafond de E_i) sur le nombre de sélections, ainsi qu'une borne inférieure (plancher de E_i).

Reproduction

Modification du modèle de sélection (de façon à réduire la variance)

(c) Stochastic Universal Sampling

- Utilise aussi le concept de la roulette, mais on sélectionne les n chromosomes "d'un seul coup", à l'aide de n nombres aléatoires.
- On suppose que $\sum_i f_i = n$.
- On choisit un nombre aléatoire entre 0 et 1, et on ajoute 1 à ce nombre $n-1$ fois, de façon à obtenir n nombres compris entre 0 et $\sum_i f_i$.
- On sélectionne les n chromosomes qui correspondent à ces n nombres sur la roulette.
- Complexité: $O(n)$.
- Introduit une borne supérieure (plafond de E_i) sur le nombre de sélections, ainsi qu'une borne inférieure (plancher de E_i).

Croisement

- L'opérateur de croisement permet à deux parents d'échanger des chaînes de bits lors de la génération des enfants.
- Taux de croisement (habituellement fixé à 0,65): lorsque le croisement n'est pas appliqué à deux chromosomes parents, ceux-ci sont copiés tels quels dans la nouvelle population.

Croisement à un point (one-point crossover)

1. Choisir une position aléatoirement et couper les deux chromosomes parents à cette position
2. Échanger les bits situés après la coupure de façon à créer deux enfants

Exemple:	parent 1	1	0		1	0	0	0
	parent 2	0	1		1	1	0	1
<hr/>								
	enfant 1	1	0	1	1	0	1	
	enfant 2	0	1	1	0	0	0	

Remarques

- Si les bits des deux parents à une position donnée ont la même valeur, alors les deux enfants hériteront de cette valeur.
- Ainsi, deux chromosomes identiques génèrent deux enfants qui leur sont identiques.

Croisement

Croisement à deux points (two-point crossover)

- Généralisation du croisement à un point, où un second point de coupe est choisi aléatoirement.

Exemple:	parent 1	1	0		1	0		0	0
	parent 2	0	1		1	1		0	1
<hr/>									
	enfant 1	1	0		1	1		0	0
	enfant 2	0	1		1	0		0	1

Croisement à M points (M-point crossover)

- M positions sont choisies aléatoirement sur chacun des deux chromosomes parents
- Les éléments entre deux positions consécutives sont alternativement échangés et maintenus (le premier segment entre le début du chromosome et la première position sélectionnée étant maintenu sur chacun des chromosomes)

Croisement uniforme (uniform crossover)

- Croisement bit par bit
- Pour chaque position, le parent qui fournira le bit au premier enfant est choisi aléatoirement. L'autre parent fournit alors son bit au deuxième enfant.

Mutation

- Les enfants générés par l'opérateur de croisement sont ensuite traités bit par bit par l'opérateur de mutation.
- Taux de mutation: en général très petit, de l'ordre de 0,001.
- L'opérateur de mutation permet de maintenir une certaine diversité dans la population
- Particulièrement utile lorsque tous les chromosomes ont la même valeur à une position donnée (allèles perdus).

Remplacement des générations

- Remplacement complet d'une population par une autre
 - Élitisme: maintien du meilleur chromosome dans une population, dans la population suivante.
 - Steady-state:
 - On crée un nombre d'enfants m inférieur à la population totale n
 - On retire les m chromosomes les moins performants de la population courante, afin de laisser la place aux nouveaux chromosomes.
- Note. Élitisme équivaut à Steady-state avec $m = n-1$. Le remplacement complet de la population équivaut à Steady-state avec $m = n$.