Analyse multi-échelles des systèmes complexes

Robin Lamarche-Perrin

Doctorat en informatique (2010-2013)

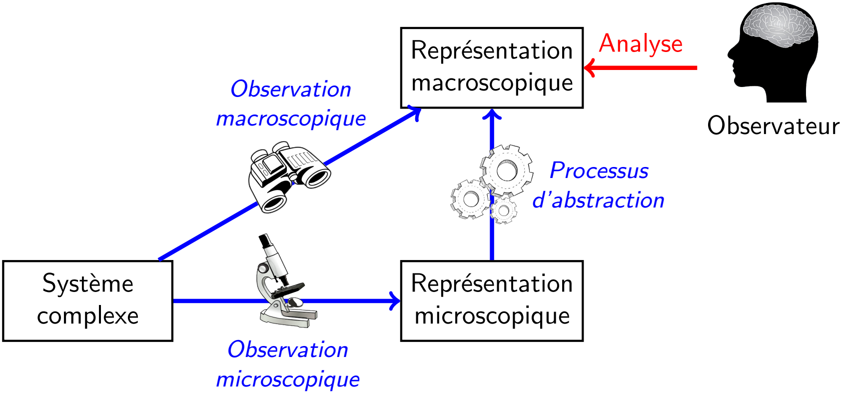
Post-doctorat en mathématiques (2013-today)

Diplôme d'ingénieur (2007-2010)

Master d'informatique (2009-2010)

Master de philosophie (2009-2012)

Problème général



Première partie I

Analyse multi-échelles des relations internationales

Hypothèse : les médias constituent un instrument d'observation adéquat du niveau national

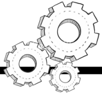
[Grasland *et al.*, 2011]

Géographe

Système international

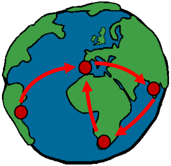
Media

Représentation microscopique

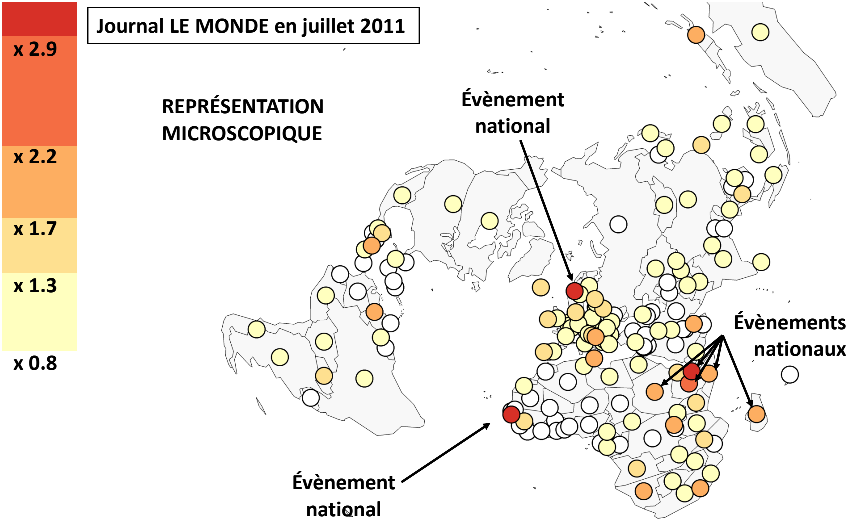


Agrégation de données

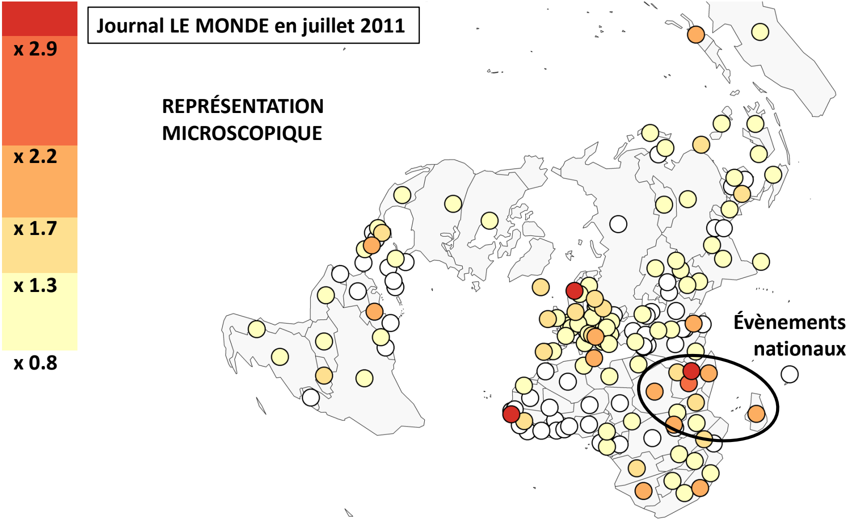
Représentation macroscopique



Détection d'évènements médiatiques



Détection d'évènements médiatiques

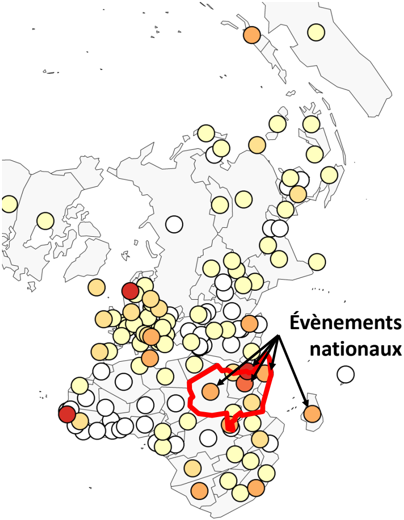


Agrégation de données géographiques

π_1 π_2 π_3 **Espace**

Temps t

	USA	Libye	Syrie	France	Israël	...	Total
2 mai	25	12	11	10	4	...	142
9 mai	14	6	12	12	5	...	108
16 mai	20	11	12	6	9	...	142
23 mai	15	9	6	13	5	...	120
30 mai	10	16	17	9	4	...	137
6 juin	14	16	16	9	4	...	114
13 juin	15	14	17	9	6	...	119
20 juin	17	13	12	12	7	...	123
27 juin	7	6	7	20	2	...	103
4 juill.	12	13	8	10	6	...	129
11 juill.	21	10	10	14	3	...	107
18 juill.	7	3	8	4	5	...	61
25 juill.	16	7	6	13	4	...	128
1 août	21	1	9	7	4	...	88
...	...	...	...	...	...	...	...
Total	423	308	260	248	153	...	3520

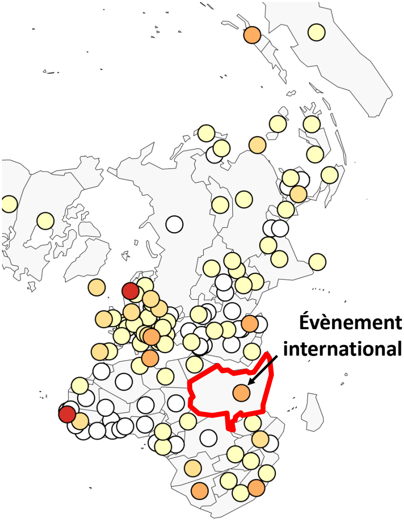


Agrégation de données géographiques

$\pi_1 \quad \pi_2 \quad \pi_3$ **Espace**

Temps t

	USA	Agrégat	Israël	...	Total
2 mai	25	13+11+10	4	...	142
9 mai	14	6+12+12	5	...	108
16 mai	20	11+12+6	9	...	142
23 mai	15	9+6+13	5	...	120
30 mai	10	16+17+9	4	...	137
6 juin	14	16+16+9	4	...	114
13 juin	15	14+17+9	6	...	119
20 juin	17	13+12+12	7	...	123
27 juin	7	6+7+20	2	...	103
4 juill.	12	13+8+10	6	...	129
11 juill.	21	10+10+14	3	...	107
18 juill.	7	3+8+4	5	...	61
25 juill.	16	7+6+13	4	...	128
1 août	21	1+9+7	4	...	88
...	...	...	...	...	...
Total	423	308+260+248	153	...	3520

