TOWARDS A CONCEPTUAL FRAMEWORK FOR
VISUAL ANALYTICS OF TIME AND TIME-ORIENTED DATA

Wolfgang Aigner, Alessio Bertone, and Silvia Miksch

Department of Information and Knowledge Engineering
Dr.-Karl-Dorrek-Strasse 30, Danube University Krems
A-3500 Krems, AUSTRIA
<www.donau-uni.ac.at/ike>

Christian Tominski and Heidrun Schumann

Department of Computer Science
Albert-Einstein-Strasse 21, University of Rostock
D-18059 Rostock, GERMANY
<wwwicg.informatik.uni-rostock.de>

ABSTRACT

Time is an important data dimension with distinct characteristics that is common across many application domains. This demands specialized methods in order to support proper analysis and visualization to explore trends, patterns, and relationships in different kinds of time-oriented data.

The human perceptual system is highly sophisticated and specifically suited to spot visual patterns. For this reason, visualization is successfully applied in aiding these tasks. But facing the huge volumes of data to be analyzed today, applying purely visual techniques is often not sufficient. Visual analytics systems aim to bridge this gap by combining both, interactive visualization and computational analysis.

In this paper, we introduce a concept for designing visual analytics frameworks and tailored visual analytics systems for time and time-oriented data. We present a number of relevant design choices and illustrate our concept by example.

1 MOTIVATION & BACKGROUND

During the last decade, capabilities to both generate and collect data have seen an explosive growth. Advances in scientific and business data collection (e.g., from remote sensors, from space satellites, or from retail and production devices as well as growingly complex simulation systems) have generated a flood of data and information. Advances in data storage technology such as faster and cheaper storage devices with higher capacity, better database management systems, and data warehousing technology have allowed us to transform this data into "mountains" of stored data. Such volumes of data and information overwhelm most traditional manual methods of data analysis such as spreadsheets, ad-hoc queries, or simple visualizations. The need for new methods and tools which can intelligently and (semi-)automatically transform data into information and furthermore, synthesize knowledge are a core area of the emerging field of *Visual Analytics*.

The basic idea of *Visual Analytics* (Thomas and Cook 2005) is the integration of the outstanding capabilities of humans in terms of visual information exploration and the enormous processing power of computers to form a powerful knowledge discovery environment. Both visual as well as analytical methods are combined intertwinedly to fully support this process. Most importantly, the user is not merely a passive element who interprets the outcome of visual and analytical methods but she is the core entity who drives the whole process.

Time is an important data dimension that is common across many application domains, like transport, call centers, retail, production, health care, police, or financial services as well as for research in medicine, biology, and economics. Particularly, in the area of simulation systems, time is central to simulating dynamic system behavior as reflected by Robinson in his definition of *computer-based dynamic simulation* as being *"an imitation (on a computer) of a system as it progresses through time"* (Robinson 2004). Exploring trends, patterns, and relationships are particularly important tasks when dealing with time-oriented data and information. In contrast to other quantitative data dimensions that are usually "flat", time has an inherent semantic structure which increases its complexity dramatically. The hierarchical structure of granularities in time, as for example minutes, hours, days, weeks, months, is unlike most other quantitative dimensions. Specifically, time comprises different forms of divisions (e.g., 60 minutes resemble one hour while 24 hours resemble one day) and granularities are combined to form calendar systems (e.g., Gregorian, Julian, Business, or Academic calendars). Moreover, time contains natural cycles and re-occurrences, as for example seasons, but also social–somehow irregular–cycles, like holidays or school breaks. Therefore, time-oriented data need to be treated differently from other kinds of data and demand appropriate visual and analytical methods to analyze them.

Conceptual frameworks and reference models exist for both, interactive visualizations, e.g., (Card, Mackinlay, and Shneiderman 1998, Chi 2000, dos Santos and Brodlie 2004)

and analytical methods, e.g., (Han and Kamber 2005). Due to the fact that, on the one hand, reference models and frameworks for analytical methods often degrade visualization to a mere postprocessing activity to represent results and, on the other hand, visualization reference models and frameworks often consider analytical methods as side issues, integrated approaches have been developed, e.g., (Fayyad, Grinstein, and Wierse 2001). However, the specifics of time are often not reflected appropriately by treating time just as any other quantitative dimension, both in visualization and analysis. Usually, sequences of simple <time-point, value> pairs resemble the data models used for analysis and visualization. These systems and frameworks could be improved by considering the specifics of the dimension time in more depth, like for example support for time intervals or built-in aggregations along time granularities. To complement this, we propose a concept for the design of visual analytics frameworks and tailored visual analytics systems that emphasizes the special role of time. We consider both, interactive visualization and analytical methods as equally important and fully integrated parts.

