Complex systems and Agent-based modelling

Jonathan Rault

E.N.S.

jrault@ens.fr

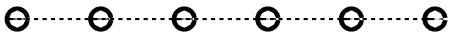
November 7, 2014

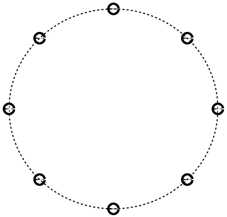
Plan

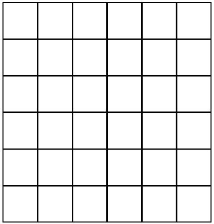
Modèles d'automates cellulaires

- Première apparition de tels models dans les années 1940
- Un automate cellulaire est un assemblage régulier de cellules (identiques) intéragissants entre elles localement
  - Modèles discrets en temps et en espace.
  - Une cellule est caractérisée à chaque instant $t \in \mathbb{N}$ par son état, appartenant à un ensemble fini d'états possibles.
  - L'état à l'instant $t + 1$ d'une cellule dépend de son etat à l'instant t ainsi que de ceux de ses voisins : la dynamique est synchrone et sans mémoire.
  - La règle de transition est indépendante de la position de la cellule dans le réseau, à part éventuellement sur les bords du domaine.

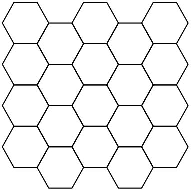
Exemples de réseaux cellulaires





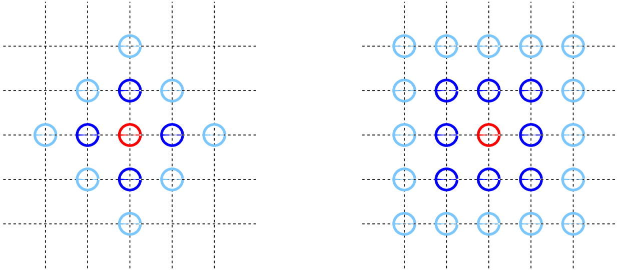


- réseau régulier de cellules
- avec / sans "bords" :
  - Dimension 1 : ligne / cercle
  - Dimension 2 : grille / tore
- grille carrée, triangulaire, hexagonale etc...



Exemples de voisinages

- Cas d'une grille carrée (dimension 2)



(a) voisinage de Von Neumann (b) voisinage de Moore

- Bleu foncé (resp. clair) : voisins à une distance de 1 (resp. 2)

Un premier exemple : modèle de l'opinion

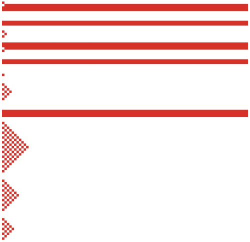
- assemblage cellulaire de dimension 1 (ligne ou cercle)
- deux états possibles pour chaque cellule (état binaire : 0 ou 1)
- choix du voisinage : les deux cellules les plus proches
  - $2^3 = 8$ configurations à envisager pour la règle locale de transition : 2 états courants possibles pour la cellule considérée et 2 etats pour chacune de ses voisines.
- exemple de règle locale :

état courant	1	1	0	0	1	1	0	0
états des voisins	1 − 1	1 − 0	1 − 1	1 − 0	0 − 1	0 − 0	0 − 1	0 − 0
configuration locale	111	110	101	100	011	010	001	000
état suivant	1	1	1	0	1	0	0	0

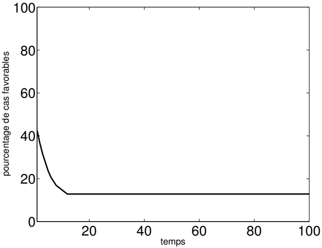
- règle de la majorité (règle 232 = 11101000 en binaire)
- nécessite l'ajout d'une règle aux bords du domaine dans le cas d'une configuration ligne

Un premier exemple : simulation

- configuration initiale à définir : état des cellules à l'instant $t = 0$.
- ici la régle est déterministe $\Rightarrow$ l'état des cellules à un certain instant $t \in \mathbb{N}$ ne dépend que de la configuration initiale.
    - On peut imaginer des règles de transitions comportant de l'aléatoire



Etat des individus en fonction du temps



Exemple de variable descriptive

Quelques questions pouvant se poser...

