Design and Experience of Generalization Tools

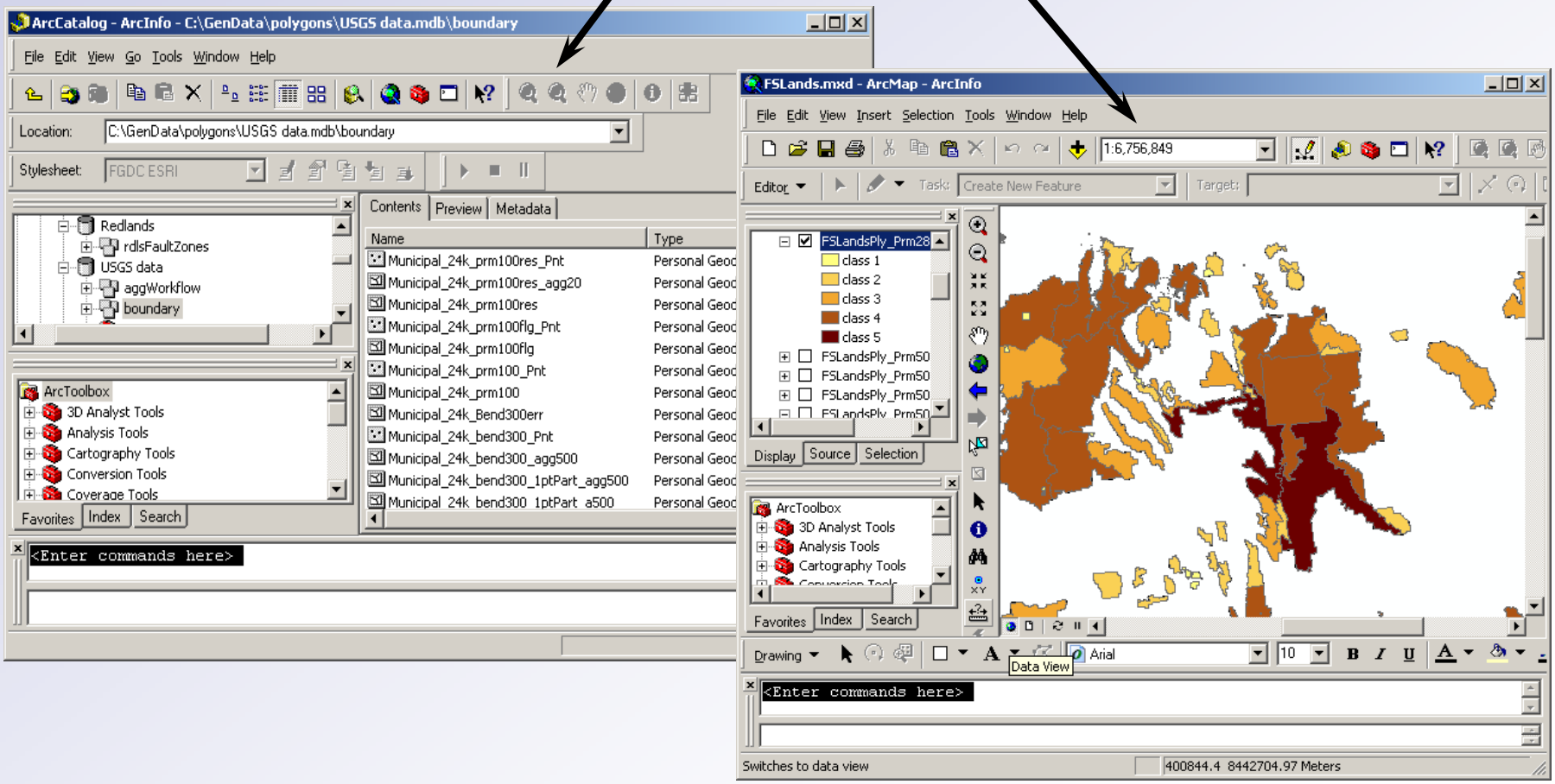
AutoCarto Conference 2006, Vancouver, WA

Dan Lee, Paul Hardy
ESRI, Redlands, USA
(dlee@esri.com; phardy@esri.com)

- Introduction to Geoprocessing
- Topology in Generalization
- TIN-based Generalization
- Quality and Status Information
- Generalization Scenarios
- Future Directions

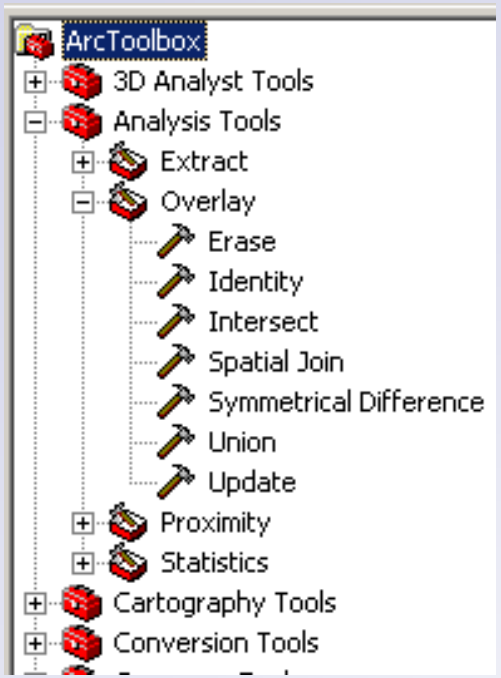
ArcToolbox – geoprocessing environment integrated in ArcGIS

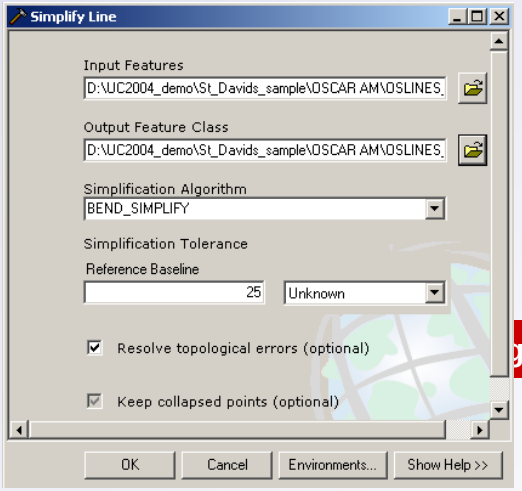
- Setting and executing batch processes
- Accessible from ArcCatalog and ArcMap

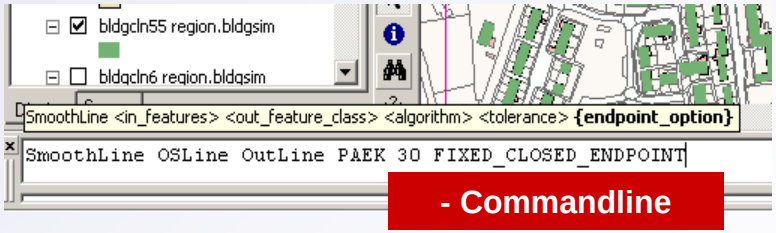


ArcToolbox – geoprocessing framework

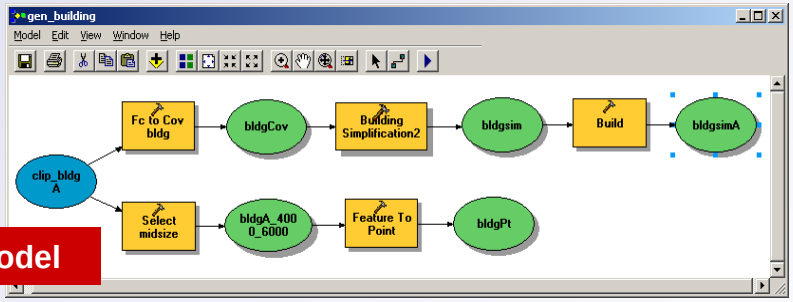
- Single processes
- Chained workflow processes