First, we present a brief overview of the structure of our conceptual framework. Then, we present a number of relevant design choices for the three main parts (1) *internal data structures*, (2) *analytical & mining components*, and (3) *visualization & interaction components*. Finally, a practical example illustrates our conceptual framework.

2 A CONCEPT FOR DESIGNING A FRAMEWORK

By proposing a concept for the development of visual analytics frameworks and tailored visual analytics systems for time-oriented data, we want to layout a basis for further discussion and implementation. However, we do not intend to present a general, all-encompassing visual analytics framework. In the following, we will describe the overall architecture along with the necessary components and their functionality in detail. Note, that we do not propose a particular framework but describe the general steps and functionality of the involved principal components. The individual components can be integrated into various application-specific (simulation) systems. To create a visual analytics framework or system for time and time-oriented data, three main tasks need to be performed:

1. Modeling of time and time-oriented data
2. Computational analysis of time-oriented data
3. Interactive visualization of time-oriented data

All these tasks are domain-specific and need to be designed according to the data characteristics and the analysis goals at hand. In the following, we will describe the basic architecture of our framework including the interplay of its main components.

2.1 Architecture

Figure 1 shows the basic blueprint of our conceptual framework which combines visual and analytical methods. In order to be able to support a seamless interplay of these methods, an appropriate data model is needed. The used *internal data structures* reflect this data model in terms of data structures to store and retrieve time-oriented data as well as associated metadata. *External data structures* in various idiosyncratic formats have to be transformed in order to comply with the framework's time model. The *visualization & interaction components* are responsible for representing the stored time-oriented data visually and providing for suitable interaction possibilities. Both of these primary functions strongly depend on users' tasks, goals, and background. Moreover, the parameterization of the representation is vital for ensuring effective visualization. Besides the visual components, non-visual elements are part of our framework. As the component name implies, they are in charge of *analytical & mining* tasks to be executed on time-oriented data and are controlled by analytical parameters.
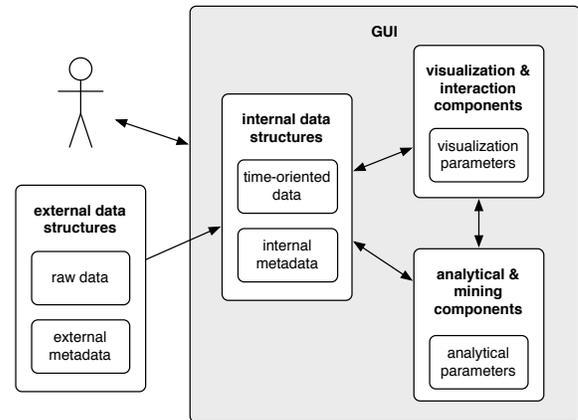

Figure 1: **Architecture of visual analytics framework for time-oriented data.**

After presenting a rough overview of the architecture's components, we will now present each of the three major tasks and their corresponding framework components in detail. We focus on aspects relevant for modeling and visually analyzing time and time-oriented data and give an overview of the corresponding design spaces. An overview of the complete design space is presented in Figure 3.

2.2 Modeling of Time and Time-Oriented Data

A useful concept for modeling data and information along cognitive principles is the *pyramid framework* (Mennis, Peuquet, and Qian 2000). The model is based on the three perspectives *where* (location), *when* (time), and *what*

(theme) on the level of data. Derived interpretations of these data aspects form *objects* on the cognitively higher level of knowledge, along with their classification (taxonomy; super-/subordinate relationships) and interrelationships (partonomy; part-whole relationships).

Depending on the phenomena under consideration and the analysis purpose, different points of view can be taken. For example, considering distinct conceptual entities that are related to time (objects in a discrete model) vs. the observation of a continuous phenomenon, like temperature over time (values in a continuous model). There cannot be a single model that is ideal of all kinds of applications. However, certain fundamental design alternatives can be identified which characterize time-oriented data models.