- Y a-t-il convergence vers un état stable ?
  - Ca sera le cas si il existe un instant $t^* \in \mathbb{N}$ tel que l'état de toutes les cellules est identique à l'instant t^* et t^*+1
- Dans le cas contraire, y a-t-il convergence vers un attracteur cyclique (de faible période) ?
- L'attracteur dépend-il de la configuration initiale de l'automate cellulaire ?
- La taille du réseau et le fait qu'il ait ou non des bords jouent-ils un rôle sur les comportements possibles ?

Le jeu de la vie

- Imaginé par John Conway en 1970
- Automate cellulaire déterministe à deux états (vivant/occupée/1, mort/inoccupée/0) sur une grille de dimension 2.
- Voisinage de Moore.
- Règle locale de changement d'état :
    - **reproduction** : une cellule inoccupée devient occupée si elle a exactement 3 cellules voisines occupées.
    - **isolement** : une cellule occupée devient vide si elle a 1 cellule voisine occupée ou moins.
    - **surpopulation** : une cellule occupée entourée d'au moins 4 cellules voisines occupée, devient inoccupée.
- Comportements complexes : structures stables / périodiques, glisseurs, canons etc...

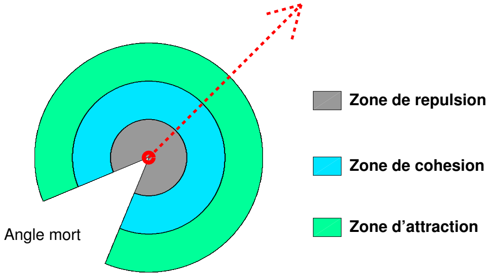
Présentation

- Pourquoi voler en groupe ou nager en banc :
  - meilleur champ de vision
  - éviter les prédateurs : taux de rencontres plus faible, coopération possible pour se défendre etc...
  - partenaires reproductifs à proximité etc...
- Premier modèle de vol d'oiseaux : Craig W. Reynolds
  - *Craig W. Reynolds. Flocks, Herds, and Schools: A Distributed Behavioral Model (1987)*
  - *Craig W. Reynolds. Steering Behaviors For Autonomous Characters (1999)*

  ↪ `http://www.red3d.com/cwr/steer/gdc99/`
- Emergence : peut-on aboutir à la formation de banc (échelle populationnelle) à partir de comportements individuels simples ?

Modèle de Craig Reynolds

- Etat d'un individu : position et vitesse courante
- Exemple de voisinage (géographique):

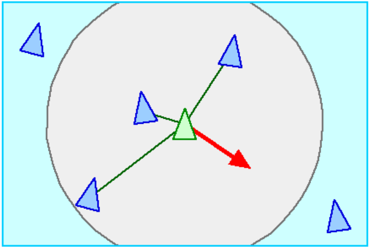


- Etablir une règle *(simple)* de mouvement dépendant de l'état (position et/ou vitesse) des individus se trouvant dans les différentes zones de ce voisinage

Comportement individuel : un exemple

- **répulsion :** Les individus s'éloignent des individus trop proches afin d'éviter la collision.

$$\vec{f}_i^{\,r} = \sum_{j \in V_i^r} \frac{1}{r_{ij}^2} (\vec{x}_i - \vec{x}_j)$$

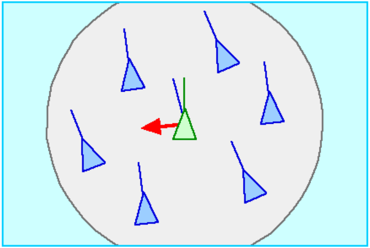


- V_i^r : ensemble des voisins de la zone de répulsion
- $\vec{x}_k$: position de l'individu k
- r_{ij} : distance euclidienne entre les individus i et j

Comportement individuel : un exemple

- **cohésion :** Les individus cherchent à aligner leur déplacement à celui des voisins de la zone de cohésion.

$$\vec{f}_i^c = \frac{1}{N_i^c} \sum_{j \in V_i^c} \vec{v}_j$$

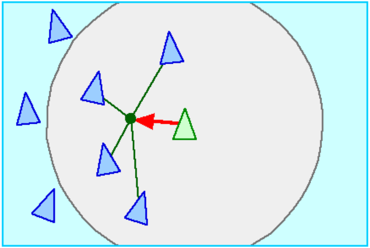


- V_i^c : ensemble des voisins de la zone de cohésion
- N_i^c : nombre de voisins de la zone de cohésion
- $\vec{v}_k$: vecteur vitesse de l'individu k