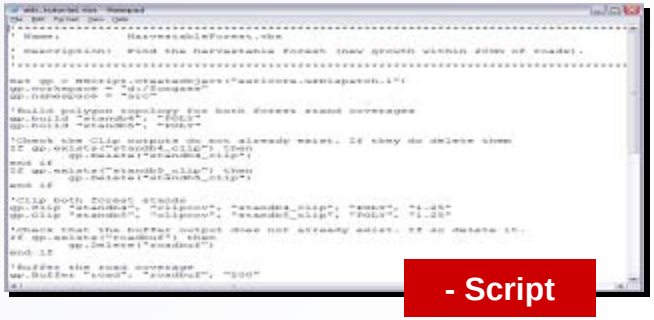




- Commandline



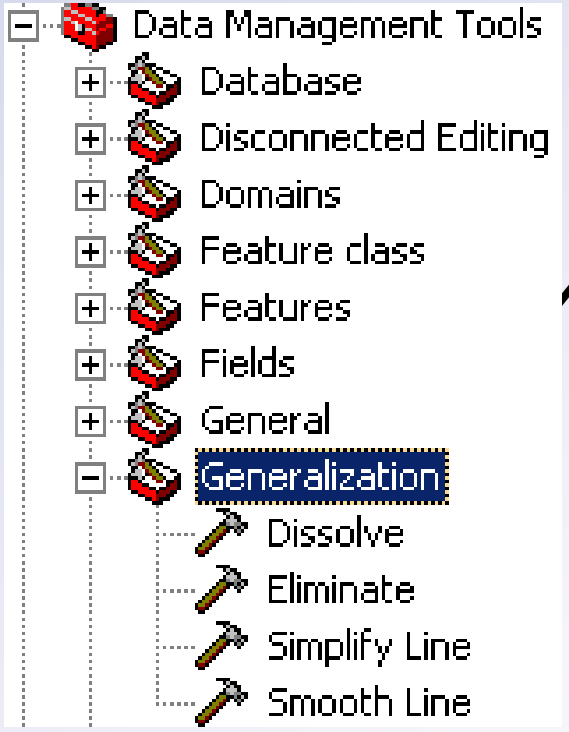
- Model



- Script

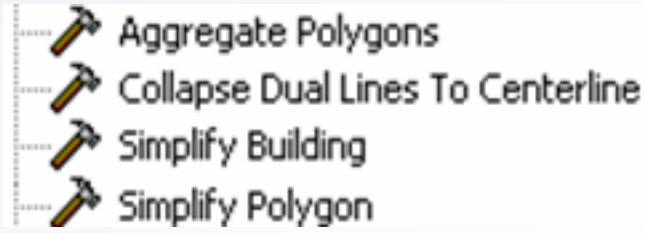
Generalization toolset in ArcToolbox

- Significant step in the provision of generalization
- Operating on geodatabase features (vs. traditional coverages)
- Discovering topological relationship on the fly
- Using TIN for neighbor feature analysis



Released in ArcGIS 9.0

For ArcGIS 9.2 release



Topology in generalization

- Geodatabase features - stored as independent geometries

  May not necessarily have vertices or endpoints at connections, intersections, and shared segments
- Different levels of requirements – speed vs. data integrity
- Tools – to analyze and preserve embedded topology
- Options – to detect and resolve introduced topology errors

Options in Simplify Line and Simplify Polygon tools:

NO_CHECK (no check for violation of topology)

PRESERVE_SHARED (preserve embedded topology and flag introduced errors)

RESOLVE_ERRORS (preserve embedded topology and resolve introduced errors)

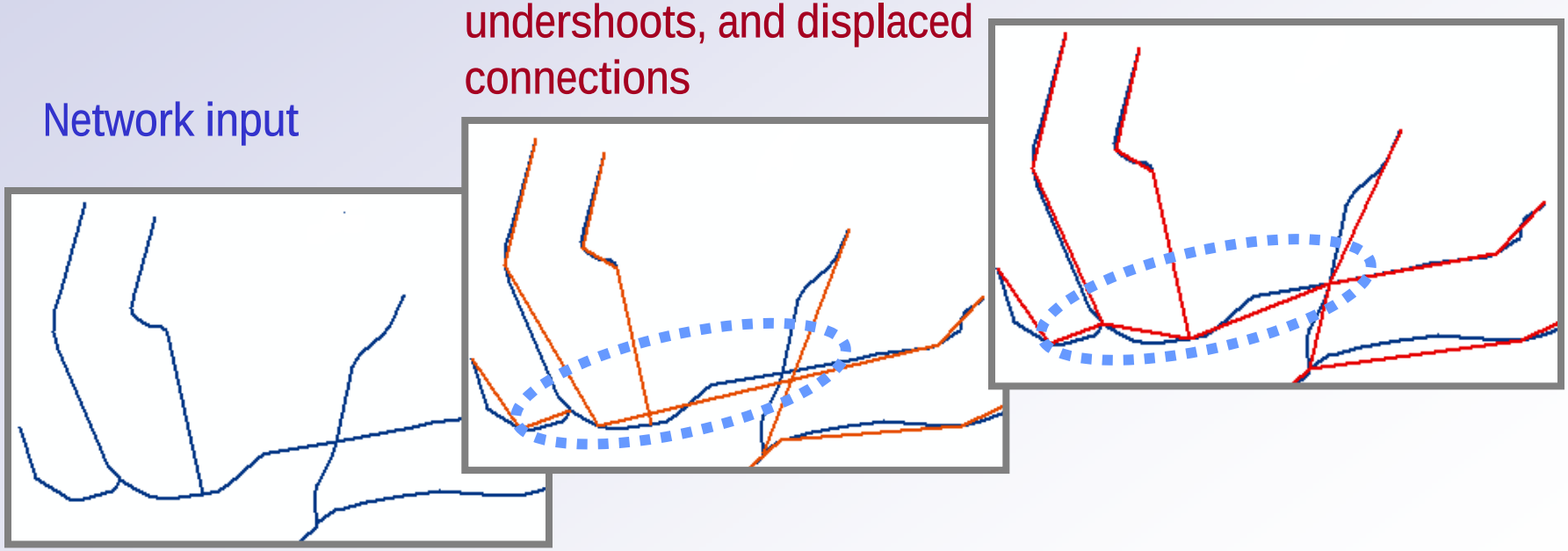
Preserving embedded (implicit) topology

- Connectivity – shared nodes in linear network

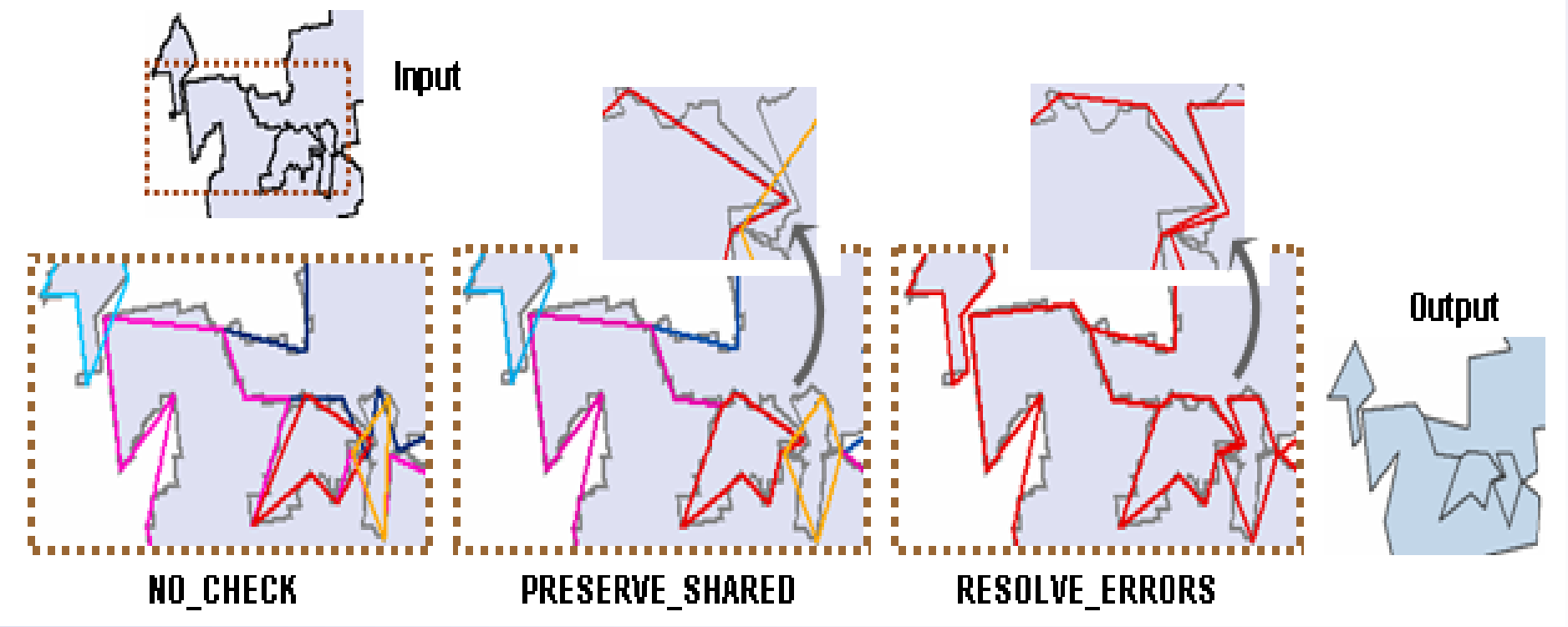
NO_CHECK
Shared nodes were ignored – causing overshoots, undershoots, and displaced connections

