



Évolution artificielle

Une introduction

Marc Schoenauer

Équipe-projet TAO – INRIA Scalay et LRI

<http://tao.lri.fr/>

12 juin 2012



Position

- Directeur de Recherche INRIA Saclay depuis 2001
CR CNRS de 1981 à 2001 au CMAP
- Co-responsable (avec Michèle Sebag) de l'équipe-projet TAO
Commun avec le LRI - UMR CNRS 8643 et Université Paris-Sud

Recherche

- TAO = Thème Apprentissage et Optimisation
- Optimum Design, Optimisation combinatoire, robotique

Responsabilités

- Editeur en chef *Evolutionary Computation* (2002-2009)
- Business Committee GECCO
- Co-fondateur de l'association *Evolution Artificielle*



Plan

- **Contexte** un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



Optimisation

Données

- Ensemble de solutions possibles
- Critère de qualité

Espace de recherche

Fonction objectif
mais aussi performance, "fitness"

But

Trouver la meilleure solution (pour le critère donné)

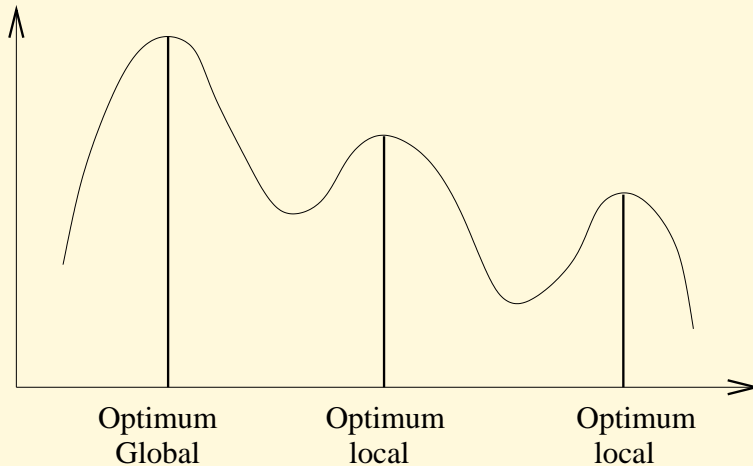
Formellement

Soit $\mathcal{F} : \Omega \mapsto \mathbb{R}$

Trouver $x^* \in \Omega$ tq $x^* = \text{ArgMax}(\mathcal{F})$



Optima (maximisation)



Optimum global : x^* t.q. $(\forall x \in \Omega) \mathcal{F}(x^*) \geq \mathcal{F}(x)$

Optimum local : x^* t.q. $(\exists \varepsilon > 0)$

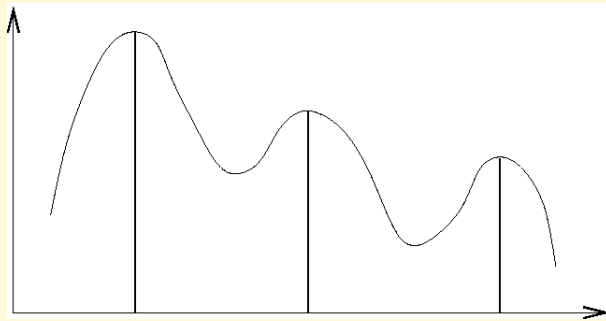
$$B(x, \varepsilon) \neq \{x\} \text{ et } (\forall x \in B(x^*, \varepsilon)) \neq \{x\}) \mathcal{F}(x^*) \geq \mathcal{F}(x)$$

Maxima **stricts** si inégalités strictes pour $x \neq x^*$

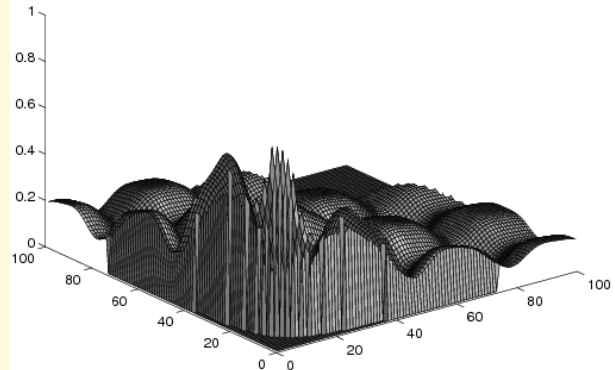


Paysages

"Graphe" de \mathcal{F} sur Ω



$\Omega = \mathbb{R}$



$\Omega = \mathbb{R}^2$



Sources de difficultés

- Espace de recherche trop grand (cas discret)

e.g. Problèmes NP-complets

- Fonction objectif "complexe"
irrégulière, non différentiable, non continue, ...
- Fonction objectif donnée via un calcul ou un expérience

Boîte noire



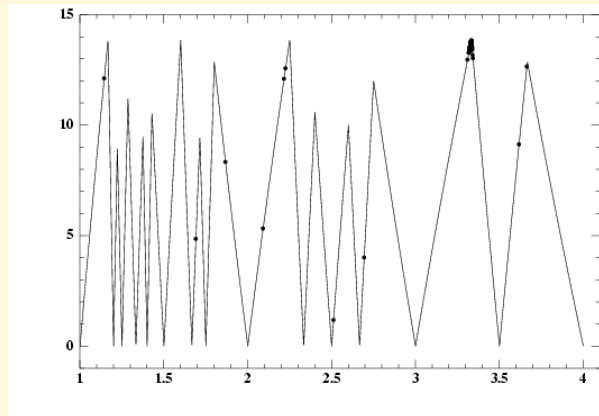
\mathcal{F} très chahutée

L. Taïeb, CMAP & Thomson

Espace de recherche : Interféromètres

Positionner des antennes

But : Maximiser la tolérance en conservant la précision.



Cas de 3 antennes, \mathcal{F} = Marge d'erreur (position 2ème antenne)

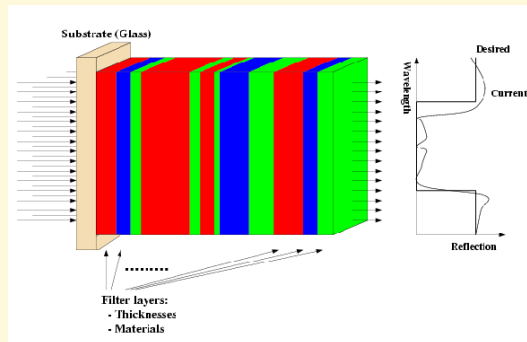


Ω mixte : discrets \times réels
 Schutz & Bäck, ICD, Dortmund.
 Martin, Rivory & Schoenauer,
 Optique des Solides Paris VI & CMAP.