Comportement individuel : un exemple

- **attraction :** Les individus cherchent à se rapprocher des individus (p.ex. de leur barycentre) se trouvant dans leur voisinage.

$$\vec{f}_i^a = \frac{1}{N_i^a} \sum_{j \in V_i^a} \vec{x}_j - \vec{x}_i$$



- V_i^a : ensemble des voisins de la zone d'attraction
- N_i^a : nombre de voisins de la zone d'attraction
- $\vec{x}_k$: position de l'individu k

Comportement individuel : un exemple

- Il convient maintenant d'établir une règle permettant, à partir de ces "forces" de répulsion, de cohésion et d'attraction, de calculer le nouvel état (i.e. position et vitesse) des individus.

  ↪ Encore une fois, plusieurs modélisation possible.

- Version possible : ajouter les différentes "forces" (éventuellement normalisées et/ou pondérées) afin de définir le nouveau vecteur deplacement de l'individu.

$$\vec{v}_i^{\,t+1} = \alpha \, \frac{\vec{f}_i^{\,r}}{\|\vec{f}_i^{\,r}\|} + \beta \, \frac{\vec{f}_i^{\,c}}{\|\vec{f}_i^{\,c}\|} + \gamma \, \frac{\vec{f}_i^{\,a}}{\|\vec{f}_i^{\,a}\|}$$

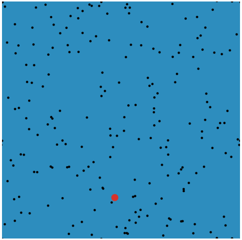
  et la nouvelle position :

$$\vec{x}_i^{\,t+1} = \vec{x}_i^{\,t} + \mu \, \vec{v}_i^{\,t+1}$$

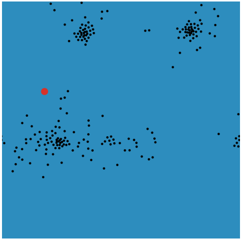
- Cas d'un domaine avec bords : gestion des collisions

Simulations

temps initial



temps intermédiaire 1



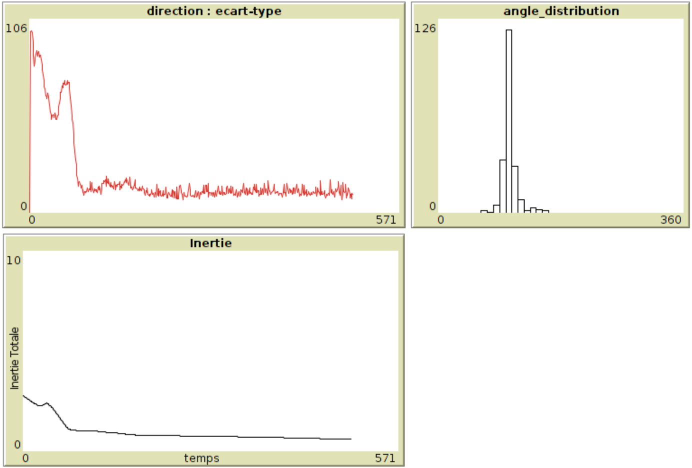
temps intermédiaire 2

temps final

● : centre d'inertie du nuage d'individus

Emergence : mise en évidence

- Utiliser des variables agrégées adaptées à la description du phénomène (global) observé.



Emergence : dépendance avec les paramètres du modèles

- Question : impacts des paramètres du modèles sur le phénomène émergent ?
- Méthode :
  - Stochasticité de la condition initiale : pour une même valeur de paramètres, les simulations diffèrent.

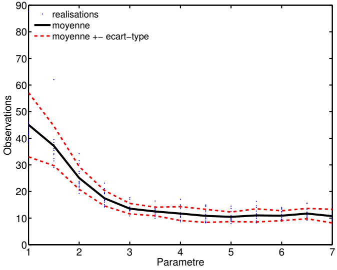
    ↪ Nécessite plusieurs simulations par jeu de valeurs des paramètres, calcul des indicateurs d'émergences pour chacune de ces simulations

    ↪ Résumer l'information (moyenne, ecart-type, intervalle de confiance etc...)

