TOSCANA
Management System for Conceptual Data

Martin Skorsky

Technische Hochschule Darmstadt
D-64277 Darmstadt
skorsky@mathematik.th-darmstadt.de

TOSCANA is a software for the visualization of data with nested line diagrams. It is a tool for navigating in databases and for the retrieval of objects in a database. It enables the user to discover relationships and implications between attributes.

TOSCANA is based on Formal Concept Analysis as introduced in [Wi82]. This method formalizes a *relation I* between a set G of *objects* and a set M of *attributes* by the *formal context* (G, M, I). A formal context can be handled as a table or a view of a relational database. The formal context determines its *concept lattice* which is the structure of the relation subconcept–superconcept. The concept lattices are graphically represented as line diagrams. For the basic notions of Formal Concept Analysis see also [DP90, chapter 11] or in [GW95].

The line diagrams for practical applications get very large. TOSCANA uses *nested line diagrams* which were introduced in [Wi84] (see also [Wi89]). We split a given context into parts, draw a line diagram of each part and then nest these line diagrams into each other. This gives a simplified diagram because groups of parallel lines are replaced by one line and the ellipses around the small diagrams (see Figure 2). TOSCANA automatizes the nesting of the diagrams and allows a flexible choice of the diagrams. As mentioned above, the formal context can be kept as a table or view in a database. TOSCANA selects the information for the nested line diagram from such a database. Thus, TOSCANA helps to analyze large sets of data and enables the user to navigate through this database.

The ideas for TOSCANA are first described in [VWW91]. The further development of these ideas is reflected in [SSVWW93], [KSVW94], [VW95]. This last paper was presented to the Graph Drawing conference in 1994 at Princeton.

We will demonstrate one basic direction of applying TOSCANA. The example (see 1, 2, 3) shows how TOSCANA is used to structure a set of documents. The documents, which take the role of the objects of the formal context, are related to search keys, which are the attributes. The documents are German law documents and technical norms. The search keys are grouped into themes. Each theme is represented by a separate line diagram. Now a users sees the documents not only for one search key but for the whole group of keys for the selected theme. The documents were unfolded from the very specific documents to the general ones. The combination of themes allow complex selections and searches.

The screen dumps in figure 1, 2, 3 show the functionallity of TOSCANA for navigating within the diagrams and selecting objects. The functions to improve
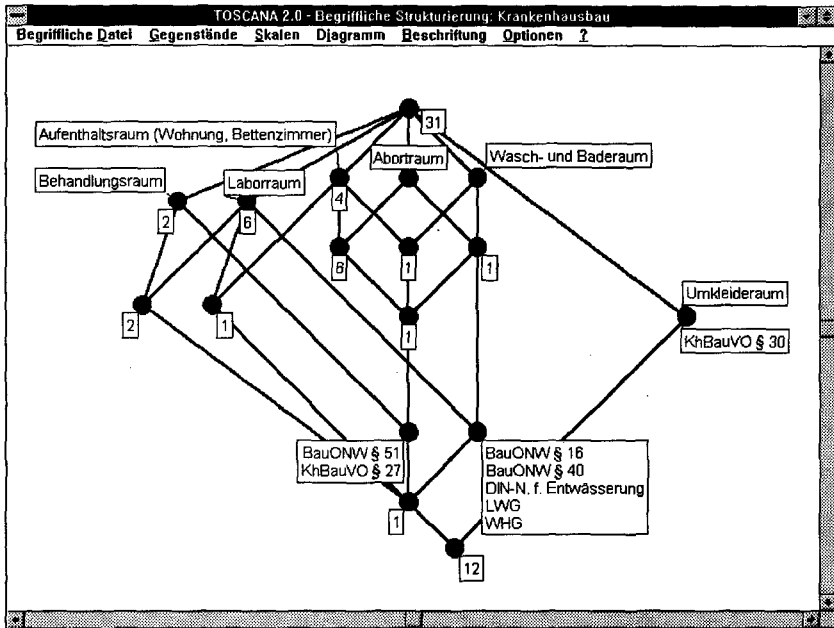
Fig. 1. line diagram about the theme 'functional rooms in a hospital'

the layout are not shown.

1. Figure 1: line diagram about the theme 'functional rooms in a hospital' ('Funktionsräume im Krankenhausbau')
2. Figure 2: nested line diagram of two layers; the first layer (ellipses) shows the theme 'functional rooms in a hospital' (same structure as in Figure 1), the second layer (small diagrams) shows the theme 'reliability in service and fire protection' ('Betriebs- und Brandsicherheit').
3. Figure 3: Zooming into the ellipse labelled with BauONW §16 BauONW §40 DIN-N. f. Entwässerung LWG WHG selects those five documents and shows how these documents concern 'reliability in service and fire protection' ('Betriebs- und Brandsicherheit')