First of all, it is important to make a clear distinction between the physical dimension time and a model of time in information systems. Many different theories for characterizing the physical dimension time have been developed and discussed over literally thousands of years in philosophy, mathematics, physics, astronomy, or biology. The most influential theories for natural sciences are probably Newton's concepts of absolute vs. relative time, and Einstein's four-dimensional spacetime. What can be extracted as bottom line across many theories is that time is *unidirectional* (arrow of time) and that time gives *order* to events.

When modeling time in information systems, our goal is not to perfectly imitate the physical dimension time, but to provide a model that is best suited to reflect the phenomena under consideration and support the analysis tasks at hand. Frank presents a taxonomy of time (Frank 1998) which we will use in a slightly modified and simplified form in order to ensure practicability.

2.2.1 Modeling the Time Domain

In this section, we will describe the most important aspects of the design space for modeling time in information systems. The different aspects that have to be considered are (see (Müller and Schumann 2003) for details):

- *Scale: ordinal vs. discrete vs. continuous time*
- *Scope: time points vs. intervals*
- *Structure: linear vs. cyclic time*
- *Viewpoints: ordered time vs. branching time vs. time with multiple perspectives*

What these four aspects do not cover is the hierarchical nature of time. However, in the light of a visual analytics framework for time-oriented data, it is mandatory to let that fact find its way into the framework design. Therefore, we describe the fifth design criterion of *granularities*.

Granularities: none vs. single vs. multiple Basically, granularities can be thought of as (human-made) abstractions of time in order to make it easier to deal with

time in every-day life (like minutes, hours, days, weeks, months). More generally, granularities describe mappings from time values to larger or smaller conceptual units. For example, 60 consecutive seconds are mapped to one minute or five time steps in a discrete simulation model might relate to one second in physical time. Combining multiple granularities leads to hierarchical structures that are often referred to as "calendars" (see Figure 2 for an example of time granularities and their relationships). Specifically in simulation systems, the concept of granularity has to be considered when describing the relationship between simulation time and physical time. In this regard, the design space ranges from not considering granularities at all and the simple case of a single granularity to modeling a fixed or flexible set of multiple granularities.
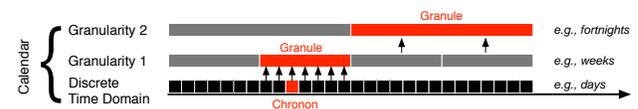

Figure 2: **Example of a discrete time domain with multiple granularities.** The smallest possible unit (chronon) is one *day*. Based on this, the granularities *weeks* and *fortnights* are defined.

After presenting the design space of modeling the time domain itself, we now move on to the question of modeling time-oriented data.

2.2.2 Modeling Time-Oriented Data

As it is the case with modeling time, the design choices for modeling time-oriented data are domain- and task-specific. The available modeling approaches are manifold and range from considering continuous to discrete data models. In the prior case, time is seen as observational space and data values that are given relative to it (e.g., time-series in form of <time, value> pairs). For the latter, data is modeled as objects or entities which have attributes that are related to time (e.g., calendar events with attributes "begin" and "end"). Moreover, certain analytic situations even demand domain transformations, as for example from the time domain into the frequency domain (Fourier Transformation).

On a general level, two decisive data aspects are the *number of variables* and the *frame of reference* (Aigner, Miksch, Müller, Schumann, and Tominski 2007b). In the first case, we differentiate between univariate and multivariate data which influences the choice of applicable visualization and analytical methods fundamentally. Compared to univariate data, where many methods have been developed, the range of applicable methods shrinks drastically in the case of multivariate data. What should be considered besides this,

is the question of the frame of reference, that is whether the data are spatial (e.g., georeferenced) or abstract.

At this point, we do not go into further detail of characterizing data themselves. Many useful modeling alternatives and reference models have been developed and can be adopted, as for example continuous models using scalars, vectors, or tensors, etc. (Wright 2007) or discrete models using structures like trees, graphs, etc. (Shneiderman 1996).

As already mentioned, we focus mainly on time-related aspects of data modeling that are particularly important in the areas of simulation and modeling and refer the interested reader to the large amount of data modeling literature available to date. In this regard, we will now discuss the two aspects of *relation to time*, and *kind of data*.

Relation to time: direct vs. indirect In the simplest case, data values are directly associated to elements of the time domain model. To be able to model more complex data to time relationships, so-called *temporal primitives* can be used that relate data to (non)continuous subsets of time. Examples for such temporal primitives are instants on varying granularities, explicitly modeled intervals (e.g., composed of two points in time or a time point and a duration), or even more complex elements modeling temporal indeterminacies where time aspects are known only to a certain degree (e.g., a one minute power outage happened on May 23rd, 2007, but it is not known when exactly). In principle, temporal primitives can be divided into anchored and unanchored ones. Anchored primitives represent elements whose absolute position in the time domain is known (e.g., September 14th, 2007) whereas unanchored temporal primitives are relative amounts of time (e.g., 3 days).

