# *Automedian sets of permutations*

Sylvie Hamel

Département d'informatique et de recherche opérationnelle (DIRO),
Université de Montréal, Québec, Canada

DIRO

Permutation Patterns 2017
June 26-30 2017, Reykjavík, Iceland

# Quebec political parties :

COALITION AVENIR QUÉBEC 10

Bloc Pot 7

Parti démocratie chrétienne du Québec 11

PMLQ 6
Parti marxiste-léniniste du Québec

PCQ 5
Parti communiste du Québec

PARTI INDÉPENDANTISTE 4

Parti Libéral du Québec 9

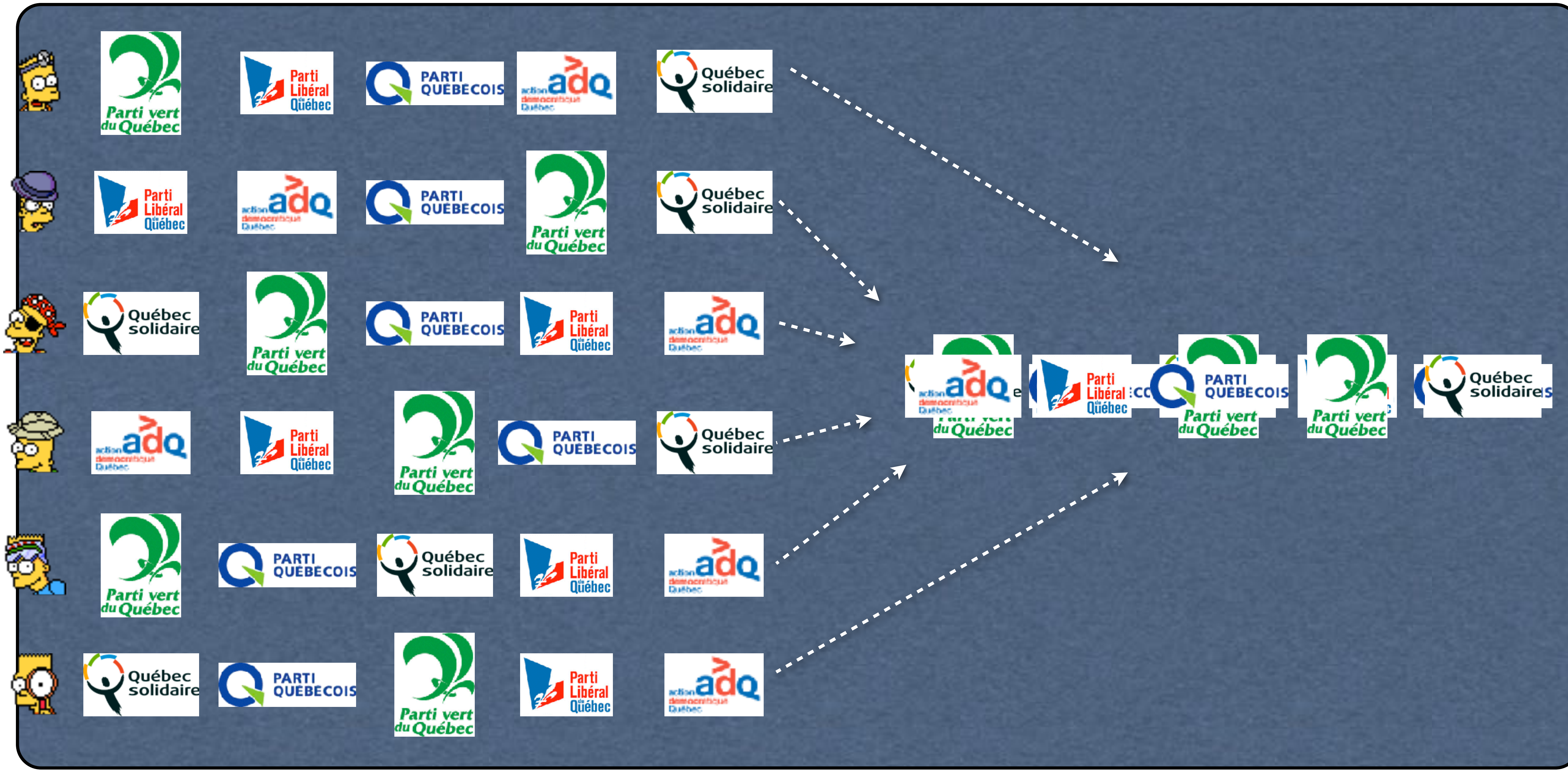
NUL 8

PARTI QUÉBECOIS 2

Parti vert du Québec 3

Québec solidaire 1 ✓

# Kemeny consensus :



# The Kendall-$\tau$ distance:

Maurice Kendall

Counts the number of order disagreements between pairs of elements in two permutations i.e

