From Multilevel Data Analysis to Multilevel Autonomous Learning

Max Planck Institute for Mathematics in the Sciences

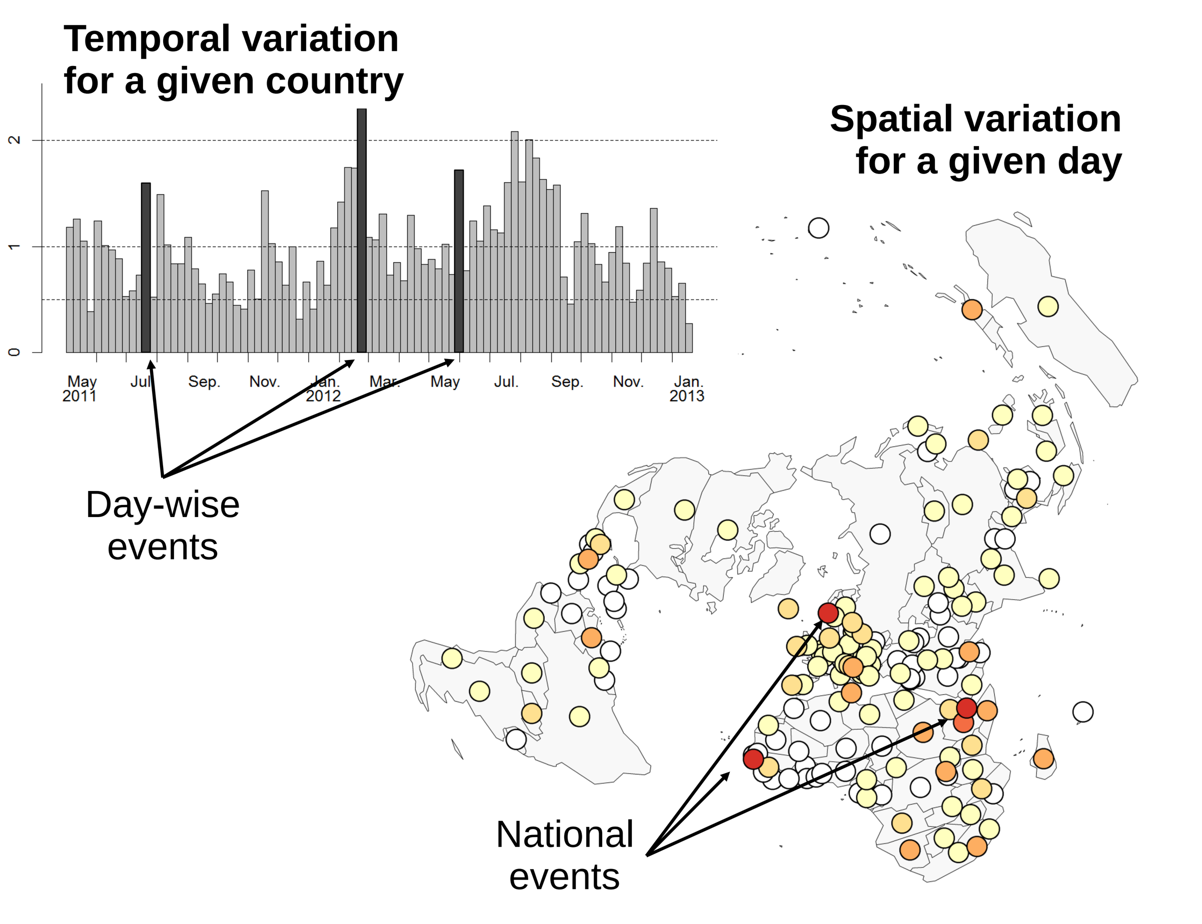
Robin Lamarche-Perrin

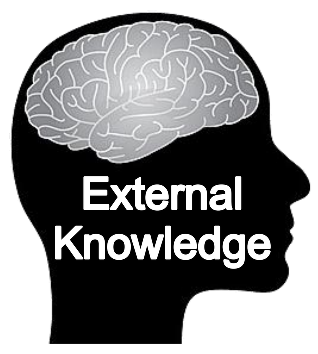
Sept. 2014