Espace de recherche : Filtres optiques

(matériau, épaisseur)₁ ... (matériau, épaisseur)_N

But : Répondre au gabarit fixé.





Ω = Circuits analogiques

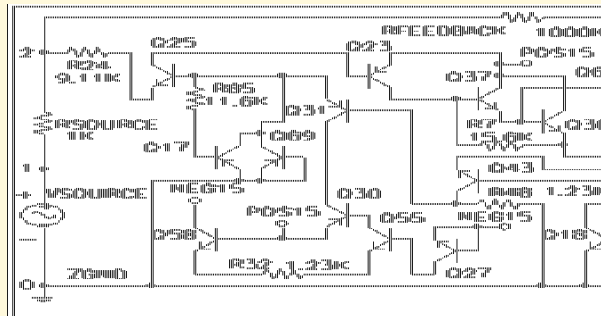
Koza et al., Stanford.

Espace de recherche : Circuits analogiques

Réseau de transistors, diodes, résistances

But : Fonctionnalités fixées

e.g. extraction de racine cubique





Ω espace de fonctions, \mathcal{F} mal connue

Robotique autonome On cherche la vitesse des moteurs en fonction des valeurs des capteurs
on veut que le robot avance en évitant les obstacles.



Espace de recherche : un espace de fonctions
But : Minimiser l'activation des capteurs

... tout en avançant !

Robotique évolutionnaire



\mathcal{F} non calculable

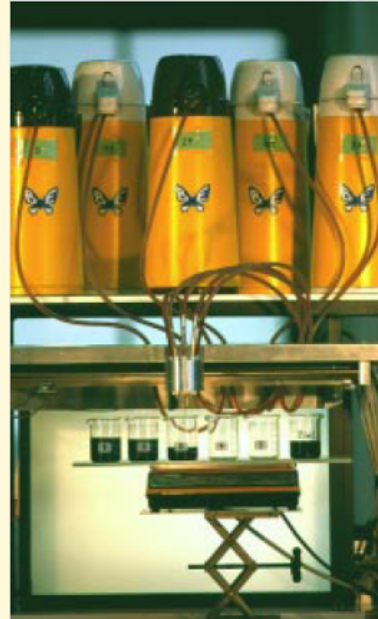
Herdy & al., Berlin, PPSN96

Espace de recherche :
Mélanges de café

But : Retrouver un arôme

\mathcal{F} = avis de l'expert

Évolution Interactive

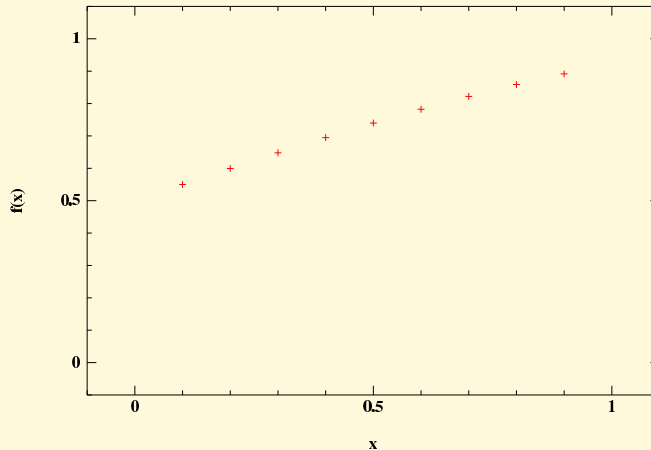




Scénario "boite-noire"

- Accès aux valeurs de \mathcal{F}
- Aucune hypothèse sur la fonction objectif

Aucune ? $\Omega = \mathbb{R}$, on a observé 10 points

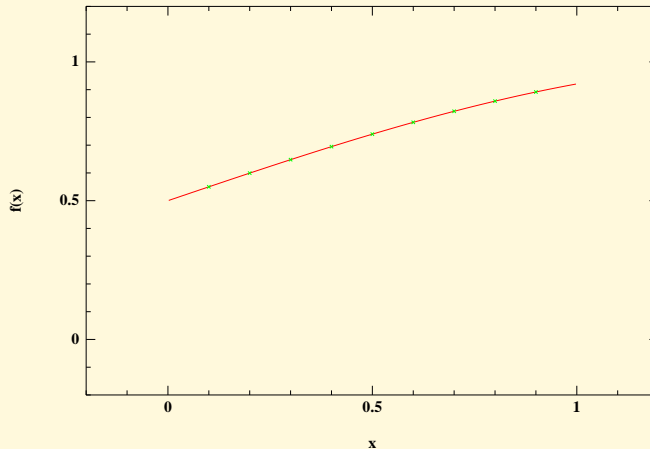




Scénario "boite-noire"

- Accès aux valeurs de \mathcal{F}
- Aucune hypothèse sur la fonction objectif

Aucune ?

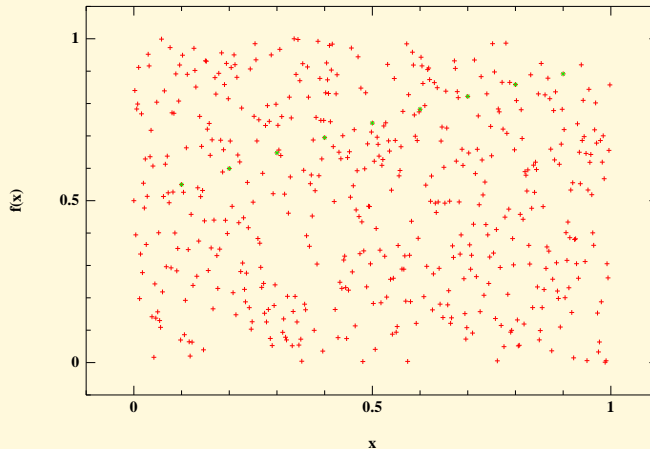




Scénario "boîte-noire"

- Accès aux valeurs de \mathcal{F}
- Aucune hypothèse sur la fonction objectif

Aucune ?





Contexte

- Optimisation globale

Recherche d'un optimum global
ou d'une bonne approximation

- Scénario boîte noire, mais ...

- Problème structuré : les voisinages de bonnes solutions peuvent contenir des solutions encore meilleures.