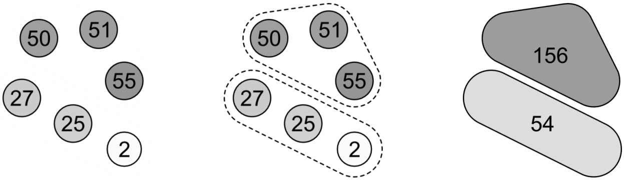


Processus d'agrégation

Population $\Omega = \{x_1, \dots, x_N\}$

Dénombrement $v : \Omega \rightarrow \mathbb{N}$

Partitionnement $\longrightarrow$ **Agrégation** $\longrightarrow$



Représentation microscopique

$$\Pr(X = x) = \frac{v(x)}{\sum_{x \in \Omega} v(x)}$$

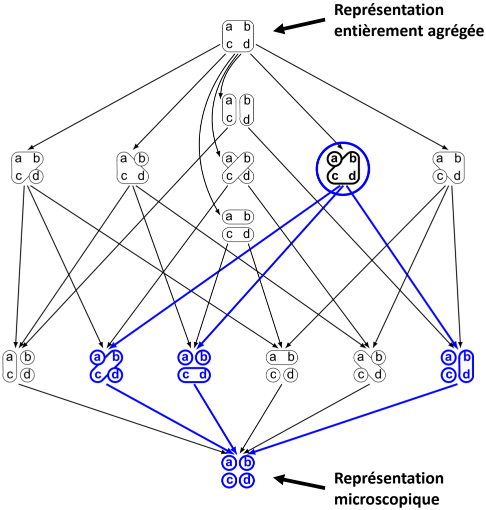
Représentation partitionnée

$$\Omega = A_1 \cup \dots \cup A_k$$
$$A_i \cap A_j = \emptyset$$

Représentation agrégée

$$\Pr(\hat{X} = A) = \sum_{x \in A} \Pr(X = x)$$

Poset des partitions admissibles

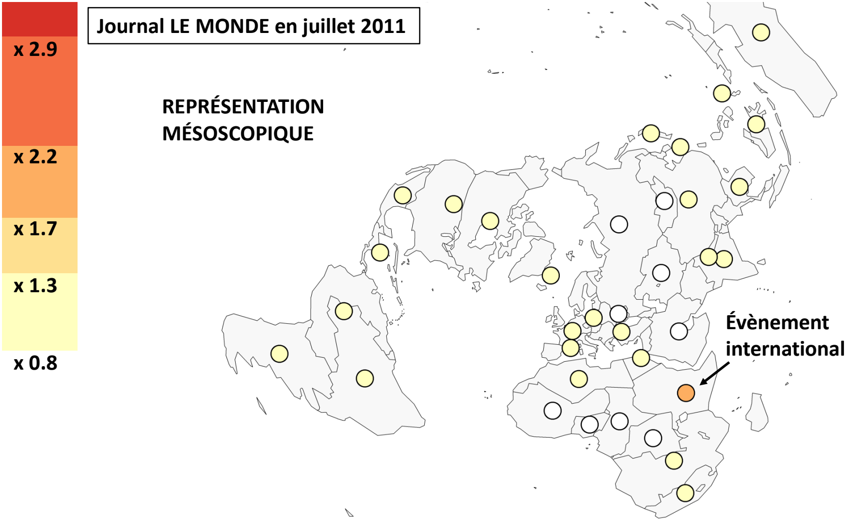


Structure algébrique

Ordre partiel sur l'ensemble des partitions possibles

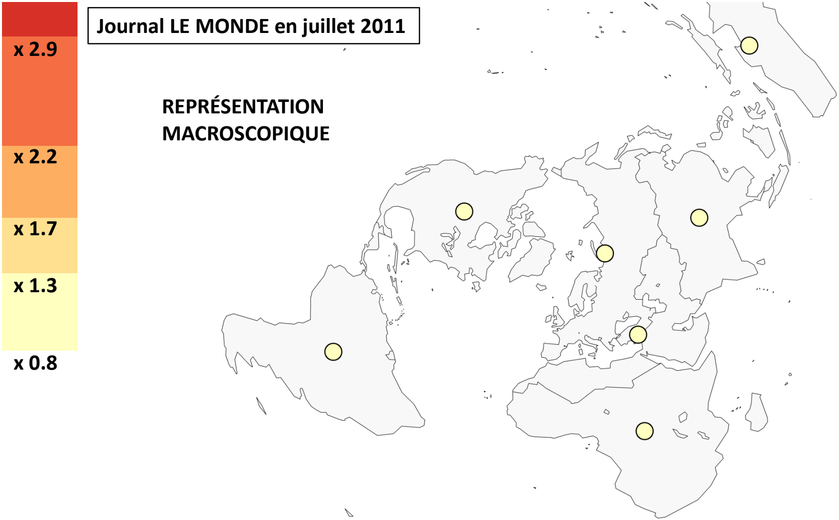
→ relations de raffinement

P1 : Granularité de l'information

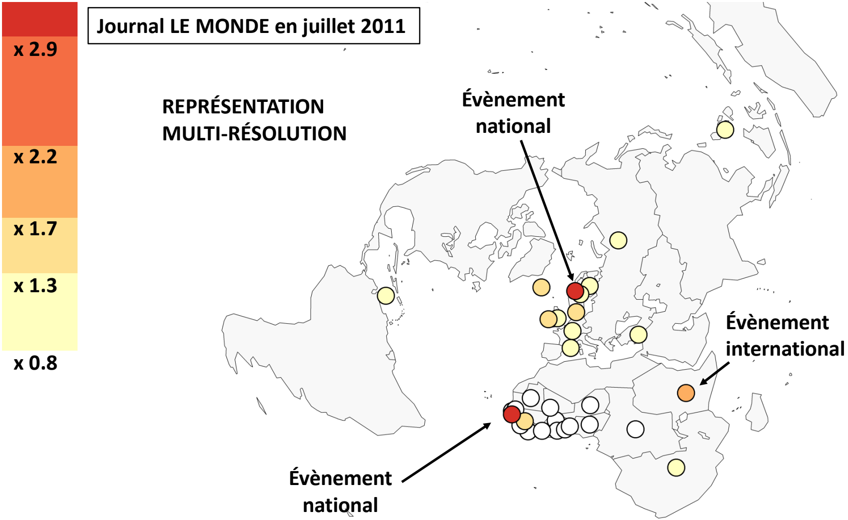


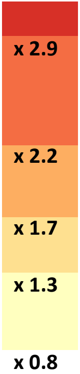
REPRÉSENTATION MÉSOSCOPIQUE

P1 : Granularité de l'information



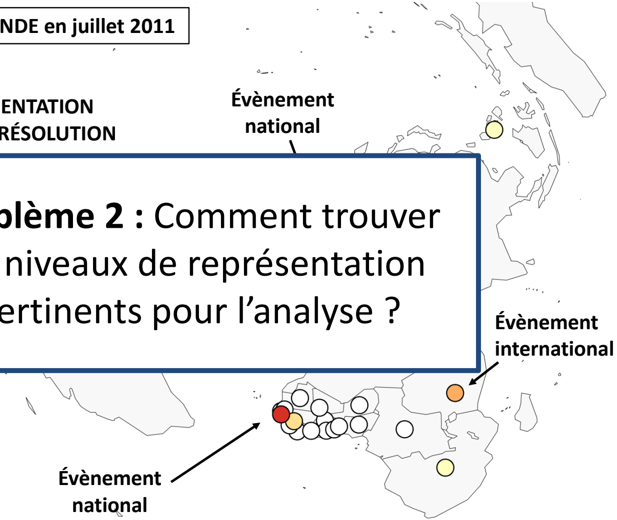
P1 : Granularité de l'information





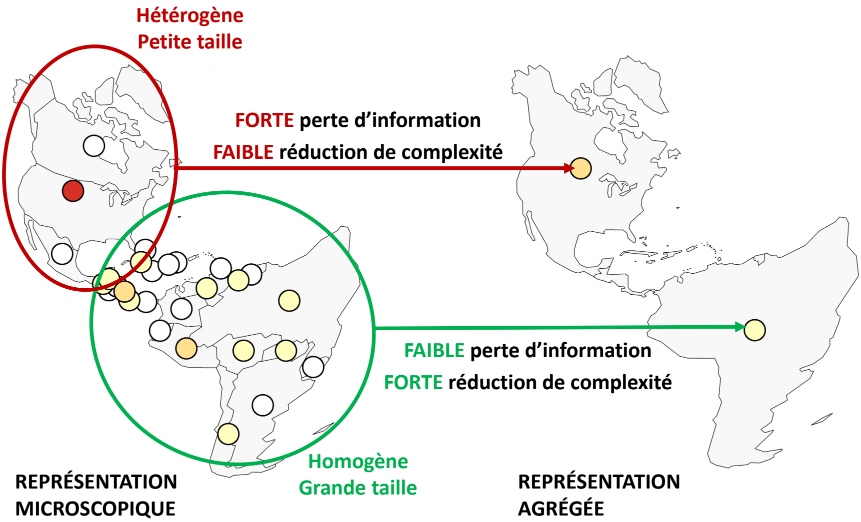
Journal LE MONDE en juillet 2011

REPRÉSENTATION MULTI-RÉSOLUTION

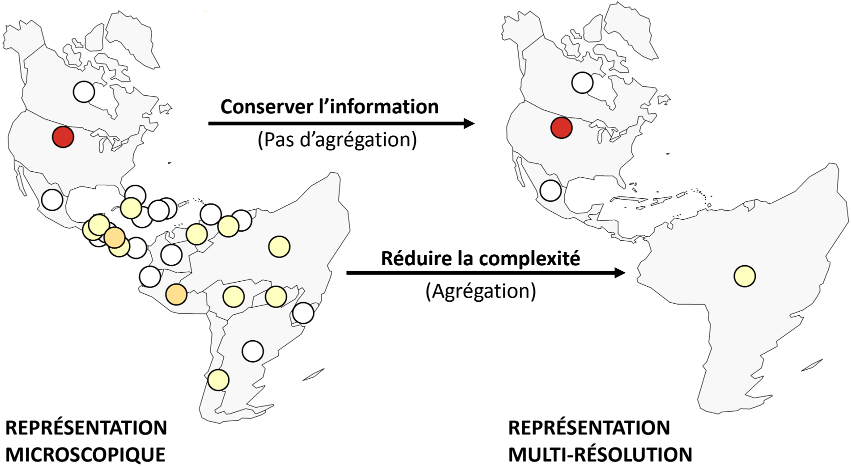


Problème 2 : Comment trouver les niveaux de représentation pertinents pour l'analyse ?

Complexité et information



Complexité et information



Quantifier la complexité et l'information

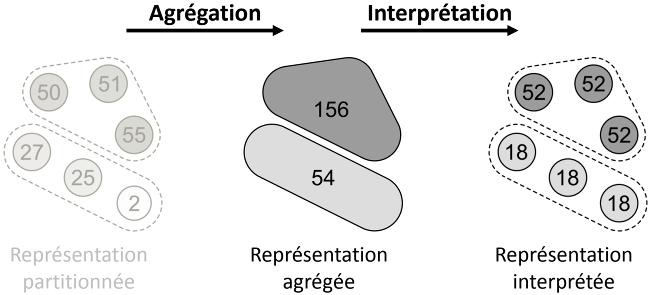
- La **complexité** dépend de la tâche à accomplir et des outils de description disponibles [Bonabeau et Dessalles, 1997]
  - Nombre de paramètres $\boxed{\mathrm{C}(A_1, \dots, A_k) = k}$
  - Entropie de Shannon $\mathrm{H}(\hat{X}) = \sum_{A_i} -\mathrm{Pr}\left(\hat{X} = A_i\right) \log_2 \mathrm{Pr}\left(\hat{X} = A_i\right)$

- La **perte d'information** est mesurée par la divergence entre deux distributions de probabilité [Kullback et Leibler, 1951]
  - Divergence de Kullback-Leibler $\boxed{\mathrm{D}(X||X') = \sum_{x \in \Omega} \mathrm{Pr}(X = x) \log_2 \frac{\mathrm{Pr}(X=x)}{\mathrm{Pr}(X'=x)}}$
  - Information de désagrégation $\mathrm{H}\left(\mathrm{X} \middle| \hat{X}\right) = \sum_{A_i} \mathrm{Pr}\left(\hat{X} = A_i\right) \mathrm{H}\left(\mathrm{X} \middle| \hat{X} = A_i\right)$

Le processus d'agrégation

Population $\Omega = \{x_1, \dots, x_N\}$

Dénombrement $v : \Omega \rightarrow \mathbb{N}$



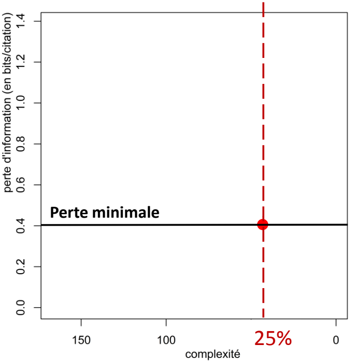
$\Pr(\hat{X} = A) = \sum_{x \in A} \Pr(X = x)$

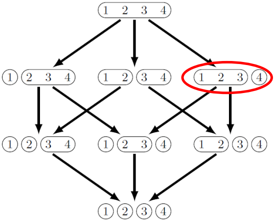
$\Pr(X' = x) = \Pr(X = x \mid \hat{X} = A_x)$

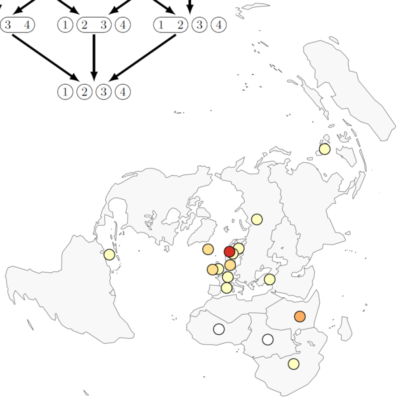
Deux objectifs à optimiser...

Compromis de qualité :

$$Q_\beta(\mathcal{X}) = \frac{C(\mathcal{X})}{C(\{\Omega\})} + \beta \frac{D(\mathcal{X})}{D(\{\Omega\})}$$



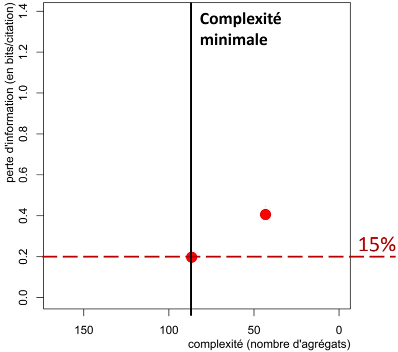


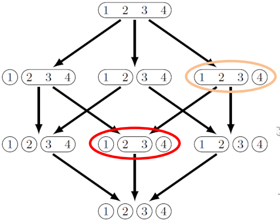


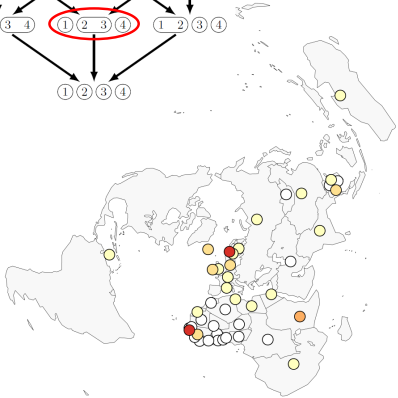
Deux objectifs à optimiser...