PRESERVE_SHARED
shared nodes were detected and preserved

Network input



- Adjacency – shared boundaries in polygons



Shared boundaries were ignored – causing inconsistent simplification

Shared boundaries were preserved; new errors flagged

Shared boundaries were preserved; new errors resolved

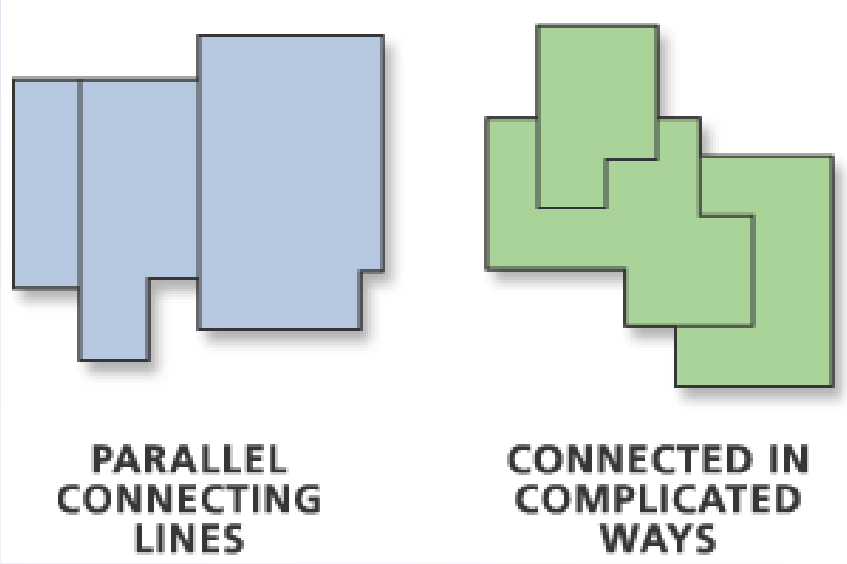
- Grouping – connected polygons to be treated as one object

Measure of the total area for a group of connected polygons

The standard query by area or size doesn't have this option.

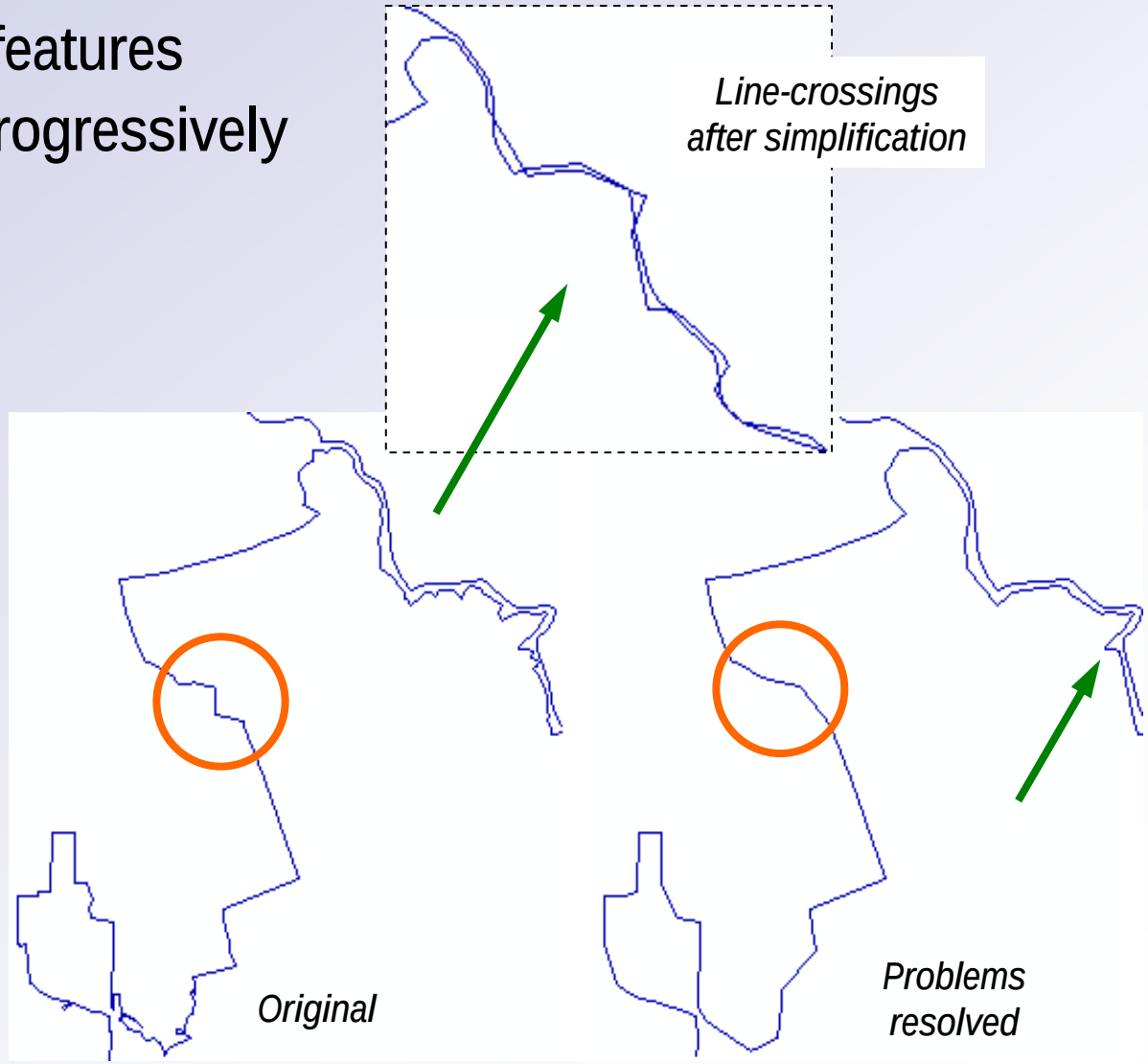
Analysis of the configuration of connected buildings

They are to be processed differently in simplification.



Handling introduced spatial conflicts

- Flagging involved features
- Resolving errors progressively



TIN-based generalization

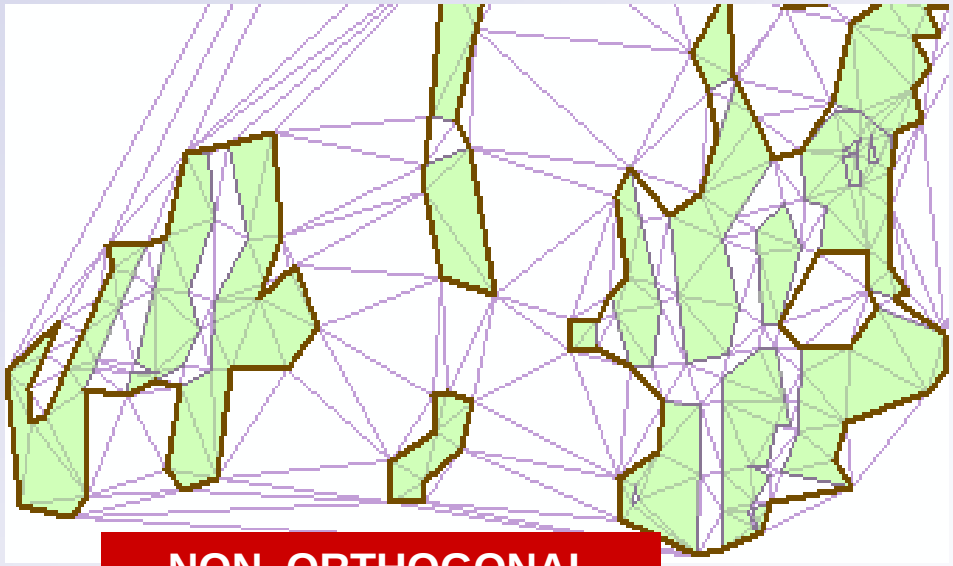
- TIN-based generalization as proven effective technique (Jones et al 1995; Peng 1997; ...)
- TIN functions accessible in ArcObjects

