Dobesova, Z.(2013): Visual language for geodatabase design.
SGEM 2013 13th International Multidisciplinary Scientific GeoConfrence,
Proceedings Volume I STEF92 Technology Ltd. Sofia, Bulgaria, 603 – 610s.
ISBN 978-954-91818-9-0. DOI: 10.5593/SGEM2013/BB2.V1/S08.040

VISUAL LANGUAGE FOR GEODATABASE DESIGN

Ing. Zdena Dobesova, Ph.D.

Palacký University, Olomouc, **Czech Republic**

ABSTRACT

Entity relation (ERA) diagrams are use in the stage of conceptual design of database. Geographic information systems are often based on a geodatabase where the spatial data and attribute data are stored together. Software ArcGIS Diagrammer is used for the design of ArcGIS geodatabase. A set of graphical symbol is used for geodatabase design. The symbols are boxes with color fill. Arrow connectors join some elements.

The aesthetic, cognitive and perceptual quality of geodatabase diagram is important for comprehension by user. The perception depends on several aspects. A scientific method called the Physics of Notation was suggested by Daniel Moody. The theory of visual notation design is focused on the physical (perceptual) properties of notations rather than their logical (semantic) properties [1]. Some of the nine principles from the Physics of Notations will be applied for ArcGIS Diagrammer. The main aim of this article is an evaluation of visual notation in ArcGIS Diagrammer.

Keywords: geoinformatics, visual programming language, diagram, cognition, perception, database

INTRODUCTION

Graphical notation is used very often in information technology. Entity-Relation (ER) modelling for database design is one of the most successful notations beside Data Flow Diagrams (DFD). ER and DFD notations were developed in the 1970s. In fact, these modelling techniques are two most commonly used techniques in practice [2]. ER modelling exists in a variety of visual dialects. Fig. 1 shows Chen notation, Information Engineering (IE) notation and Microsoft Access notation. Other dialects are associated with some CASE tools. Despite the fact, these notations have been used in practice for over 30 years, there is still no consensus on which is the best [1].

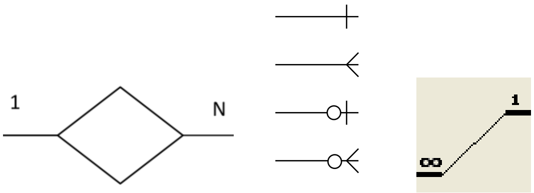

Fig. 1 Dialects of ER notation for relationship (left Chen notation, middle – IE notation, right MS Access)

In the field of visual language for geographic information systems, there are only evaluation of visual languages for data flow diagrams [3], [4], [5], [6].

PRINCIPLES OF PHYSICS OF NOTATION

The aesthetic, cognitive and perceptual quality of geodatabase diagram is important for comprehension by the author of the diagram and both another user. A scientific method called the Physics of Notation was suggested by Daniel Moody [7]. Moody presents nine principles that can be successfully used to evaluate visual programming languages. The principles are: Semiotic Clarity, Perceptual Discriminability, Semantic Transparency, Dual Coding, Visual Expressiveness, Graphic Economy, Cognitive Integration, Complexity Management and Cognitive Fit. The article takes some principles and applies it to the graphic elements of ArcGIS Diagrammer to evaluate it.

ArcGIS Diagrammer

Software ArcGIS Diagrammer is a product by Application Prototype Lab, Esri company, Redlands. ArcGIS Diagrammer is a case tool for GIS professionals to create, edit or analyze geodatabase schema. Schema is presented as editable graphics in an environment familiar to users of Microsoft Visual Studio. The utilization is strictly aimed to design geodatabase schema for Esri geodatabase format. The storing of data in geodatabase is a necessity for followed spatial analysis [8].

The offer of basic graphic element of in Palette window (Fig. 2) corresponds to the possible structures in geodatabase. Example of simple diagram is showed in Fig. 3.

