Measures for the Generalization of Polygonal Maps with Categorical Data

Beat Peter
Department of Geography
University of Zurich / Switzerland
Email: beatp@geo.unizh.ch
Web: www.geo.unizh.ch/~beatp

4th ICA Workshop on Progress in Automated Map Generalization, Beijing, 2-4 August 2001

Contents

1. Introduction

Project Approach

- Approach from the measures side
- ± Generic measures; comprehensive description of the relevant aspects of polygons and their context in a map
- Requirements of generalization operators/algorithms not explicitly considered at this point in the project
- Some measures are not explicitly connected to specific operators/algorithms (global statistical indicators)
- Approach reflects the workflow in the generalization process
- Builds on the classification developed in previous work

Properties of Categorical Data

- Plane filling polygon mosaics (no holes or unassigned polygons)
- No dedicated *background* (≠ topographic maps)
- At least 2 polygons are involved in every transformation operation
- Frequent recomputation of measures required due to permanent change of indicators
- A situation or a conflict can be looked at from different perspectives
- The character of polygons with categorical data is often fuzzy ⇨ influence on generalization methods

2. Measures for the Generalization of Categorical Data

Classification of Measures

- Spatial scope
  - Micro/Local (polygon)
  - Meso/Regional (group of adjacent polygons)
  - Macro/Global (category, partition, entire map)
- Function/Main characteristic
  - Size
  - Distance and proximity
  - Shape
  - Topology
  - Density and distribution
  - Pattern and alignment
  - Semantic

Application Scope of Measures

Three main application scopes:

1) Initial data evaluation
2) Conflict identification and transformation control
3) Evaluation of results

1) Initial Data Evaluation

- Recognition of the global map structure (e.g. total area of categories, category area relations, distribution of objects)
- Statistical measures on the global level (not linked to specific operators or algorithms)
- Identify potential problems imposed by map controls before geometric transformations take place (e.g. too many categories for the target scale ⇨ *reclassification*)
- Strategy building (e.g. delete polygons of category *x* with an area smaller than 80% of the minimum area)
- Assess the *before* state of the database