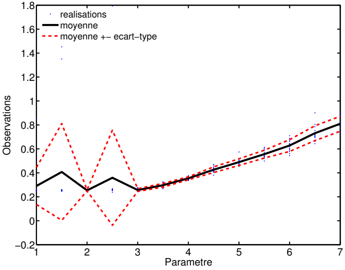
    ↪ Le faire pour différentes valeurs de paramètres
  - Necessite une condition de fin des simulations (par exemple un temps final).

Emergence : dépendance avec les paramètres du modèles

- Paramètre considéré, taille relative zone de répulsion / zone de cohésion.



orientation : ecart-type



Inertie du nuage

Algorithme d'optimisation bio-inspiré : essaim de particules

- Objectif : obtenir le minimum sur un compact de $\mathbb{R}^n$ d'une fonction réelle .
- Algorithme stochastique : ne “marche” pas tout le temps, nécessite plusieurs simulations : on est pas sûr d'obtenir le minimum.
- Les agents se promènent dans cet espace suivant des règles dépendant de leurs propres observations et de celles de leur voisins.

  $\hookrightarrow$ Voisinage social et non géographique : réseau social, plusieurs topologies possibles (anneau, étoile, rayon etc...)
- Exemple : vitesse de déplacement :

$$\vec{v}_i^{\,t+1} = \kappa\,\vec{v}_i^{\,t} + \rho_1(\vec{x}_{i,\mathrm{opt}} - \vec{x}_i^{\,t}) + \rho_2(\vec{x}_{j,\mathrm{opt}} - \vec{x}_i^{\,t})$$

  - $\vec{x}_{i,\mathrm{opt}}$: sa meilleure position
  - $\vec{x}_{j,\mathrm{opt}}$: la meilleure position de ses voisins.

Sugarscape : présentation

- Référence : *Epstein, Joshua M.; Axtell, Robert L. Growing Artificial Societies: Social Science From the Bottom Up (1996)*
- Modèle de société : intéraction environnement (ressources) - agents (individus, entreprises).
- **Sugarscape = automate cellulaire + agents**

↪ De nombreuses variantes possibles... suivant la problématique posée.

↪ Servira de base pour le projet.

L'environnement

- Environnement géographique **inhomogène** modélisé par un automate cellulaire (grille carrée)
- Chaque cellule est décrite par un attribut et une variable :
  - capacité de charge c_i = quantité maximale de ressource qu'elle peut contenir

    ↪ fixé initialement, n'évolue pas avec le temps
  - quantité actuelle de ressource r_i^t

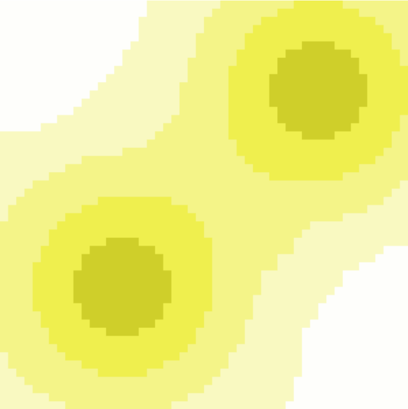
    ↪ évolue avec le temps, croit au cours du temps jusqu'à la capacité de charge et est consommée par les agents
- Règle de transition (sans agents) :

$$r_i^{t+1} = \min(c_i, r_i^t + \alpha)$$

  - α : taux de croissance de la ressource (paramètre du modèle)

    ↪ cas $\alpha = +\infty$: croissance instantanée jusqu'à c_i

L'environnement



- Grille 50×50 - configuration tore
- 5 niveaux de ressource (c_i)

↪ 2 “montagnes” (S-O et N-E)

↪ 1 “vallée” centrale

↪ 2 “déserts” (N-O et S-E)

Les agents

- Les agents se déplacent sur cet automate *(un seul agent par cellule)* et intéragissent avec l'environnement en consommant la ressource.
- Chaque agent possède deux attributs et deux variables :
  - Position sur l'automate p_j^t : il convient d'établir une règle de déplacement des individus dépendant de leur observation locale.
  - Vision v_j : l'agent j peut connaitre l'état des cellules se trouvant à une distance au plus v_j en haut, en bas, à gauche et à droite ainsi que la présence ou non d'agents sur ces cellules.
  - Métabolisme m_j : l'individu j consomme m_j unité de ressource par unité de temps. Si il ne peut pas, il meurt.
  - Quantité de ressource s_j^t : variable évoluant avec le temps, l'agent gagne de la ressource en exploitant l'environnement et en perd pour son métabolisme.