Starting Point — Multilevel Analysis of Geographical Data

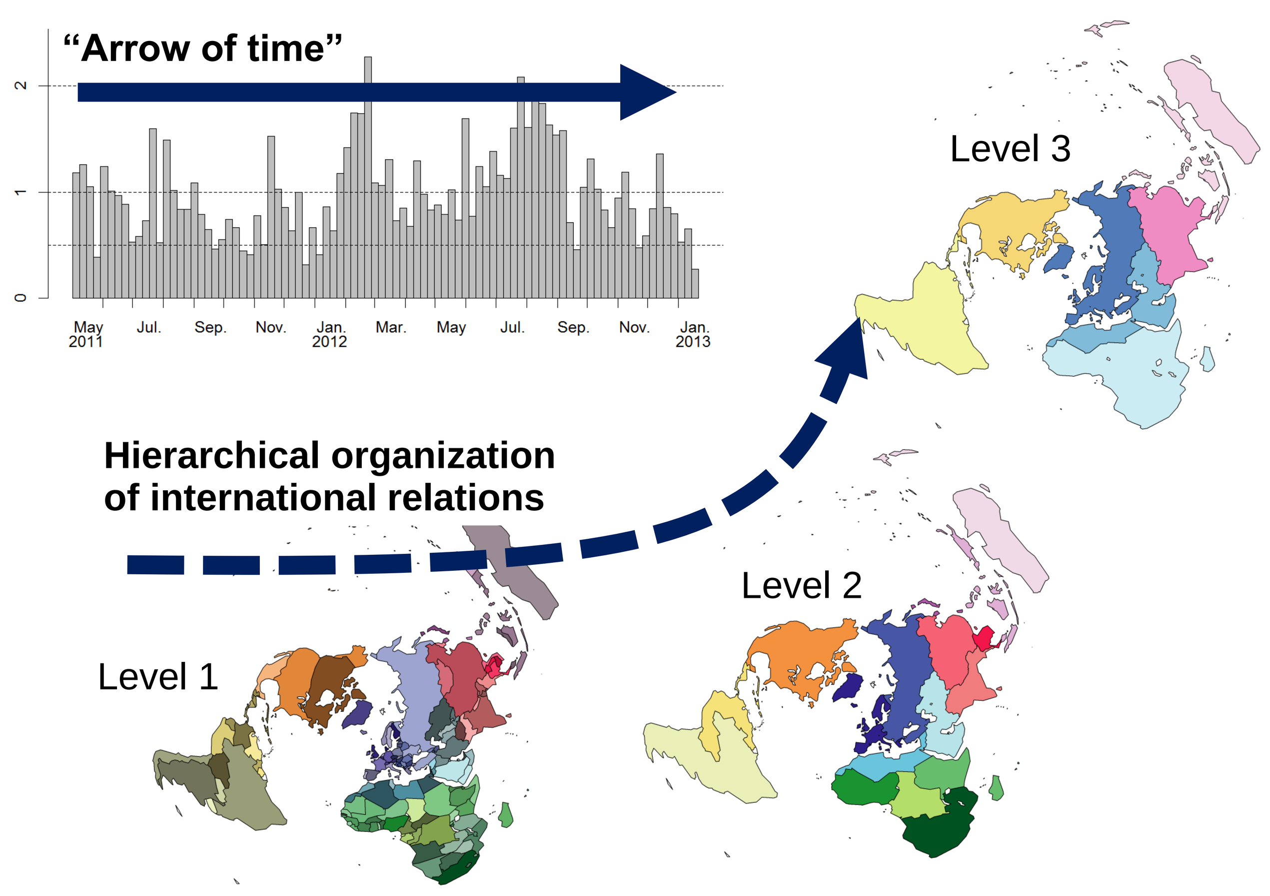
Lamarche-Perrin *et al.* Building Optimal Macroscopic Representations of Complex Multi-agent Systems. TCCI XV, LNCS 8670, 2014.