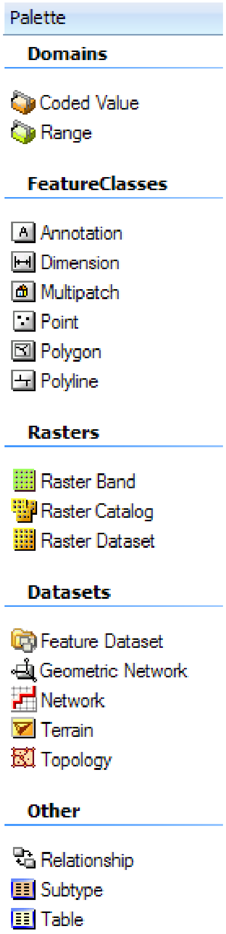

Fig. 2 Basic database objects with icons in the Palette window of ArcGIS Diagrammer

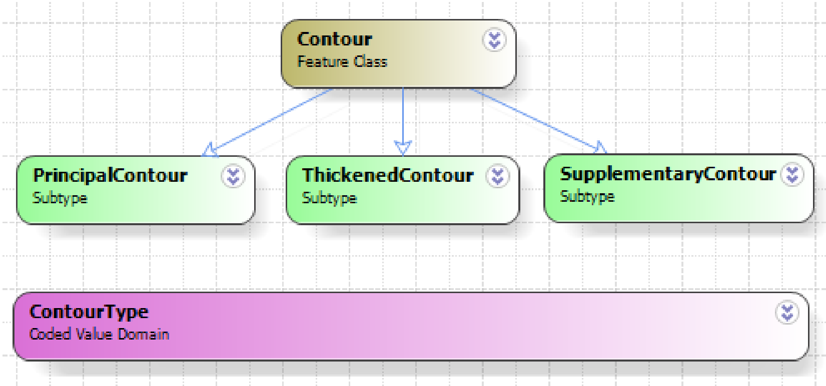

Fig. 3 Diagram with graphic symbols

The design of the model is by dragging an item from Palette window and dropping it into the diagram. All fourteen graphic elements are rounded rectangles. The size of rectangles is changeable, especially the width. The rectangles vary in color fill. The color fill is not solid. The color fill segues from dark tone on the left side to light tone of color on the right side of the rectangle.

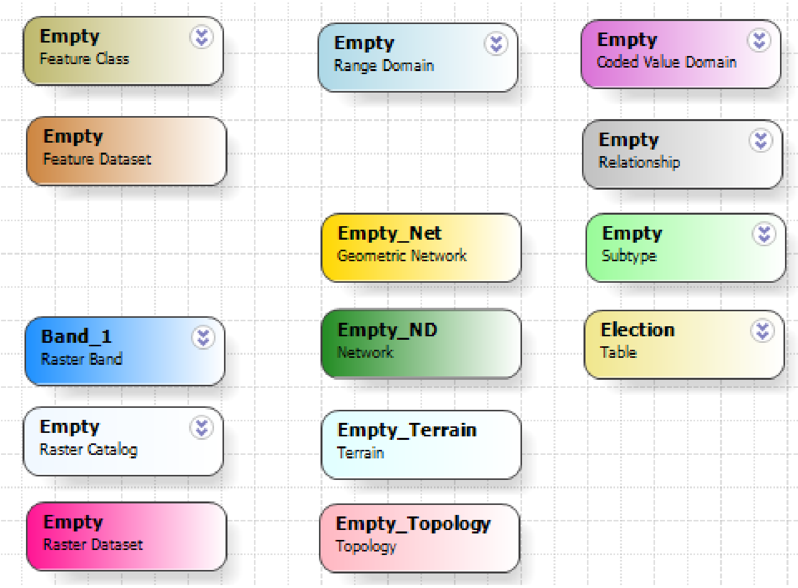

Fig. 4 List of all empty (unused) elements in diagram

Principle of Semiotic Clarity

The principle of Semiotic Clarity defines that there should be a one to one correspondence between syntactic and semantic features. There is one symbol overload. Six types of feature classes (different semantic) are encoding only by one graphic symbol (Fig.5). It is rectangle with brown fill. Feature classes differ in spatial representation (point, polygon, polyline) and meaning (Annotation, Dimension, ..). The reason is that all are feature classes and they are near semantically. There are no other beaks of that rule as a symbol deficit, symbol redundancy or a symbol excess.

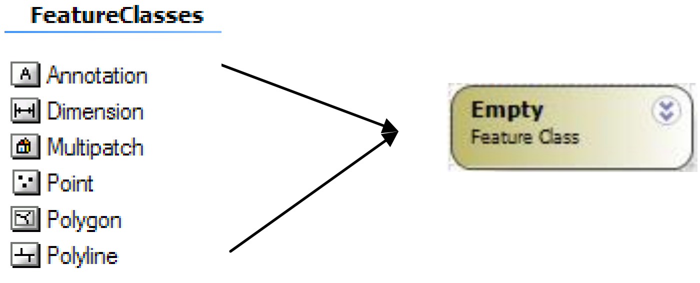

Fig. 5 Symbol overload for feature classes

Principle of Perceptual Discriminability

The principle of Perceptual Discriminability states that different symbols should be clearly distinguishable from each other [1]. The graphical symbols differ only by color fill in ArcGIS Diagrammer. Discriminability is very low, especially some tone of color are very near. Yellow table has a lighter tone that dark yellow for geometric network. There is a space for improvements.

