

Matis MIS7

Chap. 3 : Evolution et adaptation

Cyrille Bertelle

LITIS
Laboratoire d'Informatique,
du Traitement de l'Information et des Systèmes
Université du Havre

Master 2 Math-Info - Matis



Plan

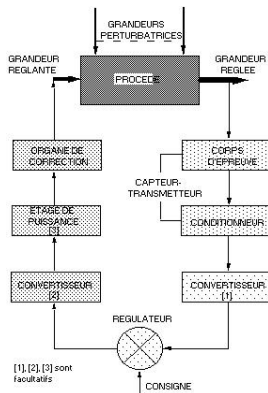
- 1 Méthodes génétiques
- 2 Programmation génétique
- 3 Classifieurs génétiques
- 4 Le modèle Echo de J. Holland

Evolution et Adaptation

Définition

Un système adaptatif s'adapte aux changements de son environnement par des modifications internes. C'est un processus de feedback (rétroaction).

Adaptation : Exemple de régulation



La métaphore biologique

S'inspire du Néo-Darwinisme, synthèse de la théorie de la sélection naturelle (Darwin) et de la théorie de l'hérédité (Mendel)

Définition

adaptation = variation + hérédité + sélection

- *Variation* : Différence entre individus ;
- *Hérédité* : Persistance temporelle ;
- *Sélection naturelle* : Avantage à ceux qui sont “le plus adapté” dans le mécanisme de reproduction.

La métaphore biologique

Une des caractéristiques du vivant : les individus n'ont pas été "programmés" pour répondre à un problème spécifique, mais ils changent en s'adaptant à leur environnement. Comment ?

- Évolution : résultat d'une altération progressive des êtres vivants au cours des générations ;
- Reproduction basée sur le caractère génétique qui subit au cours des générations des recombinaisons et des mutations ;
- Mécanisme de sélection naturelle.

Objectif

construire des systèmes informatiques sur la base de ce modèle.

Principe

La propriété d'adaptation des individus s'interprète comme une recherche d'optimum d'une fonction (la survie).

Principe

On génère initialement un grand nombre d'individus (ou solutions) et l'algorithme les fait évoluer en 3 phases.

- Reproduction
- Croisement
- Mutation

Sélection

- Le processus de sélection est basé sur une fonction d'évaluation (ou fonction objectif) qui correspond à une *performance* de l'individu.
- On en déduit une probabilité pour chaque individu de se reproduire ou de générer des clones (cette probabilité s'appelle *fitness*).
- Finalement, ce processus contribue à produire une population ayant une meilleure adaptabilité (convergence vers l'optimum).

Formalisation

- Les individus sont représentés par des *chromosomes* constitués d'*allèles* (chromosome = chaîne d'informations sur un alphabet fini).



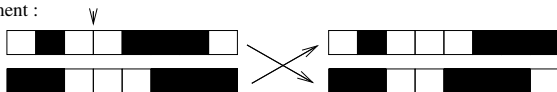
Un chromosome :



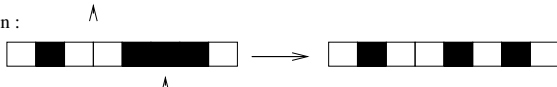
Reproduction/clonage :



Croisement :



Mutation :



Algorithme

Initialiser la population

Calculer le degré d'adaptation de chaque individu

Tant que (non fini **ou** non convergence) **faire**

Reproduction des individus

Appliquer les opérateurs génétiques

Sélectionner les survivants parmi les parents et les enfants

Calculer le degré d'adaptation de chaque individu

Fait

Conclure

Nombreuses variantes

Exemple de base

- Problème : On recherche à optimiser (le maximum) la fonction $f(x) = x^2$ sur l'intervalle $[0..31]$
- Population totale possible : nombre compris entre 0 et 31 en codage binaire sur 5 bits (00000 à 11111).
- Population initiale constituée de 4 individus choisis arbitrairement :

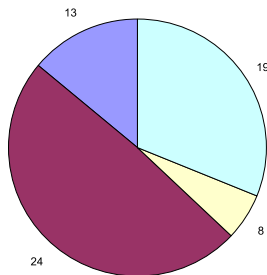
13	de code	01101
24	de code	11000
8	de code	01000
19	de code	10011

- Fonction de performance pour chaque individu : correspond à $f(x) = x^2$. On en déduit la fonction fitness qui est une répartition par pourcentage de ces performances.