2) Conflict Identification and Transformation Control

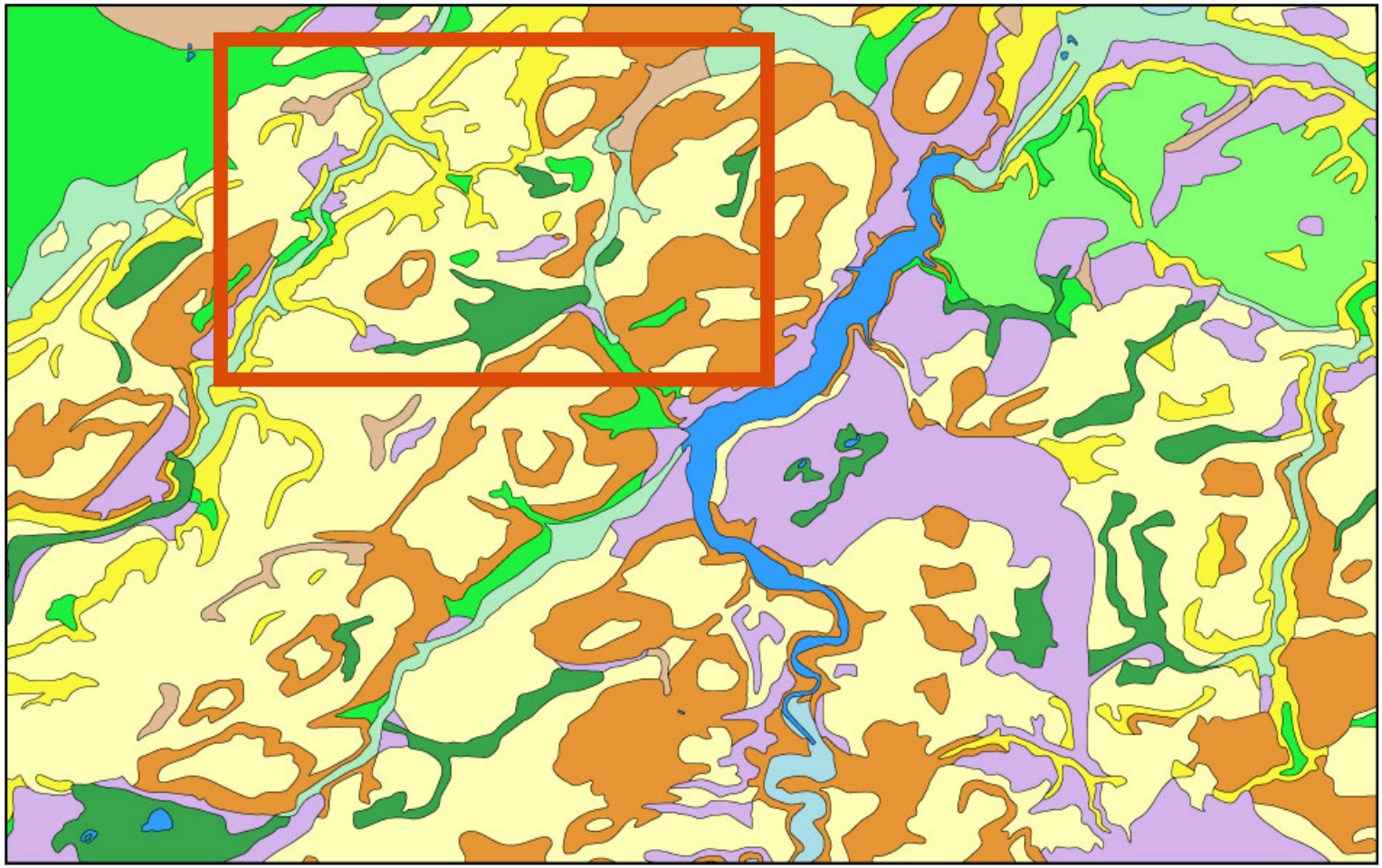
- These measures are linked to specific generalization operators (and algorithms)
- Identification of (geometric) conflicts
- Guide/Support operator selection process (e.g. do not delete this polygon)
- Guide/Support algorithm selection process (e.g. displace non-uniformly because of lack of space)
- Guide/Support parameter selection process for algorithms
- Resolve local conflicts while respecting their influence on the entire map (global indicators)

3) Evaluation of Results

- Computation of the *after* state ⇨ compare with *before* state
- Check for inconsistencies (e.g. topological error)
- Problem: visual impression ⇔ numerical values
- Accept or reject a solution (completely or partly)
- Modify strategy or exit
- Ideally, problems or potential problems are identified at run time or during initial data evaluation (appropriate measures!)

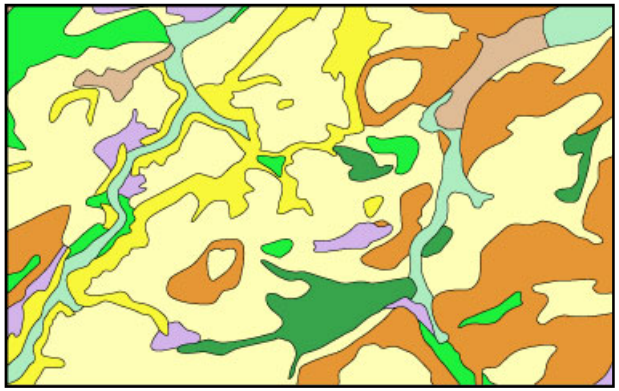
3. Application of Measures in the Generalization Workflow – Selected Examples

Examples: Geological Data (1:100'000)