Kind of data: events vs. states The third design dimension refers to the question of whether events or states are dealt with. Events, on the one hand, can be seen as markers of state changes, like for example the departure of a plane. States, on the other hand, can be characterized as phases of continuity between events (e.g., plane is in the air). As one can see, states and events are two sides of the same medal. However, depending on the modeling and analytical requirements, the one or the other, or even a combination of both, might be better suited.

In the previous section, we presented the design space of time and time-oriented data. All further components of a visual analytics framework or system are grounded upon the chosen model of time and time-oriented data. But neither is it effective, nor possible to realize a general model that encompasses each discussed aspect in its full depth. For each specific case, the characteristics of the phenomena under consideration, analysis and visualization tasks and goals, as well as issues like memory usage or performance aspects impose constraints upon applicable data models. Depending on these constraints, a system designer can use the presented design dimensions as practical catalogue for

relevant design choices. An example for a specific time model designed for the purpose of representing complex temporal specifications including time indeterminacies is used in *PlanningLines* (Aigner, Miksch, Thurnher, and Biffl 2005). Here, planning activities are specified by a set of the six time attributes earliest start, latest start, earliest end, latest end, minimum duration, and maximum duration. Furthermore, relative time attributes (e.g., "task A starts 3 days after the end of task B") and incomplete specifications are covered. In the next section, we present a set of analytical tasks particularly geared towards time-oriented data analysis.

2.3 Computational Analysis of Time-Oriented Data

Data Mining is commonly defined as the application of algorithms to extract useful structures from large volumes of data (Fayyad, Grinstein, and Wierse 2001, Han and Kamber 2005). It is a multidisciplinary field integrating work from areas including statistics, machine learning, information retrieval, database technology, and neural networks. More specifically, data mining is mainly concerned with the tasks of clustering, classification, association-rule discovery, generalization, and prediction. To adhere to the special role of time in data mining, the research branch *Temporal Data Mining* has emerged. For the scope of this paper, we will follow the grouping of temporal data mining tasks as presented in (Laxman and Sastry 2006): *prediction*, *classification*, *clustering*, *search & retrieval*, and *pattern discovery*.

Prediction An important task in analyzing time-oriented data is the prediction of likely future behavior by inferring from data collected in the past and present.

Classification The goal of classification is to automatically determine which class or category a data set, sequence, or subsequence belongs to. Examples for this are speech recognition, gesture recognition, or genome sequence classification.

Clustering Clustering is concerned with grouping data sets, sequences, or subsequences into clusters based on their similarity. For example, in the analysis of financial data, clustering could be used to group stocks that exhibit similar behavior over time. In contrast to classification, where the classes are known a priori, clusters are not defined upfront.

Search & retrieval This task encompasses searching for a priori specified queries in possibly large volumes of data and is often also referred to as "query-by-example". Search & retrieval can be applied to locate exact matches for an example query or approximate matches. In the latter case, similarity measures are needed that define the degree of exactness or fuzziness of the search (e.g., to find customers whose spending pattern over time are similar but not necessarily equal to a given spending profile).

Pattern discovery In contrast to the previously presented task of search & retrieval, pattern discovery is concerned with *automatically* discovering interesting patterns

(without the need for predefining queries). Laxman and Sartry define the term *pattern* as a local structure in the data (Laxman and Sastry 2006). Often, frequently occurring patterns are of interest, like for example to analyze whether a TV commercial generally leads to an increase in sales. In contrast to that, it is often also of interest to discover patterns that occur very rarely because these occurrences can reveal hostile behavior or failures (e.g., mining for patterns of network security breaches).

Beyond the mentioned temporal data mining tasks, other analytical methods, as for example statistical aggregation operators, are very useful in the context of time-oriented data. Examples for this are sum, average, min, max, mean, or median that are applied along time granularities (e.g., weekly or monthly mean).