Microscopic Data
Demographic data at the day or country level

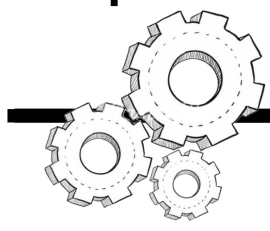




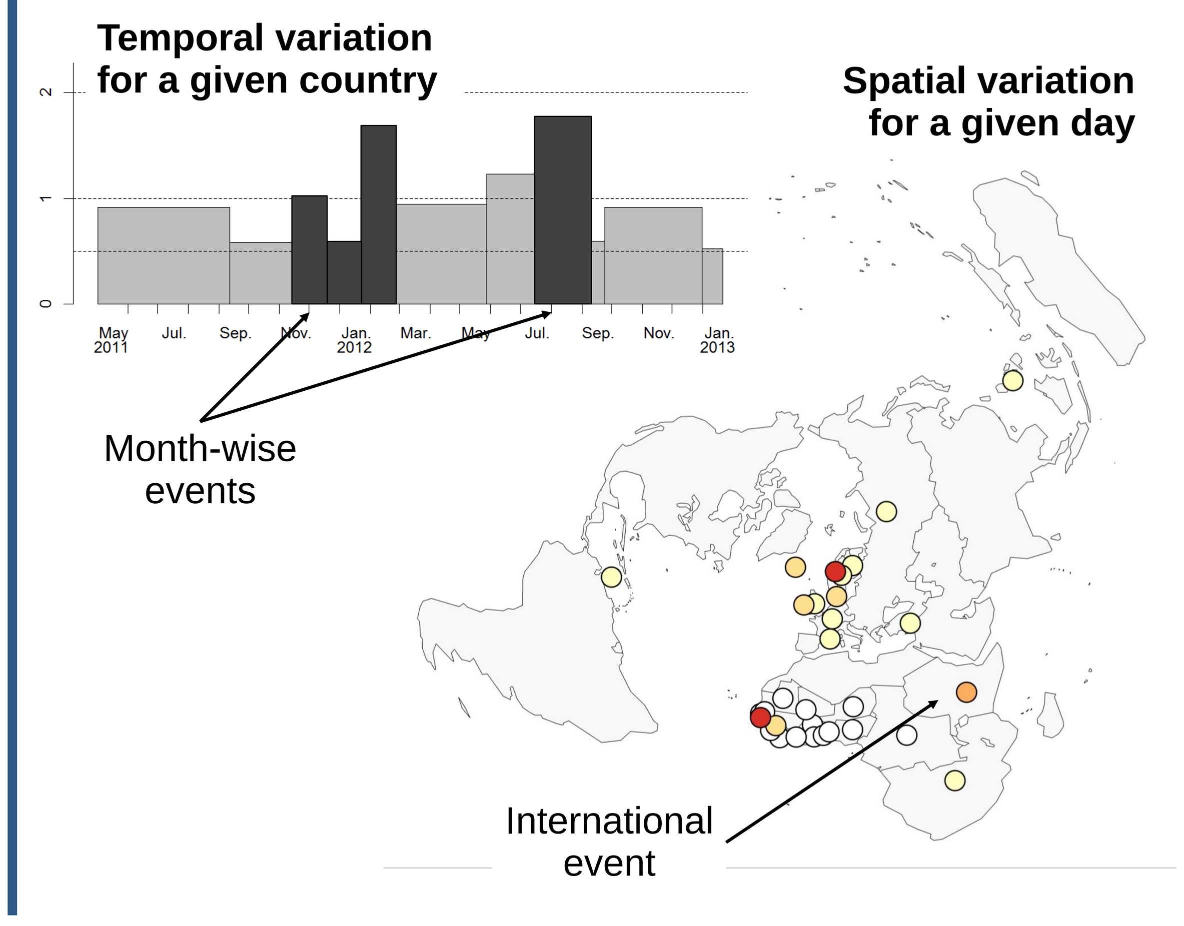
Structural Properties
Geographical and cultural relations between countries



Aggregation Operator



Aggregated Data
Demographic data at mesoscopic levels



Model and Method — Solving Special Versions of the Set Partitioning Problem

Lamarche-Perrin *et al.* A Generic Algorithmic Framework to Solve Special Versions of the Set Partitioning Problem. ICTAI 2014.

Microscopic Data is modeled by a categorical random variable X on a finite sample space $\Omega = \{x_1, \dots, x_n\}$

Aggregation Operator is modeled by a partition $\mathcal{X} = \{X_1, \dots, X_k\}$ of the state space Ω inducing a prob. distribution $p(X) = \sum_{x \in X} p(x)$

Complexity Reduction is quantified by the cardinality reduction of the partition:

$$\Delta C(\mathcal{X}) = |\Omega| - |\mathcal{X}|$$

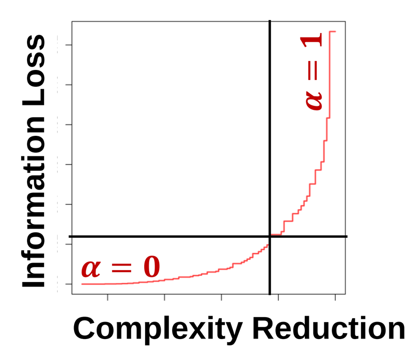
Information Loss is quantified by the KL divergence between the microscopic and the aggregated probability distributions:

$$\Delta I(\mathcal{X}) = \sum_{X \in \mathcal{X}} \sum_{x \in X} p(x) \log_2 \left(\frac{p(x)|X|}{p(X)} \right)$$