Compromis de qualité :

$$Q_{\beta}(\mathcal{X}) = \frac{C(\mathcal{X})}{C(\{\Omega\})} + \beta \frac{D(\mathcal{X})}{D(\{\Omega\})}$$





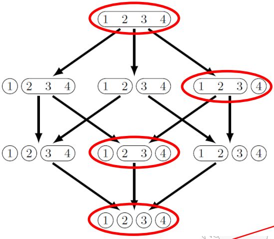


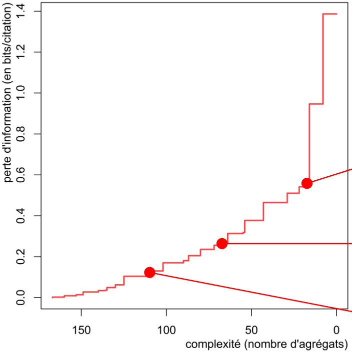
Optimisation des mesures

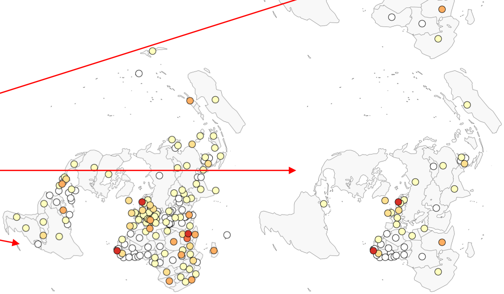
Deux objectifs à optimiser…

Compromis de qualité :

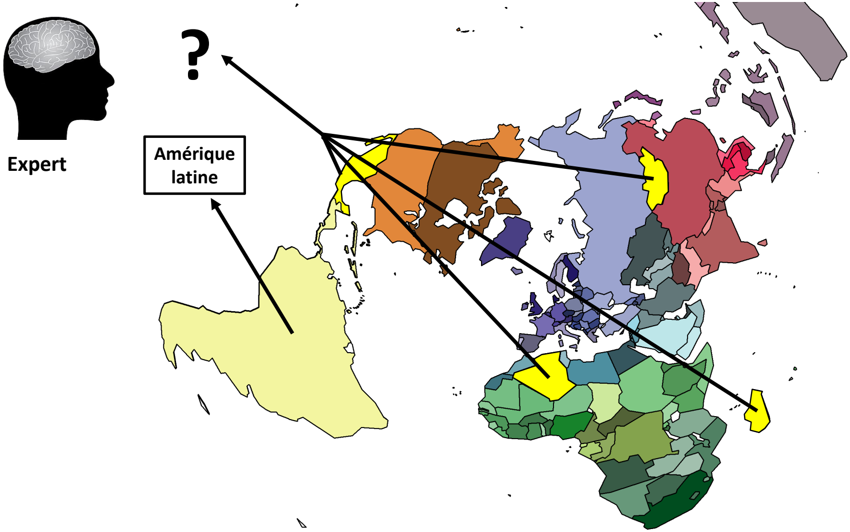
$$Q_\beta(\mathcal{X}) = \frac{C(\mathcal{X})}{C(\{\Omega\})} + \beta\,\frac{D(\mathcal{X})}{D(\{\Omega\})}$$

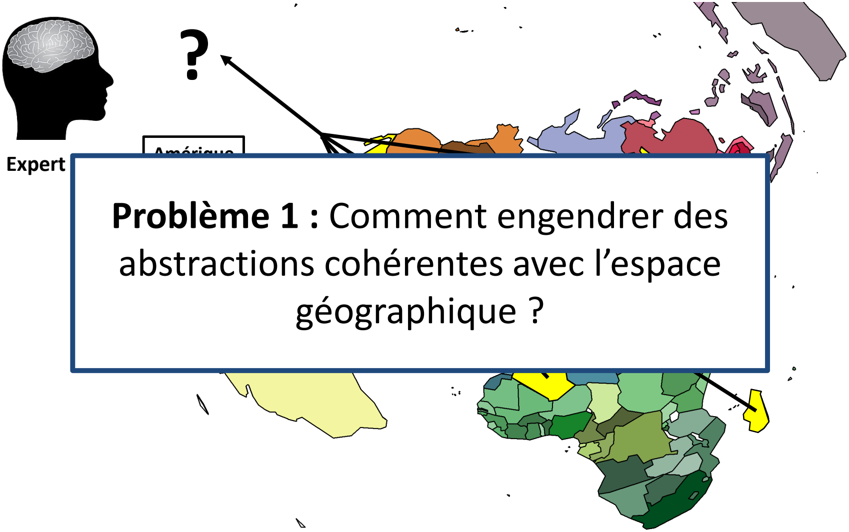






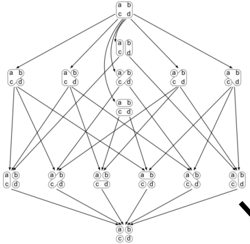
P2 : Sémantique des agrégats géographiques



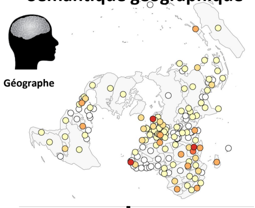


Problème 1 : Comment engendrer des abstractions cohérentes avec l'espace géographique ?

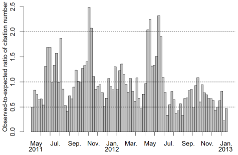
Ensemble des partitions possibles



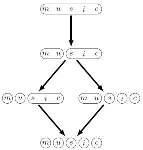
Sémantique géographique

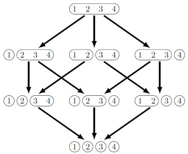


Sémantique temporelle

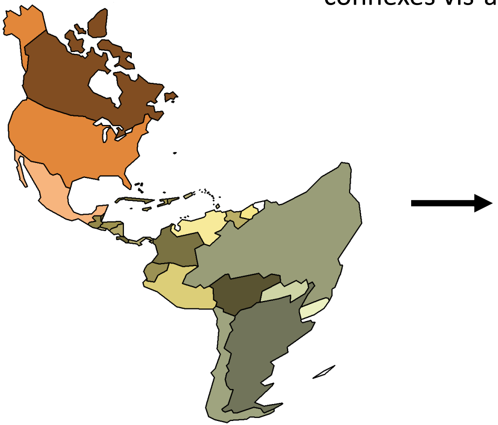


Contraindre l'ensemble des partitions admissibles

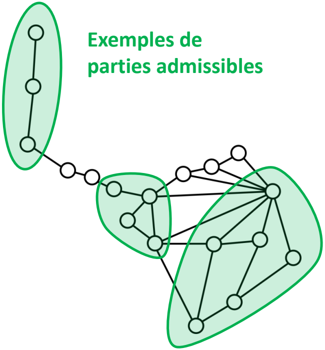




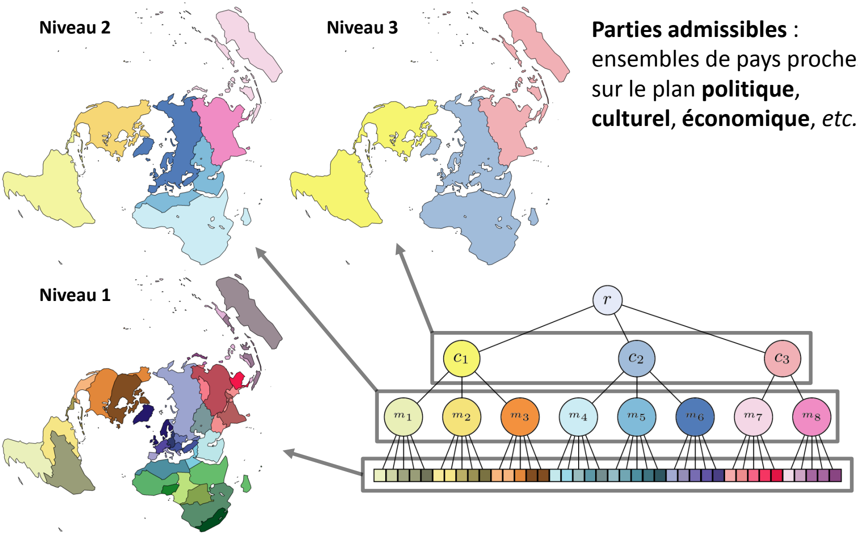
Conserver une relation de voisinage lors de l'agrégation



Parties admissibles : ensembles de pays connexes vis-à-vis du graphe de voisinage



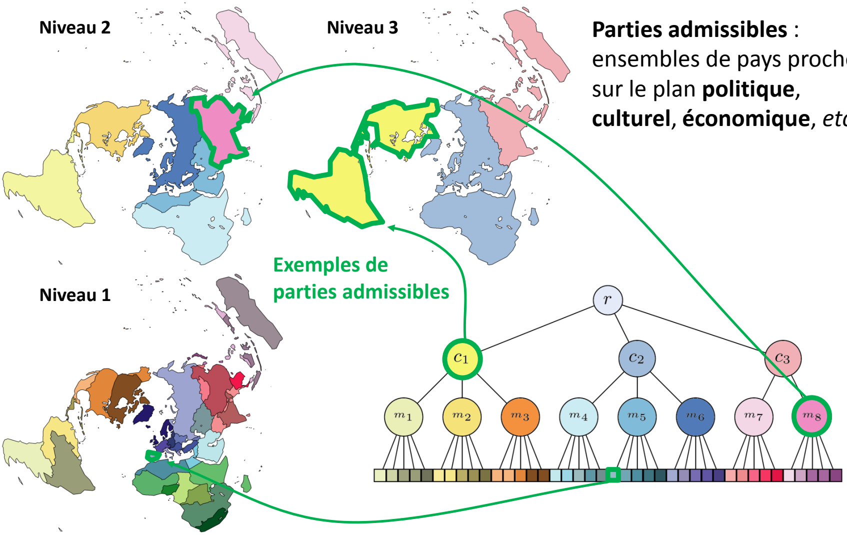
Conserver une hiérarchie territoriale lors de l'agrégation



Parties admissibles : ensembles de pays proches sur le plan **politique**, **culturel**, **économique**, *etc.*

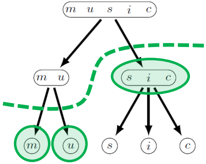
Conserver une hiérarchie territoriale lors de l'agrégation

Parties admissibles : ensembles de pays proches sur le plan **politique**, **culturel**, **économique**, *etc.*

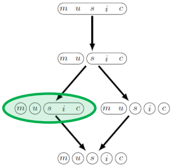


Contraintes combinatoires

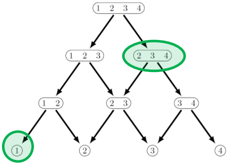
Parties admissibles
(nœuds de la hiérarchie)



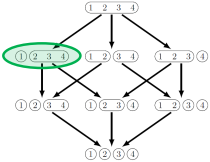
Partitions admissibles
(coupe dans la hiérarchie)



Parties admissibles
(intervalles de temps)

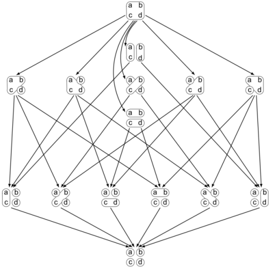


Partitions admissibles
(séquences d'intervalles)

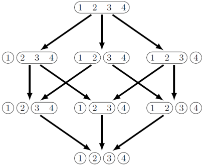


Complexité des structures algébriques

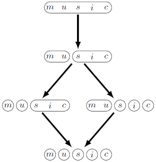
Partitions
non-contraintes



Partitions admissibles
selon un **ordre total**

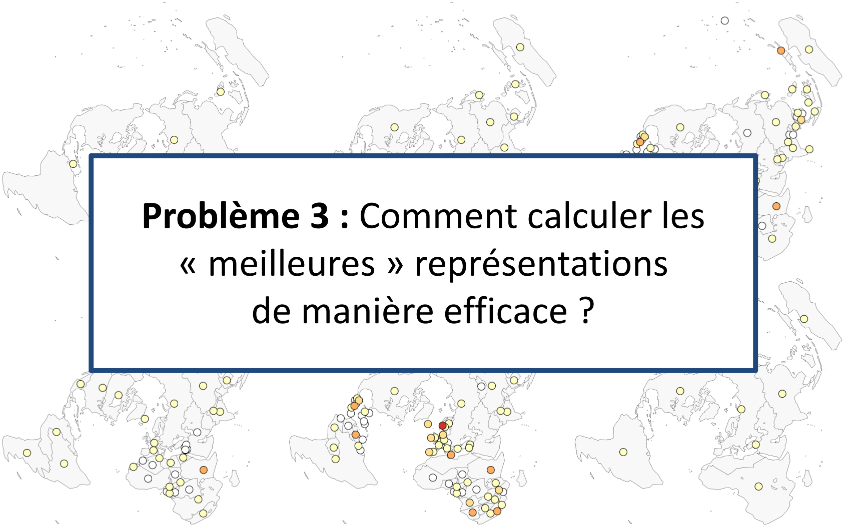


Partitions admissibles
selon une **hiérarchie**



Moins contraint ⟵⟶ **Plus contraint**

Plus complexe ⟵⟶ **Moins complexe**



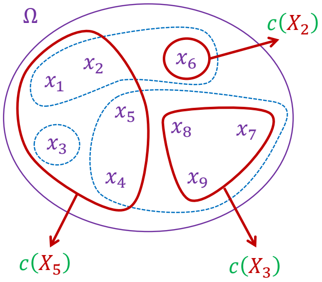
Problème 3 : Comment calculer les « meilleures » représentations de manière efficace ?