Aggregation of polygon features

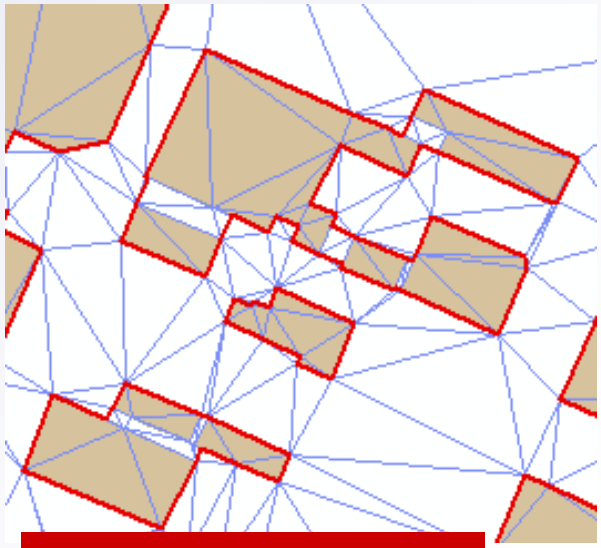
Finding clusters

Reconstructing new polygons

Excluding small areas and holes



NON_ORTHOGONAL



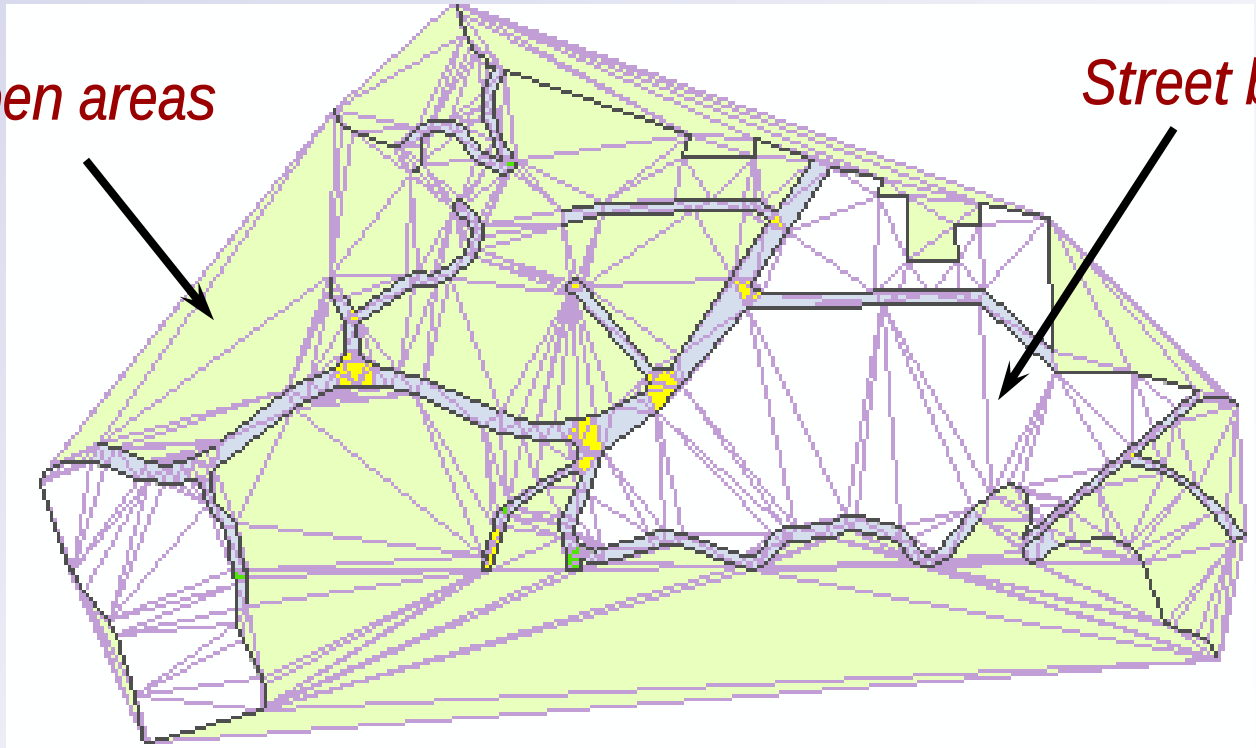
ORTHOGONAL

Derivation of road centerlines (focusing on open-ended)

Excluding street blocks

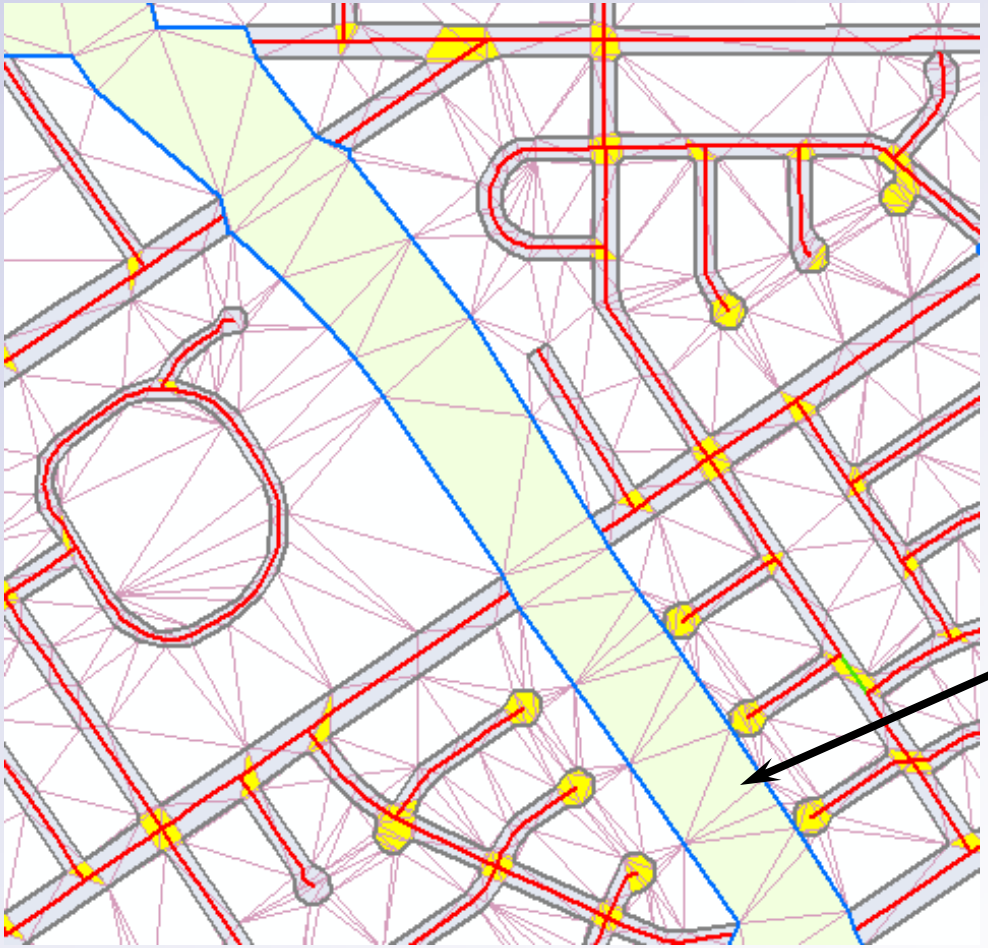
Excluding open areas

Other styles:
Closed-ended
(pavement polygons);
Mixed
(highway with crossings and ramps, railroad tracks)



Identifying and analyzing junctions and out-of-range roads

Deriving centerlines in the remaining space



Road too wide

Quality and status information

- Tracking feature conflicts

NFSLAND	ADMIN_FO	NFB_AD	Shape_Length	Shape_Area	InPoly_FID	SimPlyFlag
3536	Santa Fe	Y	87173.372750377	283730293.4164	11	1
3341	Santa Fe	Y	450491.29490264	3606204903.905	2	1
3530	Santa Fe	Y	56843.354729772	187132638.2375	10	0
3529	Santa Fe	Y	304232.87640398	2640108060.955	9	0
3479		N	58065.964633953	166854934.8164	5	0
3467		N	67295.991230405	206296014.1043	4	0

