

L2 – Vie Artificielle

Alexandre Allauzen – Michèle Sebag

LIMSI – LRI

5 dec. 2013

Overview

Automates cellulaires, applications

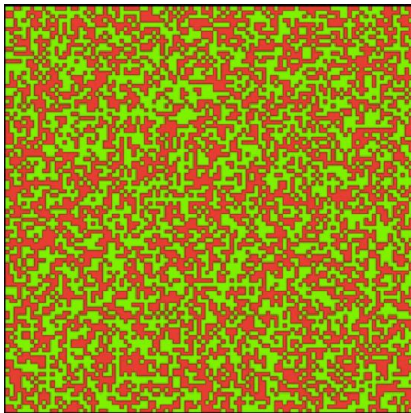
Robotique évolutionnaire
Généralités

État de l'art
Un critère
Evolution
Évolution et Transfert d'information
Expériences

Thomas Schelling

Un économiste errant

- Prix Nobel 2005 for *having enhanced our understanding of conflict and cooperation through game-theory analysis.*



Une population aléatoire de rouges et de verts.

Thomas Schelling, 2

Paramètres

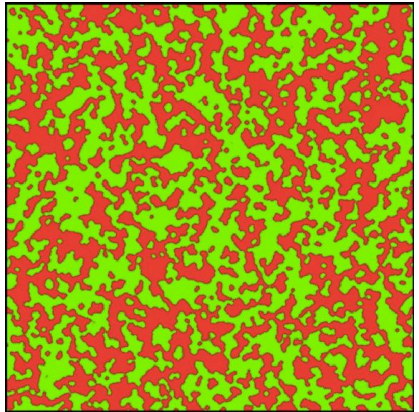
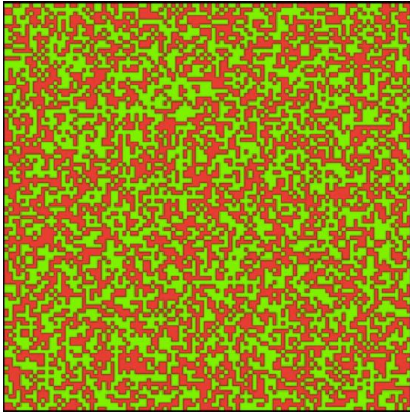
- ▶ Population n
- ▶ Voisinage: rayon w
- ▶ Seuil de tolérance τ
- ▶ Distribution ρ

Règle du jeu

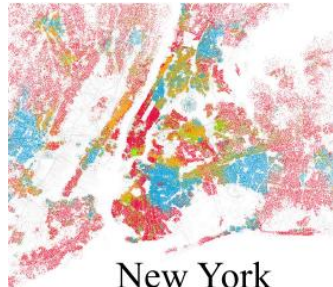
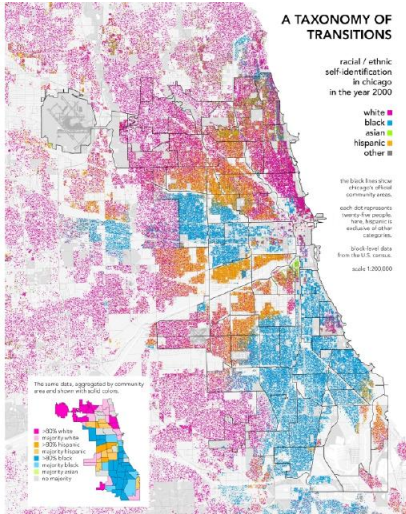
- ▶ Si dans votre voisinage, la proportion de gens différents est $> \tau$, partez ! (saut vers une position aléatoire).

