

## Algorithmes génétiques - Introduction

### Question

Comment l'évolution produit-elle des organismes de mieux en mieux adaptés à leur environnement?

### Remarques

- "Trouver un organisme dont la qualité d'adaptation à un environnement donné est maximale" peut être vu comme un problème d'optimisation. Ce problème est toutefois d'une grande complexité, et il devient encore plus complexe si l'on suppose que l'environnement nous est (partiellement ou entièrement) inconnu, qu'il est incertain et qu'il évolue de façon dynamique.
- Il faut alors que des informations sur l'environnement soient recueillies (exploration) et utilisées à bon escient (exploitation) afin de produire des organismes qui sont mieux adaptés.

### En résumé

- Exploration: permet de recueillir des informations sur l'environnement.
- Exploitation: permet d'utiliser les informations recueillis.
- L'adaptation (ou évolution) est le fruit d'un compromis entre l'exploration et l'exploitation.

## Adaptation

### Caractéristiques

- Modification progressive d'une ou plusieurs structures (*chromosomes*).
- Opérateurs (*mutation, recombinaison*) assurant la modification des structures.
- Plan d'adaptation (*algorithme génétique*) qui produit des organismes qui sont mieux adaptés à leur environnement.
- Le plan détermine dans quel ordre les opérateurs doivent être appliqués aux structures afin de produire des structures plus performantes.
- Au début, on ne possède que des informations incomplètes sur l'environnement, et sur les structures qui seraient les mieux adaptées à cet environnement. Afin de réduire cette incertitude, le plan doit tester la performance de nouvelles structures dans l'environnement.

### En résumé

- Le mécanisme d'adaptation repose sur les éléments suivants
  - un environnement
  - des structures
  - des opérateurs
  - un plan d'adaptation
  - une mesure de performance ou d'adaptation

## Adaptation

### Obstacles

- L'ensemble des structures possibles peut être très grand.
- Chaque structure peut être très complexe et se diviser en de nombreuses composantes. Il devient alors difficile de déterminer quelles sont les composantes responsables d'une bonne performance (*building blocks*).
- La mesure de performance peut elle-même être une fonction très complexe (multidimensionnelle, multimodale, non linéaire, etc.) et comporter de nombreux paramètres interdépendants.
- Enfin, la mesure de performance varie typiquement dans le temps, dû à un environnement dynamique.

## Un système adaptatif artificiel simple

- Grille de senseurs avec  $a$  rangées et  $b$  colonnes.
- Chaque senseur est activé si la quantité de lumière qui l'atteint dépasse un certain seuil; autrement le senseur demeure inactivé.
- Configuration des senseurs:  $s_i(t)$ ,  $i=1..ab$ , et  $s_i(t) = 1$  ou  $0$ .
- Si on dénote  $C$  l'ensemble des configurations possibles, alors  $|C| = 2^{ab}$
- On suppose qu'un certain sous-ensemble  $C_1$  de  $C$ , correspond aux instances d'une certaine image (e.g., une forme  $X$ , incluant toutes les rotations et translations possibles)
- Le système adaptatif est un *perceptron*:

Il traite chacun des signaux d'entrée  $s_i(t)$  en le multipliant par un poids  $w_i$ , et il somme le tout:

$$\sum_{i=1..ab} w_i s_i(t).$$

Si cette somme excède un certain seuil  $T$ , le système indique que la configuration appartient à  $C_1$ , sinon elle appartient au complément  $C_0$  (mais le système peut être dans l'erreur).

- Le système doit modifier ses poids (s'adapter) de façon à permettre la meilleure discrimination possible entre les configurations qui appartiennent à  $C_1$ , et celles qui appartiennent au complément  $C_0$ .
- Mesure de performance: à chaque vecteur de poids, on associe la proportion de configurations correctement classifiées :

$$f: W^{ab} \rightarrow [0,1].$$

## Un système adaptatif artificiel simple

- Stratégie non adaptative: on génère aléatoirement un certain nombre de vecteurs de poids, et on retient le meilleur.
- Stratégie adaptative: l'ordre dans lequel les vecteurs de poids sont générés et testés dépend de la performance des vecteurs de poids précédents.