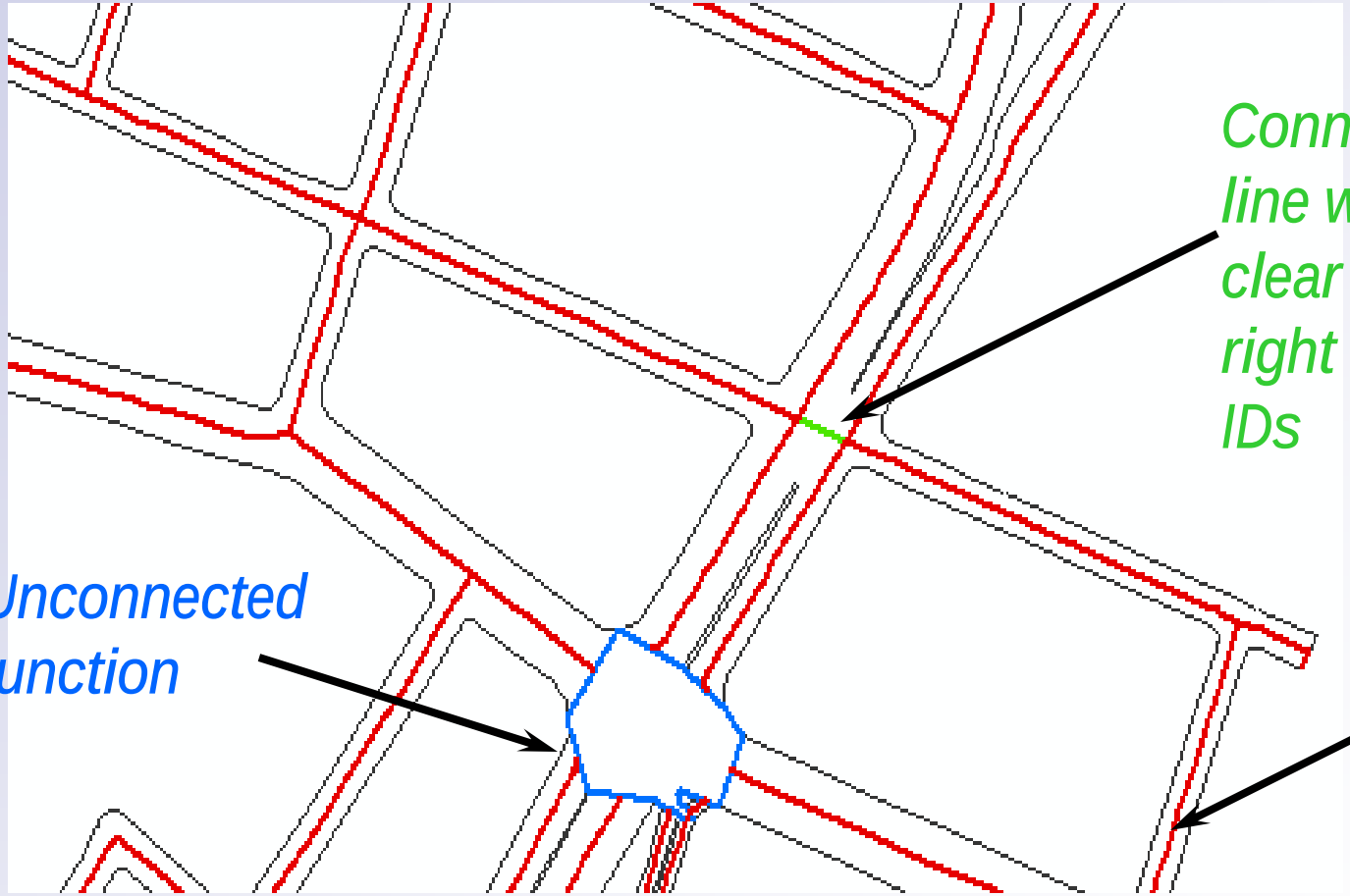
Shape_Area	InBld_FID	BLD_STATUS	BLD_GROUP	SimBldFlag
1586.31716395377	848	1	0	1
485.577059730464	345	1	0	1
1426.27573647369	346	1	0	1
822.719069886236	847	1	0	1
1017.08185538781	485	1	0	1
482.770996360363	481	1	0	1
1342.08976997045	518	1	0	0
3287.99287264462	519	1	0	0
2177.31582456626	520	1	0	0

Too many flags indicates either tolerance too large or high congestion – consider smaller tolerance or different generalization operations (elimination, aggregation, collapse, etc.)

- Categorizing generalized features and their status

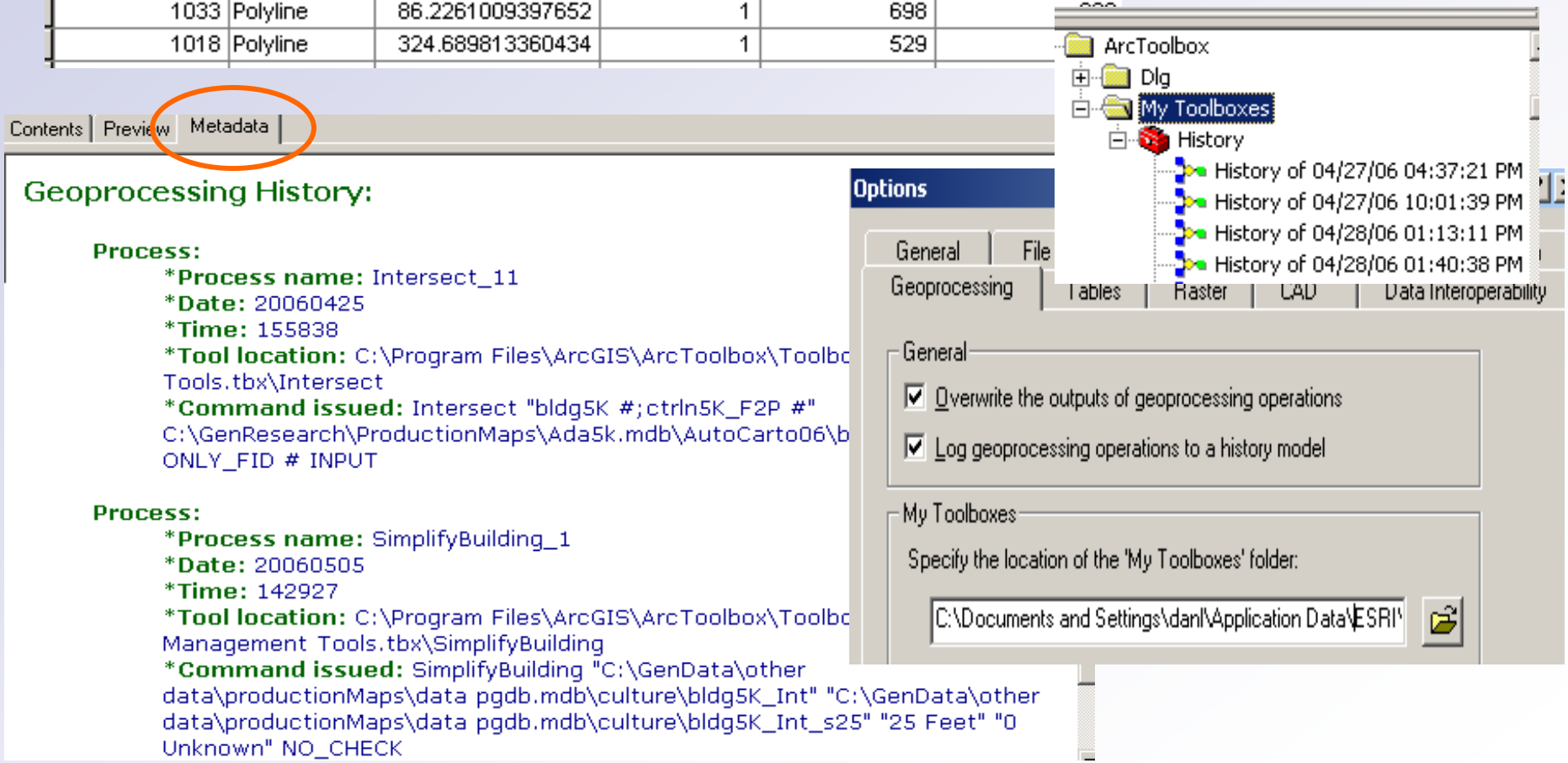
Building simplification status

Classifying line types and their status



- Linking output features to input (1 to 1; 1 to N)
- Maintaining metadata and logging processes (history models)