$$d_{KT}(\pi,\sigma) = \#\{(i,j) | i < j \; and \; [(\pi_i^{-1} < \pi_j^{-1} \; and \; \sigma_i^{-1} > \sigma_j^{-1}) \text{ or } (\pi_i^{-1} > \pi_j^{-1} \; and \; \sigma_i^{-1} < \sigma_j^{-1})]\}$$

- The Kendall-$\tau$ distance is equivalent to the "bubble-sort" distance i.e. the number of transpositions needed to transform one permutation into the other one.
- We have $d_{KT}(\pi, \imath) = inv(\pi)$

# Example:

$$\pi = [1, 4, 2, 5, 3]$$

$$\sigma = [3, 4, 1, 2, 5]$$

$$d_{KT}(\pi, \sigma) = 1 + 1 + 1 + 1 + 1 = 5$$

The Kendall-$\tau$ distance between a permutation $\pi$ and a set of permutations $\mathcal{A} = \{\pi_1, \pi_2, \ldots, \pi_m\}$:

$$d_{KT}(\pi, \mathcal{A}) = \sum_{i=1}^{m} d_{KT}(\pi, \pi_i)$$

Our problem:

Given a set of $m$ permutations $\mathcal{A} \subseteq \mathcal{S}_n$, we want to find a permutation $\pi^*$ such that

$$d_{KT}(\pi^*, \mathcal{A}) \leq d_{KT}(\pi, \mathcal{A}), \forall \pi \in \mathcal{S}_n$$

## What has been done?

Finding a median of a set of m permutations using the Kendall-$\tau$ distance

2016 Milosz et Hamel, space reduction

2015

2014 Betzler et al., space reduction

2013 Nishimura et Simjour, fixed-parameter algorithm

2012

2011 Betzler et al., fixed-parameter algorithm; Blin et al, space reduction

2010 Karpinski et Schudy, fixed-parameter algorithm

2009 vanZuylen et al., deterministic algo. approx. factor 8/5; Simjour, fixed-parameter algorithm

2008 Ailon et al., randomized algo. approx. factor 11/7

2007 Kenyon-Mathieu et Schudy, PTAS result

2006

2005 Biedl et al., first approx. algorithm + correction of Dwork et al. proof

2004

2003

2002

2001 Dwork et al., NP-complete for $m \geq 4$

# Back to our problem:

Given a set of $m$ permutations $\mathcal{A} \subseteq \mathcal{S}_n$, we want to find a permutation $\pi^*$ such that

$$d_{KT}(\pi^*, \mathcal{A}) \leq d_{KT}(\pi, \mathcal{A}), \forall \pi \in \mathcal{S}_n$$

This median is not always unique

Average number of permutations in $M(A)$ for uniformly distributed random sets $A$ *of* $m$ permutations of length $n$. Statistics generated over 100 to 1000 instances.

| $m \backslash n$ | 8 | 10 | 12 | 14 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|---|---|---|
| 3 | 2.1 | 3.0 | 3.7 | 4.8 | 5.6 | 12.2 | 23.1 | 61.4 |
| 4 | 60.6 | 331.4 | 1321.4 | 7551.4 | 14253.8 | - | - | - |
| 5 | 2.2 | 2.9 | 3.6 | 5.2 | 6.2 | 12.9 | 29.1 | 49.2 |
| 6 | 31.3 | 90.6 | 345.1 | 1506.2 | 1614.9 | - | - | - |
| 10 | 13.0 | 36.8 | 88.8 | 201.9 | 315.6 | 2947.9 | - | - |
| 15 | 1.7 | 2.2 | 2.8 | 3.5 | 3.8 | 6.3 | 12.3 | - |
| 20 | 6.3 | 11.4 | 22.2 | 39.8 | 55.5 | 256.7 | - | - |
| 25 | 1.6 | 1.9 | 2.3 | 2.6 | 2.9 | 4.6 | 7.6 | - |

