Fuzzy Generalization Inference System - the example of selection parameterization for roads and hydrographic network

Anna Fiedukowicz*

* Department of Cartography, WarsawUniversity of Technology

This work has been supported by the European Union in the framework of European Social Fund through the Warsaw University of Technology Development Programme.

Abstract. Subjective character of generalization process is difficult to be mimed by the popular algorithms and inference systems based on the classical two-valued logic, while it does not include e.g. the contextual character of generalization process. Executed studies are to prelude building the comprehensive system for generalization of geographic information based on the non-classical logics (fuzzy and rough). Using Matlab Toolbox, heuristic method was used for testing different variants of decision rules, using diverse attributes sets. Results were evaluated by visual assessment and by comparison with the existing maps at smaller scales. The established research shows that fuzzy rules can be used for selection of linear objects as roads and hydrographic network.

Keywords:generalization of geographic information, fuzzy logic, generalization parameterization, selection operator

1. Introduction

The geographic information generalization process is considered as one of the greatest challenges in contemporary cartography. During the current studies and implementation tests number of algorithms were developed, which all are able to realize certain generalization operators (like selection, simplification, object shifting). However, there still remains the problem of building the complex system, which would be able to decide about the need and sequence of particular generalization operators at each object, as well

as about the specific parameters of those operators. Choosing the proper generalization operator, algorithm and their parameters depends on number of factors, among which scale (for analog data) or level of detail (for spatial databases) are crucial ones.

Subjective character of generalization process is difficult to be mimed by the popular algorithms and inference systems based on the classical two-valued logic, while it does not include e.g. the contextual character of generalization process (Olszewski 2009). Executed studies are to prelude building the comprehensive system for generalization of geographic information based on the non-classical logics (fuzzy and rough).

The research were established using MatLab Fuzzy Logic Toolbox The goal was to develop the knowledge base containing the fuzzy rules for selection operator (within the FIS – Fuzzy Inference System) for two selected feature classes (road network and hydrological network) for the chosen test area. Fuzzy Inference System for other generalization operators will be developed during further studies. Rough sets theory is planned to be utilized for selection of significant attributes used for creation of fuzzy generalization rules.

2. Data and test area

2.1. Source data

Test area was localized in south Poland and covered about20km^2(one sheet of topographic map at scale 1: 10 000) of low urbanized region in Beskids. It included small city Dukla and some small villages.

Data chosen for generalization come from the Topographical Database (TBD) which is the basic reference database in Poland at the level of detail 1:10 000 (Gotlib, Iwaniak, Olszewski 2006). TBD contains vector representation of objects as well as its characteristic by specified attributes. Two types of linear object were chosen for the analysis: road and hydrographic network.

2.2. Data preparation

The data used as an input in FIS were modified attribute tables of the mentioned classes of objects.

Existing descriptive attributes in ordinal scale were transformed into interval scale which was needed as the first step for their fuzzification. This transformation was the first step where the subjectivity of the process can be observed. As in ordinal scale intervals between attributes values are not

specified, transition to interval scale required those intervals to be defined. Those can be done in different ways – see example in*Table 1*.

Additionally new attributes were computed using geometrical and topological dependences. Calculation of those algorithms demanded spatial analyses of exclusively considered object class or of considered object class within the dependency on other object classes. In this way the geometrical information included in a source data could be utilized. In such a way attributes as "builded area in a buffer of 100 m from the road per 1 m of road" were calculated. Computation of this attributes was realized by Model Builder tool in ArcGIS software (*Figure 1*).

oryginal attribute	attribute in interval scale			
	step = 1	normalized, equal steps	normalized, rising steps	normalized, decreasing steps
interior	1	1	1	1
other	2	20	5	45
municipal	3	40	15	70
district	4	60	30	85
provincial	5	80	55	95
state	6	100	100	100

Table 1. Different types of transformation of the same attribute (road management category) from ordinal to interval scale.