Searching and Navigating is supported in two directions. First we can combine up to 4 attributes or themes to a nested line diagram. Nested diagrams with more than 4 attributes can not be handled in a practical way. The second direction is the process of zooming in and out. The user clicks with the mouse on an ellipse and gets a more detailed diagram of the situation specified by this ellipse. This is connected with a selection in the database. There is no limit to
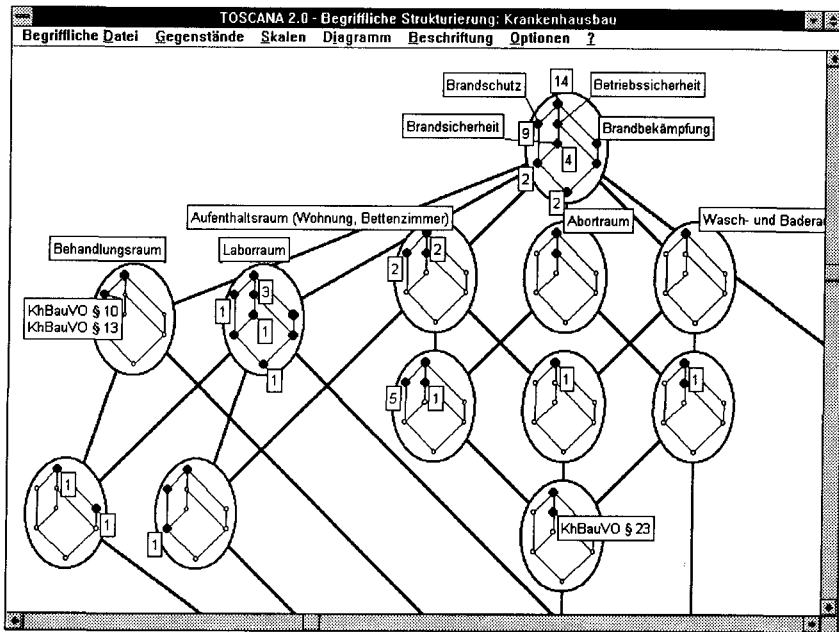
Fig. 2. nested line diagram of the themes 'functional rooms in a hospital' and 'reliability in service and fire protection'

the number of attributes or themes for the zoom process, thus we can refine the selection of the objects in the database.

The current implementation of TOSCANA runs on a PC with MS-Windows™ (486DX/50 processor or better, 8 MB RAM). It takes the data from databases managed by MS-Access™. The queries are done with SQL and transfered via DDE. The software is developed in C++ and uses a class library of concept analysis classes written by Frank Vogt.

References

[DP90]   B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order.* Cambridge University Press, Cambridge, 1990.

[GW95]   B. Ganter and R. Wille. *Formale Begriffsanalyse.* B.I.–Wissenschaftsverlag, to appear.

[KSVW94] W. Kollewe, M. Skorsky, F. Vogt, and R. Wille.   TOSCANA — Ein Werkzeug zur begrifflichen Analyse und Erkundung von Daten. In: R. Wille and M. Zickwolff, editors, *Begriffliche Wissensverarbeitung — Grundfragen und Aufgaben,* pages 267–288, Mannheim, 1994. B. I.–Wissenschaftsverlag.

[PC93]   PC Magazine, Volume 12, number 2. Ziff-Davis Publishing Company, Boulder 1993, pages 110–247.
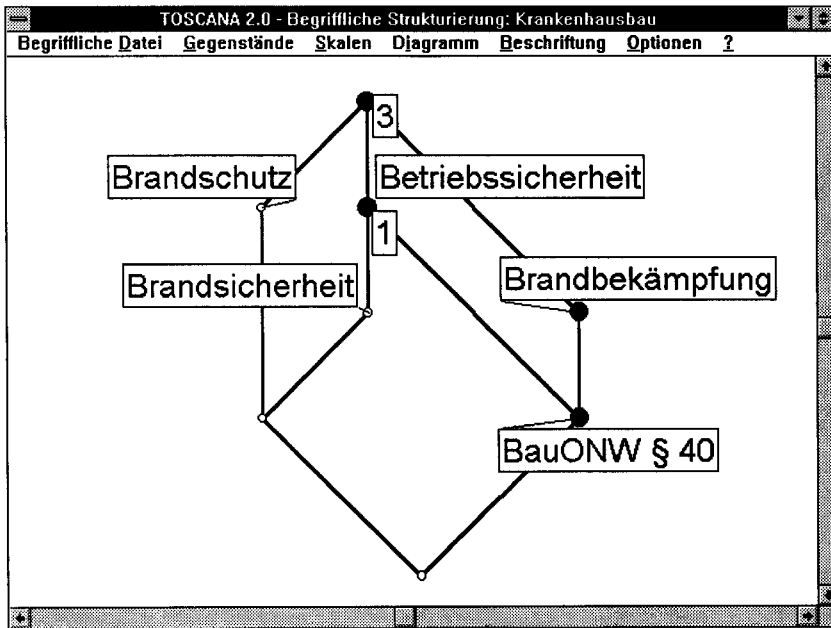
Fig. 3. line diagram about the theme 'reliability in service and fire protection'

[SSVWW93] P. Scheich, M. Skorsky, F. Vogt, C. Wachter, and R. Wille. Conceptual data systems. In: O. Opitz, B. Lausen, and R. Klar, editors, *Information and classification*, pages 72–84, Heidelberg, 1993. Springer–Verlag.

[VWW91] F. Vogt, C. Wachter, and R. Wille. Data analysis based on a conceptual file. In: H.-H. Bock and P. Ihm, editors, *Classification, data analysis, and knowledge organization*, pages 131–140, Berlin–Heidelberg, 1991. Springer–Verlag.

[VW95] F. Vogt and R. Wille. TOSCANA — A Graphical Tool for Analyzing and Exploring Data. In: R. Tamassia and I. G. Tollis, editors, *Graph Drawing*, pages 226–233, Heidelberg, 1995. Springer–Verlag.

[Wi82] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In: I. Rival, editor, *Ordered sets*, pages 445–470, Dordrecht–Boston, 1982. Reidel.

[Wi84] R. Wille. Liniendiagramme hierarchischer Begriffssysteme. In: H.-H. Bock, editor, *Anwendungen der Klassifikation: Datenanalyse und numerische Klassifikation*, pages 32–51. Indeks–Verlag, Frankfurt, 1984. Line diagrams of hierarchical concept systems (engl. Translation). *Int. Classif.* **11** (1984), 77–86.

[Wi89] R. Wille. Lattices in data analysis: how to draw them with a computer. In: I. Rival, editor, *Algorithms and order*, pages 33–58, Dordrecht–Boston, 1989. Kluwer.