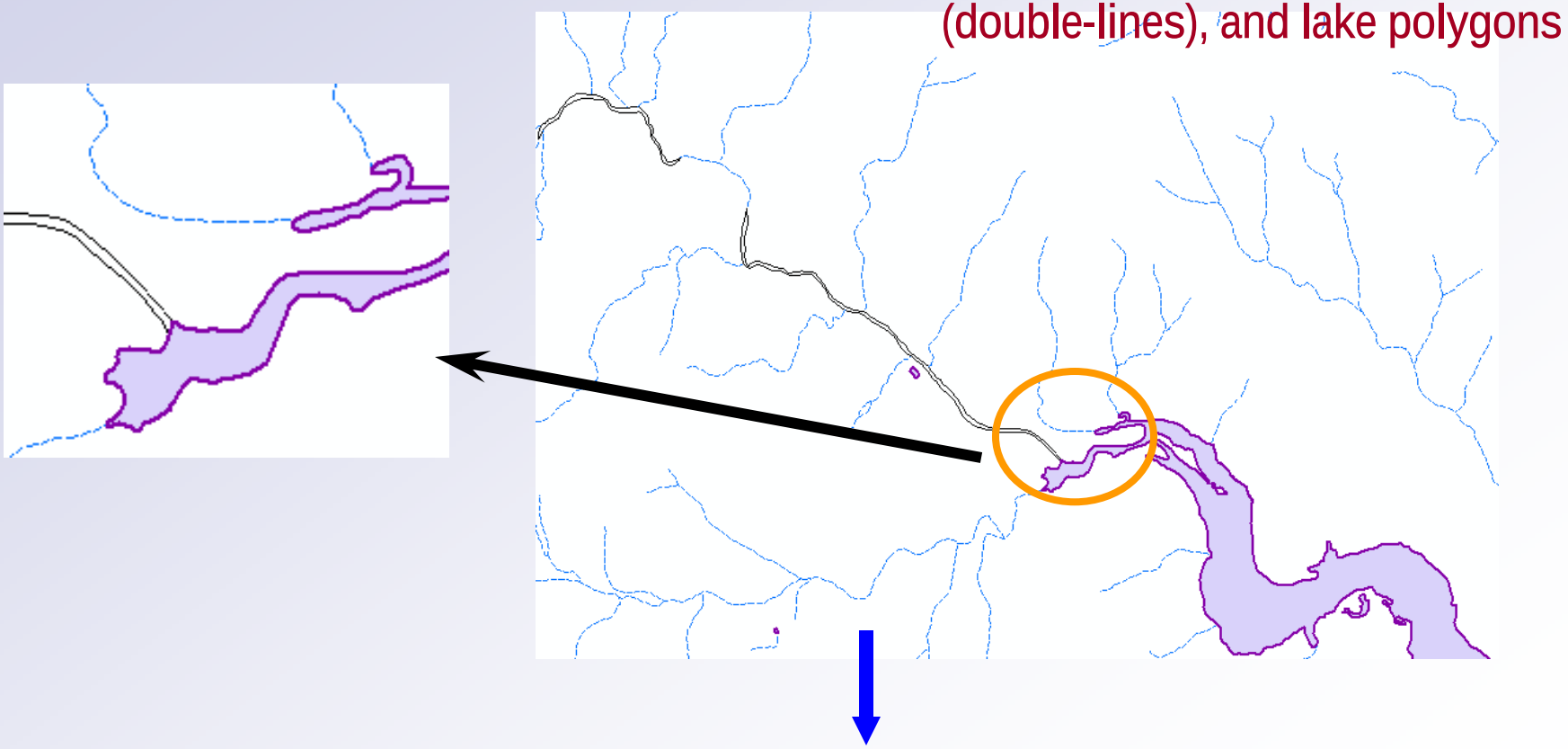
OBJECTID *	Shape *	Shape_Length	LnType	LeftLn_FID	RightLn_FID
1048	Polyline	338.201553214012	1	601	647
1049	Polyline	329.373060001993	1	552	611
1051	Polyline	333.844322718054	1	605	639
1047	Polyline	336.126794303688	1	600	639
1033	Polyline	86.2261009397652	1	698	
1018	Polyline	324.689813360434	1	529	



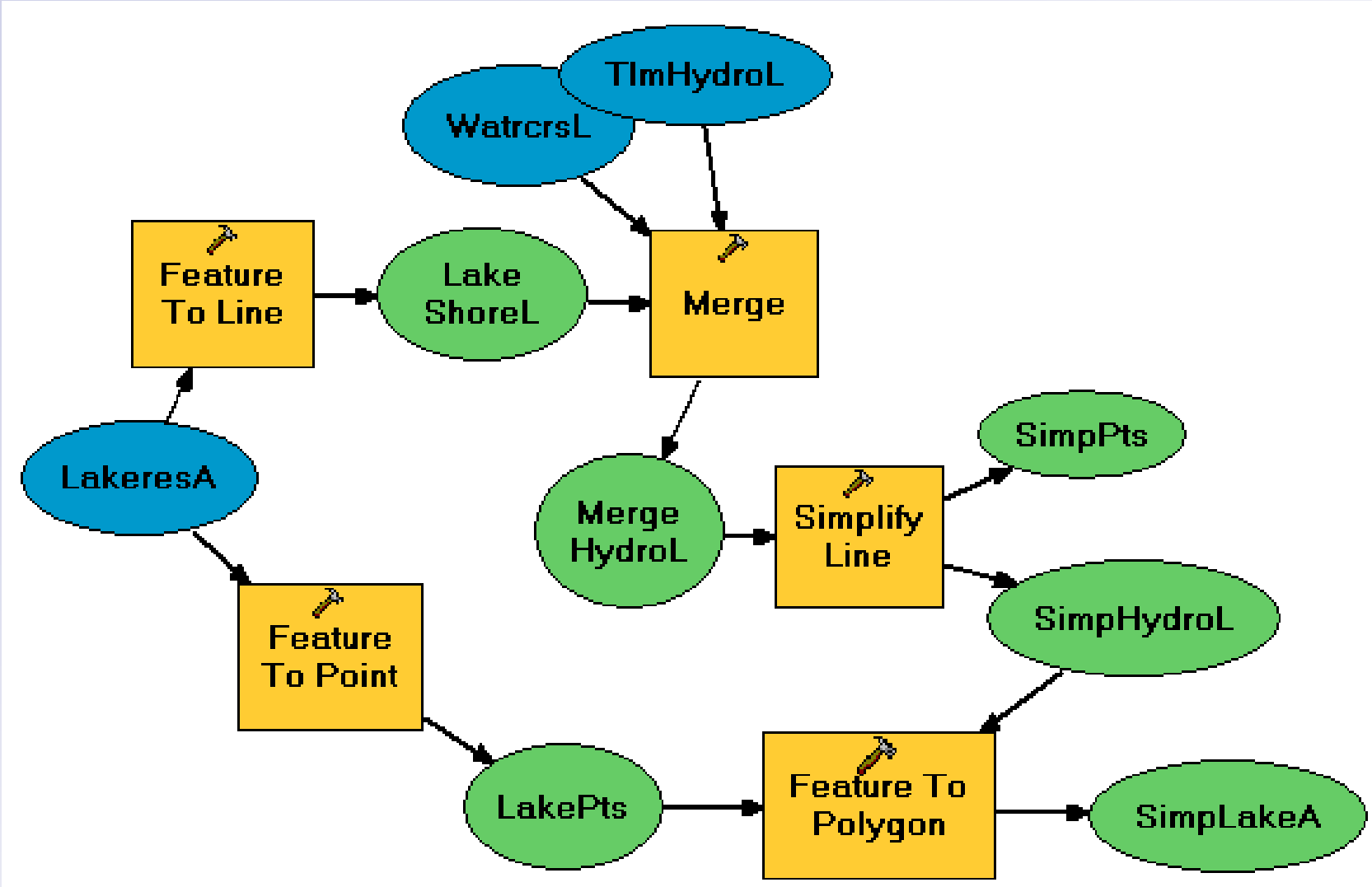
Generalization scenario 1

- Preserving connectivity while simplifying multiple hydrographic features

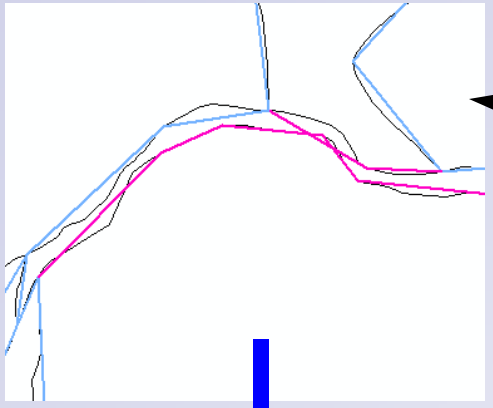
Input:
stream lines, wider rivers (double-lines), and lake polygons



