Chapter 2.29
User Interface Formalization in Visual Data Mining

Tiziana Catarci
University of Rome "La Sapienza", Italy

Stephen Kimani
University of Rome "La Sapienza", Italy

Stefano Lodi
University of Bologna, Italy

ABSTRACT

Despite the existence of various data mining efforts that deal with user interface aspects, very few provide a formal specification of the syntax of the interface and the corresponding semantics. A formal specification facilitates the description of the system properties without being concerned about implementation details and enables the detection of fundamental design issues before they manifest themselves in the implementation. In visual data mining, a formal specification can enable users to decide which interaction/operation to apply to get a desired result; help users to predict the results of their interactions/operations with the system; and enable the development of a general interaction model that designers/developers can use to understand the relationships between user interactions and their compositions. In this work, we describe an approach for formalizing the visual interface of a core data mining system, which has been employed in the development of a visual data mining system named VidaMine.

INTRODUCTION

In this day and age, data still present formidable challenges to effective and efficient discovery of knowledge. It should be acknowledged that a lot of research work has been and is being done with respect to knowledge discovery (KD). Much of the work has concentrated on the development and the optimization of data mining algorithms using techniques from other fields such as artificial intelligence, statistics, and high performance computing (Fayyad, Piatetsky-Shapiro, & Smyth, 1996b). Besides various glaring issues (such as

the need to have an overall framework that can support the entire discovery process, supporting human involvement in the entire process, etc), Mannila observes that relatively little research work has been published on the theoretical foundations of data mining (Mannila, 2000). On the same note, although there are many data mining efforts that deal with user interface issues, very few efforts provide or give a precise definition of the syntax of the user interface and the corresponding semantics.

Formal specifications enable the description of the system properties without having to be concerned about implementation details. The system properties are often specified using a precise notation. The specification can be used to construct models of the system. Formal methods make the analyst think abstractly about the problem at hand and the corresponding system thereby exposing fundamental design decisions well in advance before they manifest themselves in the implementation of the system. While it is true that the formal specification should not determine the programming aspects (e.g., algorithms and data structures), it should describe the behavior of the system in a precise or rigorous manner. Moreover, with a formal specification it is possible to transform a system model while preserving important properties of the model. In practice, a formal approach in which each and every design decision is proven to be a correct refinement step is rarely performed due to the high costs involved. However, substantial application of refinement does considerably improve the understanding of the design process. A formal specification of the visual interface of a data mining system can facilitate the gathering of information about the most useful usage patterns, which can then be used to guide the design and layout of user interfaces for visual data mining. Moreover, a formal specification of the visual interface can facilitate automated (and/or objective) evaluation of the usability of the user interface of a visual data mining system.

In information visualization, various specifications/models for characterizing visualization aspects have been proposed such as (Baudel, 2004; Chi & Riedl, 1998; Chuah & Roth, 1996). In fact, it does turn out that most of the efforts that are related to our work are mainly found in information visualization and exploration efforts rather than in core data mining. Some of the benefits of specifications/models such as the foregoing do apply to visual data mining as well, where visualization tends to be a key ingredient. Consequently and borrowing from Chi et al. (1998), a similar formal specification in visual data mining can: enable users to decide which user interaction/operation to apply in order to get a desired result; help users to predict the results of their interactions/operations with the visual data mining system; and enable the development of a general interaction model that designers/developers can use to classify and understand the relationships between user interactions and the composition of interactions. In fact, such a model could help eliminate errors caused by other imprecise or incorrect models. In this work, we describe an approach for formalizing the visual interface of a core data mining system. The proposed approach has been employed in the development of a visual data mining system named VidaMine.

BACKGROUND

Knowledge Discovery

Knowledge discovery (KD) may be defined as the process of identifying valid, novel, potentially useful, and ultimately understandable models and/or patterns in data (Fayyad, Piatetsky-Shapiro, Smyth, & Uthurusamy, 1996a; Fayyad et al., 1996b). On the whole, the knowledge discovery process may be defined as an interactive and iterative non-trivial process that entails various phases as seen in Figure 1.