Principle of Semantic Transparency

Principle of Semantic Transparency means - use visual representations whose appearance suggests their meaning. The shapes of symbols are very similar. Icons in Palette window (Fig. 2) have very well conjunction with meaning. Nevertheless, these icons are not used as a part of all graphical symbols in diagram. The using of that icon as a part of graphical symbol would increase the semantic transparency. Supplemental inner icons are also a solution for symbol overload that is mentioned in semiotic clarity (Fig.6).

Fig. 6 Suggestion for improvement the semiotic clarity, semantic transparency, perceptual discriminability and dual coding

Semantic transparency can be improved by fitting layout. The layout depends on the author of diagram. The subtypes are defined for some feature. It is valuable to put the rectangle of subtype near to the rectangle of feature and connect them by arrow. The same situation is in the definition of domain. The size of rectangle for domain could expressed the “cover” of domain (Fig. 3). The alignment of boxes helps in remembering of the base structure of geodatabase. The good example is an arrangement of feature classes in tree datasets in geodatabase for the book “Atlas of election” [9]. Each dataset is for separate year of election - Election2000, Election2004, Election2008.

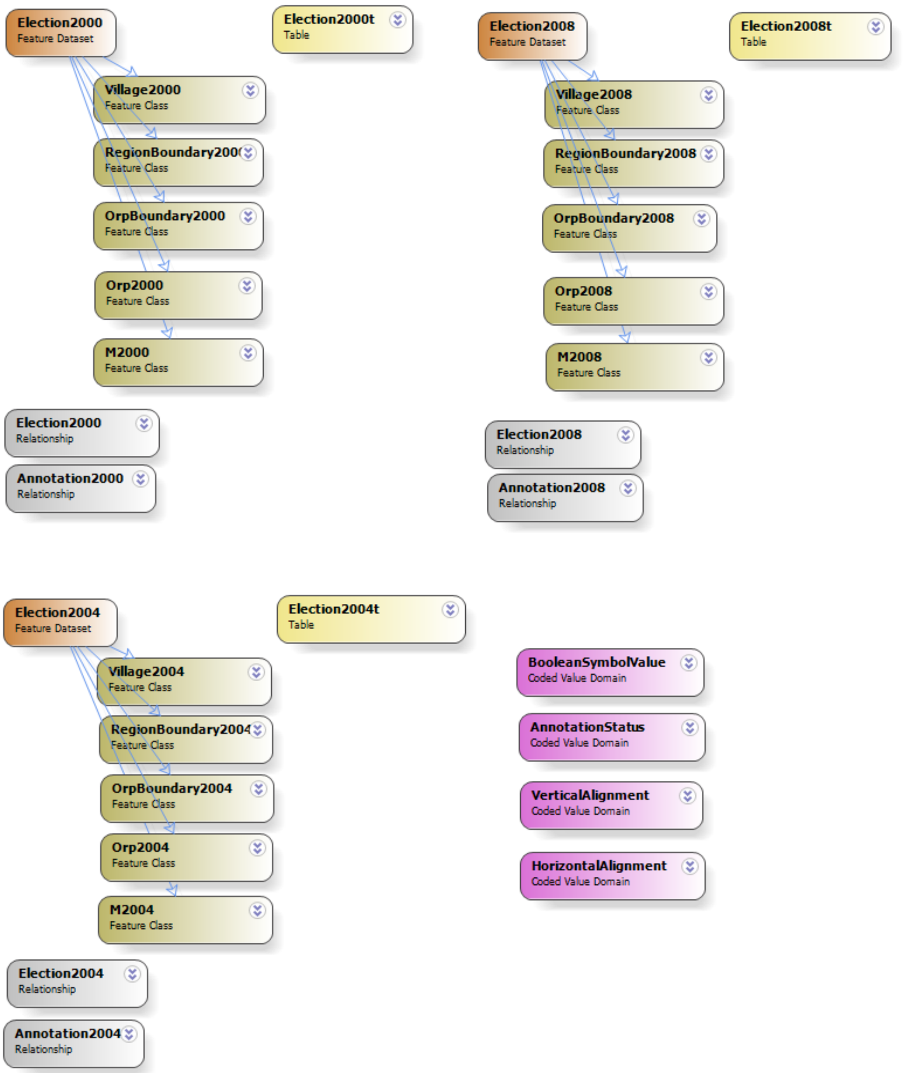

Fig. 6 Database model for the data for the Atlas of Election in ArcGIS Diagrammer [9]