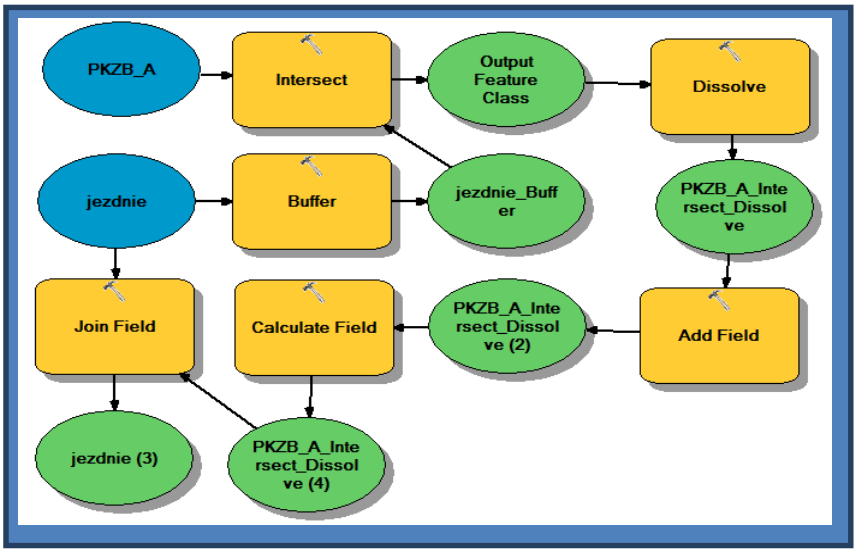

Figure 1. An example of model in Model Builder containing computation of built-up area (in square meters) in a buffer of 100 m from the road per 1 m of road.

Because diversified data range is not recommended for fuzzy reasoning and can perturb its results for some of the tests all of the data were normalized to the interval 0 to 100.

3. Methods and tools

3.1. Fuzzification

The first step in fuzzy reasoning is mathematical definition of linguistic variables by membership functions. This step was done using MatLab Fuzzy Logic Toolbox predestined for this goal. This tool allows user to design input and output variables of FIS (Fuzzy Interference System). At this step for each of the variable it is defined: the range of values it can take, the number of linguistic variables assigned to this values, shape and parameters of membership function for each of linguistic variables (*Figure 2*). All this decisions have strictly subjective character, demand cartographer experience and expert knowledge and strongly influence the further steps of fuzzy reasoning.

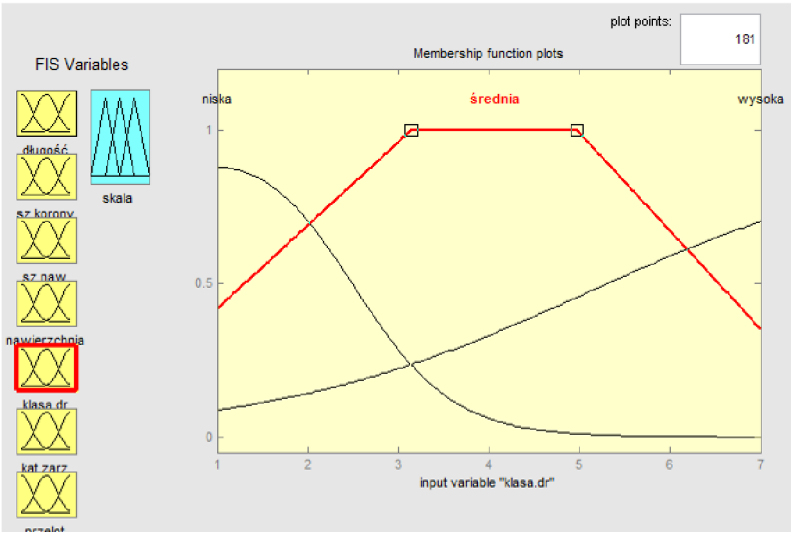

Figure 2. Membership function definition.

At this step it is also needed to decide what the decisive attribute of fuzzy reasoning is. For the operator of selection the natural answer has binary character – it selects or does not select this object. However such an answer is restricted to one, predefined level of detail, while the great advantage would be to have solution which helps in generalization for different scales. To meet this assumptions the proposed decision attribute was the scale (or strictly saying the scale denominator) above which specified object should

be deleted. Here also the fuzzification of scale denominator values was done (for small, medium and high).

3.2. Fuzzy Inference System – rule designing

The crucial step is designing fuzzy rules based on the linguistic variables. While the classical rule can read as follows: „if the length of the stream is shorter than 100 m delete it for scale 1: 50 000“, the rule based on linguistic variables would be rather „if the stream is **short** the minimal scale of display is **big**“(or „maximal scale denominator is **small**“). “Short” and “big/small” are in this case linguistic variables which were subjectively defined in previous step.