Les agents : comportement

- Un après l'autre *(ordre aléatoire)*, chaque agent :
  1) **observe :** L'agent identifie dans son voisinage (qui dépend de sa vision) les cellules inoccupées ayant le plus de ressource r_i^t.
  2) **analyse et définit un objectif :** Si ces cellules possèdent plus de ressource que leur propre cellule, il choisit pour destination la (une des) plus proche(s) d'entre elles.
  3) **agit :** Il se déplace sur cette nouvelle cellule si elle existe.
  4) **agit :** Il consomme toute la ressource présente sur sa *(nouvelle)* cellule.

  *) **mise à jour de son état:** Il paye son coût métabolique, si il ne peut pas, il meurt.
- Couplage agents / environnement :

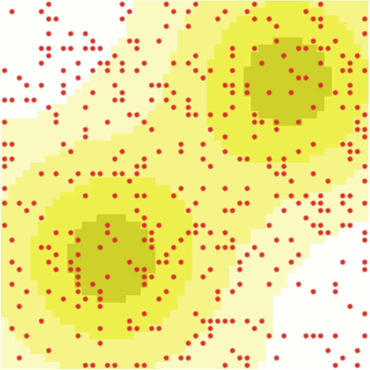
$$\mathbf{A}^{\mathbf{t+1}} = f(\mathbf{A}^{\mathbf{t}}, \mathbf{E}^{\mathbf{t}})$$

$$\mathbf{E}^{\mathbf{t+1}} = g(\mathbf{A}^{\mathbf{t}}, \mathbf{E}^{\mathbf{t}})$$

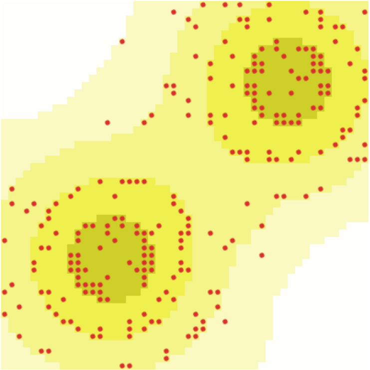
Condition Initiale

- Environnement:
  - $r_i^0 = v_i$
- Agents:
  - v_j tirés selon une loi uniforme $\mathcal{U}([\![1, v_{\max}]\!])$
  - m_j tirés selon une loi uniforme $\mathcal{U}([\![1, m_{\max}]\!])$
  - p_j^0 : les individus sont initialement disposés aléatoirement sur l'automate
  - Les quantités initiales de ressource s_j^0 sont tirées selon une loi uniforme $\mathcal{U}([\![k_{\min}, k_{\max}]\!])$
- 5 paramètres du modèle : $\alpha, v_{\max}, m_{\max}, k_{\min}, k_{\max}$

Sugarscape : simulation ($\alpha = +\infty$)

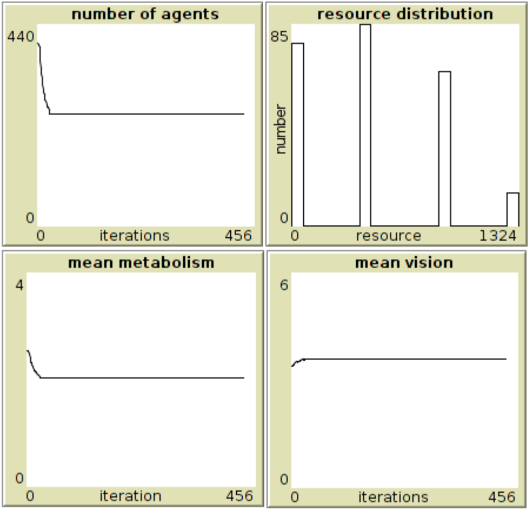


temps initial



temps final *(après convergence)*

Analyse des simulations



Analyse des simulations

- On constate:
  - Une diminution de la population totale : une partie de la population est morte, ne pouvant pas exploiter suffisamment de ressource pour assurer le cout du métabolisme.