Principle of Dual Coding

Principle recommended - use text to complement graphics. This principle is fulfilled by text labels on the second line of each rectangle. Bold text on the first line is the name of the object. It is changeable to express the meaning of feature (e.g. River is a name of feature class for storing lines of rivers). The text on the second line is fixed and expresses the type of object. Textual differentiation of symbols is only one way of ArcGIS Diagrammer in semiotic clarity. Text labels are cognitively ineffective way of dealing with excessive graphic complexity (principle of Graphic Economy). Text processing relies on less efficient cognitive processes [1].

Principle of Visual Expressiveness

Principle of Visual Expressiveness states that the full range of visual variables and their full capacity should be used to represent the notational elements. Basic visual variables are color, shape and size. Variable color is used as fill of rectangles in a good way. The shape and size are not used. The shape is the same for all graphic elements. There is also space for improvement of notation. The dataset is a directory for feature classes. The change of size can be added manually by the author of diagram (Fig. 7). The emphasis for dataset can be by double outline. It is only suggestion for future development. Another visual variable is orientation. It is used only for connectors. Last two visual variables texture and brightness are not used. Visual expressiveness is far from the full range of visual variables.

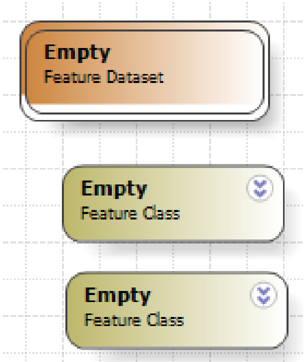

Fig. 7 Double outline for Feature Dataset and bigger symbol for Feature Dataset

Principle of Graphic Economy

That principle detects if the number of different graphical symbols should be cognitively manageable. The number of graphical symbols is 14. From that starting information can be assumed that memorizing of all symbols is hard and graphic economy is bad. Human working memory is limited to 7 ± 2 graphical symbols. In fact, all symbols are used very seldom in one diagram. Different tasks need topology, terrain or geometric network object in geodatabase. The number of different graphic symbols is 5 in average in one diagram. From that point of view, the graphic economy is cognitively manageable.

Principle of Cognitive Integration

This principle states – include explicit mechanism to support integration from different diagrams. Moody points out that a user need to be able to address four basic questions with respect to navigation: Where am I? Where can I go? Am I on the right path? Am I there yet?[10]. Cognitive integration only applies when multiple diagrams are used. ArcGIS Diagrammer creates one separate diagram without connection to another diagram. The existence of overview window with diagram (long- shot, context diagram) for diagram belongs to that principle. ArcGIS Diagrammer has no this overview diagram.

The principle of **Cognitive Fit** states that different representations of information are suitable for different tasks and different audiences. There is one of visual dialects. There is one visual representation for all purposes.

In the Fig. 1 are some examples of notation for relationship. There is no line symbol for expressing relationship. Relationship in geodatabase is a separate object relation feature class. Fig. 8 shows a simple example of two tables in geodatabase that have relationship OneToMany. Information are set and stored as properties of that object. The setting is on Properties window on the right side. The graphic expression of relationship by connect line is missing. It is a pity that this classical graphic notation in the database modelling is missing.