Reproduction

Individu	perf.	fitness
13	169	0,14
24	576	0,49
8	64	0,06
19	361	0,31

Roulette biaisée



On tire au hasard 4 nouvelles chaînes parmi les existantes en tenant compte de la valeur de répartition. On obtient, par exemple : 13, 24, 24 et 19.

Croisement

On prend les individus 2 par 2. On coupe les chromosomes à une position aléatoire et on croise les parties coupées

$$\left\{ \begin{array}{l} 13 : 0110 \parallel 1 \\ 24 : 1100 \parallel 0 \end{array} \right. \begin{array}{l} \searrow \nearrow \\ \nearrow \searrow \end{array} \begin{array}{l} 01100 : 12 \\ 11001 : 21 \end{array}$$

$$\left\{ \begin{array}{l} 24 : 11 \parallel 000 \\ 19 : 10 \parallel 011 \end{array} \right. \begin{array}{l} \searrow \nearrow \\ \nearrow \searrow \end{array} \begin{array}{l} 11011 : 27 \\ 10000 : 16 \end{array}$$

Mutation

On opère par tirage aléatoire à partir d'une probabilité de mutation définie initialement. On prend ici 0,05 donc sur les 20 bits des 4 chromosomes, on suppose qu'un seul allèle a été sélectionné et mute.

$$12 : 01\underline{1}00 \rightarrow 01000 : 8$$

Nouvelle génération

8, 21, 27, 16.

Si on somme les performances, on obtient 1490 au lieu de 1170 à la génération précédente

Programmation génétique : définition

Programmation automatique

- Génération automatique de programmes (J. Koza)
- Génération automatique de comportements représentés par des programmes exécutables (P. Angeline).

Approche évolutionniste

Une majorité de travaux s'inspirent de la philosophie des algos génétiques \Rightarrow **Programmation génétique.**

Le modèle de Koza

- Programmes structurés en s-expressions ;
- Définition d'un ensemble de fonctions primitives et de terminaux, seuls constituants autorisés des expressions.
- Type de retour unique pour toutes les expressions.

Constat

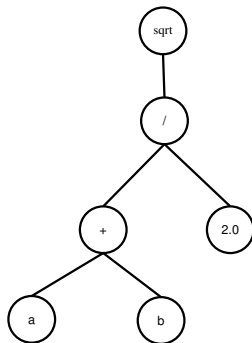
Programmation génétique démarche très voisine des algorithmes génétiques

La donnée est le programme

s-expression ou arbre

Une s-expression :

- Issue du LISP ;
- Représentée par une liste ;
- Notation préfixée ;
- Ex : (sqrt (/ (+ a b) 2.0))
- Implantation possible dans d'autres langages (C++, C, Java ...).
- Programme = Arbre syntaxique



Les mots du langage

- Les terminaux :
 - Ce sont les feuilles de l'arbre syntaxique ;
 - Pseudo-variables contenant les entrées du programme et variables ordinaires ;
 - Constantes, fixées d'après la connaissance préliminaire du problème, ou générées aléatoirement (random ephemeral constants) ;
 - Fonctions sans arguments mais avec effets de bord ;

Les mots du langage

- Les fonctions et les opérateurs :
 - Nœuds internes de l'arbre ;
 - Exemple : fonctions booléennes, arithmétiques, transcendantales, à effet de bord (assignation de variable, déplacement d'un robot, ...), fonctions implantant des structures de contrôle : alternative, boucle, appel de routines, ?
 - Préférer un ensemble de fonctions petit et bien ajusté au domaine du problème, pour réduire l'espace de recherche.
 - Attention à ne pas le réduire trop, sous peine de perdre la possibilité de trouver des solutions intéressantes !

Algorithme

4 étapes principales :

- Générer une population aléatoire initiale à l'aide des fonctions, opérateurs et terminaux qui composent le problème.
- Exécuter chaque programme de la population et affecter à chacun d'entre-eux une fitness en fonction de la façon où il résout le problème.
- Créer une nouvelle population de programmes.
 - Copier les meilleurs programmes ;
 - Créer des nouveaux programmes par mutation ;
 - Créer des nouveaux programmes par reproduction sexuée (crossing-over) ;
- Déterminer quel est le meilleur programme apparu dans chaque génération,