Partition Quality is quantified by a parameterized trade-off between complexity reduction and information loss:

$$Q_\alpha(\mathcal{X}) = \alpha \frac{\Delta C(\mathcal{X})}{\Delta C(\{\Omega\})} - (1-\alpha) \frac{\Delta I(\mathcal{X})}{\Delta I(\{\Omega\})}$$

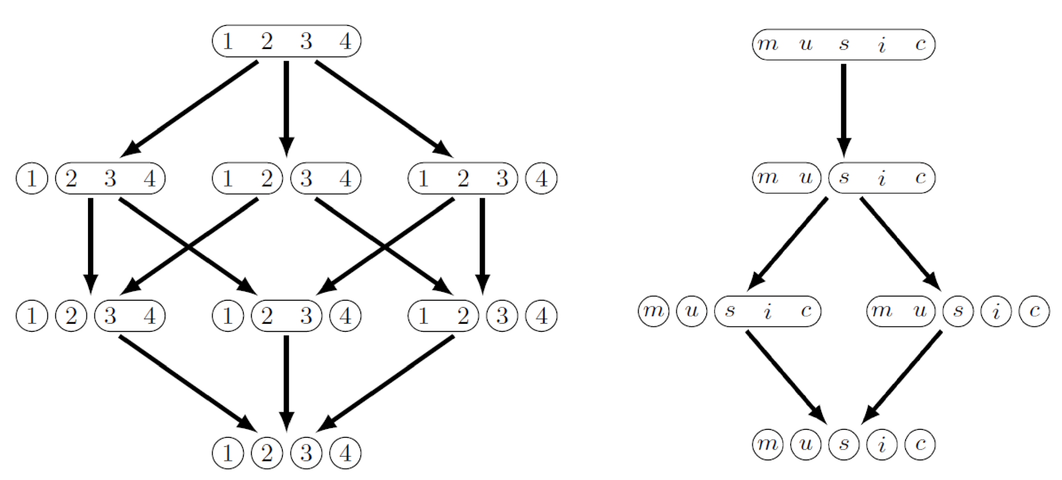
where $\alpha \in [0,1]$



$\alpha = 0$ → microscopic partition
$\alpha \in]0,1[$ → multilevel partition
$\alpha = 1$ → macroscopic partition

Structural Properties are modeled by posets of feasible partitions $\mathfrak{P}$

Posets of feasible partitions $\mathfrak{P}$ in the case of a **hierarchical** space / an **ordered** space

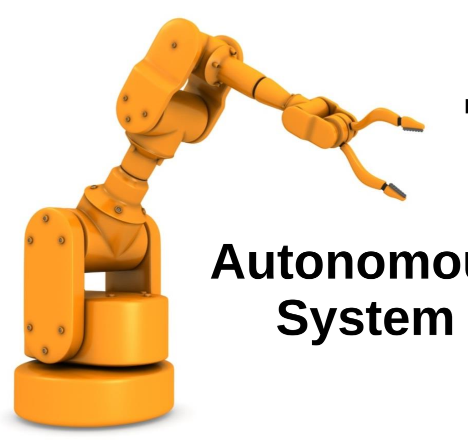


Multilevel Data Aggregation is achieved by solving the following combinatorial optimization problem, the **Set Partitioning Problem**:

$$\arg\max_{\mathcal{X} \in \mathfrak{P}} Q_\alpha(\mathcal{X})$$

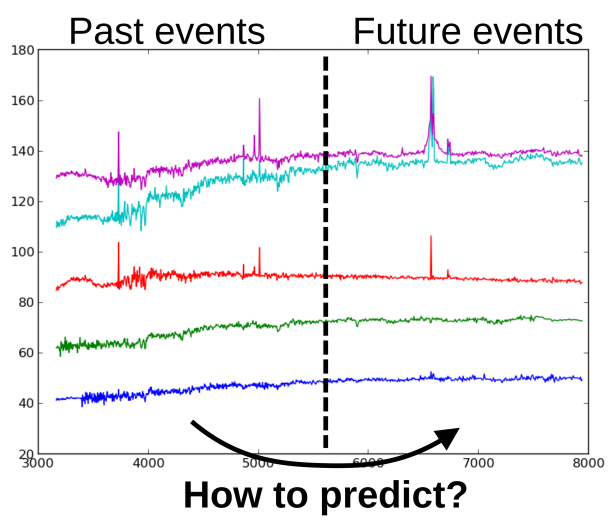
To do so, one might use dynamic programming… but this is another story! **(see ICTAI 2014 paper)**