Again, the design and applicability of analytical & mining components in a visual analytics framework or system are application- and task-dependent. An example for such a component is *temporal data abstraction* which can be considered as classification task. Here, qualitative values or patterns are derived from current and past situations, as for example the quantitative variable of body temperature can be abstracted to the qualitative value "moderate fever during nighttime" if $37 <$ body temperature < 38.5 between 8pm and 6am. *VIE-VENT* (Miksch, Horn, Popow, and Paky 1996) is an example for a temporal data abstraction system that goes far beyond this and is applied in the medical domain to compute context-sensitive and expectation-guided abstractions of ICU measurements.

Note, that we do not aim to provide an exhaustive and detailed listing of possible analytics or mining tasks, but give a brief overview of major task groups in the analysis of time-oriented data to aid system designers in assessing analytical methods needed for the tasks and goals of users. Apart from the presented techniques, other, mainly statistical methods to analyze time-oriented data exist, as for example extensions of principal component analysis (PCA) (Aigner, Miksch, Müller, Schumann, and Tominski 2007a). Next, we will discuss the third part of our conceptual framework that is concerned with interactive visualization.

2.4 Interactive Visualization of Time-Oriented Data

When building a visual analytics system for time-oriented data, we have to deal with the aspects of visualizing and communicating *time*, *data related to time*, *analysis and visualization parameters*, as well as the *analysis process*. While the first two tasks are quite obvious, visualization and analysis parameters together with the analysis process also need to be regarded appropriately in systems that aim to fully integrate interactive visualizations and analytical methods. When we are using the term *visualization* in this section, we always refer to *interactive visualization* that facilitates active exploration and control by the user.

2.4.1 Visualizing Time and Time-Oriented Data

Visualizing Time   First, the time domain itself must be represented visually in order to create a reference system with respect to which time-oriented data is visualized. To make the abstract dimension time visible, some kind of mapping has to be applied. For visual representation, two fundamental ways of mapping exist:

1. Time → Space (e.g., 2D time-series plot)
2. Time → Time (e.g., animation)

When space is used to convey time, time-oriented data is represented in static images that do not change automatically over time. The most common choice is using position and length on a common time scale that represents a certain portion of the time domain. However, not only straight lines (time axes) can be used, but also layouts in rows, columns, or both (like the typical monthly calendar layout), circular or spiral layouts, and other kinds of patterns as for example in pixel-oriented visualizations (Keim 2000). Besides these explicit positional layouts, time can also be represented using other visual variables, as for example angle, slope, line thickness, transparency, color, connection, or textual labels. In contrast to that, representations that follow the second mapping of *time→time*, use the physical dimension time to convey the time dependency of the data (representations that change automatically over time such as slide shows or animations). Distinguishing between these mappings is crucial for Visual Analytics, because different tasks and goals are supported. Dynamic representations are well suited to convey the dynamics of the general development of the analyzed data over time. Static representations show information of different points of time on one screen, which is advantageous to fully concentrate on the data and to compare different parts of the time domain directly. Both kinds of mappings may provide means for interactive manipulation as for example widgets to control animation speed or to change cycle lengths in a spiral representation. Note that there might also exist hybrid forms which combine both, mappings of time→time and time→space.

Visualizing Data Related to Time   Second, besides defining an appropriate representation of time itself, visualizing the data that are related to time is a key concern. Here, we have taken into account the characteristics of the data, as referred to in Section 2.2.2. An example for this is the distinction of qualitative and quantitative data that is particularly interesting when designing and selecting visualization methods. *Qualitative data* refers to data typically gathered on nominal scales where no ordering or arithmetic can be applied. In this case, often layout algorithms are applied to position elements corresponding to some defined criteria, for example to avoid overlaps. A well-known example for this kind of data are project plans that are visualized via

Gantt charts where individual project tasks are represented as bars along time. *Quantitative data* values can be directly mapped to visual variables such as position (e.g., time-series plot), length (e.g., bar plot), and orientation (e.g., glyphs).

A design choice for visualizations that carries a lot of implications upon implementation is whether 2D or 3D presentation spaces are used (the *dimensionality* of the representation). Again, whether or not it makes sense to exploit three dimensions for visualization depends heavily upon the data to be represented and the analysis tasks to be performed.

Interaction and the Role of the User   Interaction methods and navigation are essential to explore the data as well as the parameter space. Moreover, it is important that these methods are designed according to user's demands. Interacting directly with the visual representation and the analytical & mining methods provides more control and tighter feedback for the human analyst. This must also include interactive parameterization of both, visual and analytical methods as discussed in the next section.