Algorithmes d'optimisation

Méthodes exactes

- Algorithmes de gradient
- Méthodes énumératives
- Hill-Climbing

cas continu

cas discret

Méthodes approchées

- Heuristiques
- Recherche locale
- Méthodes stochastiques (méta-heuristiques)



Méthodes de gradient

$$X_{i+1} = X_i + dX \times \nabla \mathcal{F}(X_i)$$

Contexte :

- Espace continu : $\Omega \subseteq \mathbb{R}^N$
- Fonction \mathcal{F} dérivable

dérivation numérique

Conditions nécessaires :

- Fonction régulière,
- ou Connaissance a priori

\mathcal{F} convexe .. optimum unique

X_0 bien choisi

Méthode locale, pour des problèmes continus et réguliers



Méthodes de type énumératif

Parcourir l'espace suivant un ordre déterministe

Contexte :

- Espace fini : $\Omega \equiv [1..N]$ **Attention** : $\text{opt.}_{discret} \neq \text{arrondi}(\text{opt.}_{continu})$
- Toute fonction \mathcal{F} , mais ...
- Ordre de parcours
 - fixé
 - dépend du problème **Branch-and-Bound, A*, contraintes, ...**

Conditions nécessaires :

- Taille de l'espace limitée
- Discrétisation bien choisie **Méthode des intervalles**

Méthode globale, coûteuse, pour des problèmes discrets



Hill-Climbing

$$X_{i+1} = \text{" Meilleur Voisin " de } X_i$$

Contexte :

- Toute fonction \mathcal{F}
- Tout espace $\Omega \dots$
- pour lequel on sait définir des voisinages

Trouve l'optimum local (pour les voisinages choisis) le plus proche du point de départ

Conditions nécessaires :

- X_0 doit être bien choisi

Méthode locale, coûteuse



Exemples de voisinages

- $\Omega = \{0, 1\}^N$
- $d_H((x_i), (y_i)) = \sum |x_i - y_i|$
- $V_d(x) = \{y \in \{0, 1\}^N; d_H(x, y) \leq d\}$

distance de Hamming

Le choix de d détermine la "localité" de la recherche

- $d = N$: recherche globale (énumération !)
- $d = 1$: recherche très locale



Heuristiques

- Dépend **fortement** du problème
- Permet d'obtenir rapidement une solution
- pas très bonne en général ...

Exemple : le TSP

Heuristique gloutonne: choisir la ville la plus proche.

Méta-heuristiques :

Heuristiques que l'on peut appliquer à plusieurs classes de problèmes

Ex : Hill-Climbing ... si on n'explore pas tout le voisinage



Méthodes stochastiques

Les points suivants sont choisis à l'aide de tirages aléatoires

- Monte-Carlo • X_i tiré avec une loi uniforme
• Marche aléatoire sans mémoire
- Recherche locale • Hill-climbing stochastique
- Métropolis, Recuit Simulé • Kirkpatrick, Gelatt and Vecchi, 1983
- Recherche Taboue • F. Glover – 1977 & 1989
- Algorithmes évolutionnaires ... • depuis 1965

Méthodes globales, mais TRÈS coûteuses.



Recherche locale (stochastique)

- Choisir x_0 uniformément dans Ω
- Calculer $\mathcal{F}(x_0)$
- Boucle e.g. Tant que améliorations ...
 1. Choisir y (uniformément) dans $V(x_t)$ voisinage
 2. Calculer $\mathcal{F}(y)$
 3. Si $\mathcal{F}(y) > \mathcal{F}(x_t)$, alors $x_{t+1} = y$ Acceptation
sinon $x_{t+1} = x_t$



Voisinages, exploration et exploitation

- Grands voisinages (e.g. $d = N$ dans $\{0, 1\}^N$)
≡ Monte-Carlo: **exploration** sans tenir compte du passé
- Petits voisinage (e.g. $d = 1$)
Recherche "très" locale: **exploitation** sans exploration

Solutions

- Étendre la notion de voisinage Voisinages stochastiques
- Accepter aussi des points moins bons Item 3 de la boucle
Metropolis et Recuit simulé
- Interdire de revenir sur ses pas
Recherche taboue
- Utiliser plusieurs points simultanément et accepter des moins bons
Algorithmes évolutionnaires



Recherche (locale?) stochastique (2)

- Choisir x_0 uniformément dans Ω
- Calculer $\mathcal{F}(x_0)$
- Boucle e.g. Tant que améliorations ...
 1. $y = \text{Move}(x_t)$ opérateur \equiv voisinage \equiv distribution sur Ω
 2. Calculer $\mathcal{F}(y)$
 3. Si $f(y) > f(x_t)$, alors $x_{t+1} = y$ Acceptation
sinon $x_{t+1} = x_t$



Recherche locale stochastique (3)

- Choisir x_0 uniformément dans Ω
- Calculer $\mathcal{F}(x_0)$
- Boucle
 1. $y = Move(x_t)$
 2. Calculer $\mathcal{F}(y)$
 3. $x_{t+1} = select(y, x_t)$

e.g. Tant que améliorations ...

e.g. Boltzman \rightarrow Metropolis



Recherche stochastique (4)

- Choisir x_0 uniformément dans Ω
- Calculer $\mathcal{F}(x_0)$
- Boucle

e.g. Tant que améliorations ...

1. Faire λ fois

- $y_i = \text{Move}(x_t)$
- Calculer $\mathcal{F}(y_i)$

2. $x_{t+1} = \text{select}(y_1, \dots, y_\lambda, x_t)$

Exemple : $\Omega = \mathbb{R}^n$, $\text{Move}(x) = \mathcal{N}(x, \sigma)$, sélection déterministe

→ $(1 + \lambda)$ -**Evolution Strategy**



Recherche stochastique (5)

- Choisir x_0^1, \dots, x_0^μ uniformément dans Ω
- Calculer $\mathcal{F}(x_0^1), \dots, \mathcal{F}(x_0^\mu)$
- Boucle e.g. Tant que améliorations ...
 1. Faire λ fois
 - $y_i = \text{Move}(x_t^1, \dots, x_t^\mu)$
 - Calculer $\mathcal{F}(y_i)$
 2. $(x_{t+1}^1, \dots, x_{t+1}^\mu) = \text{select}(y_1, \dots, y_\lambda, x_t^1, \dots, x_t^\mu)$