The Set Partitioning Problem

Given:

- a set of individuals $\Omega = \{x_1, \dots, x_n\}$
- a set of admissible parts $\mathcal{P} = \{X_1, \dots, X_m\} \subset 2^{\Omega}$
- a cost function $c : \mathcal{P} \to \mathbb{R}$
- the corresponding set of admissible partitions $\mathfrak{P} = \{\mathcal{X} \subset \mathcal{P}$ such that $\mathcal{X}$ is a partition of $\Omega\}$

Additional assumptions



Problem: Find an admissible partition that minimizes the cost function:

$$\mathcal{X}^* = \underset{\mathcal{X} \in \mathfrak{P}}{\arg\min} \left(\sum_{X \in \mathcal{X}} c(X) \right)$$

→ NP-complete!

Applications and Special Versions of the SPP

Multilevel Geographical Analysis

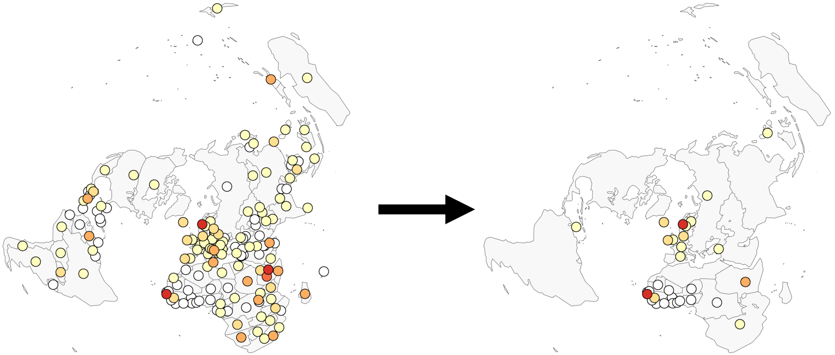
- Ω = territorial units
- $\mathcal{P}$ = admissible aggregates
- c = compression rate
- $\mathfrak{P}$ = aggregated representations

Hierarchical SPP

- Assumption: $\mathcal{P}$ forms a hierarchy
- Result: $\mathcal{O}(n)$ depth-first search
  [Pons *et al.*, 2011] [Lamarche-Perrin *et al.*, 2014]

Graph SPP

- Assumption: $\mathcal{P}$ are connected parts of a graph
- Result: NP-complete [Becker *et al.*, 1998]



Applications and Special Versions of the SPP

Multilevel Geographical Analysis

Time Series Analysis

- Ω = ordered data points
- $\mathcal{P}$ = time intervals
- c = compression rate
- $\mathfrak{P}$ = aggregated time series

Hierarchical SPP

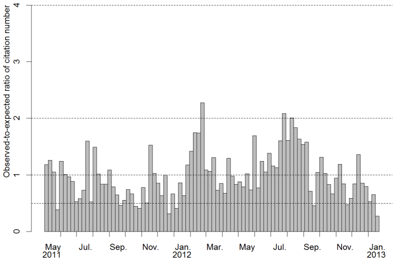
- Assumption: $\mathcal{P}$ forms a hierarchy
- Result: $\mathcal{O}(n)$ depth-first search
  [Pons *et al.*, 2011] [Lamarche-Perrin *et al.*, 2014]

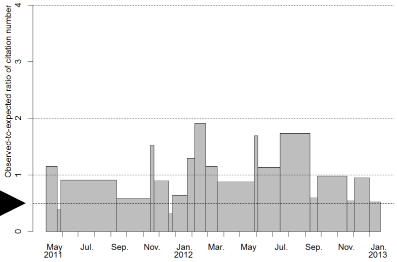
Graph SPP

- Assumption: $\mathcal{P}$ are connected parts of a graph
- Result: NP-complete [Becker *et al.*, 1998]

Ordered SPP

- Assumption: $\mathcal{P}$ are intervals
- Result: $\mathcal{O}(n^2)$ dynamic programming
  [Anily *et al.*, 1991] [Jackson *et al.*, 2005]





Applications and Special Versions of the SPP

Multilevel Geographical Analysis

Time Series Analysis

Coalition Structure Generation

Community Detection

Distributed System Monitoring

Load Balancing Problem

Database Optimization

Image Processing

Combinatorial Auctions

Hierarchical SPP
- Assumption: $\mathcal{P}$ forms a hierarchy
- Result: $\mathcal{O}(n)$ depth-first search
  [Pons *et al.*, 2011] [Lamarche-Perrin *et al.*, 2014]

Graph SPP
- Assumption: $\mathcal{P}$ are connected parts of a graph
- Result: NP-complete [Becker *et al.*, 1998]

Ordered SPP
- Assumption: $\mathcal{P}$ are intervals
- Result: $\mathcal{O}(n^2)$ dynamic programming
  [Anily *et al.*, 1991] [Jackson *et al.*, 2005]

Complete SPP
- Assumption: $\mathcal{P}$ contains all parts
- Result: $\mathcal{O}(3^n)$ dynamic programming
  [Yeh, 1986] [Lehmann *et al.*, 2006]

Ordered x Hierarchical SPP [Dosimont *et al.*, 2014]

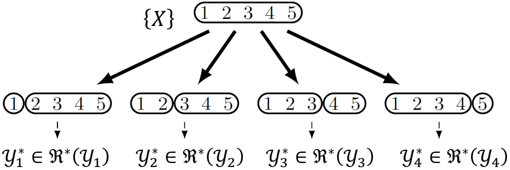
Array SPP [Muthukrishnan *et al.*, 2005]

SPP with Size Bounds [Rothkopf *et al.*, 1998]

Cyclic SPP [Rothkopf *et al.*, 1998]

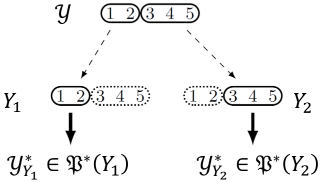
DIVIDE...

Branching the state space according to the covering relation

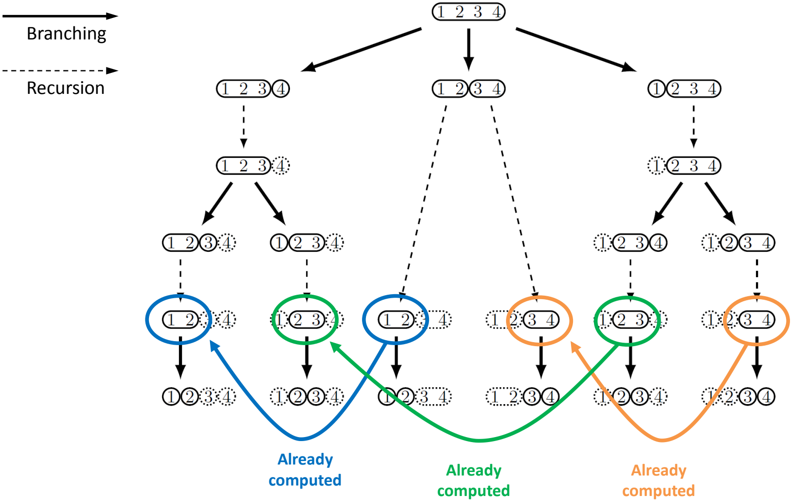


...AND CONQUER

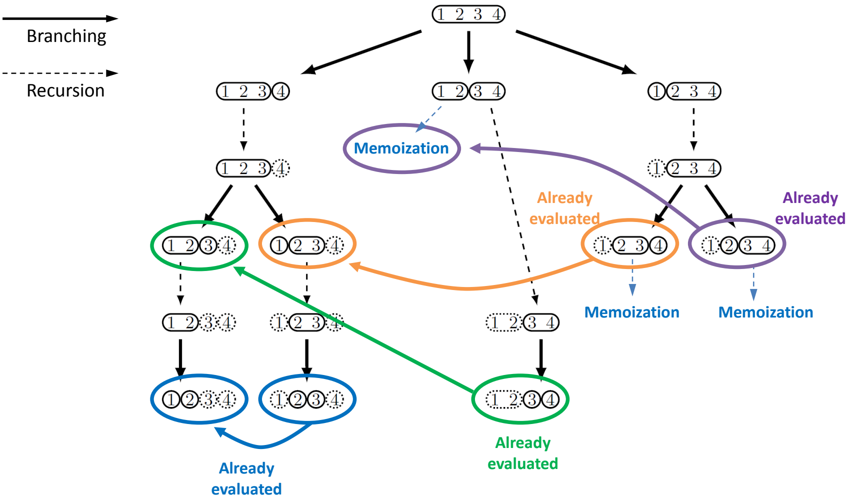
Recursion according to the principle of optimality



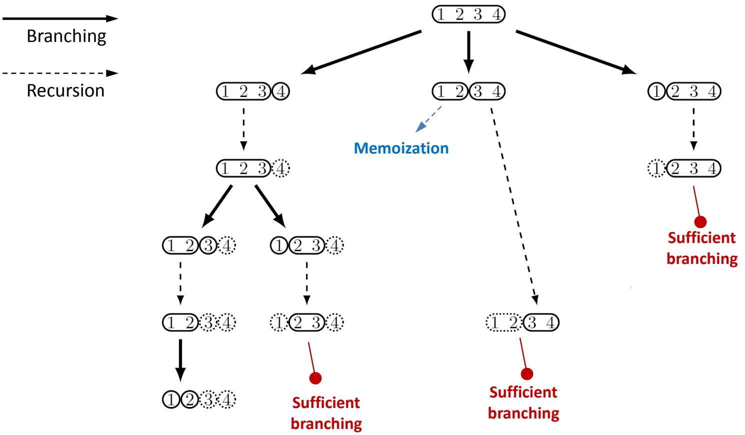
Memoization



Non-redudant Branching



Non-redudant Branching



The Generic Algorithm

A Generic Algorithm to Solve the SPP

Global Inputs:

c a cost function;

$\mathcal{P}$ a set of admissible parts defining admissible partitions;

$\mathfrak{L}$ a set of locally-optimal admissible partitions of parts on which the algorithm has already been applied.

Local Inputs:

X an admissible part;

$\overline{\mathcal{X}}$ the complementary partition of X inherited from the "higher" call ($\overline{\mathcal{X}}$ is a partition of $\Omega\backslash X$);

$\mathfrak{D}$ the set of admissible partitions which refinements have already been evaluated during "higher" calls.