Geoprocessing model – Simplify Multi-Hydro Features

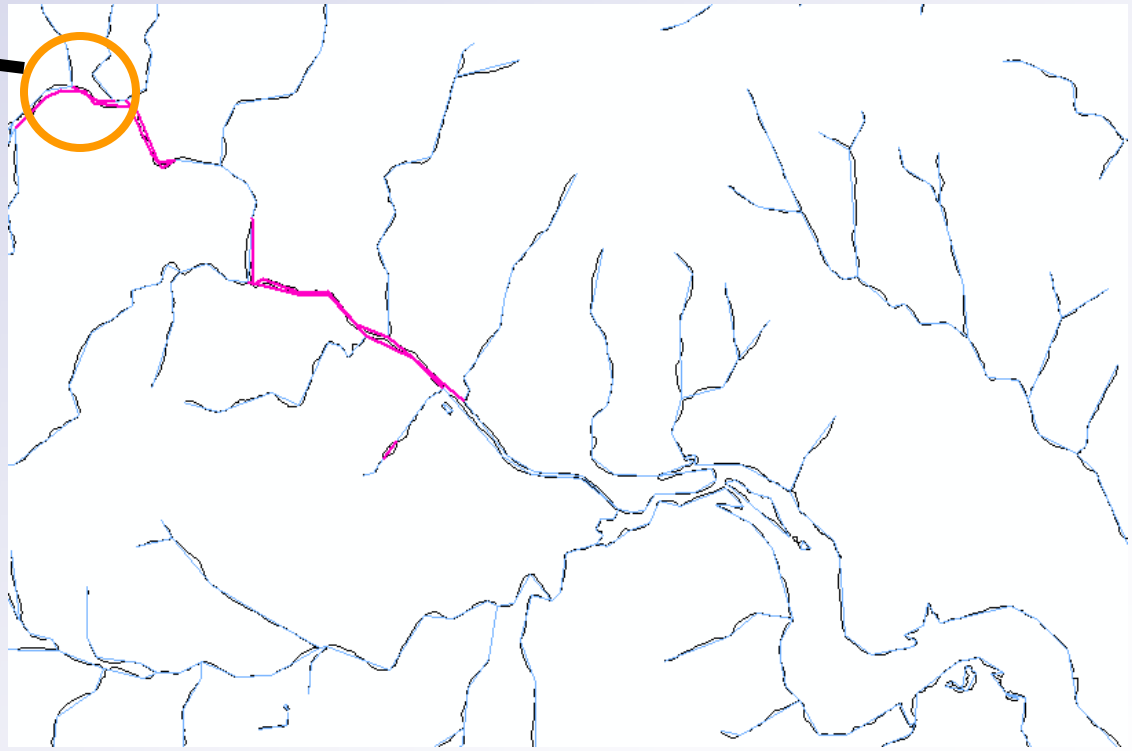


Line crossing flagged

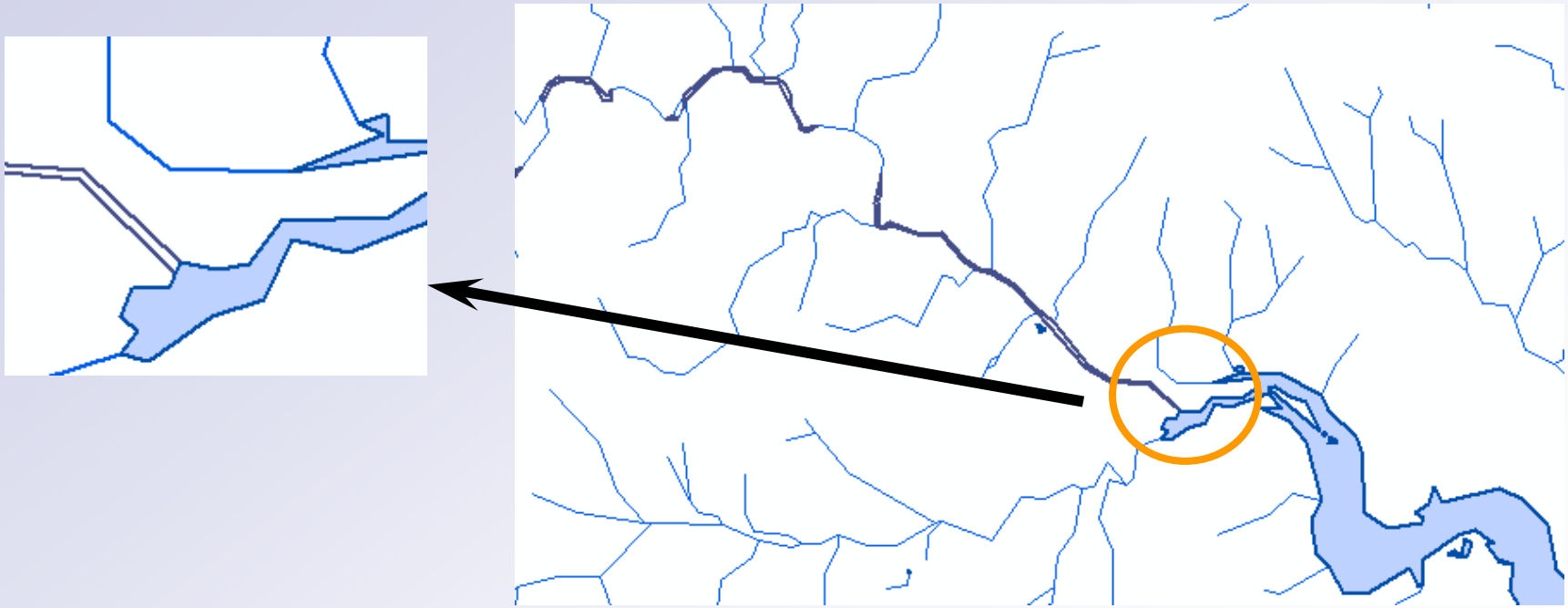


Result of RESOLVE_ERRORS

Merged lines simplified with PRESERVE_SHARED



Output:
simplified stream and river lines
and restored lake polygons



Generalization scenario 2

- Partitioning and generalizing buildings

Selecting roads and forming polygons as partitions (Feature To Polygon);

Overlaying large buildings with partition polygons (Intersect)

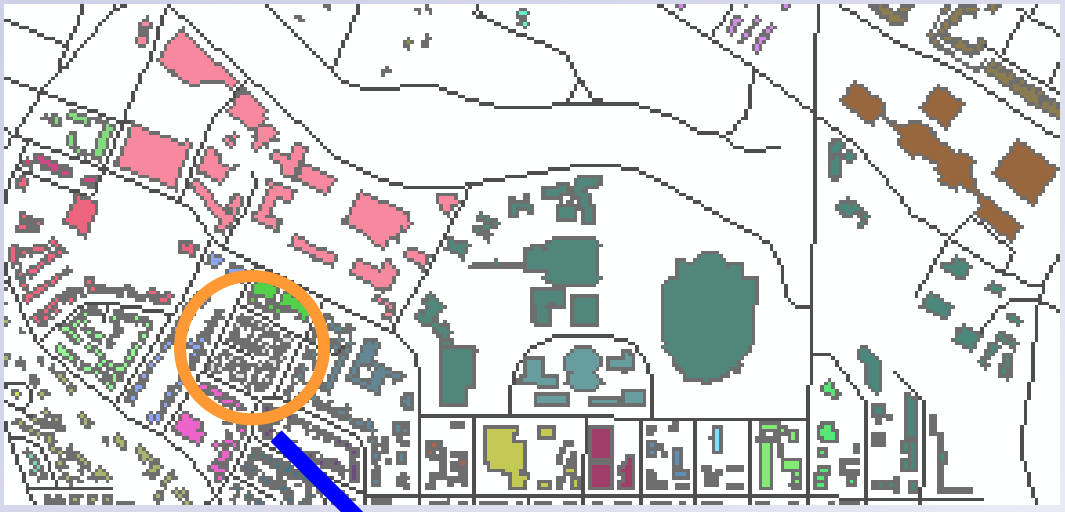
– buildings in each partition carry a unique partition ID;

Selecting buildings by partition ID;

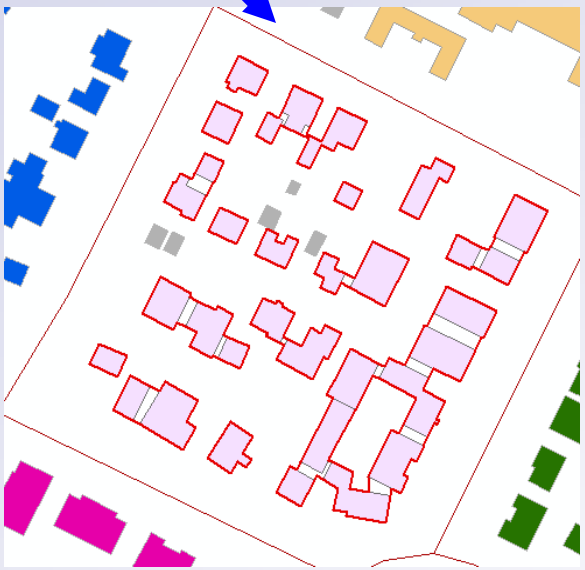
Aggregating buildings;

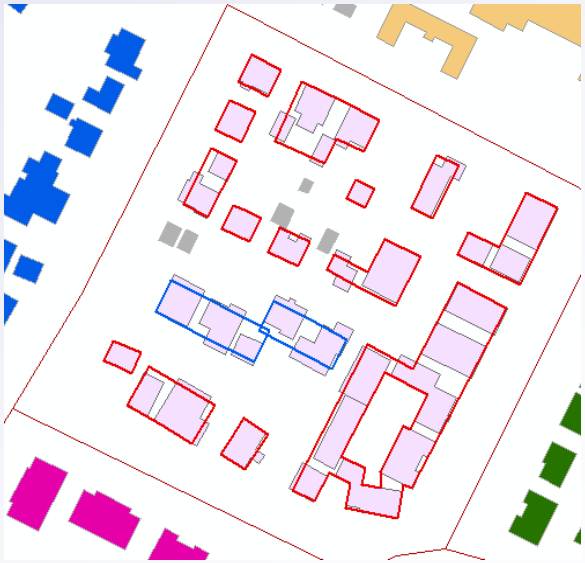
Simplifying aggregated buildings;

Evaluating results (flagged conflicts)



Buildings are grouped by street block partitions.