Exemple : $\Omega = \mathbb{R}^n$, $\text{Move}(x) = \mathcal{N}(x, \sigma)$, sélection déterministe

→ $(\mu + \lambda)$ -**Evolution Strategy**



Plan

- Contexte un peu d'humilité
- **L'algorithme** le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



Paradigme Darwinien

- Sélection naturelle avantage aux espèces adaptées à leur environnement
- + Variations aveugles
 parents → enfants par petites déviations apparemment non dirigées.
- = Adaptation apparition d'espèces (e.g. bactéries résistantes).
- “Objectif” capacité de survivre et de se reproduire

Mais

- Source d'inspiration
- Aide à l'explication
- **Pas justification**



Algorithmes Évolutionnaires : La Métaphore

Modèle : L'évolution darwinienne des populations biologiques.
Les individus les plus adaptés survivent et se reproduisent

Vocabulaire :

Individu	Élément X de Ω
Performance	Valeur de $\mathcal{F}(X)$
Population	Ensemble de P éléments de Ω
Génération	Passage de la population Π_i à Π_{i+1}

Processus :

- 1) Sous la pression du milieu,
- 2) Les individus se croisent, mutent et se reproduisent.
- 3) Au bout d'un nombre certain de générations, les individus les plus performants apparaissent dans la population.

≡ les **optima** de \mathcal{F} ...



Parallèle biologie/algorithmique

Différences

Evolution	
Biologique	Artificielle
Environnement changeant mécanismes spécifiques ?	Généralement fixe
Performance inconnue finalité des plumes du paon* ?	C'est le point de départ
Le sélection élimine les nuls	Tentation eugénique Lamarckisme possible
Adaptation/sélection boîte noire	Études théoriques

* Richard Dawkins, *Le gène égoïste*



Parallèle biologie/algorithmique

Points communs

Diversité génétique essentielle

Maladies fatales
Multi-racial utile

Convergence prématurée
Solutions multiples utiles

Lenteur du processus

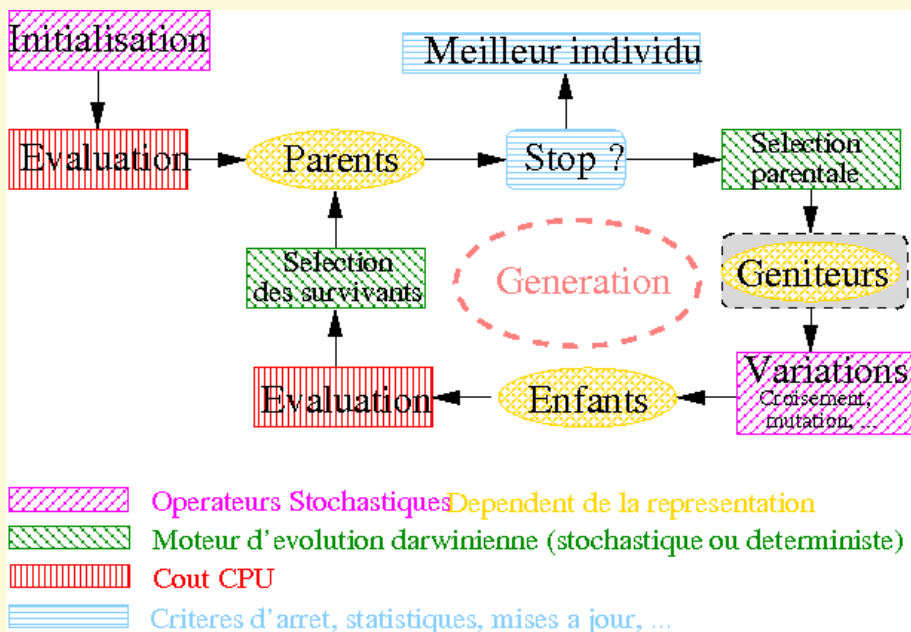
Néanderthal: -150 000 à -35 000
 \simeq 6000 générations
Cro-Magnon: -30 000 à nous
 \simeq 1200 générations

Vous verrez!

On est peut-être pas à l'optimum,
mais on a des solutions assez adaptées :-)



Algorithmes d'évolution : Le Squelette





Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- **Un exemple jouet** sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



La feuille de papier qui tombe

P. Bentley

Le problème

- Trouver la forme d'un morceau de papier
- qui met le plus de temps possible à tomber

Difficultés

- Pas de simulation
- Pas d'a priori sur la forme de la solution

Quoique ...



Un algorithme sans ordinateur

Matériel nécessaire

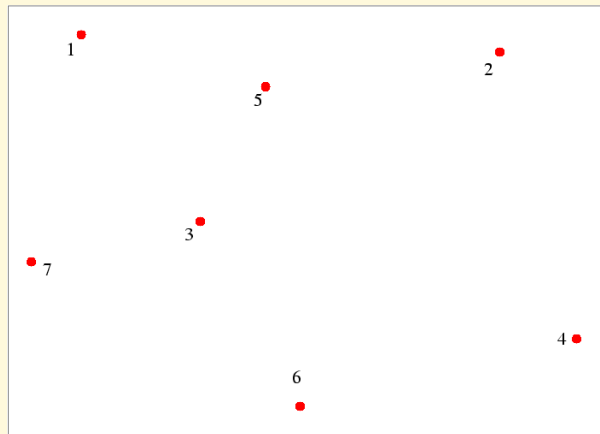
- Une ramette de feuilles A4
- Des ciseaux
- Un bac à sable format A4
- Une dizaine de petites cailloux numérotés
- Un chronomètre
- Une fléchette
- ... de la patience



Initialisation

- Lancer les petits cailloux en l'air au dessus du bac à sable
- Reporter les positions des cailloux tombés dans le bac sur une feuille A4