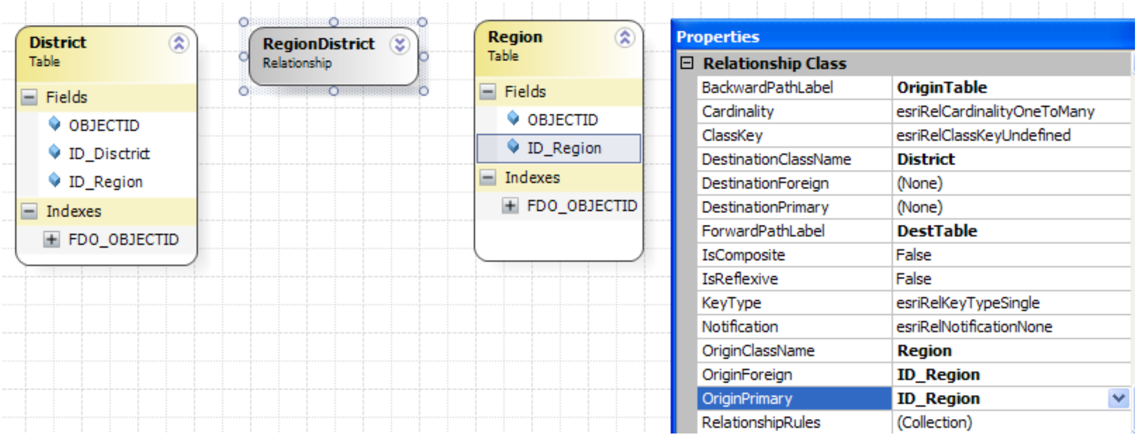

Fig. 8 Database relation

CONCLUSION

Method "Physics of Notation" by Daniel Moody was applied on database modeler ArcGIS Diagrammer. This article tries the utilization of some principles on visual language for Esri geodatabase design. Semiotic Clarity is medium in ArcGIS Diagrammer. There is one symbol overload (brown rectangle) for feature classes. Perceptual Discriminability is also medium. It corresponds to the principal of Visual Expressiveness. Only one visual variable color is used. Dual coding by text is used very good. The evaluation according to physics of notation results to the suggestion of one improvement. The adding of inner icons to the rectangles increases the semiotic clarity, semantic transparency, perceptual discriminability and dual coding. The idea of adding inner icon is the same as in the evaluation of ModelBuilder [12].

Future work expects empirical studies by eye tracking systems Popelka [11]. We plan to prepare set of geodatabase diagrams and verify the hypothesis in evaluation according to principals of physical notation.

ACKNOWLEDGEMENT

Work was supported by the project CZ.1.07/2.3.00/20.0166.

REFERENCES

1. Moody, D.L., The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. Ieee Transactions on Software Engineering, 2009. 35(6): p. 756-779.
2. Davies, I., et al., How do practitioners use conceptual modeling in practice? Data & Knowledge Engineering, 2006. 58(3): p. 358-380.

3. Dobesova, Z. Visual programming language in geographic information systems. in Recent Researches in Applied Informatics, 2nd International Conference on Applied Informatics and Computing Theory, AICT 11. 2011. Prague: NAUN/IEEE, WSEAS Press.
4. Dobesova, Z. Programming language Python for data processing. in Electrical and Control Engineering (ICECE), 2011 International Conference on. 2011.
5. Dobesova, Z. and P. Dobes. Comparison of visual languages in Geographic Information Systems. in IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). 2012.
6. Dobesova, Z., Visual programming for novice programmers in geoinformatics, in 12th International Multidisciplinary Scientific GeoConference. 2012: Bulgaria, Albena. p. 433- 440.
7. Moody, D. Theory development in visual language research: Beyond the cognitive dimensions of notations. in Visual Languages and Human-Centric Computing, 2009. VL/HCC 2009. IEEE Symposium 2009.
8. Sedlak, P., J. Komarkova, and A. Piverkova, Identification of Movement Barriers for Physically Impaired People in the City Environment by means of Spatial Analyses, in Acs'09: Proceedings of the 9th Wseas International Conference on Applied Computer Science, R. Revetria, B. Mladenov, and N. Mastorakis, Editors. 2009, World Scientific and Engineering Acad and Soc: Athens. p. 247-252.
9. Dobesova, Z., Database modelling in Cartography for the “Atlas of Election”. Geodesy and Cartography, 2012. 38(1): p. 20-26.
10. Thomas, J.C., et al. Using the Physics of notations to analyze a visual representation of business decision modeling. in Visual Languages and Human-Centric Computing (VL/HCC), 2012 IEEE Symposium 2012.
11. Popelka, S. and V. Voženílek, Specifying of requirements for spatio-temporal data in map by eye-tracking and space-time-cube, in International Conference on Graphic and Image Processing (ICGIP 2012). 2013: Singapore p. 87684N-87684N.
12. Dobesova, Z. Using the “Physics of notation to analyse ModelBuilder diagrams, 13th International Multidisciplinary Scientific GeoConference. 2013: Bulgaria, Albena