Navigation methods for large information spaces are decisive for analysis environments that support exploration. In parallel, they should allow for visual overviews as well as the ability to drill down into areas of interest while preserving orientation within the information space. A helpful principle concerning user interaction tasks is Shneiderman's "visual information seeking mantra" *overview first, zoom and filter, then details-on-demand* (Shneiderman 1996).

Moreover, user's tasks and goals determine the adequate choice of visualization methods. For example, if we want to identify cycles in the data, suitable representations that reinforce the visual detection of periodic behavior need to be chosen. The SpiralGraph technique (Weber, Alexa, and Müller 2001), for instance, maps time-oriented data onto a spiral and uses interaction and animation to detect previously unknown cycles in the data.

2.4.2 Visualizing Parameters and the Analysis Process

As discussed and shown in (Aigner, Miksch, Müller, Schumann, and Tominski 2007a), appropriate *parameterization* is needed in order to ensure expressive and effective visualizations. Similarly, parameters of analytical methods have an immense influence on the generation of meaningful outcome. Often, detailed knowledge about underlying algorithms is necessary in order to be able to tune these parameters properly. Obviously, this limits their usability and poses a barrier for broader usage by non-experts (i.e., people that are experts in their domain, as for example criminal investigators, but are not experts in statistics or data mining methods). Similarly, simulation system behavior is determined by a set of input parameters that need to be configured appropriately. Suitable visual communication and interaction methods regarding the management of these parameters can enhance the usability of such systems. Be-

yond that, proper communication of the functionality of the simulation or *analysis process* as a whole, beyond regarding the system as black box, can greatly improve the knowledge gain and understanding for the user.

A number of practical examples of visualization and interaction methods for time- and time-oriented data can be found in (Aigner, Miksch, Müller, Schumann, and Tominski 2007b, Müller and Schumann 2003, Silva and Catarci 2002). As it is the case with the described data model and analytics components, it is neither feasible nor desirable to include every mentioned aspect in a visualization & interaction component. The application designer has to assess the domain- and task-specific constraints adhering to the problem under consideration and create or select appropriate methods. An example for specific visualization & interaction components are *VisAxes* (Tominski, Abello, and Schumann 2004). VisAxes comprises implementations of reusable visualization & interaction components for time-oriented, axes-based visualizations. They provide means to navigate through different time granularities (hierarchical axis), select time intervals to display (scroll axis), or apply Focus+Context views (focus within context axis).

2.5 Conceptual Visual Analytics Framework

To bring together the framework's basic structure and the discussed design aspects, Figure 3 summarizes our conceptual visual analytics framework. With this conceptual framework at hand, two goals have been achieved: First, existing frameworks and applications can be characterized systematically and second, the design and implementation of specific components and frameworks is guided. Now, we present a practical example by characterizing an existing visual analytics system to illustrate the individual design choices we have put forward.

3   EXAMPLE

To demonstrate the applicability of our conceptual framework, we will systematically analyze the application *TimeSearcher* (Buono, Aris, Plaisant, Khella, and Shneiderman 2005). We have selected this application because it covers all design dimensions (interactive visual as well as analytical aspects), is well described, and easily accessible[1]. TimeSearcher is a visualization and exploration tool for multiple time-series (see Figure 4). Its goal is to identify and find familiar patterns in the investigated data. Three application scenarios, (1) stock price data, (2) gene expression levels in DNA microarray experiments, and (3) online auction data are discussed in related publications. For brevity, we will focus on the first application scenario of stock price data

---

[1]See <www.cs.umd.edu/hcil/timesearcher/> for more information and a downloadable prototype. Accessed at June 5, 2007.

internal data structures

Modeling the time domain

Scale	ordinal	discrete	continuous
Scope	point-based	interval-based	
Structure	linear	cyclic	
Viewpoints	ordered	branching	multiple perspectives
Granularities	none	single	multiple