Génotype : Liste ordonnée des points





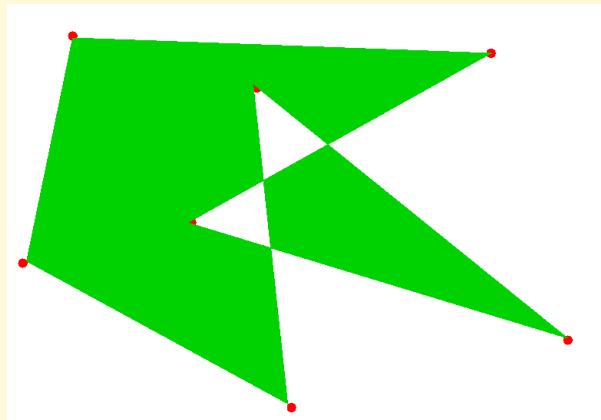
Morphogénèse

passage du génotype au phénotype

- Tracer le polygone rempli correspondant
- Découper la forme obtenue

à la MacPaint

Phénotype : La forme de papier





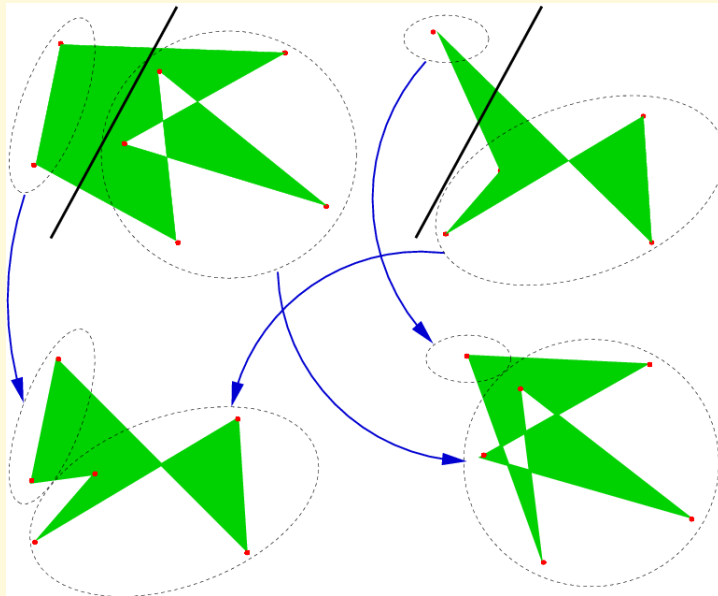
Évaluation

Pour chaque forme

- Faire 5 fois
- Lâcher la forme de 2m de haut
- Chronométrer le temps d'atteinte du sol
- Faire la moyenne des 5 temps



Croisement

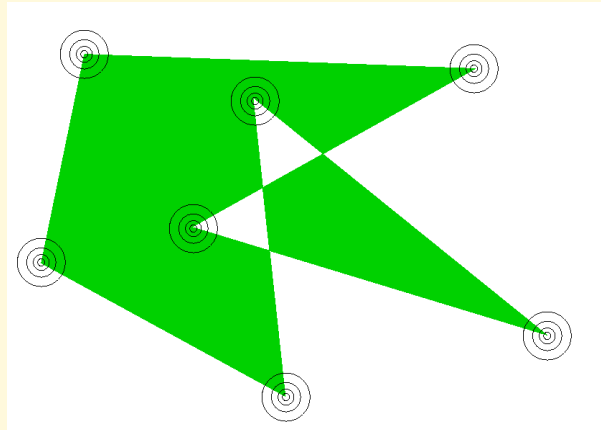


Un croisement possible, pour lequel les ciseaux et le scotch ne suffisent pas.



- Viser avec la fléchette chaque point tour à tour

Distance à ajuster selon votre habileté



- Tracer sur une nouvelle feuille les impacts

Si la fléchette sort du cadre, enlever le point

Si votre portable sonne, ajouter un point



Darwinisme

- Pas de sélection parentale

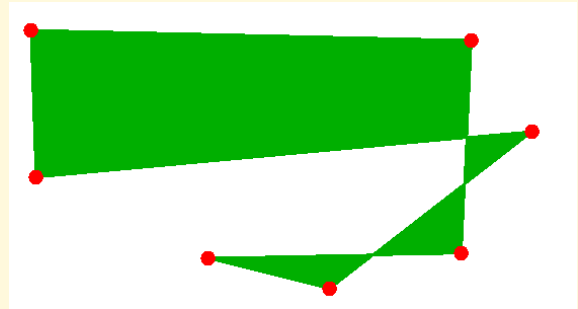
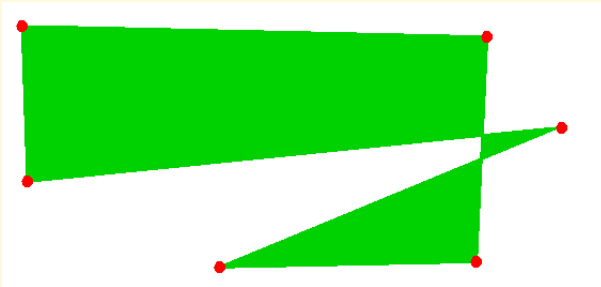
Tous les parents sont géniteurs

- Sélection pour la survie déterministe :
Déchirer la moitié des formes – les plus rapides



Résultats

Avec 5 formes et 10 générations



Deux des meilleures formes obtenues

Meilleure forme aléatoire : 0.8s

Toutes les formes après 10 générations de 10 formes : $> 2s$

Comportement “hélicoptère”





Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- **Points-clé** représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



Représentation / moteur d'évolution

- Espace **phénotypique** : Évaluation, sélection
La feuille de papier
- Espace **génotypique** : Variations (croisements, mutations)
Les petits cailloux
- Le **Darwinisme** ne dépend que de la **performance**
- **L'initialisation et les opérateurs de variation** ne dépendent que de la **représentation**.

Trois exemples les plus courants:

- Représentation “binaire” Chaînes de bits, Algorithme Génétique
- Représentation “réelle” Stratégies d'évolution – voir exposé ES
- Représentation par arbres GP – voir exposé GP

Le choix de la représentation est **crucial**



Perte de la Diversité Génétique

Si les individus d'une population se ressemblent trop,

1. Les populations suivantes deviennent de plus en plus homogènes
fragilité au changement
2. évolution d'une population → évolution d'un individu
3. Découverte du plus proche optimum local et enlèvement de la recherche

Dans la pratique, **la population ne se rediversifie pas.**

⇒ Convergence Prématuration



Le Dilemme Exploitation vs Exploration*

- **Exploitation** des bons individus.
recherche locale: chercher dans le voisinage des meilleurs individus de la population.
- **Exploration** des zones inconnues de Ω .
recherche globale: il faut pouvoir aller partout.