During rules definition linguistic variables can be precede by “not” operator and the rules can be combined using standard logical operators as “and” or “or”. There are few ways mathematically deal with this operators in fuzzy logic and it can be also decide by the designer of knowledge base.

The number and content of the rules should be the result of cartographer knowledge and experience. Here the heuristic method was used to test different variants. MatLab Fuzzy Logic Toolbox has two useful visualizations of rules and impact of specific attributes for the decisive one (*Figure 3*).

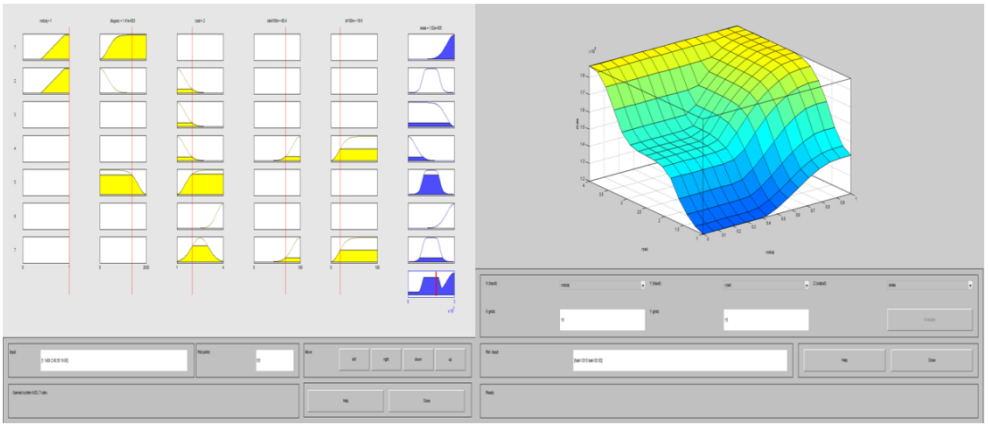

Figure 3. Left: impact of all attributes on decision can be tested by fluent, graphical changing its values – system shows the results for defined attribute values. Right: decisive surface showing influence of two chosen attributes on decision for the whole extent of their values. Both based on currently defined rules.

3.3. Results and their evaluation

The result of fuzzy reasoning has, as well as the input variables, fuzzy character. Therefore the process of defuzzification is needed to get final, crisp parameter. There are different methods of defuzzification which can be ap-

plied depending on expectations. The result of this process is crisp value of scale denominator below which specified object is not selected.

The evaluation of results is done by displaying the objects above certain scale value. First step is visual evaluation which can detect i.e. problems with topology and the general tendencies of parameterized process (Mackaness WA, Ruas A 2007). Secondly the comparison with existing maps at similar scales is done. The maps content is treated as the hidden expression of cartographical knowledge and experience (Fiedukowicz, Olszewski 2011).

4. Results and discussion

4.1. Hydrographic network

Data about hydrographic network do not have very detailed attributes itself as they come from topographic database. Specifically there is the distinction between river and stream (0-1) and the watercourse width (which value is frequently missing) therefore the first trials on FIS were based mainly on attributes calculated in spatial analysis. First system employed information such as:

- length of other watercourses per 1 m of each object,
- length of roads sections per 1 m of eachwatercourse,
- the watercourse order appointed according to the Horton–Strahler number (see *Figure4*) .

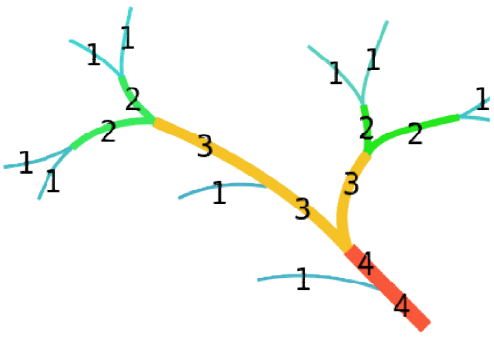

Figure 4. Horton-Strahler stream order – method of determining (source: http://en.wikipedia.org/)

Figure 5 shows the result: *B* - is for the scale 1: 150 000 and *C* - 1: 200 000. For the same sets of attributes but different rules we have *E* (1: 100 000)

and *F* (1: 150 000). *D* employs also default attributes for the database and shows result of selection for the scale 1: 150 000.