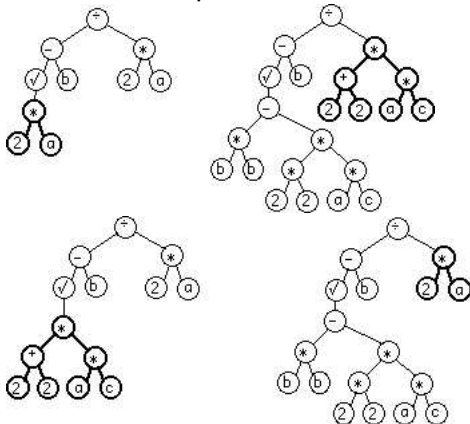
Tirage aléatoire

- Problème un arbre construit aléatoirement a peu (aucune) chance d'être syntaxiquement correct ;
- On peut partir de la racine :
 - Le mot est terminal, on s'arrête ;
 - C'est une fonction ou un opérateur, on tire ses descendants jusqu'à obtenir un arbre complet ;
- On impose une profondeur minimale pour éviter les arbres réduits à une variable ;
- On impose une taille maximale ;
- ...

Crossing-Over

Deux parents différents

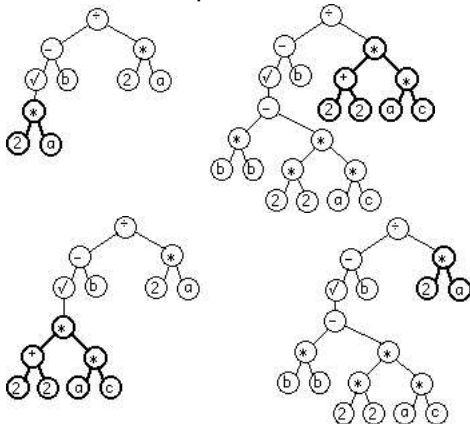
- Choisir les nœuds aléatoirement
- Échanger les sous-arbres.



Crossing-Over

Deux parents différents

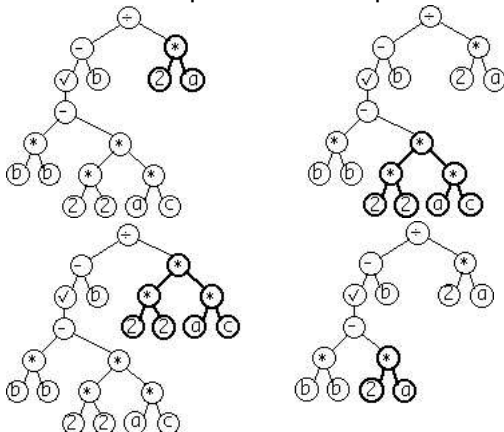
- Choisir les nœuds aléatoirement
- Échanger les sous-arbres.



Crossing-Over

Deux parents identiques

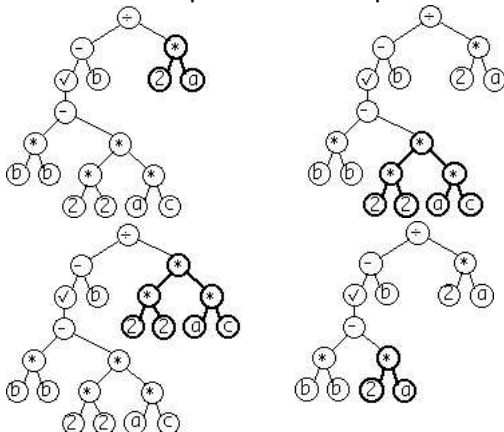
- Choisir les nœuds aléatoirement
- Échanger les sous-arbres.



Crossing-Over

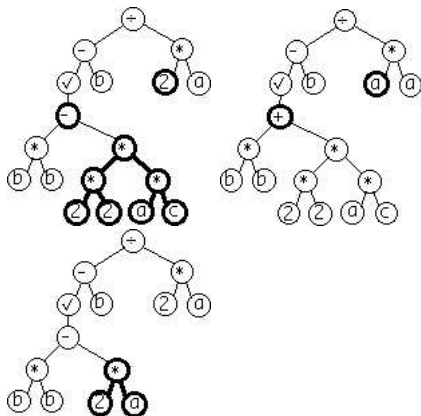
- Choisir les nœuds aléatoirement
- Échanger les sous-arbres.