Thomas Schelling, 3

Une ségrégation se produit



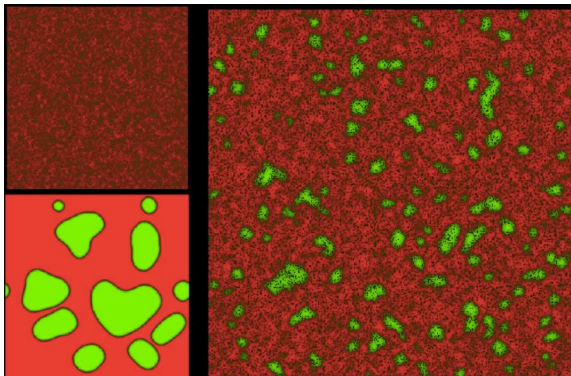
Dans la vraie vie



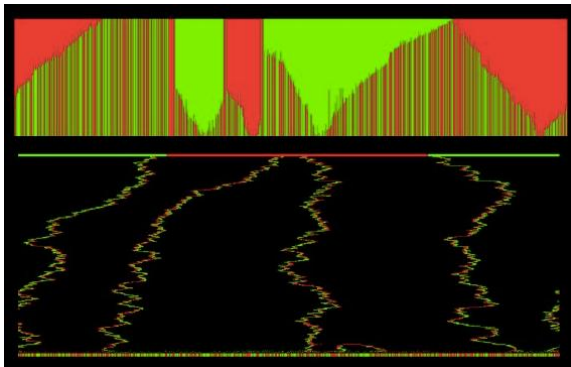
Thomas Schelling, 4

Observer / prédire

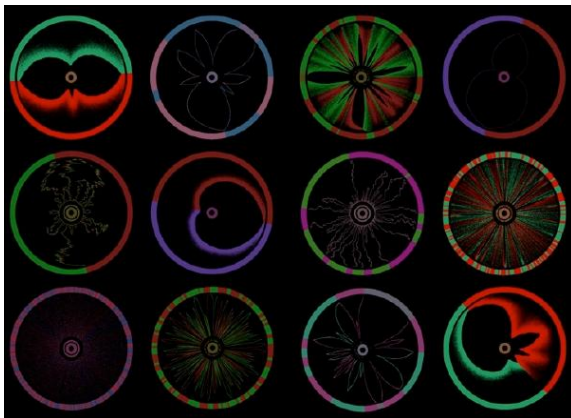
- ▶ Le type de ségrégation (caractériser la situation d'équilibre)
- ▶ Savoir quand on y arrive (temps de convergence).



Le cas 1D



Le cas du cercle



A propos

- ▶ Souvent des logiciels multi-agents (Repast, Net-logo, online java applets)
- ▶ Lents !
- ▶ C++:
 - ▶ Graphique rapides, OpenGL
 - ▶ Tableaux dynamiques: coûteux en temps
 - ▶ Tableaux statiques: prise en compte des entrées vides
 - ▶ Optimisation du compilateur

Qu'est-ce que c'est ?

1. Systèmes dynamiques
2. Structure spatiale
3. Propriété de Markov (d'ordre 1) : i.e. le futur dépend uniquement de l'état courant
4. Pas réversible
5. Beaucoup d'états stationnaires.

Exemples de résultats

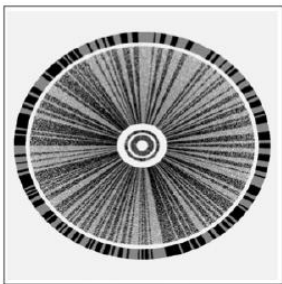
Thm 1

Si $\tau = 1/2$, avec une distribution initiale uniforme, en dimension 1, les blocs homogènes sont de longueur polynomiales en w .

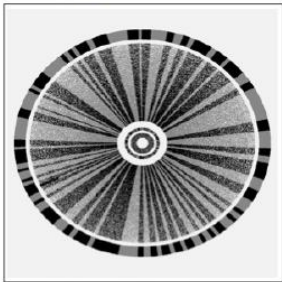
Thm 2

Le temps moyen pour arriver à l'équilibre est $O(w^2)$

Quels outils ?