Output:

$\mathcal{X}^*$ a locally-optimal admissible partition of X.

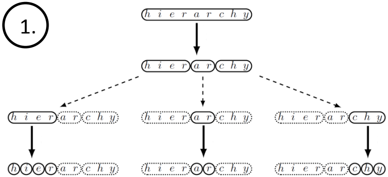
- If the algorithm has already been applied to part X, return the locally-optimal partition recorded in $\mathfrak{L}$.
- Initialization: $\mathcal{X}^* \leftarrow \{\{X\}\}$ and $\mathfrak{D}' \leftarrow \mathfrak{D}$.
- For each $\mathcal{Y} \in \mathfrak{C}(\{X\})$ such that $\overline{\mathcal{X}} \cup \mathcal{Y}$ does not refine any partition in $\mathfrak{D}$, do the following:
  - For each part $Y \in \mathcal{Y}$, call the algorithm with local inputs $X \leftarrow Y$, $\overline{\mathcal{X}} \leftarrow \overline{\mathcal{X}} \cup \mathcal{Y}\backslash\{Y\}$, and $\mathfrak{D} \leftarrow \mathfrak{D}'$ to compute a locally-optimal partition $\mathcal{Y}^*_Y \in \mathfrak{P}^*(Y)$.
  - $\mathcal{Y}^* \leftarrow \bigcup_{Y\in\mathcal{Y}} \mathcal{Y}^*_Y$.
  - If $c(\mathcal{Y}^*) > c(\mathcal{X}^*)$, then $\mathcal{X}^* \leftarrow \mathcal{Y}^*$.
  - $\mathfrak{D}' \leftarrow \mathfrak{D}' \cup \{\mathcal{Y}\}$.
- Return $\mathcal{X}^*$ and record this result in $\mathfrak{L}$.

Generic: solve any instance of the SPP
→ but inefficient for special versions

Designing dedicated implementations:

1. Analysing the generic execution
2. Building appropriate data structures
3. Deriving a specialized algorithm

1.



2.

Data Structure

- Set of parts: rooted tree
- Optimal partition: cut of the tree
- Algorithm: depth-first search

3.

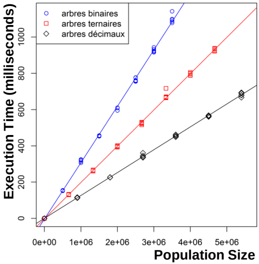
Algorithm 1 for the HSPP

Require: A tree with a label *cost* on each node representing the cost of the corresponding admissible part.

Ensure: Each node of the tree has a Boolean label *optimalCut* representing an optimal partition (see above).

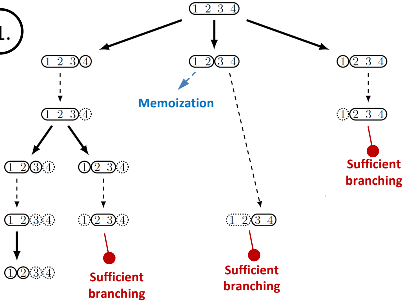
```
procedure SOLVEHSPP(node)
    if node has no child then
        node.optimalCost ← node.cost
        node.optimalCut ← true
    else
        MCost ← node.cost
        µCost ← 0
        for each child of node do
            SOLVEHSPP(child)
            µCost ← µCost + child.optimalCost
        node.optimalCost ← max(µCost, MCost)
        node.optimalCut ← (µCost < MCost)
```

Linear Complexity



Application to the Ordered SPP

1.



2.

Data Structure

- Set of parts: triangular matrix
- Optimal partition: array of cuts
- Algorithm: dynamic programming

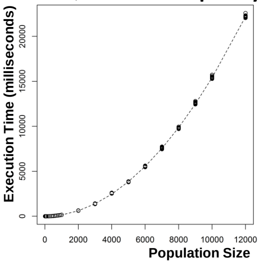
3.

Algorithm 2 for the OSPP

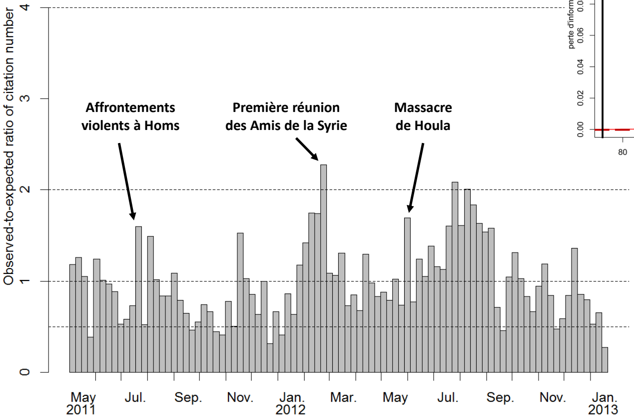
Require: A matrix *cost* recording the costs of intervals.
Ensure: The vector *optimalCut* represents an optimal partition (see text above).

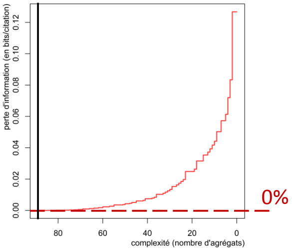
for $j \in [\![1, n]\!]$ **do**
  $optimalCost[j] \leftarrow cost[1, j]$
  $optimalCut[j] \leftarrow 1$
  for $cut \in [\![2, j]\!]$ **do**
    $\mu Cost \leftarrow optimalCost[cut - 1] + cost[cut, j]$
    if $\mu Cost > optimalCost[j]$ **then**
      $optimalCost[j] \leftarrow \mu Cost$
      $optimalCut[j] \leftarrow cut$

Quadratic Complexity

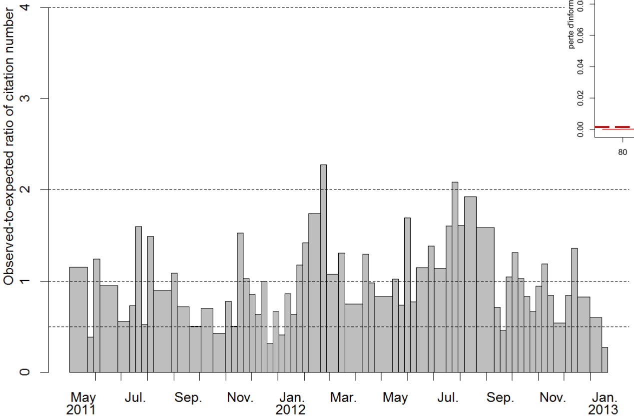


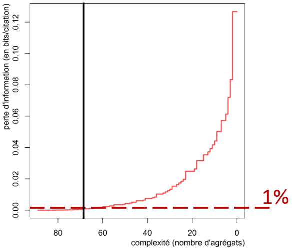
La Syrie vue par Le Monde en 2011 et 2012



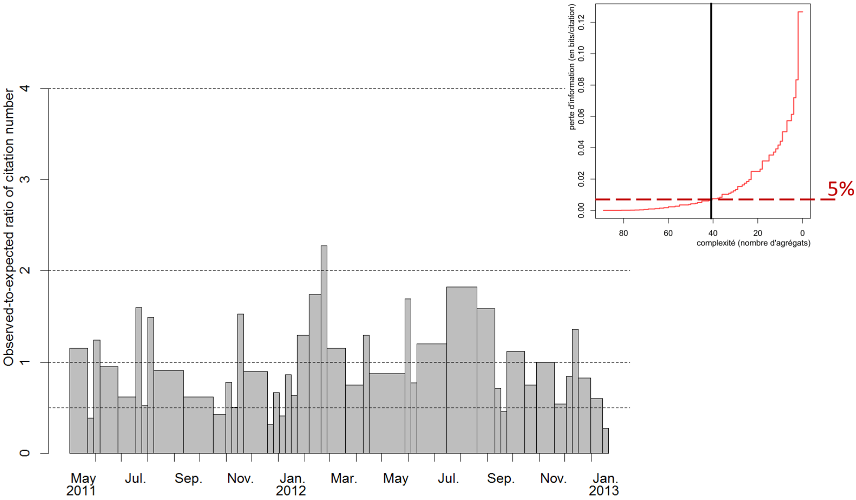


La Syrie vue par Le Monde en 2011 et 2012

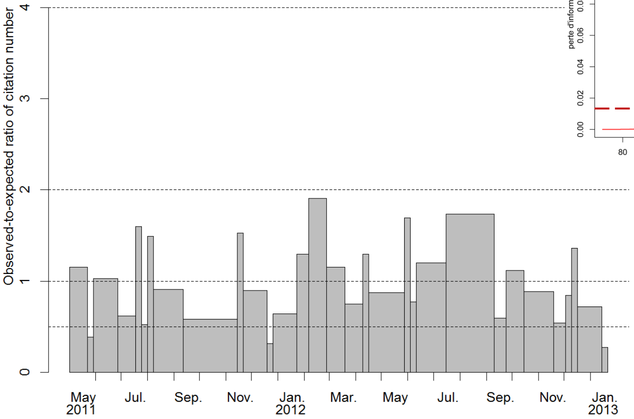


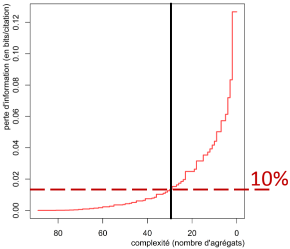


La Syrie vue par Le Monde en 2011 et 2012

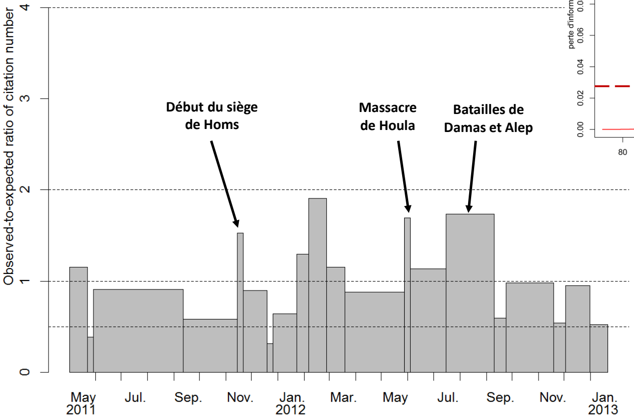


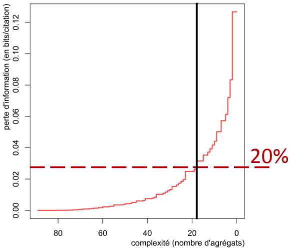
La Syrie vue par Le Monde en 2011 et 2012





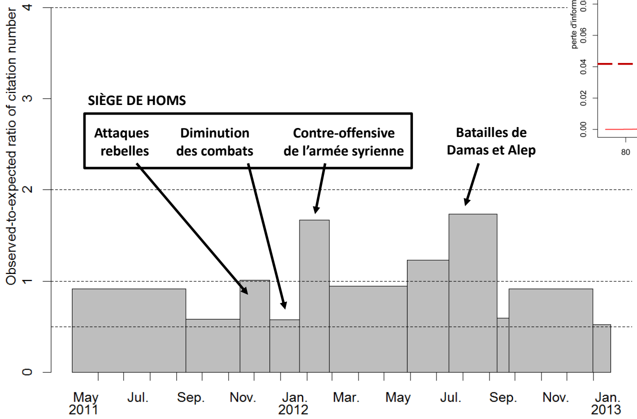
La Syrie vue par Le Monde en 2011 et 2012