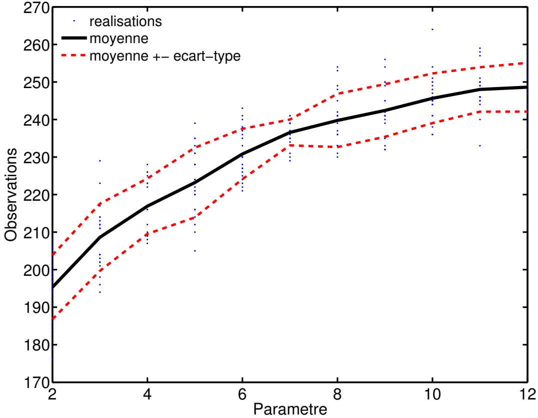
    ↪ "Capacité de charge" du milieu en agents.
  - Une augmentation de la vision moyenne dans la population, diminution du métabolisme moyen des agents.

    ↪ Sélection naturelle : les démons darwiniens (forte vision / petit métabolisme) ont une probabilité plus faible de mourir.
  - Répartition géographique : densité plus forte sur les montagnes. Les individus se trouvent essentiellement aux bords des paliers

    ↪ Leur vision ne leur permet pas de voir le palier suivant : ils ne se déplacent donc plus.

Dépendance avec les paramètres

- Evolution de la "capacité de charge" du milieu en fonction de la vision maximale des agents pour un nombre initial d'agents de 400.



Extension du modèle : morts et naissances

- On suppose maintenant que les agents peuvent mourir de "vieillesse". Ajout d'attributs et variables aux agents :
  - âge de l'agent a_j^t : l'individu nait avec un âge valant 0, cet âge est incrémenté de 1 à chaque itération.
  - âge de mort d_j : fixé à la naissance de l'agent. Tiré suivant une loi uniforme $\mathcal{U}([\![d_{\min}, d_{\max}]\!])$
  - Lorsqu'un agent meurt, il donne naissance à un nouvel agent initialisé aléatoirement ⇒ Conservation du nombre d'agents : évite l'extinction de la population. Plus de convergence : il y a toujours des agents en mouvement.
- D'autres modèles naissance-mort envisageables (probabilité de mourir à chaque itération dépendant éventuellement de leur âge, niveau de ressource etc...)

Inégalités

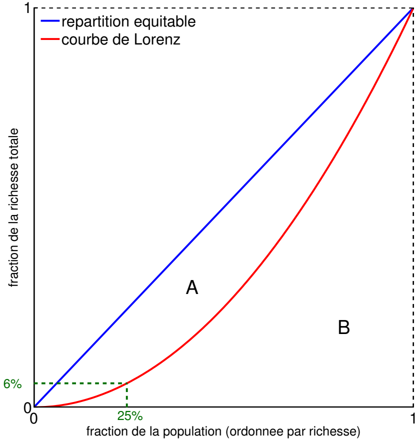
- **courbe de Lorenz :** associe à la fraction des plus pauvres individus de population, la fraction de la richesse totale qu'ils possèdent.

↪ courbe concave, valant 0 en 0 et 1 en 1.

- **Coefficient de Gini :** indice permettant d'évaluer les inégalités de richesse dans une population.

$$G = \frac{A}{A+B} = 2A = 1 - 2B$$

- $G \in [0, 1[$
- $G = 0$: équi-répartition des richesses entre les individus
- $G \approx 1$: un individu (ou une fraction epsilonnienne de la population) possède toutes les richesses



Extension du modèle : pollution

- **Causes** possibles :
  - Mouvements des agents, recolte ou utilisation de la ressource par les agents, croissance de la ressource etc...
- **Conséquences** possibles :
  - agents : dégradation de leur vision, augmentation de leur métabolisme, augmentation des probabilités de mort, désagrément etc...
  - environnement : diminution du taux de renouvellement des ressources, contamination des sites voisins etc...

Extension du modèle : pollution

- Un exemple simple : la pollution est provoquée par la récolte des ressource.

  ↪ ajout d'une variable *quantité de pollution* q_i^t aux cellules de l'automates avec la règle de transition :

  $$q_i^{t+1} = q_i^t + \gamma\ c_i^t$$

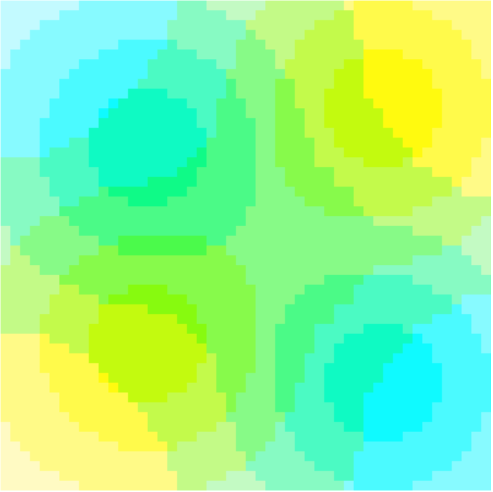
  - γ : quantité de pollution produite par exploitation d'une unité de ressource.
  - c_i^t : quantité de de ressource exploitée sur cette cellule à l'instant t.