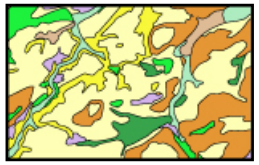


Task: Generalization from 1:100'000 to 1:250'000

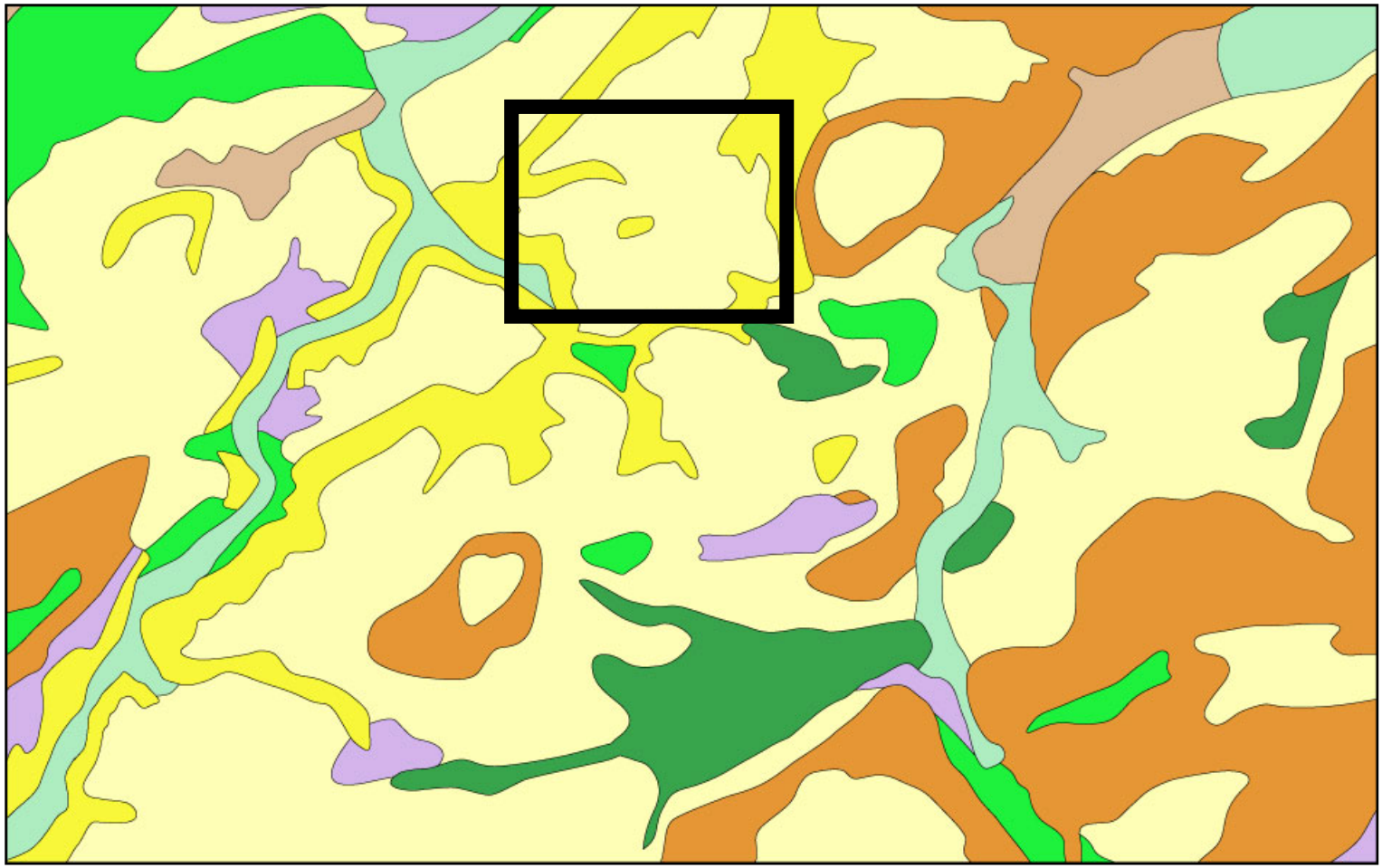
1:100'000



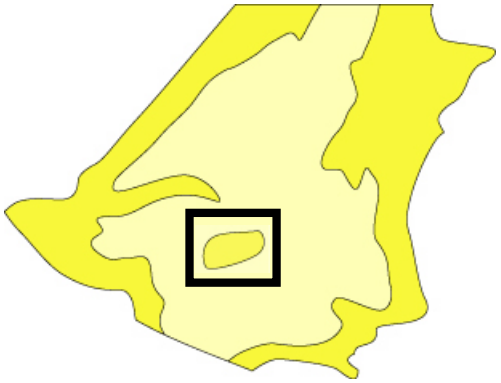
1:250'000



Selected Examples 1



Minimum Area Conflict



1.

Id: 319
Category: 7
Area: 3.7mm^2
Diff. to a_{min}: 0.3mm^2/7.5%