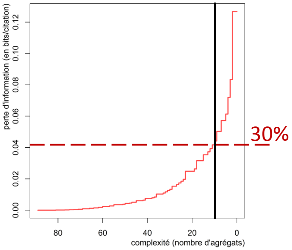




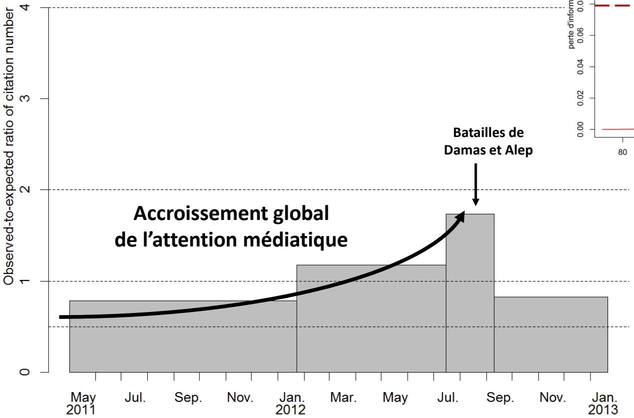
La Syrie vue par Le Monde en 2011 et 2012

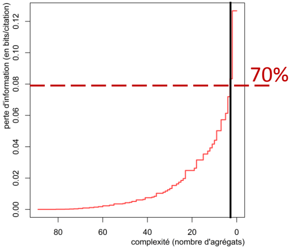
Source : Wikipedia
Timeline of the Syrian civil war
Siege of Homs



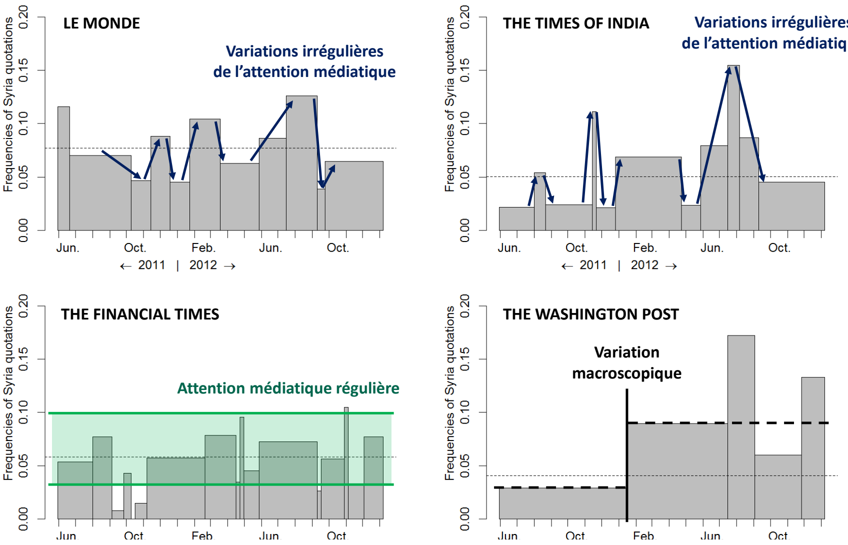


La Syrie vue par Le Monde en 2011 et 2012

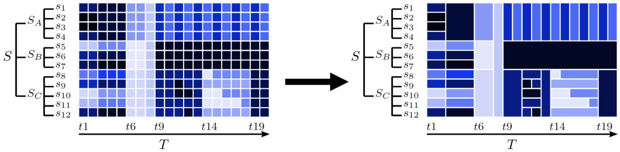




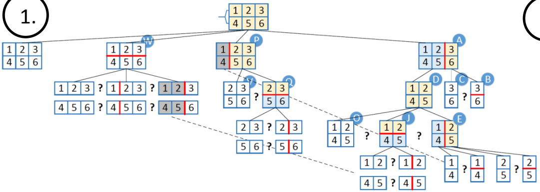
La Syrie vue par 4 journaux différents



Application to a Bidimensional Version of the SPP



1.



2.

Data Structure

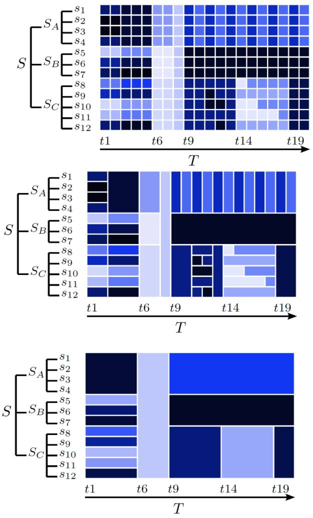
- Set of parts: rooted tree of triangular matrices
- Optimal partition: cut of the tree and arrays of cuts
- Algorithm: depth-first search and dynamic programming

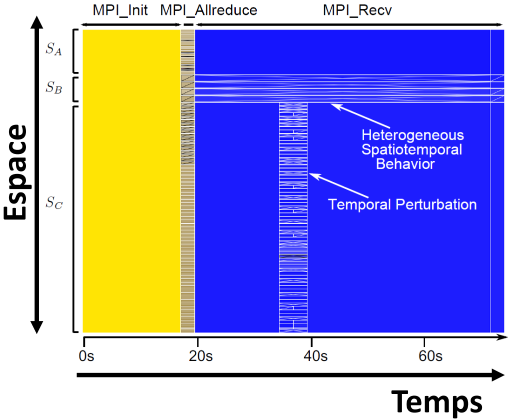
3.

Algorithm 1 computes a hierarchy-and-order-consistent partition that maximizes the parametrized information criterion

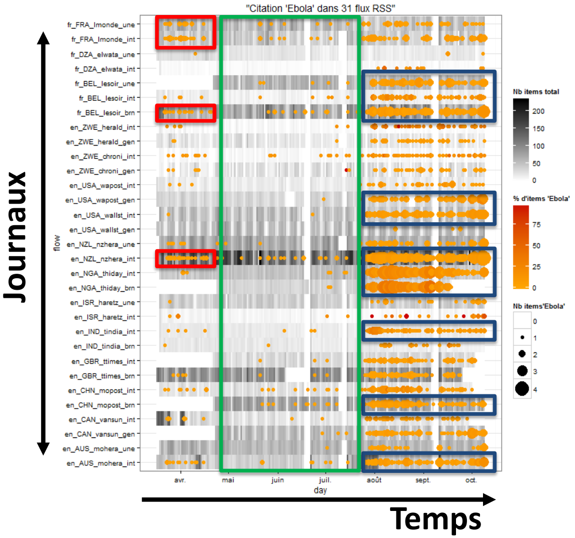
procedure *node*.COMPUTEOPTIMALPARTITION(p)
  for each *child* **do** ▷ Recursion
    child.COMPUTEOPTIMALPARTITION(p)
  for $i = |T| - 1, \dots, 0$ **do** ▷ Iteration
    for $j = i, \dots, |T| - 1$ **do**
      $cut[i,j] \leftarrow j$ ▷ No cut
      $pIC[i,j] \leftarrow p.gain[i,j] - (1-p).loss[i,j]$
      if has children **then** ▷ Spatial cut?
        $pIC_s \leftarrow 0$
        for each *child* **do**
          $pIC_s \leftarrow pIC_s + child.pIC[i,j]$
        if $pIC_s > pIC[i,j]$ **then**
          $cut[i,j] \leftarrow -1$
          $pIC[i,j] \leftarrow pIC_s$
      for $cut_t = i, \dots, j-1$ **do** ▷ Temporal cut?
        $pIC_t \leftarrow pIC[i, cut] + pIC[cut+1, j]$
        if $pIC_t > pIC[i,j]$ **then**
          $cut[i,j] \leftarrow cut_t$
          $pIC[i,j] \leftarrow pIC_t$

Agrégation spatio-temporelle



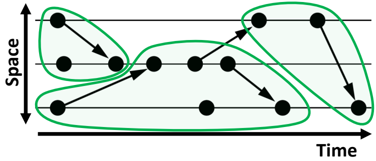


Agrégation médiatique

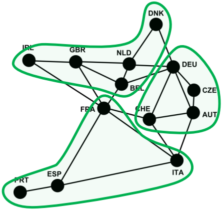


Application Perspectives

Partitioning of Interaction Diagrams [Mattern, 1989]



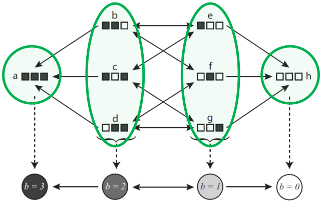
Partitioning of Graphs



Partitioning of Interaction Matrices

	ESP	FRA	GBR	BEL	CHE
ESP	X	12	11	10	4
FRA	14	X	12	12	5
GBR	20	11	X	6	9
BEL	15	9	6	X	5
CHE	10	16	17	9	X

Partitioning the State Space of Dynamical Systems [Banisch *et al.*, 2013]



Deuxième partie II

Prédiction optimale de de processus stochastiques multi-échelles

General Setting

$$
\begin{tikzcd}
 & \phi(X^t) \arrow[r, dashed, "\text{prediction}"] & \psi(X^{t+\tau}) \\
X^0 \arrow[r, "T^t"] & X^t \arrow[u, "\phi"] \arrow[r, "T^\tau"] & X^{t+\tau} \arrow[u, "\psi"']
\end{tikzcd}
$$

• Initial State	$X^0 \in \Sigma$	within a finite state space Σ
• Current State	$X^t \in \Sigma$	with current time $t \in \mathbb{N}$
• Future State	$X^{t+\tau} \in \Sigma$	with prediction horizon $\tau \in \mathbb{N}$
• Transition Kernel	$T(X^{t+1} \vert X^t)$	(Markovian and time-homogeneous)
• Pre-measurement	$\phi : \Sigma \to \mathcal{S}_\phi$	defined by $\Pr(\phi(X) \vert X)$
• Post-measurement	$\psi : \Sigma \to \mathcal{S}_\psi$	defined by $\Pr(\psi(X) \vert X)$

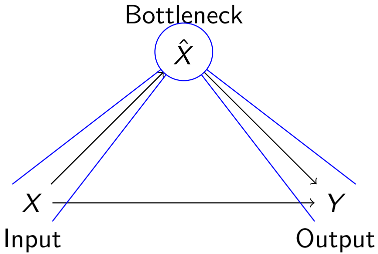
General Setting

$$
\begin{tikzcd}
 & \phi(X^t) \arrow[r, dashed, "\text{prediction}"] & \psi(X^{t+\tau}) \\
X^0 \arrow[r, "T^t"] & X^t \arrow[u, "\phi"] \arrow[r, "T^\tau"] & X^{t+\tau} \arrow[u, "\psi"]
\end{tikzcd}
$$

The Optimal Prediction Problem

Given a transition kernel T, an initial distribution X^0, a time t, an horizon τ, and a post-measurement ψ, **find** a pre-measurement ϕ such that $\phi(X^t)$ is "efficient" to predict $\psi(X^{t+\tau})$.