$w = 1500, \tau = 0.50$



$w = 3000, \tau = 0.50$

$$\sqrt{n} \left(\left(\frac{1}{n} \sum_{i=1}^n X_i \right) - \mu \right) \xrightarrow{d} N(0, \sigma^2).$$



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Summary

- > If $\tau < 0.353$ no segregation
- > If $0.353 < \tau < 1/2$ exponential segregation
- > If $\tau = 1/2$ polynomial segregation (Brant et. al.)
- > If $\tau > 1/2$ total segregation

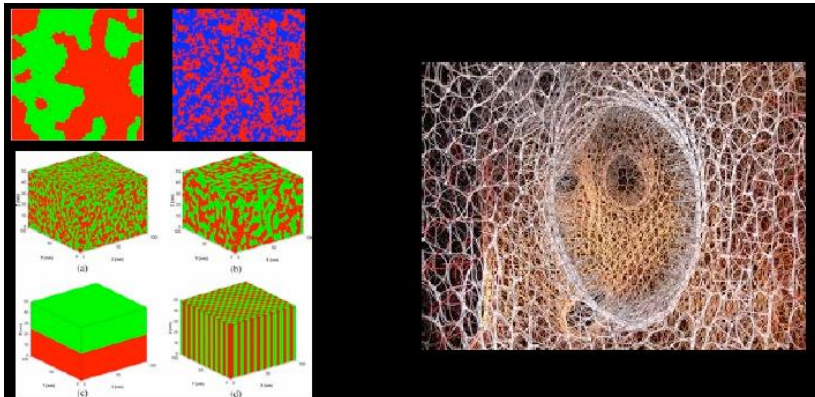
Paradox: in the interval $[0.353, 1/2]$ increased tolerance leads to increased segregation

(Schelling: Micromotives and Macrobehaviour)

Vie artificielle

Les domaines sont différents

- ▶ Physique: modèles d'Ising
- ▶ Biologie: morphogenèse
- ▶ Sociologie/économie



Vie artificielle, 2

Phénomènes similaires:

Effets macroscopiques des choix microscopiques

Plus de détails

- ▶ L'histoire des gauchers
- ▶ <http://www.youtube.com/watch?v=EbN596buqKo>

Crédit:

George Barmpalias (Chinese Academy of Sciences), Richard Elwes (University of Leeds), Andy Lewis-Pye (London School of Economics).

Overview

Automates cellulaires, applications

Robotique évolutionnaire
Généralités

État de l'art

Un critère

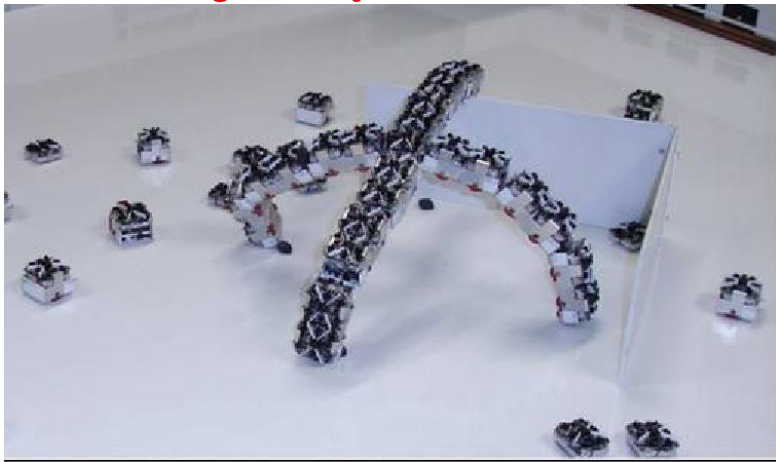
Evolution

Évolution et Transfert d'information

Expériences

Contexte: Robotique en essaim

SYMBRION Integrated Project



<http://symbrion.org/>

Contexte: Robotique en essaim, 2



Objectif et Motivations

1. Robotique Evolutionnaire
2. Apprentissage de contrôleur autonome
3. Quel critère d'optimisation ?
 - ▶ Optimisable bootstrap
 - ▶ Efficace
 - ▶ Embarquable

État de l'art

- ▶ Robotique Évolutionnaire

Nolfi-Floreano 2000

- ▶ Evaluation par simulateur / in situ

Jacobi 96, Lipson et al. 07

- ▶ Espaces de recherche

Gomez-Miikkulainen 97

- ▶ Tâches

Lanzi 00, Godzik et al. 04, Hartland et al. 09

- ▶ Critères

Lehman-Stanley 08, Mouret-Doncieux 09

Quelle fitness pour la Robotique Évolutionnaire

Dilemme du bootstrap (amorçage)

- ▶ fitness précise optimum == comportement voulu.
- ▶ optimisable comportement **très** mauvais < mauvais.