It is visible that the result varies depending on the rules and attributes chosen. There is a high sensitivity on scale denominator for which we visualize the data and the meaning of the same scale value is different depending on FIS. The following step of the research will be to scale the process in the way that the selection results are adequate to the scale of display. It can be done by changing the range and membership functions of input and output variables and modifying fuzzy rules.

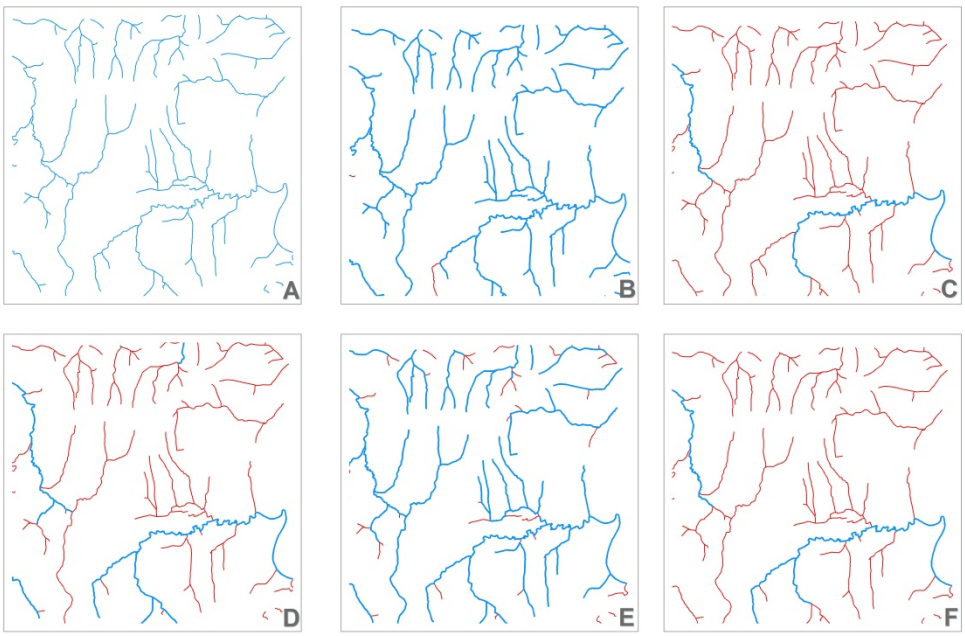

Figure 5. Hydrographic network selection operator: A - source data visualization (in other figures in red); B,C, E, F are based on the same sets on geometric attributes (the same rules but different scales: B&C and E&F); D - also attributes from source data.

4.2. Road network

For the road segments the number of attributes in database is much higher than in case of hydrographic network. Therefore the firs trials of FIS for roads selection were based on the TBD's original attributes.

The effort was done to check how the attributes values scale transition mentioned in *Section 2.2* influences the reasoning (with unchanged other para-

meters like attributes taking part in reasoning, types and parameters of membership functions as well as the read of rules). The results are presented in *Figure 6*. It is visible that in this case normalizing interval scale, while maintaining equal steps, did not influenced results (*Figure 6 B* and *D* are identical). However there were no big differences between original scale ranges of the attributes (there varied between 6 and 12). If the scale ranges are more diversified normalization could probably influence FIS answers. Contrary the influence of steps size into generalization results was proven. Parts *D*, *E* and *F* of *Figure 6* show result for the same attributes set, the same membership function and scale which differs only with the step chosen. For rising step ("better" roads were much better than the others) the selectivity of the system is greater, for decreasing intervals ("worse" roads are much worse than others), more roads left. It should be noted that in this case the same approach was used for all four attributes. It can be however customized by combining them.