The Information Bottleneck Method (Tishby et al., 1999)

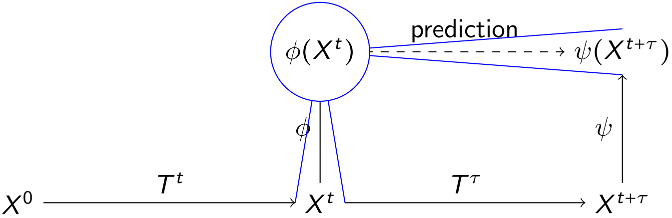


- **Maximize** Bottleneck Capacity $\quad max_{\Pr(\hat{X}|X)}\ I(\hat{X};Y)$
- **Minimize** Bottleneck Complexity $\quad min_{\Pr(\hat{X}|X)}\ I(X;\hat{X})$

- **Minimize** the IB-variational

$$min_{\Pr(\hat{X}|X)}\ I(X;\hat{X}) - \beta\, I(\hat{X};Y) \quad \text{with } \beta \in \mathbb{R}^+$$

The Optimal Prediction Problem



- **Minimize** Measurement Complexity $min_\phi \; I(X^t; \phi(X^t))$
- **Maximize** Predictive Capacity $max_\phi \; I(\phi(X^t); \psi(X^{t+\tau}))$

The Optimal Prediction Problem

Given a time t, an horizon τ, a post-measurement ψ, and a trade-off parameter $\beta \in \mathbb{R}^+$, **find** a pre-measurement ϕ with minimal complexity and maximal anticipatory capacity :

$$\min_\phi \; I(X^t; \phi(X^t)) \; - \; \beta \; I(\phi(X^t); \psi(X^{t+\tau}))$$

Constraining the Set of Feasible Measurements

$$
\begin{tikzcd}
\{\phi_1(X^t), & \phi_2(X^t), & \dots, & \phi_k(X^t)\} \arrow[rr, dashed, "\text{prediction}"] & & \psi(X^{t+\tau}) \\
X^0 \arrow[rrr, "T^t"] & & & X^t \arrow[ulll, "\phi_1"] \arrow[ull, "\phi_2"] \arrow[u, "\phi_k"] \arrow[rr, "T^\tau"] & & X^{t+\tau} \arrow[u, "\psi"]
\end{tikzcd}
$$

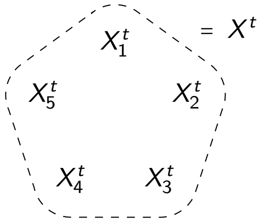
The Constrained Optimal Prediction Problem

Given a time t, an horizon τ, a post-measurement ψ, a trade-off parameter $\beta \in \mathbb{R}^+$, and a set of feasible pre-measurements $\Phi = \{\phi_1, \dots, \phi_k\}$, **find** a feasible pre-measurement ϕ with minimal complexity and maximal anticipatory capacity :

$$\min_{\phi \subset \Phi} \; I(X^t; \phi(X^t)) \; - \; \beta \; I(\phi(X^t); \psi(X^{t+\tau}))$$

Agent-based Systems

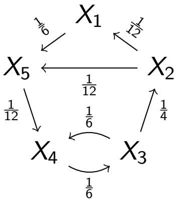
- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in S$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \Sigma = S^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$



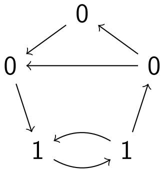
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j, \quad X_k^{t+1} = X_k^t$



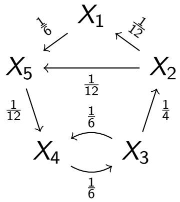
$t = 0$



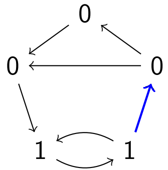
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



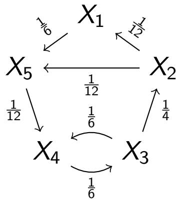
$t = 0$ $\rightarrow$ arc $(3, 2)$



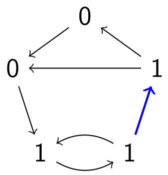
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



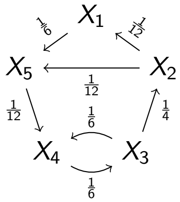
$t = 1$



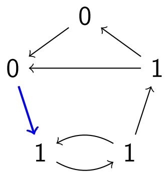
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



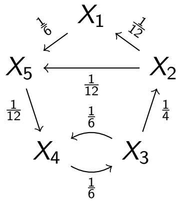
$t = 1$ $\rightarrow$ arc $(5, 4)$



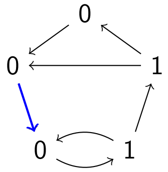
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

$$\text{arc } (i, j) \text{ selected} \quad \Rightarrow \quad X_j^{t+1} = X_i^t \quad \text{and} \quad \forall k \neq j, \quad X_k^{t+1} = X_k^t$$



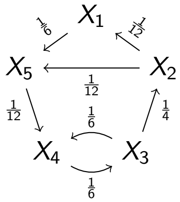
$t = 2$



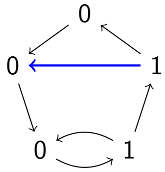
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0,1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0,1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

$$\text{arc } (i,j) \text{ selected} \quad \Rightarrow \quad X_j^{t+1} = X_i^t \quad \text{and} \quad \forall k \neq j, \quad X_k^{t+1} = X_k^t$$



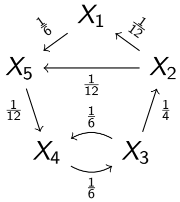
$t = 2 \quad \rightarrow \quad \text{arc } (2,5)$



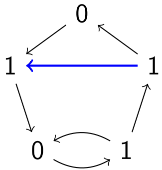
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



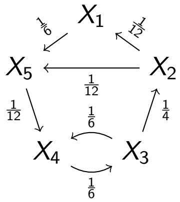
$t = 3$



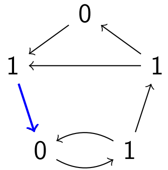
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j,$ $X_k^{t+1} = X_k^t$



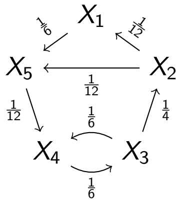
$t = 3 \quad \rightarrow \quad$ arc $(5, 4)$



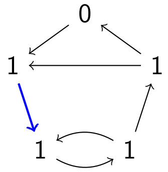
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

$$\text{arc } (i,j) \text{ selected} \quad \Rightarrow \quad X_j^{t+1} = X_i^t \quad \text{and} \quad \forall k \neq j, \quad X_k^{t+1} = X_k^t$$



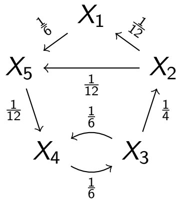
$t = 5$



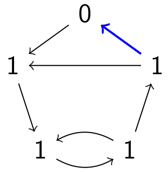
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \ldots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \ldots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

arc (i, j) selected $\Rightarrow$ $X_j^{t+1} = X_i^t$ and $\forall k \neq j$, $X_k^{t+1} = X_k^t$



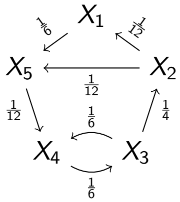
$t = 5$ $\rightarrow$ arc $(2, 1)$



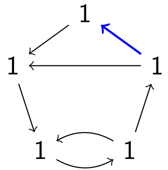
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

$$\text{arc } (i, j) \text{ selected} \quad \Rightarrow \quad X_j^{t+1} = X_i^t \quad \text{and} \quad \forall k \neq j, \quad X_k^{t+1} = X_k^t$$



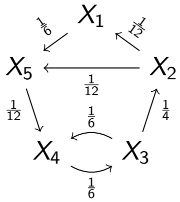
$t = 6$



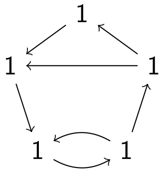
The Voter Model (Banisch & Lima, 2012)

- Set of Agents $\Omega = \{1, \dots, N\}$
- State of ith Agent $X_i^t \in \{0, 1\}$
- System State $X^t = (X_1^t, \dots, X_N^t) \in \{0, 1\}^N$
- Transitions Kernel $T(X_i^{t+1} | X^t)$ given by an interaction graph :

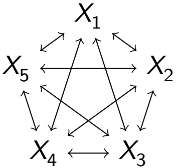
$$\text{arc } (i, j) \text{ selected} \quad \Rightarrow \quad X_j^{t+1} = X_i^t \quad \text{and} \quad \forall k \neq j, \quad X_k^{t+1} = X_k^t$$



$t = 6$



The Complete Graph

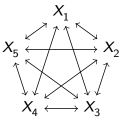


- All arcs are equally likely $\quad \forall (i,j) \in \Omega^2, \quad \Pr(\text{arc } (i,j)) = \frac{1}{N\,(N-1)}$
- Uniform Initial State $\quad \forall x \in \{0,1\}^N, \quad p(X^0 = x) = 2^{-N}$
- Two Equiprobable Attractors $\quad X^\infty \in \{(0,\ldots,0),(1,\ldots,1)\}$

- Aggregated-state (Generic) Measurement $\quad \forall A \subset \Omega, \quad \eta_A(x) = \sum_{i \in A} x_i$

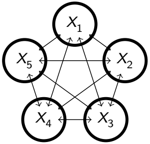
EMPTY

$\eta_{\emptyset}(x) = 0$



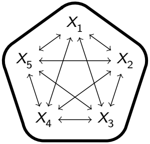
MICRO

$(\eta_{\{1\}}, \ldots, \eta_{\{N\}})(x)$



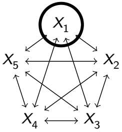
MACRO

$\eta_{\Omega}(x) = x_1 + \ldots + x_N$



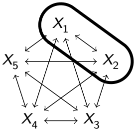
AGENT

$\eta_{\{1\}}(x) = x_1$



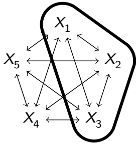
SIZE2

$\eta_{\{1,2\}}(x) = x_1 + x_2$

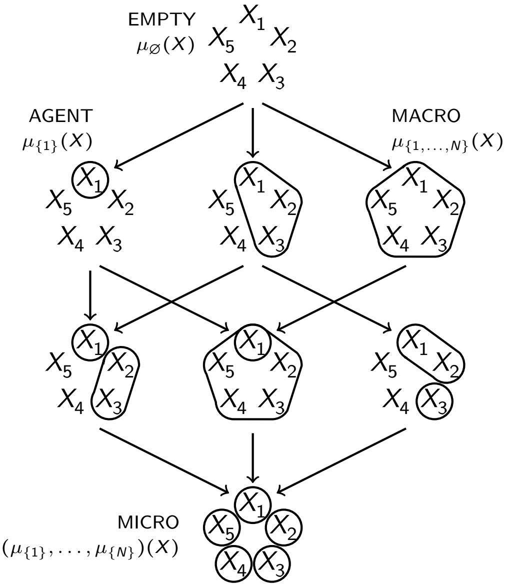


SIZE3

$\eta_{\{1,2,3\}}(x) = x_1 + x_2 + x_3$



Example of Feasible Measurements



Definition 1 (Refinement)

$\phi_1 < \phi_2$
iff $X \to \phi_1(X) \to \phi_2(X)$ is aMarkov chain
iff $I(X;\phi_2(X)|\phi_1(X)) = 0$

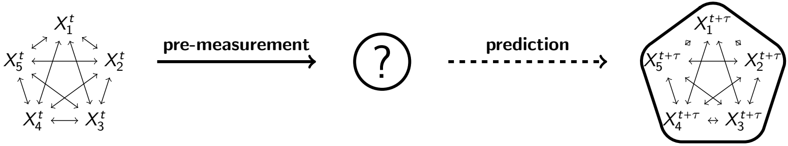
Theorem 2 (Monotonicity)

$\phi_1 \subset \phi_2 \;\Rightarrow$
$I(X^t;\phi_1(X^t)) \geq I(X^t;\phi_2(X^t))$ *and*
$I(\phi_1(X^t);\psi(X^{t+\tau})) \geq I(\phi_2(X^t);\psi(X^{t+\tau}))$

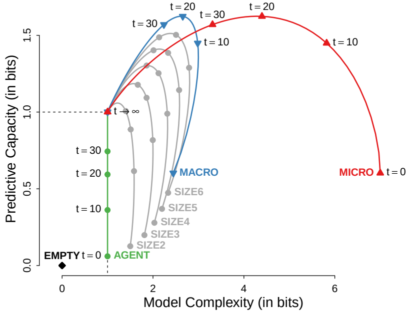
Definition 3 (Additivity)

μ additive iff $\forall A \cap B = \varnothing$,
$H(\mu_{A\cup B}(X) \mid \mu_A(X), \mu_B(X)) = 0$
and $H(\mu_A(X) \mid \mu_{A\cup B}(X), \mu_B(X)) = 0$