Modeling time-oriented data

Number of variables	univariate	multivariate
Frame of reference	spatial	abstract
Relation to time	direct	indirect
Kind of data	events	states

visualization & interaction components

Mapping	space	time
Type of data	qualitative (coconut, banana, apple)	quantitative (3.14, 3.27, 4.88)
Dimensionality	2D	3D
Analysis & visualization parameters	yes/no	
Analysis process	yes/no	

analytical & mining components

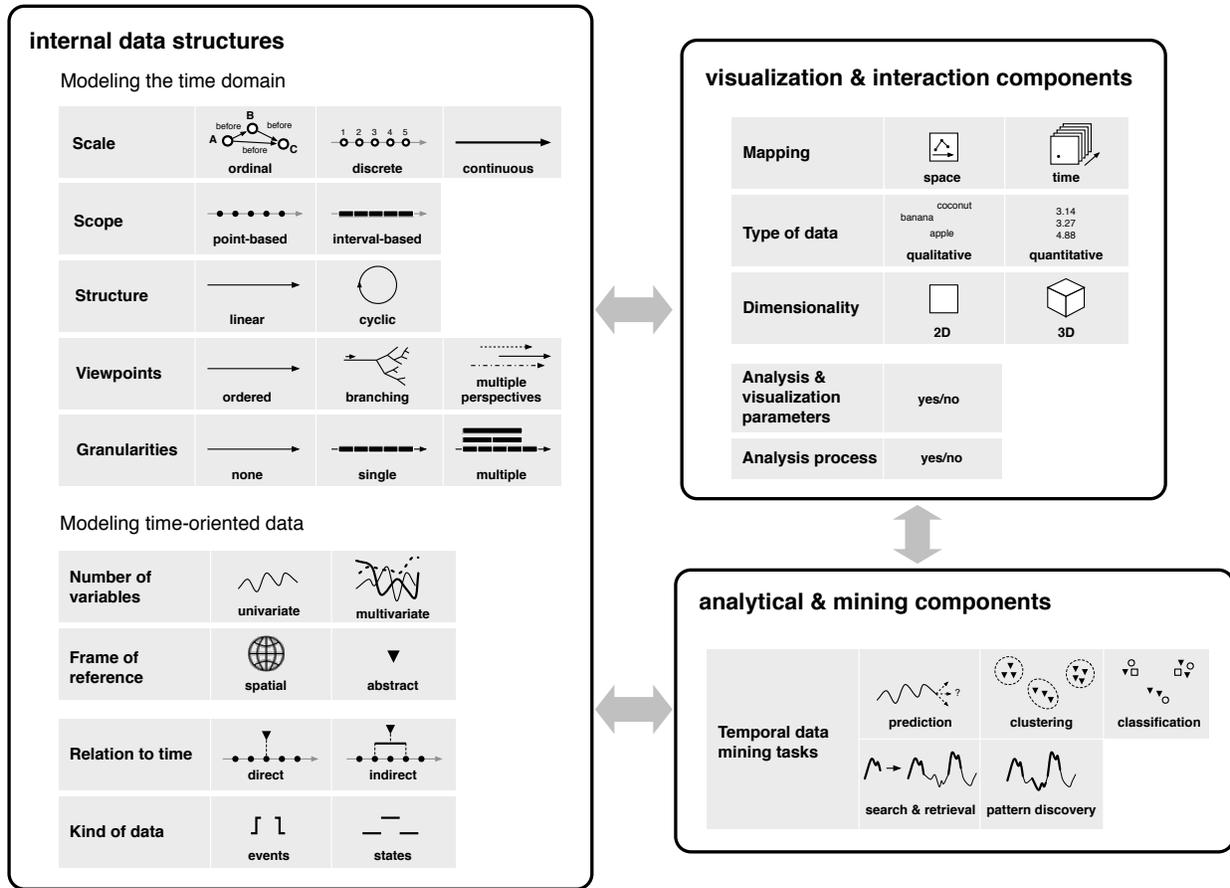|			
Temporal data mining tasks	prediction	clustering	classification
	search & retrieval	pattern discovery	

Figure 3: **Design space of the conceptual visual analytics framework for time-oriented data.**

in our discussion. In the following, we will present how TimeSearcher instantiates our conceptual framework.

Internal Data Structures: How do the models of the time-domain and time-oriented data look like?

Scale: discrete   Measurements are stored in a regular temporal manner in form of <time-point, value> pairs.

Scope: point-based   Each value corresponds only to a certain moment in time (e.g., closing price of a stock at a particular day) and nothing is known about the time in between measurement time-points.

Structure: linear   Time is structured linearly from the past to the present.

Viewpoints: ordered   Only one temporal perspective is used (valid time).

Granularities: multiple   Values are measured in definable steps of days, weeks, months, or years that are also used in the visual representation.

Number of variables: multivariate   Multiple stocks and multiple attributes per stock (e.g., opening and closing prices) can be dealt with.

Frame of reference: abstract   The modeled and visualized data have no spatial attributes.