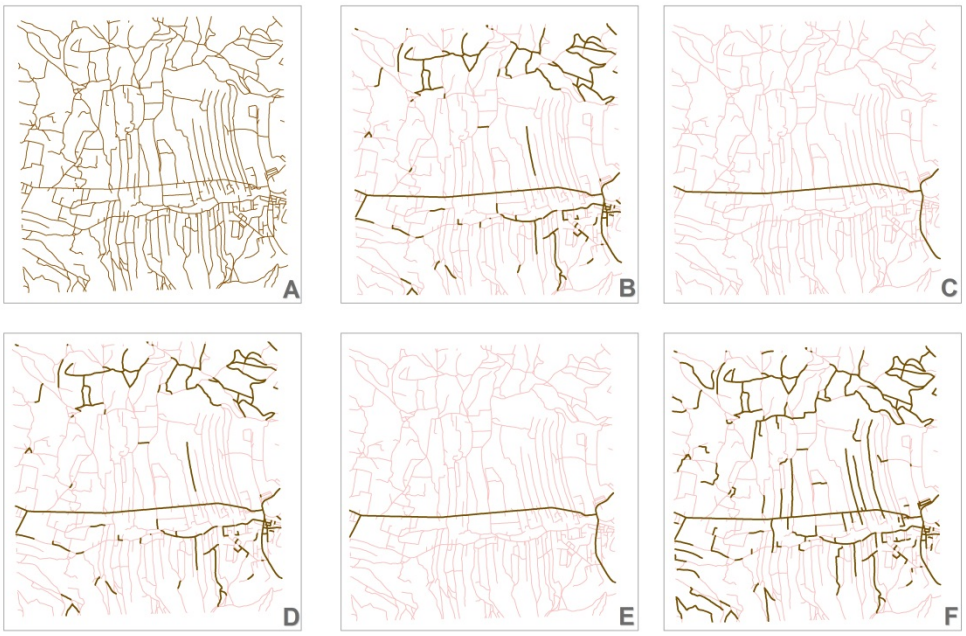

Figure 6. Road network selection operator based on original attributes from TBD: A - source data visualization (in other figures in brighter color); B,C - transformation to interval scale with step=1 (B - scale 1: 100 000, C - 1: 250 000); D, E, F - normalized interval scale (1 to 100), scale 1: 100 000; D - equal intervals; E - rising intervals, F - decreasing intervals.

Next trialswere prepared using only geometric attributes calculated by analysis (all of them were normalized to the range 0 to 100):

- built-up area in 100 m buffer per 1 m of road,
- built-up area in 50 m buffer per 1 m of road,
- number of buildings in 100 m buffer per 1 m of road,
- length of other roads in 100 m buffer per 1 m of road.

While for the hydrographic network proposed geometric attributes seemed reliable, here it is clearly visible that they are insufficient(*Figure 7A* and *B*). The results could be surely improved by better rule designing but for the two presented examples of FIS, based on the same set of arguments, gives undesirable results.In those cases insignificant roads were selected (and there were too many or not enough of them).

Last but not least the original attributes and attributes derived from geometry were combined in FIS, which gave quite interesting results (*Figure 7 C*). This kind of solution seems most valuable and will be developed in further studies.

Figure 7. Road network selection operator employingattributes calculated during analysis: A, B - the same sets of attributes coming exclusively from analysis; C - results of combination of two types of algorithms - original and derived from geometry.

5. Conclusion

The above example evidences that non-classical (specifically fuzzy) reasoning has a high potential for generalization process. However they still remain some challenges while using such methods.

The main problem which can be seen is the selection of attributes needed for the reasoning process. Too high number of attributes is not desirable as it makes reasoning process complicated and extends computation time

(both for attributes computation and for the reasoning process itself). Therefore author plans to apply rough logic and reduct concept for selection of significant attributes.

The further analyses are also to point which variant of attributes scale transformation should be applied, how to choose the proper membership function and how to formulate the fuzzy rules. All this parameters are needed for efficient and automate generalization process taking advantage of fuzzy logic. The approach seem valuable for automation of generalization process as it may mimic the way of human thinking (by linguistic variables) and deal with its ambiguity.

References

Fiedukowicz A, Olszewski R (2011) Ewaluacja statystyczna jako miara poprawności generalizacji informacji geograficznejna przykładzie opracowania komponentów pochodnych BDG in Zastosowanie statystyki w GIS i kartografii, Główne problemy współczesnej kartografii,p. 104-126

Gotlib D, Iwaniak A, Olszewski R (2006), Budowa Krajowej Infrastruktury Danych Przestrzennych w Polsce. Harmonizacja baz danych referencyjnych. Wrocław: Wydawnictwo Akademii Rolniczej we Wrocławiu.

Gotlib D, Iwaniak A, Olszewski R (2007) GIS. Obszary zastosowań, Wydawnictwo Naukowe PWN, Warszawa

Mackaness W A, Ruas A (2007) Generalization of geographic information: cartographic modeling and applications, Evaluation in the Map Generalization Process, p. 89-111, ICA

Olszewski R (2009) Kartograficzne modelowanie rzeźby terenu metodami inteligencji obliczeniowej