- Les agents sont génés par la pollution :

  ↪ ils peuvent par exemple évaluer les cellules en considérant le ratio ressource / pollution.

Extension du modèle : ressources multiples et commerce

- Deux ressources disponibles : le sucre et les épices



"bien-être" et mouvement des agents

- Les agents ont maintenant un métabolisme m_s pour le sucre et un métabolisme m_e pour les épices et des quantités accumulées de ressources valant respectivement r_s^t et r_e^t.

  $\hookrightarrow$ Comment évaluer la qualité des sites de ressources ? Un compromis dans l'acquisition des deux ressources est nécessaire.

- Fonction de "bien-être" *(welfare)* possible :

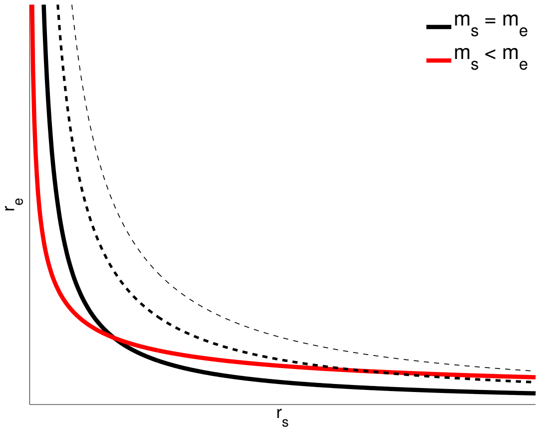
$$W(r_s^t, r_e^t) = (r_s^t)^{m_s/m_T} \; (r_e^t)^{m_e/m_T} \qquad \text{avec} \;\; m_T = m_s + m_e$$

  - fonction de type Cobb-Douglass
  - fonction croissante par rapport aux deux variables

$$\log[W(r_s^t, r_e^t)] = \frac{m_s}{m_T} \log(r_s^t) + \frac{m_e}{m_T} \log(r_e^t)$$

"bien-être" et mouvement des agents

- Isovaleurs de la fonction de bien-être :



"bien-être" et mouvement des agents

- Règle du mouvement : l'agent choisit la cellule inoccupée de son voisinage lui permettant d'accroitre au plus son bien-être, i.e. celle maximisant la quantité :

  $$W(r_s^t + s_i, r_e^t + e_i)$$

  où s_i et e_i sont les quantités respectives de sucre et d'épices de la cellule i.

Echanges de ressources entre agents

- Estimation du temps restant à vivre τ en se basant uniquement sur les ressources possédées :

$$\tau = \min(\tau_s\,,\,\tau_e) \qquad \text{avec} \;\; \tau_s = \frac{r_s^t}{m_s} \;\; \text{et} \;\; \tau_e = \frac{r_e^t}{m_e}$$

- Taux marginal de substitution (MRS): pour un agent, étant donné ses reserves actuelles de ressources, détermine la quantité d'épices ayant la même valeur qu'une unité de sucre.
    - ↪ $MRS > 1$: une "unité" de sucre est plus rentable qu'une "unité" d'épices ⇒ le sucre vaut plus cher que les épices.

Echanges de ressources entre agents

- Les agents peuvent réaliser des échanges avec leur voisins (voisinage de Von-Neumann). Le prix d'échange entre un agent A et un agent B dépend du MRS des deux agents. Une possibilité de prix

$$p = \sqrt{MRS_A \, MRS_B}$$

  - supposons $MRS_A > MRS_B$, alors $p \in [MRS_B, MRS_A]$.
  - B donne une quantité q de sucre à A et A donne une quantité $p\,q$ d'épices à B.
  - les deux agents font une "bonne affaire"
  - Après itérations successives de ce procédé, convergence des MRS vers une valeur commune.

Illustration : capacité de charge du milieu