Diff. to d_{min}: 0.6mm^2/20%
Semantic sign.: yes
Perimeter, Shape, ...

2.

Category 7:
Total no. of Polys: 834
Total relative area: 9.2%
No. of polys below a_{min}: 32
Area below amin: 6.8%
...

No. of neighbors: 1
Cat. of neighbors: 4
2nd Order dist.: > d_s
...

3.

"Background" cat.: 1,4,12

4.

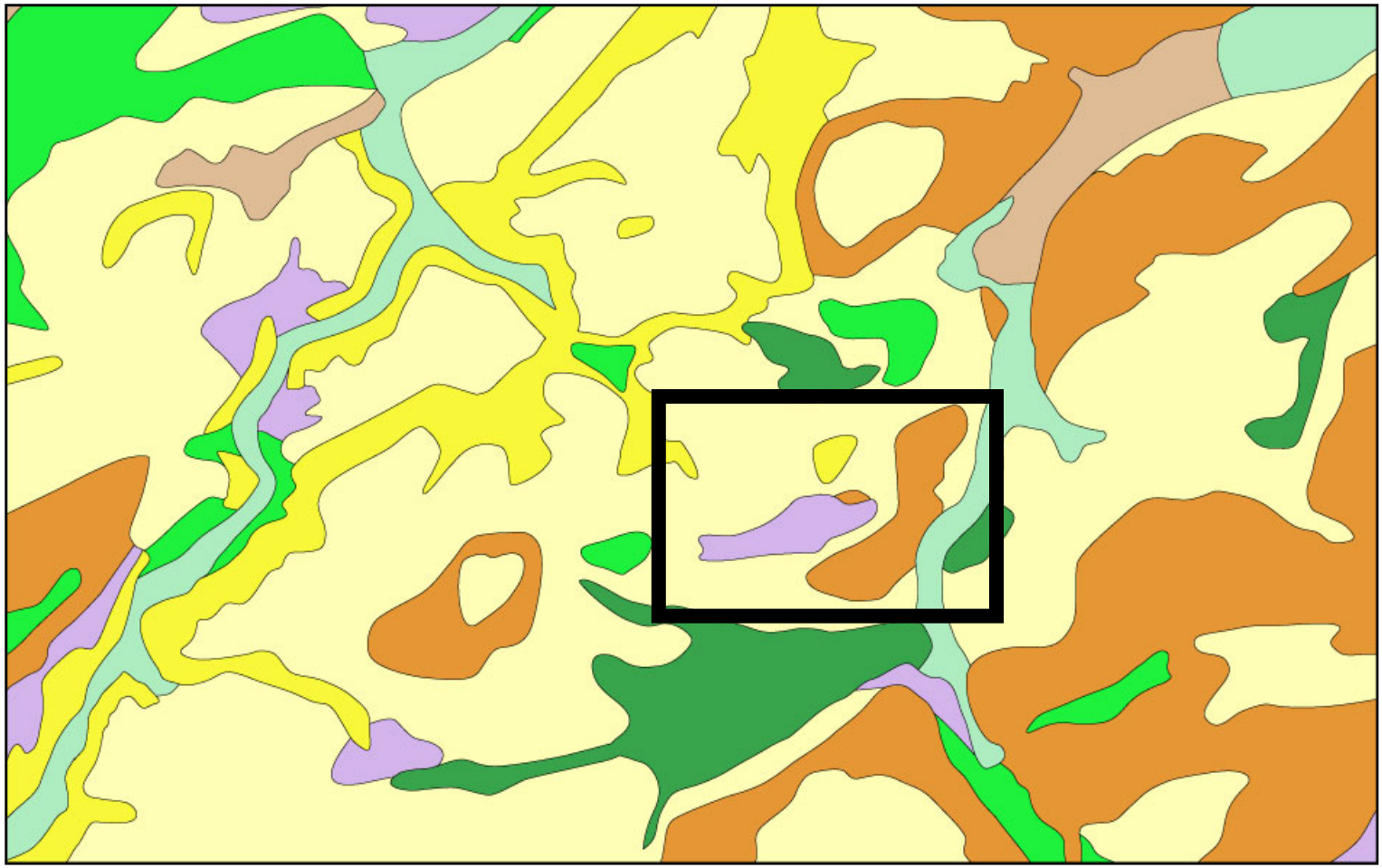
5.