Approches

- ▶ Séquences de fitness de Garis 90
Harvey et al. 94, Gomez et al. 97
- ▶ Diversité Gomez et al. 07.
- ▶ Multi-objectif Mouret-Doncieux 09.

Quelle fitness (2)

Résister à l'opportunisme de l'évolution

- ▶ Evitement d'obstacles: en restant immobile...
- ▶ Vitesse: Aller-retours rapides dans un couloir...

La fitness proposée par Nolfi & Floreano (2001)

$$\mathcal{F} = \sum_{i=1}^n V * (1 - \Delta V) * (1 + \text{SensorMax})$$

Quelques autres approches

- ▶ Apprentissage par renforcement Sutton et Barto 98
- ▶ Apprentissage par imitation Bain-Sammut 95; Billard et al. 05
- ▶ AR Inverse Abbeel-Ng 04
- ▶ Théorie des jeux Syed-Schapire 07, 08

Le saut de la réalité (reality gap)

Apprentissage en simulation (in silico)

- ▶ Simulateur du robot/environnement (modèle du monde)
- ▶ Simulateur cher ou imprécis (ou les deux)
- ▶ Modèle et bruit des senseurs/moteurs

peaufinage in situ requis

Lipson et al. 06

Apprentissage in situ

- ▶ Le seul, pour les roboticiens
- ▶ Cher en temps (+ risque matériel)
- ▶ Pas de vérité terrain
- ▶ Besoin d'une optimisation frugale (tps, mémoire)

Overview

Automates cellulaires, applications

Robotique évolutionnaire
Généralités

État de l'art

Un critère

Evolution

Évolution et Transfert d'information

Expériences

Fitness Entropique

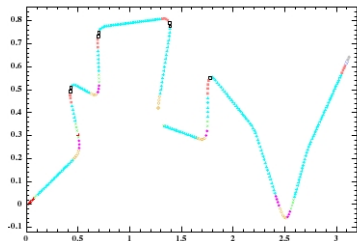


Figure: Trajectoire du robot

Log: T pas de temps $\times \{8 \text{ senseurs}, 2 \text{ moteurs}\}$

Information disponible “for free”

- ▶ Le robot est une source d'informations.
- ▶ Log robotique :

$$t \rightarrow \text{sensations}[t], \text{moteurs}[t]$$

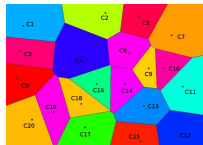
Fitness Entropique (2)

Conjecture

Un comportement riche $\equiv$ forte diversité des données sensori-motrices

Approche

- Fitness: Entropie
- Requis: Discrétiser les données sensori-motrices
- Algorithme: Clustering



Trajectoire $-i$
 $C_1 C_3 C_6 C_{17} C_2 \dots$

Clusters dans l'espace sensori-moteur

Clustering

k-Means *Alg. aléatoire: garder la meilleure solution sur p tirages*

1. Tirer k -centres c_i (sites de Voronoï)
2. Partitionner les points $x_t, t \in 1 \dots T$ (cellules de Voronoï)

$$x_t \rightarrow c_i = \operatorname{argmin}\{d(x_t, c_j), c_j \in \mathcal{C}\}$$

3. Centre $c_i \rightarrow$ moyenne des x_t ds la cellule i
4. Goto 2

ϵ -Means

1. Init : $\mathcal{C} = \{\}$ *Init liste des centres*
2. Pour $t = 1$ to T *boucle sur les points*
 - ▶ $c_i = \operatorname{argmin}\{d(x_t, c_j), c_j \in \mathcal{C}\}$
 - ▶ si $d(x_t, c_i) < \epsilon$, incrémenter $n(c_i)$
 - ▶ sinon, $\mathcal{C} = \mathcal{C} \cup \{x_t\}$; $n(x_t) = 1$ *créer un autre centre*

Fitness Entropique

Définition

- ▶ Contrôleur X + environnement $\rightarrow$ Trajectoire
- ▶ Clustering des données du log
- ▶ Centre $c_i \rightarrow p_i = \frac{1}{T} |\{x_t \in C_i\}|$

Log

$$\mathcal{F}(X) = - \sum_{i=1}^n p_i * \log(p_i)$$

Propriétés :