Excès d'exploitation \implies

Convergence prématurée
enlèvement dans un optimum local

Excès d'exploration \implies

Pas de convergence
 \approx marche aléatoire

* aka Intensification vs Diversification



Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- **Darwinisme artificiel** et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



Sélections darwiniennes

- Darwinisme: Biais en faveur des plus adaptés
- Règle l'intensité de l'exploitation

Sélection parentale : Un individu peut être sélectionné plusieurs fois

- Stochastique, basée sur les **valeurs** de la fitness **Roulette**
- Stochastique, basée sur des **comparaisons** de fitness **Tournois**
- Déterministe **basée sur des comparaisons de fitness**

Sélection pour la survie : un individu est sélectionné 1 fois ou disparaît

- Déterministe ou stochastique (tournoi)
- Choix parmi les enfants seulement, ou conflit de générations

Opérateurs d'exploitation



Sélection Déterministe

Les moteurs $(\mu \ddagger \lambda)$ -ES

- Sélection parentale uniforme ≡ pas de sélection parentale
- les μ parents donnent λ enfants (**opérateurs de variation**)
- Sélection des survivants : Choix des prochains μ parents
 (μ, λ) -ES : μ survivants = meilleurs parmi les λ enfants
Pour: meilleurs résultats de convergence
Contre: on peut perdre les meilleurs

- $(\mu + \lambda)$ -ES : μ survivants =
meilleurs parmi les μ parents + les λ enfants
Pour: robustesse pratique
Contre: on peut converger vers un opt. local

Paramètres : μ, λ



Sélection Stochastique

Les tournois

- **Tournoi (déterministe)**

$$ps \approx T$$

- Tournoi de taille $T \in \mathbb{N}$
- Choix uniforme de T individus
Rendre le meilleur

avec ou sans remplacement?

- **Tournoi stochastique (binaire)**

$$ps \approx 2t$$

- Taux $t \in [0.5, 1]$
- Choix uniforme de 2 individus
Rendre le meilleur avec probabilité t

- **Avantages** : Ne dépend que de l'**ordre** des valeurs de fitness
Robuste par rapport aux erreurs sur \mathcal{F}
Facile à paramétrer T or t



Moteur AG standard

Historiquement :

- Sélectionner P individus par sélection stochastique

avec remplacement

- Générer P enfants à l'aide des **opérateurs de variation**

- Remplacer les parents par les enfants

remplacement générationnel

Paramètres: La sélection – et ses paramètres

- Tournoi : T ou t

- Proportionnelle (roulette):
 P pression sélective et mise à l'échelle

A éviter



Moteur “Steady State GA”

Historiquement (SSGA)

- Sélectionner 1 individu par tournoi
- Générer 1 enfants à l’aide des **opérateurs de variation**
- Remplacer 1 parent par l’enfant par “anti-tournoi”

Paramètres: Tailles des tournois sélection et remplacement

Note: Autres stratégies de remplacement – le pire, le plus vieux, ...



Sélection Multi-objectif

- Il "suffit" de remplacer les comparaisons de fitness ... par des comparaisons multi-objectif
- pour obtenir ... un magnifique **algorithme d'optimisation multi-critère**

Voir un des exposés suivants



Évaluation de la population courante

DE TRÈS LOIN l'étape la plus coûteuse

- Ne pas recalculer $\mathcal{F}(X)$ inutilement
- Utiliser une estimation de \mathcal{F} voire la construire à la volée
Response Surface Method ou *SURrogate Models*
- ... mais pas trop longtemps Optimum approché \neq optimum réel



Critères d'arrêt

Pas si simple !

- Quand on a trouvé l'optimum ...:-)
- Quand on n'espère pas trouver mieux Perte de diversité
- Quand le ratio $\frac{\text{gain espéré}}{\text{surcoût de calcul requis}}$ est trop élevé Rationalité limitée
- Quand on a épuisé ses ressources Nombre fixé d'évaluations

Règle heuristique: après un nombre d'évaluations sans amélioration
compté après un nombre minimum d'évaluations



Évaluation des résultats

Point de vue informatique

Ne **jamais** tirer de conclusions d'**un seul** essai !

Utiliser des mesures statistiques

Moyennes et écarts-types, médianes, ...

Pour **comparer** des algorithmes, utiliser des tests statistiques

Paramétriques (e.g. Student) ou non-paramétriques (e.g. Kolmogorov, Wilcoxon)

Voir exposé à venir

Point de vue de l'application

Contexte de **conception**

Trouver au moins une fois une très bonne solution

Contexte de **production**

Trouver en moyenne une solution assez bonne



Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- **Opérateurs de variation** ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- Conclusions et références



Initialisation de la population

Choix de $\Pi_0 = \{X_1, \dots, X_P\}$

- Par tirage uniforme dans Ω $\Omega = \{0, 1\}^N$, $X_i^j = 0$ ou 1 équiprobables
 $\Omega = [0, 1]^N$, $X_i^j = \text{random}()$
- ... mais attention au critère d'uniformité
- En tenant compte des connaissances a priori
Ajout de bonnes solutions manuelles
Mais pas de biais vaut mieux qu'un mauvais biais
- Comme résultat d'une évolution précédente.
Utilisation de plusieurs "milieux" $\mathcal{F}_1, \mathcal{F}_2, \dots$

La diversité génétique est essentielle !

Opérateur d'exploration



Opérateurs de variation

Suivant l'arité on définit

- opérateur unaire → **mutation**
- opérateur binaire → **crossover** Peut modifier l'un des deux, ou les deux
- opérateur N-aire → **orgie**



Reproduction

- Géniteurs (*mating pool*) \longrightarrow enfants ...
- Par application d'opérateurs de variation Généralement stochastiques
- Application des opérateurs
 - **Séquentielle**, avec une certaine **probabilité**
AG historique
 - * Croisement avec probabilité p_c
 - * Mutation avec probabilité p_m par individu
 - * Exemple : avec $p_c = 0.8$ et $p_m = 0.3$,
14% des individus sont inchangés
 - **Proportionnelle** selon des **poids** relatifs tirage de roulette
GP historique: Croisement (6), Mutation(1), Copie(3)
 - ou toute combinaison



Reproduction : le cas d'école

On dispose de

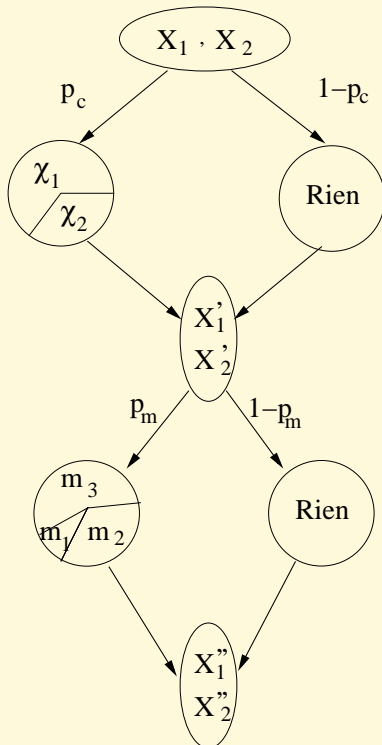
- 2 croisements χ_1 et χ_2 ,
- 3 mutations m_1, m_2, m_3

On définit

- $p_c = 0.6$ et $p_m = 0.3$,
- $pds_{\chi_1} = 2$ et $pds_{\chi_2} = 1$
- $pds_{m_1} = 1$, $pds_{m_2} = 3$ et $pds_{m_3} = 6$

Exercice : quel pourcentage

- subira χ_2 puis m_1 ?
- subira seulement m_3 ?
- ne sera pas modifié ?