⇨ Operator: *enlarge*

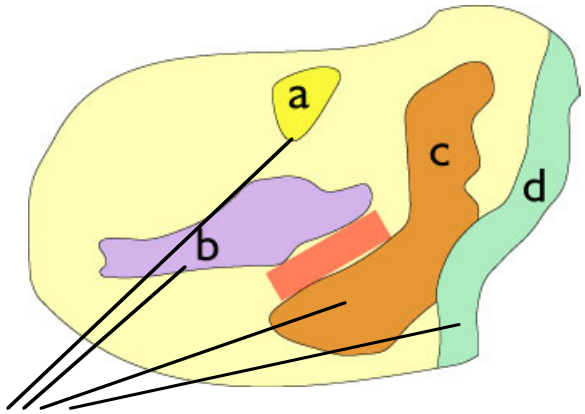
⇨ Algorithm: *scale*

6.

Selected Examples 2



Distance Conflict



Id: ...
Category: ...
Area, Perimeter,
Shape, Semantics, ...

Information about the inv. categories

1.

"Background" cat.: 1,4,12

Def. min. Dist.: 2.5mm

⇨ Operator: *displace*

4.

2.

a: No. of neighbors: 1
Category of neighbor: 4
2nd Order Dist.: 2.8mm to poly b, Dir.: 180°

b: No. of neighbors: 1
Category of neighbor: 4
2nd Order Dist.: 2.8mm to poly a, Dir.: 0°
1.8mm to poly c, Dir.: 215°

c: No. of neighbors: 2
Category of neighbor: 4,9
2nd Order Dist.: 1.8mm to poly b, Dir.: 45°
2.7mm to poly d, Dir.: 260°

3.

Unwanted neighborhood:
cat(a) and cat(b), cat(b) and cat(c), ...

5.

⇨ Displace poly *b* by 0.7mm, do not displace poly *c*

6.

⇨ Displace poly *b* by 0.3mm

4. Conclusions and Outlook

Conclusions

- Work in progress; much is still unclear (e.g. complex situations)
- Measures are a key element in the generalization process; a holistic approach allows better controlling transformations
- Semantic information and expert knowledge is very important
- Acquisition of generic knowledge for different types of categorical data can be a problem
- Every dataset has its individual properties
- Interpretation of values (changes of values) may be a problem (e.g. shape and global/macro indicators)

Outlook – Next Steps

- Harmonization with requirements of operators and algorithms (M. Galanda)
- Step by step implementation from simple to complex on the Laser-Scan/Lamps2 platform
- Find appropriate methods for the implementation of complex measures
- Find ways for integrating semantic information
- Extensive empirical testing
- Modification of measures, introduce new ones, drop others