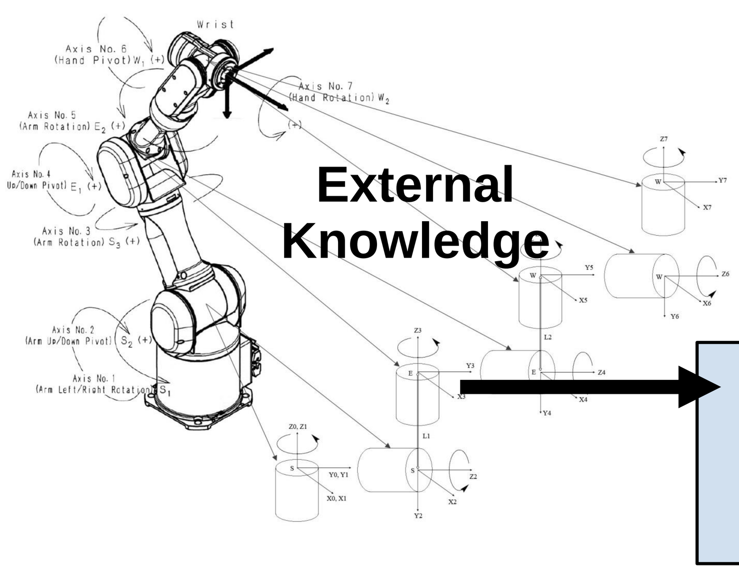
Perspective — An Application to Multilevel Autonomous Learning



Autonomous System

Microscopic Data
Inputs from sensors
Outputs from actuators

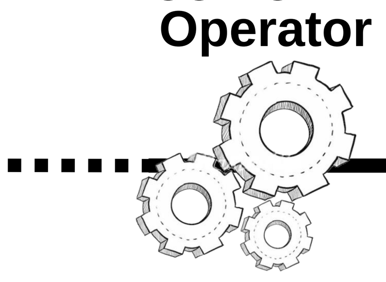




External Knowledge

Structural Properties
External knowledge regarding sensor-actuator articulations

Aggregation Operator



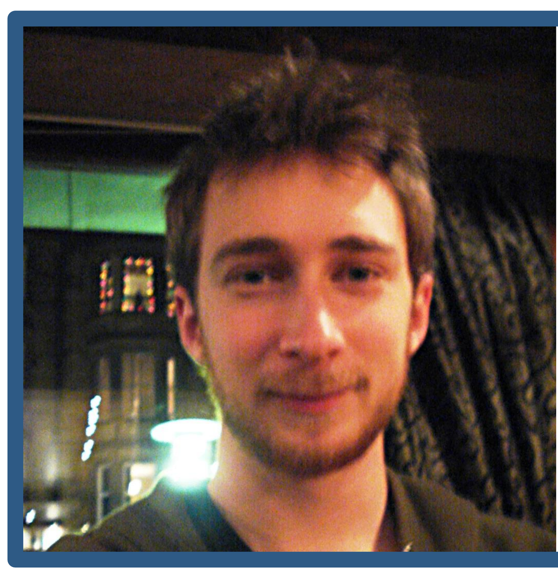
Aggregated Data
Learned predictive patterns between sensors and actuators

Objective: How to predict sensory inputs according to actuation outputs?

Temporal Aggregation: The system might more easily predict
(1) future *macro*-events than future *micro*-events;
(2) according to past *macro*-events than past *micro*-events.

Sensor and Actuator Aggregation: The system might more easily predict
(1) *high-level* sensory patterns than *low-level* sensory events;
(2) according to *high-level* actuation patterns than *low-level* actuation events.

Multidimensional Aggregation:
Which *actuation macro-patterns* (temporally-aggregated actuation outputs) might efficiently predict *sensory macro-patterns* (temporally-aggregated sensory inputs)?



Dr Robin Lamarche-Perrin

Max-Planck-Institute for Mathematics in the Sciences
Office C10 / Inselstraße 22 / 04103 Leipzig / Germany

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