Le croisement

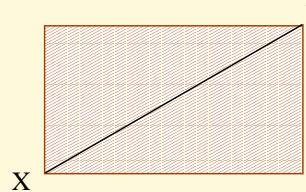
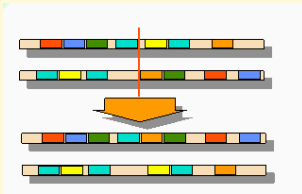
opérateur: $\Omega \times \Omega \rightarrow \Omega$ (ou Ω^2)

- **Analogie**: reproduction sexuée.
- **Intuition** :
les enfants héritent des qualités de leurs deux parents...
- **Débat** : le croisement est
L'opérateur majeur des AGs
Un opérateur mineur (inutile) pour ES et EP
- **Recommandation** : **essayez !**
l'intérêt dépend du problème
notion de fragments de solution



Opérateurs de variation : Le croisement

Exemples classiques



Échange de **gènes** Croisement de paramètres réels

Exemple orgiaque: Cinq parents

La foule subjuguée **boira** ses paroles enflammées
 Ce plat **exquis** enchantait leurs papilles expertes
 L'aube aux doigts de roses se leva sur un jour **nouveau**
Le cadavre sanguinolent encombra la police nationale
 Les coureurs assoiffés se jetèrent sur **le vin** pourtant mauvais
pour un enfant surréaliste



Croisement: discussion

Propriétés:

- **Opérateur d'exploitation**
- Recombinaison des “bonnes” parties
- Effets destructeurs

Choix du partenaire:

Aveugle en général, mais on peut introduire des préférences “sexuelles”

Un exemple :

- Des parents sur des pics différents d'une fonction multi-modale donneront sans doute des enfants peu performants
- Croisement “restreint” : croiser X avec Y ssi $d(\vec{X}, \vec{Y}) < \text{seuil}$
- Prise en compte des contraintes: your brain and my beauty



La mutation

opérateur: $\Omega \rightarrow \Omega$

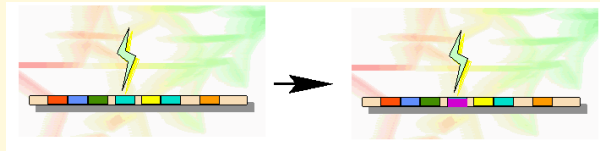
- **Analogie**: reproduction asexuée.
- **Intuition** : les enfants du croisement sont limités par Π_t
le seul contrepoids: la mutation
- **Grandes lignes** :
L'enfant doit être en général **proche** du parent
opérateur d'exploitation
L'enfant doit pouvoir être **n'importe où**
opérateur d'exploration – ergodicité
- **Quelle force?**
De nombreuses mutations heureuses
sinon
mutation plus faible
mutations plus forte
- **Débat** : Un opérateur homéopathique des AGs
L'opérateur majeur pour ES et EP



Opérateurs de variation : La mutation

Exemples classiques

- Mutation d'un gène



- Ajout de bruit Gaussien aux paramètres réels

Un exemple sans queue ni tête

La terre est comme un orange **bleue**

La terre est **bleue** comme une orange



Le débat croisement – mutation

Le croisement:

- permet de grandes modifications
- les modifications dépendent de la population
- plutôt opérateur d'exploitation
- effets décroissants avec l'évolution

La mutation:

- nécessaire
- nécessité de pouvoir faire de grands pas
- plutôt opérateur d'exploration
- effets destructeurs augmentent avec l'évolution



Le débat croisement – mutation (2)

Efficacité empirique du croisement

- **Hypothèse constructive** :
Assemblage de **briques de base**, i.e. pseudo-linéarité de la fonction fitness par rapport à des parties des individus
- **Hypothèse opportuniste** :
Le croisement n'est qu'une **macro-mutation**

Headless-chicken crossover

Besoin de mutation

- Ergodicité Résultats théoriques
- Relaxé en cas de **très grandes** population GP originel



Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- **Sujets d'actualité** paramétrage, représentations
- Conclusions et références



Réglage des paramètres

- De **très nombreux** paramètres
sélection, taille population, probabilités/poids des opérateurs
- Pas de méthode basée sur les premiers principes :-)

Off-line: plans d'expériences pour trouver le jeu optimal

pour une classe de problèmes donnée

- ANOVA et Racing Valeurs possibles des paramètres fixées
- Sequential Parameter Optimization
Processus Gaussiens pour le paramétrage
- REVAC, ParamILS Méta-EAs depuis Grefenstette 89 !

Méthodes **très coûteuses**



Réglage des paramètres (2)

On-line: les valeurs changent pendant l'évolution

- Paramètres dynamiques à un rythme fixé a priori par l'utilisateur
e.g. $p_c \searrow 0$
- Paramètres adaptatifs en fonction de l'évolution
mais jusqu'à quel point peut-on extrapoler
- Paramètres auto-adaptatifs les paramètres sont modifiés aléatoirement
l'évolution sélectionne indirectement les bons paramètres

Un seul succès notable : les stratégies d'évolution voir exposé à venir



Un point de vue orthogonal

Parents + opérateurs de variation \equiv échantillonnage d'une distribution

- | | |
|---|---|
| • Choisir Π_0 | • Choisir \mathcal{D}_0 |
| • Répéter | • Répéter |
| – $\mathcal{E}_t = \text{Variation}(\Pi_t)$ | – $\mathcal{E}_t = \text{Echantillonne}(\mathcal{D}_t)$ |
| – $\text{Evaluation}(\mathcal{E}_t)$ | – $\text{Evaluation}(\mathcal{E}_t)$ |
| • $\Pi_{t+1} = \text{Select}(\mathcal{E}_t, \Pi_t)$ | • $\hat{\mathcal{E}}_t = \text{Select}(\mathcal{E}_t)$ |
| | • $\mathcal{D}_{t+1} = \text{Update}(\hat{\mathcal{E}}_t, \mathcal{D}_t)$ |

Evolutionary **A**lgorithm
Un exemple mixte : les stratégies d'évolution

Estimation of **D**istribution **A**lg.
 voir exposé à venir



Vers des représentations développementales

● Représentations directes

- Paramétrique vs Non-paramétrique Taille fixe vs variable
- OneMax, Stratégies d'Evolution vs arbres GP, graphes, ...

