Structured Visualization of Search Result List

U. Preiser
Siemens AG ZT IK 1
D-81739 München, Germany

Abstract: Information highways are being built around the world. The impact on the business of companies like Siemens is tremendous. New techniques for information processing are required. The integration of features for navigation, in addition to the static presentation, enables the user to find the relevant information by browsing through the information space. By structured visualization, which means to represent the information in form of a 3-D net, the “Kontextgestalt”, the user gets an overview of the relevant topics the information set is concerned about.

1 Introduction

During the past years the growing networks (Internet and Intranet) resulted in a large number of potential possibilities to access information sources. This availability produces a vast amount and also a large variety of information. Therefore, it is necessary to provide specified functions for information systems, which will visualize the complex information bases and their structure. Although while today large bases of data and knowledge are being maintained, the man-machine interface is, by its textual command orientation, still not sufficiently adjusted to human perception (visual sense !) and orientation. The approach by “Kontextgestaltgebung” realizes an easy and intuitive way to visualize large data and knowledge bases and therefore results in an easier way to access information by retrieval and processing.

Remark: The German word “Kontextgestalt” is very hard to translate into English. Its meaning is something like context shape and its plural is “Kontextgestalten”. The meaning of the German word “Kontextgestaltgebung” is something like building a “Kontextgestalt”.

The techniques of contextual information processing by 3D-visualization are supported by the principle “contextual correlation ≈ spatial proximity”. Because of their visual sense, human beings are most familiar with the three-dimensional euclidian space (in the following called 3D-space). Therefore spatial proximity as a metaphor for contextual coincidence or similarity is especially memorable. Any electronic source of information can be described in a suitable way by means of an application specific terminology in the form of formal contexts. In this case, a formal context consists of the information units called “objects and attributes” and a binary relation (incidence relation) between objects and attributes. It is possible to deduce specific contextual measures for objects and attributes from such a formal context by using

I. Balderjahn et al. (eds.), Classification, Data Analysis, and Data Highways
© Springer-Verlag Berlin · Heidelberg 1998
the formal incidences. These measures are either called formal similarity of two objects or formal equivalency of two attributes. In a scaled space, information units (objects respectively attributes) can be represented as points in the sphere by using suitable norms respectively by distance measures. They are arranged in a way that the spatial distances, which are defined by the suitable norm, will equal the contextual distances (spatial proximity). The embedding of a context in such a space, will therefore maintain similarity and equivalency. The arrangement of the context in the 3D-space is made by an approximate transformation which represents contextual proximity by a useful spatial representation (i.e. contextual correlation \( \cong \) spatial proximity). Thereby a complex formal context is represented in an easily conceivable way, and its processing by means of search and navigation in space is made possible. This metaphor is used in a demonstrator for obtaining a structured visualization of search results. Thereby the search results can be received either from Internet-based search engines or from local data bases. The information is domain specific, a fact which influences the selection of the attributes used. The goal is to offer the user a structured preparation of the search results in the 3D-sphere instead of a pure listing. Therefore the user will be able to quickly identify relevant information, without having to view all search results.

2 Method and Metaphor

In most situations of data analysis the data records (obtained by measurements, interviews, collections ...) can be described by an application specific conceptual language. In this case, the data records will be represented by objects and their observed attributes in the form of so called "formal contexts", which are known from the formal concept analysis. A (single valued) formal context \( K = (G, M, I) \) consists of the following information units (Wille (1984)):

1. a set \( G \) of objects
2. a set \( M \) of attributes
3. a binary relation (incidence relation) \( I \subseteq G \times M \) between objects and attributes

The relation \( gIm \) between an object \( g \) and an attribute \( m \) is interpreted as "the object \( g \) has the attribute \( m \)" or "\( m \) belongs to \( g \)". Such single valued formal contexts exist for example in literature data bases: The attribute set consists of the relevant descriptors and the assignment of the descriptors to supply the incidence relation. The data records which are to be evaluated, often exist in the form of a table whose rows are seen as objects and whose columns are qualified or quantified attributes. This situation is described by an multivalued Kontext \( K = (G, M, W, I) \), where \( W \) is called the attribute expression. \( I \subseteq G \times M \times W \) is a ternary relation, so \( (g, m, w) \)