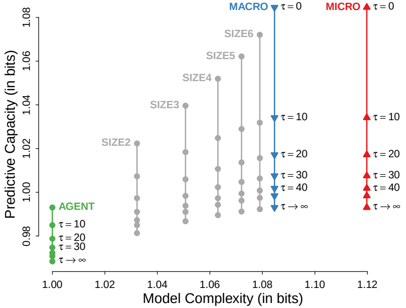
Predicting the Macroscopic Measurement



Fixed size $N = 7$
Fixed horizon $\tau = 3$



Fixed size $N = 7$
Fixed time $t = 100$



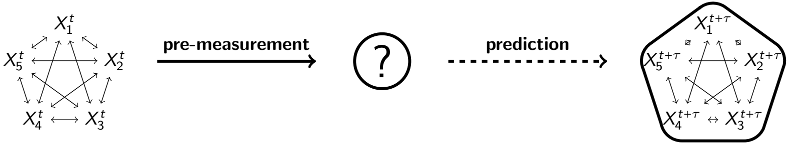
Information Bottleneck Diagram

- Tridimensional Parameter Space $(t, \tau, \beta) \in \mathbb{N} \times \mathbb{N} \times \mathbb{R}^+$
- Given two pre-measurements ϕ_1 and ϕ_2 :

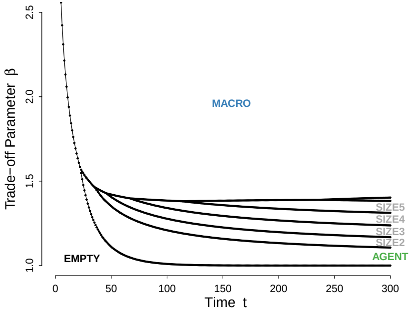
The border is...	when...	ϕ_1 opt. region	ϕ_2 opt. region
(1) strictly positive	$H_1 < H_2$ and $I_1 < I_2$	$[0, \beta^{t,\tau}_{\phi_1,\phi_2}]$	$[\beta^{t,\tau}_{\phi_1,\phi_2}, +\infty[$
	$H_1 > H_2$ and $I_1 > I_2$	$[\beta^{t,\tau}_{\phi_1,\phi_2}, +\infty[$	$[0, \beta^{t,\tau}_{\phi_1,\phi_2}]$
(2) null	$H_1 = H_2$ and $I_1 < I_2$	$\{0\}$	$[0, +\infty[$
	$H_1 = H_2$ and $I_1 > I_2$	$[0, +\infty[$	$\{0\}$
(3) infinite	$H_1 < H_2$ and $I_1 = I_2$	$[0, +\infty[$	$\{+\infty\}$
	$H_1 > H_2$ and $I_1 = I_2$	$\{+\infty\}$	$[0, +\infty[$
(4) defined nowhere	$H_1 < H_2$ and $I_1 > I_2$	$[0, +\infty[$	$\varnothing$
	$H_1 > H_2$ and $I_1 < I_2$	$\varnothing$	$[0, +\infty[$
(5) defined everywhere	$H_1 = H_2$ and $I_1 = I_2$	$[0, +\infty[$	$[0, +\infty[$

with $H_i = I(X^t; \phi_i(X^t))$, $I_i = I(\phi_i(X^t); \psi(X^{t+\tau}))$ and $\beta^{t,\tau}_{\phi_1,\phi_2} = \frac{H_2 - H_1}{I_2 - I_1}$.

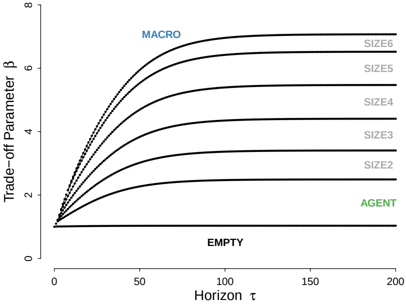
Predicting the Macroscopic Measurement



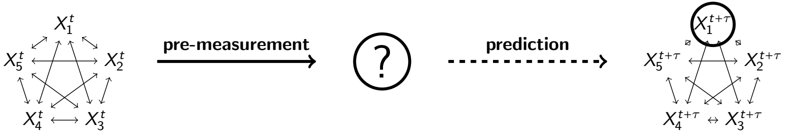
Fixed size $N = 7$
Fixed horizon $\tau = 3$



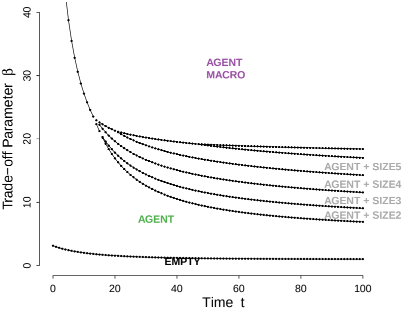
Fixed size $N = 7$
Fixed time $t = 100$



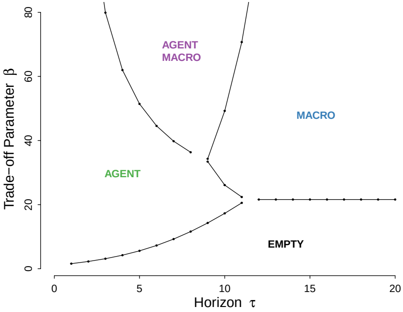
Predicting the Agent Measurement



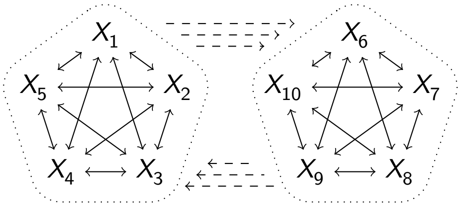
Fixed size $N = 7$
Fixed horizon $\tau = 3$



Fixed size $N = 7$
Fixed time $t = 0$



The Two-community Voter Model



First Community Ω_1

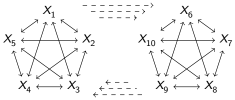
Second Community Ω_2

Coupling Parameter $\quad \rho = \frac{\Pr(\text{inter edge})}{\Pr(\text{intra edge})} < 1$

The Two-community Voter Model

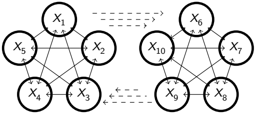
EMPTY

$\eta_\varnothing$



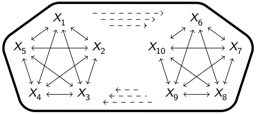
MICRO

$\{\eta_{\{1\}}, \dots, \eta_{\{N\}}\}$



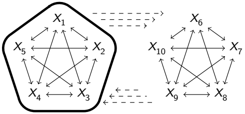
MACRO

η_Ω



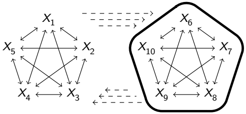
MESO1

η_{Ω_1}



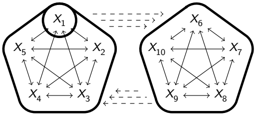
MESO2

η_{Ω_2}

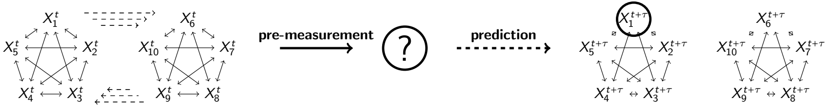


AGENT+MESO1+MESO2

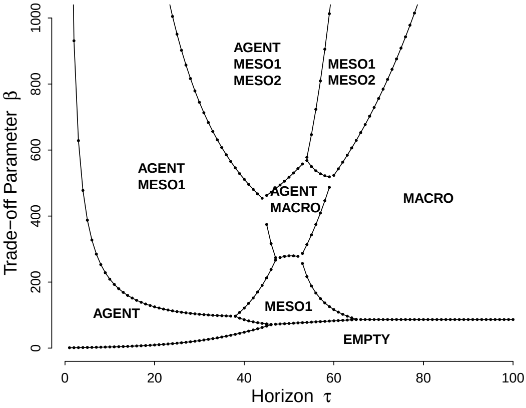
$(\eta_{\{1\}}, \eta_{\Omega_1}, \eta_{\Omega_2})$



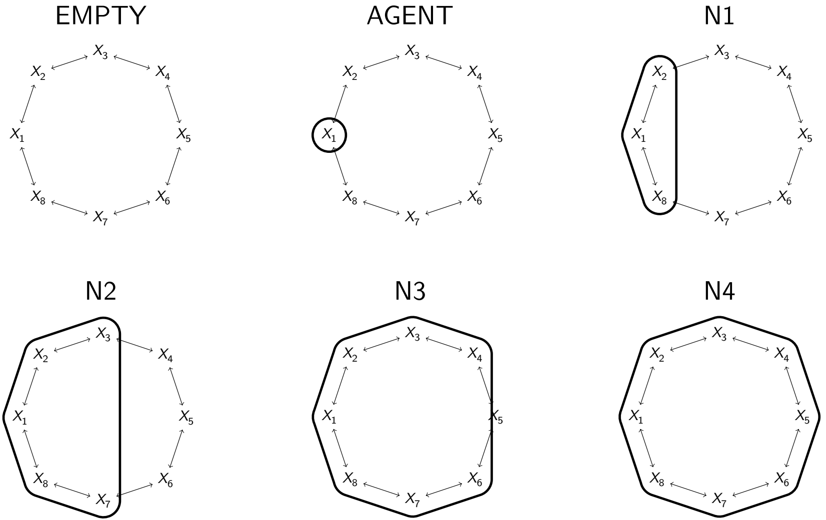
Predicting the Agent Measurement



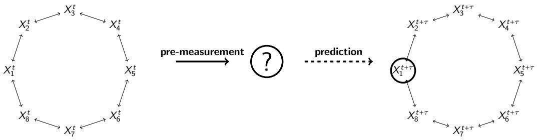
Fixed size and coupling $N_1 = 10$, $N_2 = 10$, $\rho = \frac{1}{5}$
Fixed time $t = 0$



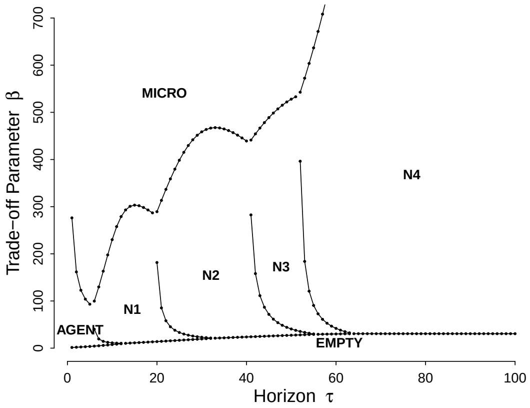
The Ring Voter Model



Predicting the Agent Measurement in the Ring



Fixed size $N = 9$
Fixed time $t = 0$



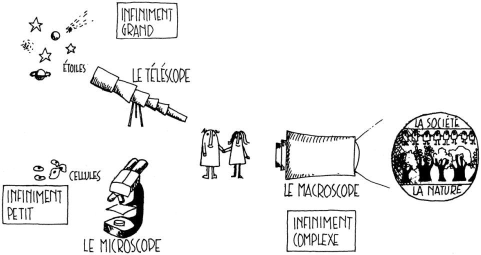
Application Perspectives

- Application to efficient prediction of **population dynamics in ecology :** different species compete for survival in the environment
  - Multilevel measurements regarding the geographical space (related to geographical data aggregation)
  - Multilevel measurements regarding the structure of inter-species mutual dependencies (webs of food, reproduction dynamics)
  - Complexity should be used to model data collection costs

- Application to efficient prediction of **trade flows in economy** : countries exchange products on a global scale
  - Multilevel measurements regarding the network of international trade (related to graph theory and community modelling)
  - Multilevel measurements regarding the structure of products (production chains, economic fields)
  - Complexity should also be used to model data collection costs

Merci pour votre attention

Courriel: Robin.Lamarche-Perrin@mis.mpg.de
Web: www.mis.mpg.de/jjost/members/robin-lamarche-perrin.html



« Aujourd'hui nous sommes confrontés à un autre infini : l'infiniment complexe. Mais cette fois, plus d'instrument. »

Joël de Rosnay, *Le macroscope*, 1975