● Représentations indirectes

- Genotype 'simple', morphogénèse complexe codée à la main
- Optimisation combinatoire, Optimisation de formes

● Représentations développementales

- Genotype : la morphogénèse
- Circuits analogiques GP, Koza, 1998
Réseaux de régulation génétiques Bio-inspiration



Inspirations indispensables pas forcément suffisantes

- Apprentissage
 - Choix offline des paramètres → apprentissage statistique
 - Choix online des opérateurs → apprentissage par renforcement
 - Modèles *surrogate* → problème de régression

Quelle interaction apprentissage/optimisation ?

- Évolution biologique
 - Puissance de l'évolution évolution ouverte, nouveauté
 - Mécanismes de l'évolution Morphogénèse, sélection multi-niveaux, métabolisme artificiel
 - Interactions avec l'environnement ... apprentissage et évolution :-)



Plan

- Contexte un peu d'humilité
- L'algorithme le paradigme biologique
- Un exemple jouet sans ordinateur
- Points-clé représentation, diversité, exploitation vs exploration
- Darwinisme artificiel et autres points indépendants de la représentation
- Opérateurs de variation ne dépendant que de la représentation
- Sujets d'actualité paramétrage, représentations
- **Conclusions** et références



Conclusions provisoires

Échecs:

- J'ai essayé en boîte noire, ça ne marche pas...
- J'ai essayé sur un problème facile, c'était ridiculement lent, comparé à ...

Contextes recommandés:

- Problèmes non résolus fonctions chahutées, contraintes chahutées
- Plusieurs optima critères implicites, multi-critères
- Problèmes (très) mal posés validation de l'utilisateur
- A coupler avec des méthodes locales avec mesure

Choix crucial : la représentation et les opérateurs de variation



LIVRES: Si vous n'en lisez qu'un :-)

- A.E. Eiben and J.E. Smith *Introduction to Evolutionary Computing*, Springer Verlag, 2003 (2nd edition soon)
- W. Banzhaf, P. Nordin, R.E. Keller, F.D. Francone, *Genetic Programming — An Introduction On the Automatic Evolution of Computer Programs and Its Applications*, Morgan Kaufmann, 1997.

LIVRES: Les ancêtres

- L.J. Fogel and A.J. Owens and M.J. Walsh, *Artificial Intelligence Through Simulated Evolution*, John Wiley & Sons, 1966
- J. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, 1975
- H.-P. Schwefel, *Numerical Optimization of Computer Models*, John Wiley & Sons, 1981 (Nelle édition 1995).
- D. E. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Addison Wesley, 1989 Classique, mais dépassé
- J. Koza, *Genetic Programming*, I, II (& III), MIT Press, 1992, 1994 (& 1999) ... + videos!



Les référence

- Z. Michalewicz, *Genetic Algorithms+Data Structures = Evolution Programs*, Springer Verlag, 1992–1996. Toujours jeune !
- D. B. Fogel, *Evolutionary Computation, Toward a New Philosophy of Machine Intelligence*, IEEE Press, 1995.
- Melanie Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, 1996.
- T. Bäck, *Evolutionary Algorithms in theory and practice*, New-York:Oxford University Press, 1995.
- Th. Bäck and D.B. Fogel and Z. Michalewicz, *Handbook of Evolutionary Computation*, Oxford University Press, 1997.
- Michael D. Vose, *The Simple Genetic Algorithm: foundations and theory*, MIT Press, 1999.
- D. E. Goldberg, *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*, Kluwer Academic Publishers, 2001.

Journaux



MIT Press

- Evolutionary Computation IEEE
- Transactions on Evolutionary Computation IEEE
- Genetic Programming and Evolvable Hardware Journal Kluwer (now Springer)
- Theoretical Computer Science – TCS-C – Theory of Natural Computing Elsevier
- Natural Computing Kluwer (now Springer)
- Applied Soft Computing Elsevier
- BioSystems Elsevier
- Journal of Heuristics Kluwer (now Springer)
- Journal of Global Optimization Springer
- Evolutionary Intelligence Springer – 2008
- Artificial Evolution and Applications Hindawi – 2008 – *Open Content****
- Swarm Intelligence ..., 2010
- Memetic Algorithms Elsevier, 2010



Conférences

Historiquement

- **ICGA** Int. Conf. on Genetic Algorithms
1987-1997, tous les 2 ans, USA
- **PPSN** Parallel Problems Solving from Nature
Depuis 1990, tous les 2 ans, Europe
- **EP** Annual Conf. on Evolutionary Programming
1992-1998, annuelle, USA
- **ICEC** IEEE Int. Conf. on Evolutionary Computation
1994-1998, annuelle, World-wide
- **GP** Genetic Programming Int. Conf.
1996-1998, annuelle, USA



Aujourd'hui

- GP + ICGA = **GECCO** (annuelle depuis 1999)
Genetic and Evolutionary Computation Conference
Conférence **ACM** depuis 2005
Philadelphie (VA, USA), Juillet 2012
- ICEC + EP + Galesia = **IEEE CEC** (annuelle depuis 1999)
Congress on Evolutionary Computation
Brisbane (Australie), Juin 2012
- **EuroGP, EvoCOP, EvoBIO** et **EvoApplications** (annuels depuis 1998)
Anciennement Evonet Workshops
Vienne, Avril 2013
- **EA – Evolution Artificielle** (tous les 2 ans depuis 94-95) Toulouse 1994, Brest
1995, Nîmes, Calais, Dijon, Marseille, Lille, Tours, Strasbourg, Angers 2011
Bordeaux, Octobre 2013
- ... et toujours **PPSN** ! Taormina, Sept. 2012