- ▶ $\mathcal{F}$ croît avec le nombre n de clusters
- ▶ pour n donné, $\mathcal{F}$ maximale si $p_i = 1/n$

Discussion

Avantages de $\mathcal{F}$

- ▶ Pénalise l'inaction $\equiv$ un seul état
- ▶ Peu sensible au bruit (senseurs, capteurs)
- ▶ Favorise la diversité sensori-motrice

Limites

- ▶ Récompense comportements périodiques dégénérés
- ▶ Cas $k - means$:
 - ▶ Fitness stochastique non reproductible.
 - ▶ pas de garantie sur le volume sensori-moteur exploré.

Overview

Automates cellulaires, applications

Robotique évolutionnaire
Généralités

État de l'art

Un critère

Evolution

Évolution et Transfert d'information

Expériences

Évolution darwinienne

seule transmission: génotype

Evolution classique

- ▶ Evolution embarquée possible = fitness frugale
- ▶ Algorithme (1+5)ES
 1. Un parent génère 5 enfants
 2. Evaluation des 5 enfants
 3. Nouveau parent: le meilleur (parent, enfants)

Résultats préliminaires

- ▶ Si l'arène est assez complexe *condition générale...*
- ▶ Si la durée de vie de l'individu est assez longue,
Apparition d'individus compétents, explorateurs, bonne survie du robot

Évolution "culturelle"

transmission: génotype + expérience

Capitalisation des acquis

- ▶ Expérience d'un contrôleur $\rightarrow$ les centres des clusters
- ▶ Information frugale : peut être transmise aux enfants
- ▶ Condition: info reproductible ϵ -Means

Fitness "culturelle"

- ▶ Etant donné l'histoire du passé liste des (c_i, p_i)
- ▶ Lancer ϵ -Means ($\mathcal{C} = \{c_i\}$, $\log(\text{contrôleur } X \text{ courant})$)
- ▶ Mise à jour de p_i
- ▶ $\mathcal{F}$ mesure l'apport d'information par rapport au passé:

$$\mathcal{F}(X) = IQ(p_i)$$

- ▶ Récompense le fait de trouver de nouveaux états, ou visiter un ensemble d'états rares

Overview

Automates cellulaires, applications

Robotique évolutionnaire
Généralités

État de l'art

Un critère

Evolution

Évolution et Transfert d'information

Expériences

Validation Expérimentale

Objectifs de l'expérimentation

- ▶ Tester $\mathcal{F}$: frugalité, résistance au bruit, surmonter le pb de bootstrap
- ▶ Lien entre entropie et diversité des contrôleurs/comportements,
- ▶ Lien entre entropie et exploration de l'environnement
- ▶ Les bons individus exhibent-ils des comportements intéressants ?

Etude de sensibilité

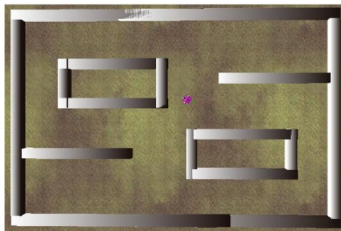
- ▶ Impact de la fitness (individuelle ou avec passé)
- ▶ Impact de l'environnement (simple ou complexe)

Protocole expérimental

Espace de recherche: réseaux de neurones

- ▶ 8 entrées, 2 sorties
- ▶ 1 couche cachée
- ▶ Init: poids ds $[-1,1]$

Deux arènes: 1 simple, 1 complexe



Résultat: moyenne sur 15 runs indépendants.

Évolution darwinienne, k -Means

Fitness individuelle: pas de passé

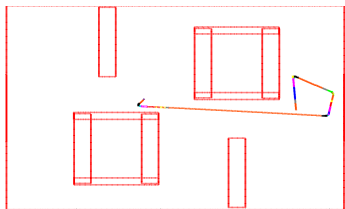


Figure: $k=15$, arène simple

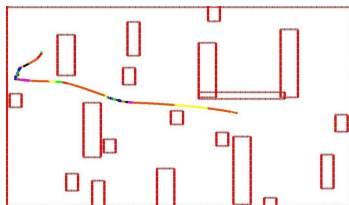


Figure: $k=30$, arène complexe

Trajectoires d'individus représentatifs après 200 générations.