Deux parents identiques



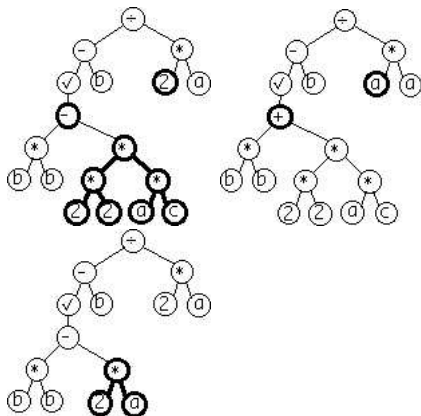
Mutation

- Deux types de mutations sont possibles
- On remplace uniquement des nœuds de l'arbre.
- Un sous-arbre est remplacé par un sous-arbre généré aléatoirement.



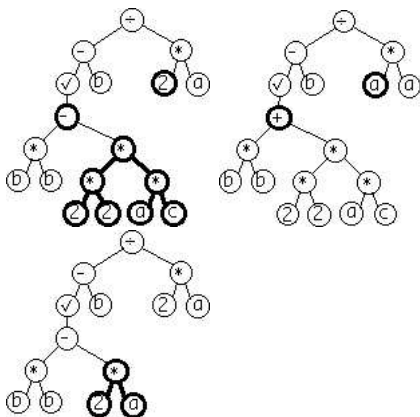
Mutation

- Deux types de mutations sont possibles
- On remplace uniquement des nœuds de l'arbre.
- Un sous-arbre est remplacé par un sous-arbre généré aléatoirement.



Mutation

- Deux types de mutations sont possibles
- On remplace uniquement des nœuds de l'arbre.
- Un sous-arbre est remplacé par un sous-arbre généré aléatoirement.



Classifieurs génétiques

Système capables d'apprendre des règles pour optimiser leurs performances dans un environnement perçu sous forme de messages.

Il est constitué de 3 parties principales :

- Un système de règles ;
- Un système d'attribution de crédit ;
- Un algorithme génétique.

Classifieurs génétiques

Systeme de règles

Une règle peut s'écrire sous la forme :

$\langle \textit{cond} \rangle : \langle \textit{resultat} \rangle$

exemple :

01#0 : 0111

- Sélection d'un message binaire de 4 bits commençant par 01 et se terminant par 0 ;
- Si acceptation d'un message, on renvoie le message 0111.

Classifieurs génétiques

Système d'attribution de crédits

Système de compétition entre les règles qui possèdent chacune un crédit variable. Lorsqu'un message arrive :

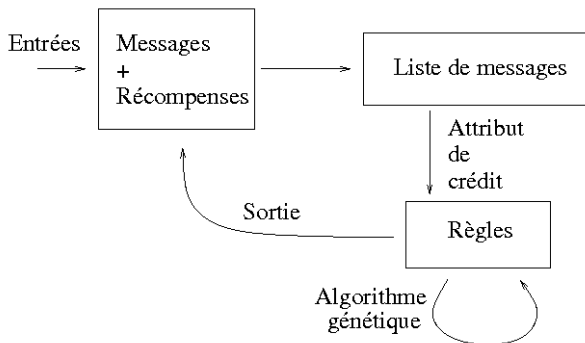
- Les règles qui peuvent le sélectionner font une offre proportionnelle à leur crédit actuel ;
- Le système d'attribution sélectionne les règles qui offrent le plus et les active ;
- Une règle activée renvoie son message de résultat et offre une récompense correspondant à son crédit pour les règles qui auront sélectionnées son message émis.

Classifieurs génétiques

Algorithme génétique

- Un algorithme génétique est mis en place sur la population des règles ;
- La fonction de performance est directement le crédit associé à chaque règle ;
- On sélectionne, reproduit, croise et fait muter les règles les plus performantes.

schéma de fonctionnement d'un classifieur génétique



Le modèle Echo de J. Holland

- Ensemble d'agents situés dans un environnement nutritif ;
- Comportement des agents déduit de leur génomes ;
- Interactions entre agents de nature variée : reproduction, compétition, coopération ;
- Performance/fitness des agents : leur capacité de survie.

Caractéristiques

- Temps discret et monde représenté par une grille torique ;
- Déplacement possible d'un agent dans une case voisine à chaque pas de temps ;
- Interactions possibles entre agents que s'ils sont sur des cases voisines ;
- Les comportements des agents sont contrôlés par des gènes spécifiques appelés *conditions* ;
- Les agents ont des signes extérieurs apparents ou "tags" représentatifs de leur état social. Ces tags sont représentés par des gènes spécifiques ;

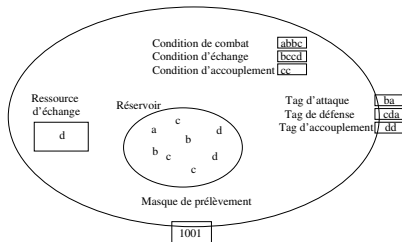
Caractéristiques

- Reproductions sexuelles ou asexuelles permettant l'évolution du génotype ;
- Chaque site produit une quantité de ressources spécifiques à chaque pas de temps. Il y a 4 variétés de ressources : a, b, c et d sur lesquelles est basée la survie de chaque agent ; Ce sont les “molécules” de base du monde Echo. Elles sont les constituants du génôme des agents et forment pour cela des chaînes (simple chaînage, pas de réaction chimique). Il n'y a pas d'interaction entre les gènes d'un génome.