# Reformulation of our problem:

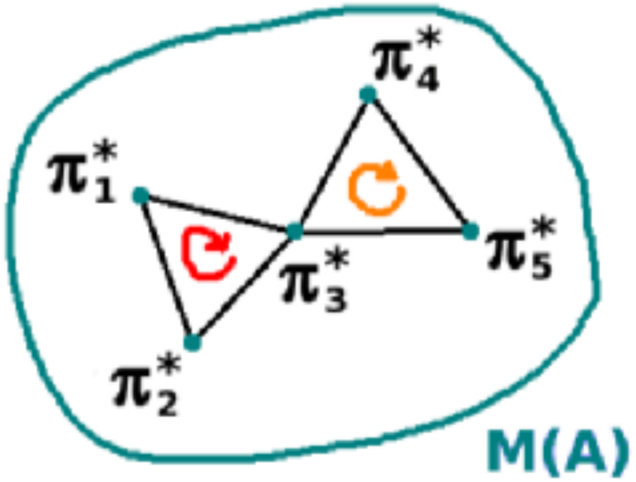
Given a set of $m$ permutations $\mathcal{A} \subseteq \mathcal{S}_n$, we want to find the set $\mathcal{M}(\mathcal{A})$ of all the permutations $\pi^*$ satisfying

$$d_{KT}(\pi^*, \mathcal{A}) \leq d_{KT}(\pi, \mathcal{A}), \forall \pi \in \mathcal{S}_n$$

# Properties of $\mathcal{M}(\mathcal{A})$:

$$\mathcal{A} = \{[7,8,2,3,6,1,5,4],[3,5,1,7,8,6,2,4],[5,8,3,4,1,2,7,6]\}$$

$$\begin{aligned} \mathbf{M(A)} = \{\, & \pi_1^* = [5, 8, 3, 1, 7, 2, 6, 4], \\ & \pi_2^* = [8, 3, 5, 1, 7, 2, 6, 4], \\ & \pi_3^* = [3, 5, 8, 1, 7, 2, 6, 4], \\ & \pi_4^* = [3, 5, 1, 7, 8, 2, 6, 4], \\ & \pi_5^* = [3, 5, 7, 8, 1, 2, 6, 4] \,\} \end{aligned}$$



# Properties of $\mathcal{M}(\mathcal{A})$:

Let us define the following left group action:

$$\begin{aligned} \mathcal{S}_n \times \mathbb{P}(\mathcal{S}_n) &\longrightarrow \mathbb{P}(\mathcal{S}_n) \\ \pi \cdot \mathcal{A} &\longrightarrow \pi\mathcal{A} = \{\pi \circ \sigma \mid \sigma \in \mathcal{A}\} \end{aligned}$$

We can show that $\mathcal{M}$ is a group morphism i.e. that

$$\pi \cdot \mathcal{M}(\mathcal{A}) = \mathcal{M}(\pi\mathcal{A})$$

# Automedian sets:

**Definition 1:** A permutation $\pi \in \mathcal{S}_n$ will be called $a$-decomposable if $i > a \iff \pi_i > a$, $\forall i \in \{1, 2, \ldots, n\}$.

**Example:** Let $\pi = [3, 2, 1 | 4, | 5]$ then $\pi$ is 3;4-decomposable

**Definition 1′:** A set of permutations is $a$-decomposable if all of its permutations are $a$-decomposable.

**Definition 1″:** A permutation or set will be called indecomposable if it is not $a$-decomposable for any $a \in \{1, 2, \ldots, n-1\}$.

# Automedian sets:

**Definition 2:** Let $\mathcal{A} \subseteq \mathcal{S}_n$ be an $a$-decomposable set and let $\sigma \in \mathcal{S}_n$ be any permutation. Then the set $\sigma\mathcal{A}$ is called **$a$-separable**.

If $\mathcal{A} \subseteq \mathcal{S}_n$ is not separable for any $a \in \{1, \ldots, n-1\}$, it is called **inseparable**.

**Example:** $\mathcal{A} = \{[3, 2, 1, 5, 4], [3, 1, 2, 4, 5], [1, 2, 3, 5, 4]\}$

is a 3-decomposable set. Let $\sigma = [4, 2, 1, 3, 5]$, then

$$\sigma\mathcal{A} = \{[1, 2, 4, 5, 3], [1, 4, 2, 3, 5], [4, 2, 1, 5, 3]\},$$

which is 3-separable.