### Exemple de stratégie adaptative

On suppose que les valeurs des poids doivent être choisies dans un ensemble ne contenant que  $k$  valeurs différentes (e.g., si  $k = 10$ ,  $a = b = 10$ , alors  $|W^{ab}| = 10^{100}$ )

On suppose que les  $k$  valeurs sont triées en ordre croissant, et sont uniformément espacées.

On considère un algorithme qui itère sur toutes les configurations possibles et qui, pour chaque configuration, obéit à la règle suivante:

0. Si la configuration courante est correctement reconnue, alors on ne fait rien.
1. Si la configuration courante est affectée à  $C_0$ , alors qu'elle aurait dû être affectée à  $C_1$ :  
pour tout  $i$  tel que  $s_i(t) = 1$ , remplacer le poids  $w_i$  courant par le poids suivant  $w_i + \Delta$ . Les poids associés aux senseurs inactifs demeurent inchangés.
2. Si la configuration courante est affectée à  $C_1$ , alors qu'elle aurait dû être affectée à  $C_0$ :  
pour tout  $i$  tel que  $s_i(t) = 1$ , remplacer le poids  $w_i$  courant, par le poids suivant  $w_i - \Delta$ . Les poids associés aux senseurs inactifs sont inchangés.

## Un système adaptatif artificiel simple

- Environnement: sous-ensemble de configurations  $C_1$
- Structures: vecteurs de poids
- Opérateurs: modification des vecteurs de poids (opérateur 1: même poids; opérateur 2 : poids plus petit; opérateur 3: poids plus grand; )
- Plan: algorithme de décision permettant de choisir un opérateur de modification des poids.
- Mesure de performance:  $f: W^{ab} \rightarrow [0,1]$

## Un système adaptatif naturel complexe

### Système génétique

- Les chromosomes (structures) sont constitués de gènes. Chez les vertébrés, chaque chromosome est constitué de dizaines de milliers de gènes, chacun capable de prendre différentes valeurs (allèles). Le nombre de structures possibles est donc immense.
- Génotype: code génétique
- Phénotype: expression physique du code génétique (organisme)
- Établir une relation directe entre une caractéristique physique de l'organisme et un gène ou groupe de gènes particulier est très complexe (dû à de nombreuses interactions entre les gènes).

## Un système adaptatif naturel complexe

- Structures: chromosomes
- Opérateurs: reproduction (copie), mutation, croisement
- Si on utilise seulement la reproduction, on maintient la performance (i.e., on conserve les acquis), mais on ne peut espérer d'évolution. Si on utilise seulement des mutations aléatoires, on crée de nouveaux organismes, mais ceux-ci ont de fortes chances d'être moins performants que leurs prédécesseurs.
- D'où l'utilité d'un opérateur de croisement qui permet de redistribuer des groupes de gènes performants au sein de la population.

### Exemple de Plan

Répéter pour un nombre fini de générations:

- Phase 1.** Les organismes ayant les mesures de performance les plus élevées se reproduisent (i.e. le taux de reproduction est proportionnel à la performance).
- Phase 2.** Les opérateurs de croisement et de mutation sont ensuite appliqués aux copies, de façon à obtenir des descendants qui diffèrent de leurs parents.

### Remarques

- Les groupes d'allèles associés aux organismes les plus performants auront tendance à se propager dans les générations suivantes.
- Si l'association entre un groupe d'allèles et la mesure de performance est fortuite, cette association tendra à disparaître dans les générations suivantes.
- Ainsi, on peut espérer qu'au fil des générations, seuls les groupes d'allèles qui sont véritablement responsables d'une meilleure performance, seront conservés.
- De plus, en combinant sur un seul chromosome des groupes d'allèles performants provenant de deux chromosomes différents (via l'opérateur de croisement), on peut espérer créer des descendants qui seront plus performant que leurs parents.