Selected by a partition ID (small buildings already eliminated); then aggregated.

Simplified with conflicts flagged – hint for displacement or adjustment of parameters.

Future directions

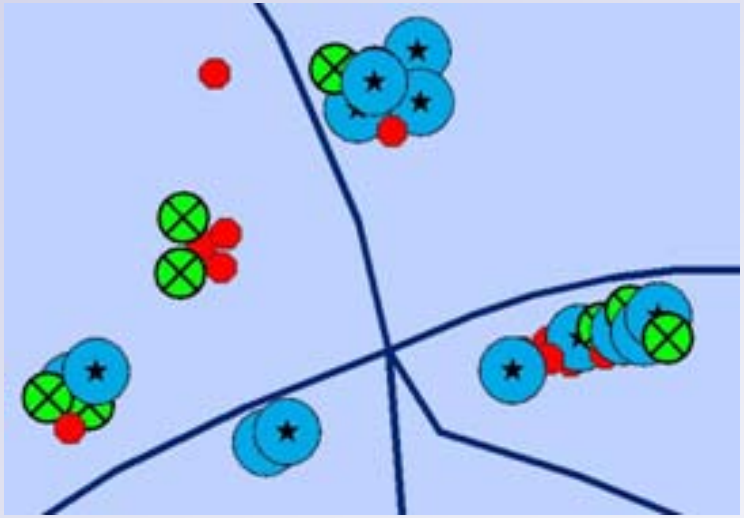
- Extend self-contained processes to involve additional spatial context and constraints
- Introduce an optimization mechanism for generalization

Optimizer – a research and development project underway

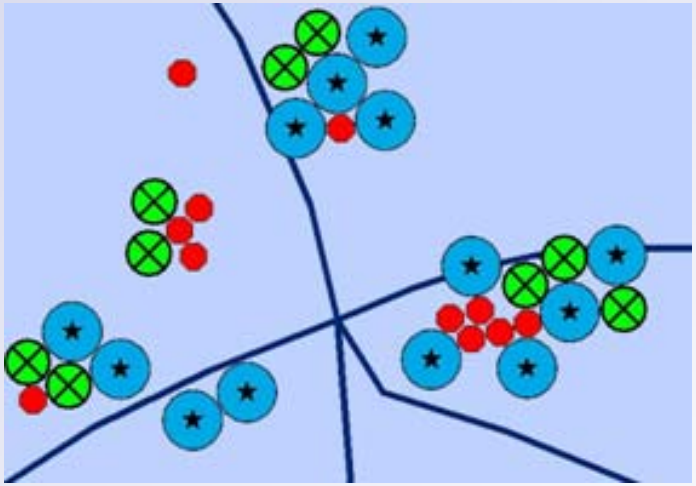
- *Optimizing both map content and legibility while conserving map accuracy, geographic characteristics, and aesthetic quality*
- *Constraints with associated actions*
- *Iterations against the measures of satisfaction*
- *Simulated Annealing technique – gradually lowering notional "temperature"*

Optimization of displacement

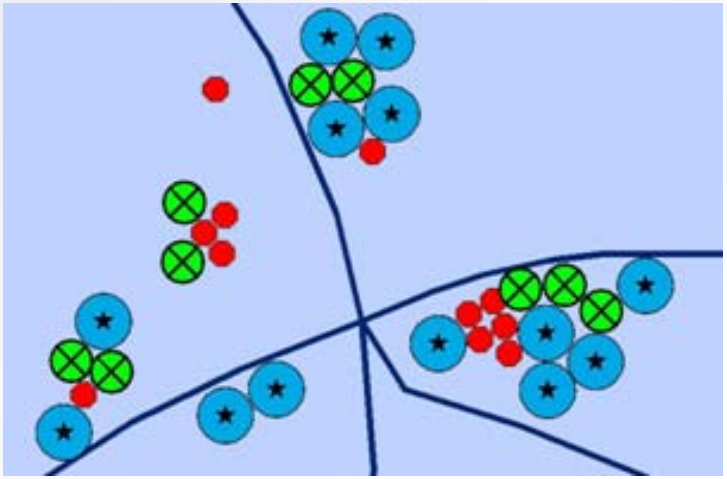
Initial stage



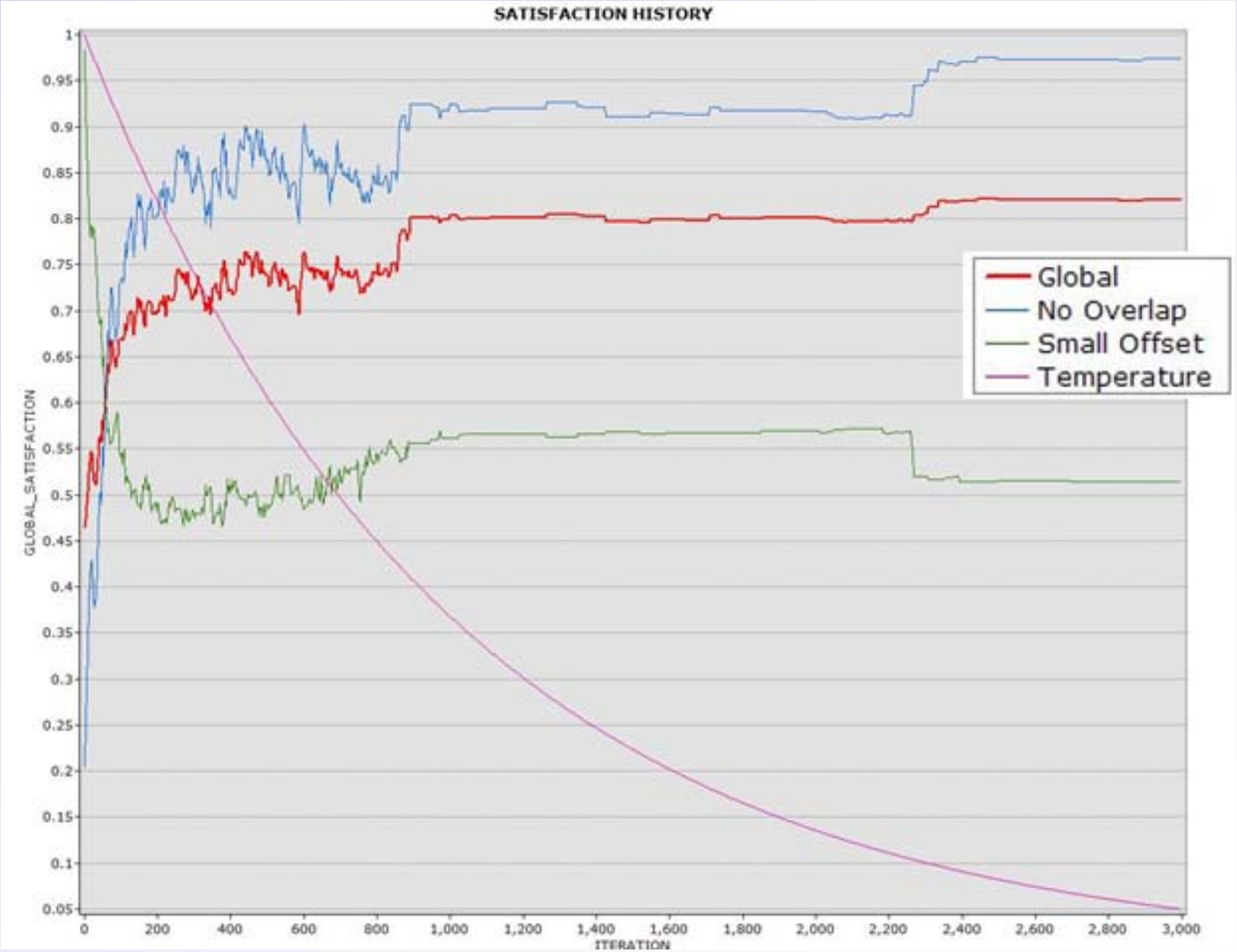
Displacement without barriers



Displacement with barriers



Satisfaction graph – constraints and "temperature"



Conclusions

- Continue designing and enhancing generalization tools with more flexibility, feedback, and integration (batch and follow-up)
- Derive and formalize generalization workflows (models) to reach higher productivity
- Address contextual generalization through optimization
  - Develop methods to derive context space
  - Establish constraints and measures of satisfaction
  - Define and prioritize rules and actions