Commentaires

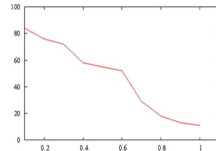
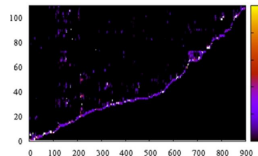
- ▶ Clustering k -Means et ϵ -Means comparables
- ▶ Complexité environnement : facteur favorable
- ▶ Ajuster k et ϵ

Évolution culturelle

Evolution des découvertes

Abscisse: cluster/états (ordonnés par instant de découverte)

Une ligne: un individu: états visités

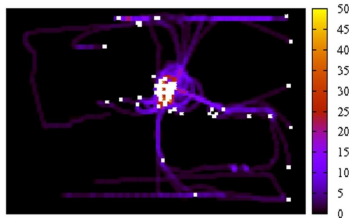


Influence de ϵ

- ▶ Augmentation exponentielle du nombre d'états avec ϵ
- ▶ ϵ : trop grand ou trop petit, pénalise l'évolution
tout se ressemble, tout est différent

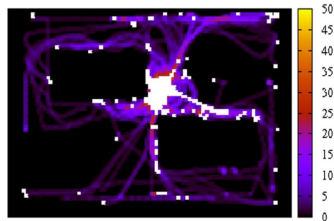
Évolution culturelle (2)

Exploration de l'arène



après 50 gen.

Zones noires: non visitées



après 200 gen.

zones d'obstacles

Les robots vont partout: exploration collective

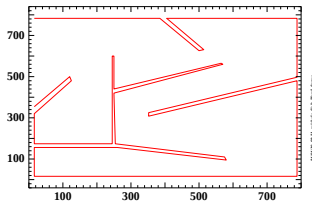
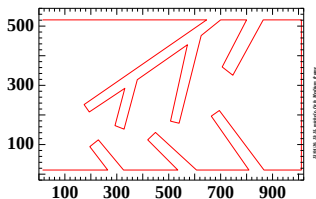
Validation

Protocole experimental

Robot = Cortex M3, 8 infra-red sensors, 2 motors.

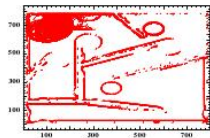
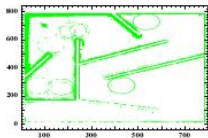
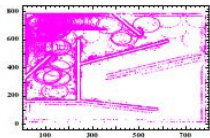
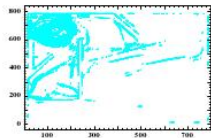
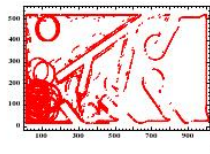
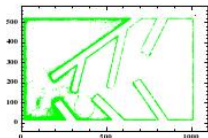
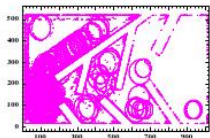
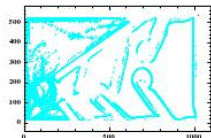
Controlleurs = Multi-layer Perceptron, 10 neurones cachés.

Deux arènes (moyenne et dure)



Validation, 2

Plot points visités 10 fois ou plus par les meilleurs 100 individus.



Nolfi & Floreano

Lehman & Stanley

Curiosity

Discovery

PPSN 2010

Discussion

The approach fits the requirements: computable on-board for swarm robot architecture; no need of prior knowledge/ground truth.

The robot only aims at gathering information, as opposed to, building a world model. (unsupervised as opposed to supervised learning).

The robot is rewarded for what it gets (sensory information) not what it does (e.g. going fast and not circling).

Caveat: needs a stimulating environment.

Vision and Perspectives

Requirements



- X No simulation
- X On-board training
 - ▶ Frugal (computation, memory)
 - ▶ No ground truth
- ? Providing “interesting results”

Human Robot Interaction

- ▶ Sensori-motor states can be interpreted, enabling the designer to take control, e.g. for emergency situations.
- ▶ Display trajectories: The designer can emit preferences
Interactive optimization

Preference learning: associate “values” to sensori-motor states