Analytical Visualization Framework - a visual data processing and knowledge discovery system

Ivan Denisovich, Serge Shumilov

Department of Computer Science III, University of Bonn
Römerstraße 164, 53117 Bonn, Germany
divan@iai.uni-bonn.de

Study of physical, economical and social phenomena in modern city is a multi-disciplinary process, which often includes extensive simulation and modeling, statistical analysis, data exploration and visualization. Usually a user is forced to open several applications and manually exchange data between them repeating the same routine several times. For other tasks, the sequence of applied tools may change every time depending on the acquired intermediate results. In this case the user has to remember which actions were made that led to successful outcomes or to which step to come back to try another strategy. To make knowledge discovery and the following decision making effective, it is important to trace the user actions and grant support by automating routine sequences of actions.

One of the best ways to quickly evaluate computation results is to see data in a visual form. Visual data mining techniques have proven to be of high value in exploratory data analysis and they also have a high potential for mining large databases. Visual data mining is especially useful when little is known about the data and the exploration goals are vague. Since the user is directly involved in the exploration process, shifting and adjusting the exploration goals is automatically done if necessary.

The proposed Analytical Visualization Framework (AVF) provides its user with an instrument for effective knowledge discovery through visualization and interactive data manipulations. The paper presents the base architecture of the framework and describes its functionality on several practical examples. The AVF helps to organize and manage the knowledge discovery process decomposing it on small data processing steps implemented by individually replaceable modules. With a visual programming workspace provided by the AVF, the modules can be easily instantiated, connected and grouped into complex programs. The framework provides a rich palette of general-purpose modules, as well as modules specially designed for import/export, processing of data of specific types or production of custom visualizations. Moreover, using a convenient programming interface, advanced users can wrap external systems as modules or develop their own modules according to their needs.

Exploring a visual representation of data, the user can interactively change parameters, add or replace some modules at runtime, resulting in immediate update of the computed results and their visual representations. Chains of transformation modules can be branched to allow exploration and comparison of alternative scenarios. This interactivity significantly simplifies the finding of optimal policies providing the decision maker with an instrument for easy adjusting of the exploration goals. The structures of resulting programs lay a basis for customized decision-making tools and interactive reports.

The framework was successfully applied to monitor execution and analyze results of hydrological and economical models for integrated simulation of water distribution in the Volta basin. Data exchanged between models and other model output data is gathered and aggregated into convenient tabular form. To explore gathered data and compare different model runs, the user is provided with predefined interactive visualization tools. Using visual programming workspace, one can change this default representation, compose other visualizations or channel aggregated data to external systems.