# Automedian sets:

**Definition 3:** Let $\pi \in S_k$ and $\sigma \in \mathcal{S}_\ell$ be two permutations of length $k$ and $\ell$, respectively. The **direct sum** of $\pi$ and $\sigma$, denoted $\pi \oplus \sigma$, is defined as

$$\pi \oplus \sigma = \pi_1 \pi_2 \ldots \pi_k (\sigma_1 + k)(\sigma_2 + k)\ldots(\sigma_\ell + k)$$

**Example:** Let $\pi = [3, 2, 1, 4, 5]$ and $\sigma = [1, 3, 2]$

then $\pi \oplus \sigma = [3, 2, 1, 4, 5, 6, 8, 7]$

# Automedian sets:

**Definition 4:** Let $\mathcal{A} \subset \mathcal{S}_k$ and $\mathcal{B} \subset \mathcal{S}_\ell$ be two set of permutations. The **direct sum** of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} \oplus \mathcal{B}$, is defined as

$$\mathcal{A} \oplus \mathcal{B} = \{\pi \oplus \sigma \mid \pi \in \mathcal{A} \text{ and } \sigma \in \mathcal{B}\}$$

$\mathcal{A} \oplus \mathcal{B}$ is $k$-decomposable

# Automedian sets:

Example: Let $\mathcal{A} = \{[1,3,2], [3,1,2]\}$ and $\mathcal{B} = \{[2,1,4,3], [2,3,1,4] [2,4,3,1]\}$

then $$\mathcal{A} \oplus \mathcal{B} = \{[1,3,2,5,4,7,6], [1,3,2,5,6,4,7], [1,3,2,5,7,6,4], [3,1,2,5,4,7,6], [3,1,2,5,6,4,7], [3,1,2,5,7,6,4]\}$$

# Automedian sets:

**Theorem 1:** $\mathcal{M}(\mathcal{A} \oplus \mathcal{B}) = \mathcal{M}(\mathcal{A}) \oplus \mathcal{M}(\mathcal{B})$

**Theorem 2:** $\mathcal{A} = \mathcal{M}(\mathcal{A})$ and $\mathcal{B} = \mathcal{M}(\mathcal{B})$

$$\Longleftrightarrow$$

$$\mathcal{A} \oplus \mathcal{B} = \mathcal{M}(\mathcal{A} \oplus \mathcal{B})$$

**Theorem 3:** If $\mathcal{C} \subset \mathcal{S}_n$ is an $a$-decomposible automedian set, then $\exists$ automedian sets $\mathcal{A} \subset \mathcal{S}_a$, $\mathcal{B} \subset \mathcal{S}_{n-a}$ such that $\mathcal{C} = \mathcal{A} \oplus \mathcal{B}$.

# Counting automedian sets:

**Definition:** $\mathcal{AM}_n = \{\mathcal{A} = \mathcal{M}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathcal{S}_n\}$

**Definition:** $\mathcal{I}_n = \{\mathcal{A} = \mathcal{M}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathcal{S}_n \text{ and } \mathcal{A} \text{ inseparable}\}$

$$|\mathcal{AM}_n| = |\mathcal{I}_n| + \sum_{i=1}^{n-1} \binom{n}{i} \times |\mathcal{I}_i| \times |\mathcal{AM}_{n-i}|$$

$$|\mathcal{I}_n| : 1, 1, 3, 27, \ldots$$

$$|\mathcal{AM}_n| : 1, 3, 15, 117, \ldots$$

# What's left to do:

A lot!

- Completely solve the automedian case
- Investigate the shuffle of sets of permutations
- Complexity in the case where m=3
- Do the same kind of investigation for the problem of finding a median of other kind of combinatorial objects