Agents

Constitues de 2 éléments :

- Chromosome codant le comportement et l'apparence (tag) des agents : ce sont des chaînes de ressources (appelés aussi *nucléotide*) ;
- Un réservoir de ressources stockées dans chaque agent.



Chromosomes d'un agent

Constitués de 8 gènes au total

- 3 gènes de tag : tag d'attaque, tag de défense et tag d'accouplement ;
- 3 gènes de conditions (comportementales et sociales) : condition de combat, condition de négociation et condition d'accouplement ;
- 2 gènes de tags pour les échanges de ressources :
 - Ressources d'échanges : type de ressource que peut donner l'agent à un autre au cours d'une négociation ;
 - Masque de prélèvement : masque binaire d'un bit par type de ressources. Une ressource peut être acquise dans l'environnement que si le bit correspondant vaut 1.

D'autres possibilités d'échanges de ressources : suite à un combat ou par transmission filiale.

Auto-reproduction

Lorsque la quantité de ressources du réservoir de l'agent est suffisamment importante,

- il y a génération d'un clone auquel on transmet une quantité des ressources ;
- le chromosome du clone est obtenu par copie + mutation.

Mutation

Dans Echo, les mutations agissent sur les gènes.

Sur les tags et les conditions, 3 types de mutation sont permis :

- suppression du dernier nucléotide du gène ;
- insertion d'un nucléotide à la fin du gène ;
- modification d'un nucléotide dans le gène.

Sur les ressources d'échanges, seule la modification est permise. Sur le masque de prélèvement, la mutation correspond à la modification d'un bit.

Interactions entre les agents

Elles sont contrôlées par les tags et les conditions. Deux agents proches peuvent interagir :

- par indifférence ;
- par échange ;
- par reproduction et génération de nouveaux agents avec croisements génétiques.

Les interactions sont contrôlées par un mécanisme de préfixe : elles ne peuvent se produire que si la condition est le préfixe du tag associé.

Interactions entre les agents

Agent A	Agent B
Condition de combat : <u>ab</u>	Tag d'attaque : <u>abb</u>
Tag d'attaque : <u>bcd</u> Condition d'échange : <u>ab</u>	Condition d'échange : <u>b</u> Tag d'attaque : <u>abb</u>
Tag d'accoupl. : <u>dcd</u> Condition d'accoupl. : <u>a</u>	Condition d'accoupl. : <u>dc</u> Tag d'accoupl. : <u>adb</u>

- L'agent A peut attaquer l'agent B ;
- Les deux agents peuvent procéder à des échanges ;
- Les deux agents peuvent s'accoupler.

Combat

- Les combats se produisent si les conditions sont réalisées et avec une certaine probabilité ;
- Dans le cas d'un combat, on compare 1 à 1 les nucléotides du tag d'attaque d'un agent avec le tag de défense de l'autre afin d'obtenir un score. La probabilité d'être le vainqueur correspond au rapport de son score sur la somme des 2 scores.
- Le vainqueur remporte toutes les ressources du vaincu, y compris celles qui constituent son génôme. Le vaincu est retiré du monde.

Accouplement

Reproduction par croisement de deux sous-chaînes :

$$\left\{ \begin{array}{l} \text{agent } A : ab||bcbd||cd \\ \text{agent } B : b||bd||ca \end{array} \right. \begin{array}{l} \searrow \nearrow \\ \nearrow \searrow \end{array} \begin{array}{l} ab||bd||cd \\ b||bcbd||ca \end{array}$$

Les longueurs des gènes des ressources d'échange et du masque de prélèvement doivent être constantes.

Cycle de base

- 1 interaction entre des agents sélectionnés ;
- 2 alimentation des agents par distribution des ressources de l'environnement ;
- 3 mort accidentelle des agents ;
- 4 renouvellement des ressources de l'environnement.