# Election issues :

health


taxes

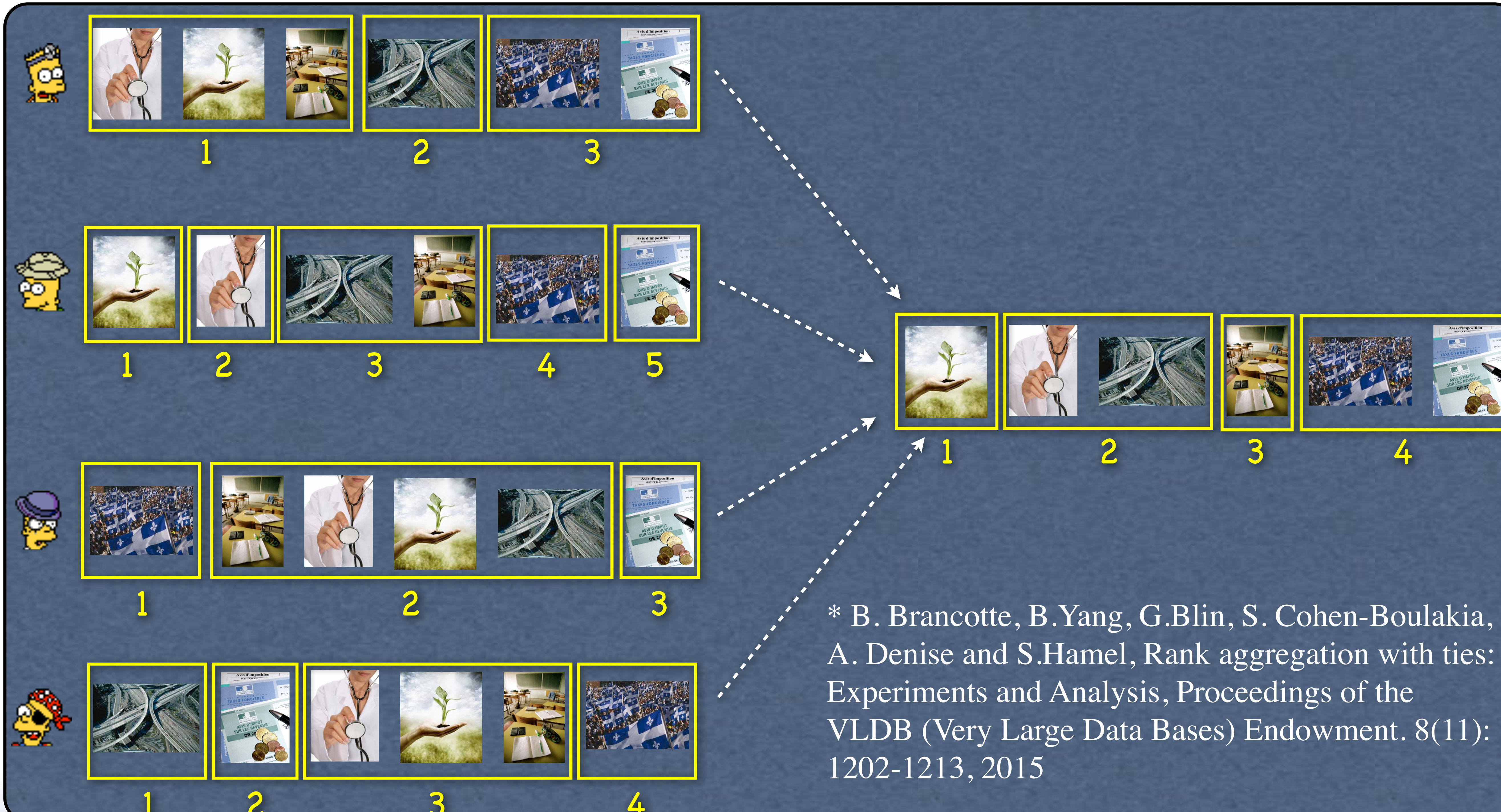
infrastructures

environment

independence

education

# Generalized Kemeny consensus :




* B. Brancotte, B.Yang, G.Blin, S. Cohen-Boulakia, A. Denise and S.Hamel, Rank aggregation with ties: Experiments and Analysis, Proceedings of the VLDB (Very Large Data Bases) Endowment. 8(11): 1202-1213, 2015

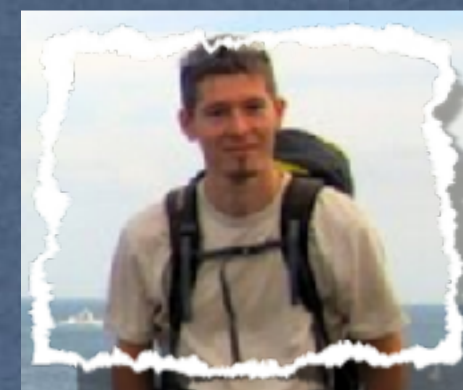
# Collaborators:

## Automedian:

Charles Desharnais

Robin Milosz

## Generalized problem:



Bryan Brancotte

Guillaume Blin

Alain Denise

Sarah Cohen-Boulakia

+ all the summer research trainees

# Questions ?

# Spurningar?