Relation to time: direct   Measurements are directly related to values of the time domain (each measurement belongs to a single value of the time domain).

Kind of data: state   The state of the observed variable is measured at discrete steps in time.

Analytical & Mining Components: What kind of analytical methods are applied?

Temporal data mining tasks: search & retrieval Basically, two options for data querying are offered by TimeSearcher. First, the "TimeBox query" allows for the specification of a rectangular query region that defines both, a time period and value range of interest. All time-series that comply to this query are shown whereas all others are filtered out. Moreover, multiple timeboxes can be combined to refine the query further. Second, using a "SearchBox query" the user can highlight an interesting part (pattern) of a selected time-series which is then searched for in all other time-series (see Figure 4). In contrast to the first
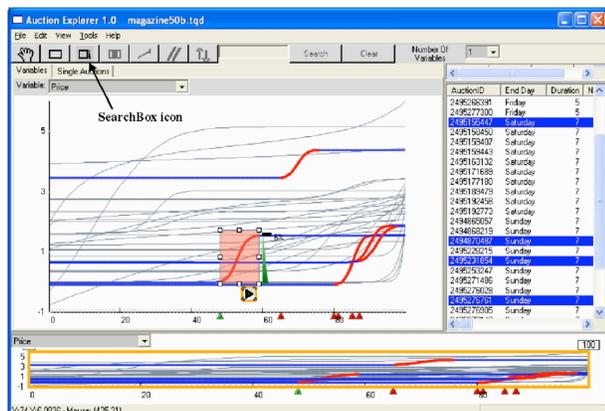
Figure 4: **TimeSearcher** (Buono, Aris, Plaisant, Khella, and Shneiderman 2005). Visualization and exploration tool for multiple time-series. The above screenshot shows a query-by-example of online auction bid data over time. A query pattern is defined by the shaded red box and the found instances of this pattern are highlighted in red.

query method, no restrictions are imposed on the position and extent of the pattern in time but it is searched across the full time and value range of each time-series. Moreover, the tolerance of the pattern search can be defined by the user and allows for varying degrees of matching exactness.

Visualization & interaction components: How are time and data visualized?

Mapping: space   TimeSearcher uses a mapping of time to space and represents data via time-series plots where horizontal, linear time axes are used.

Type of data: quantitative   Numerical values (float) are dealt with in the application case of stock price data.

Dimensionality: 2D   The visual representation uses a two-dimensional representation space.

Visualizing and communicating analysis parameters and the analysis process   Analysis parameters for both presented query methods (TimeBox and SearchBox) are set via direct manipulation by the user. This also includes the tolerance settings for matching exactness.

In the previous section, we have illustrated the applicability of our conceptual framework via analyzing an example system. Similarly, other visual analytics systems can be characterized along these design dimensions. Furthermore, application designers can use our conceptual framework as guide when developing application-specific modeling and simulation systems.

4   CONCLUSION AND FUTURE WORK

It is widely acknowledged that time is a unique data dimension with distinct characteristics. Many visualization,

analysis, and simulation systems deal with time-related aspects. However, these systems could be improved by considering the specifics of time and their implications in a broader sense. On top of that, an improved integration of interactive visualization and computational methods would lead to more powerful and more usable methods.

We introduced a conceptual visual analytics framework to give practical hands-on the most important issues when designing visual analytics frameworks and systems. Our main aim is to support the special characteristics of time and time-oriented data. By developing our conceptual framework, two goals have been achieved: First, to enable a systematic characterization of existing frameworks and applications, and second, to guide the design and implementation of specific components and frameworks. Note that our conceptual framework outlines a broad design space and mostly only subsets of the entire spectrum are relevant for certain application-areas, data, and tasks.

Next steps to be done are the implementation of the described components according to domain and task-characteristics and their integration into simulation systems. Each of the presented parts offers various opportunities for further work and especially the fruitful and seamless integration of all components are big challenges. More powerful and more usable data exploration and knowledge discovery systems are the goal of these efforts.

As a matter of fact, powerful, human-centered knowledge discovery environments, obtained from the integration of the best of visual methods and analytical methods seem to be not only challenging, but also a very promising approach.

REFERENCES

Aigner, W., S. Miksch, W. Müller, H. Schumann, and C. Tominski. 2007a. Visual Methods for Analyzing Time-Oriented Data. *IEEE Trans. on Visualization and Computer Graphics*. to appear.

Aigner, W., S. Miksch, W. Müller, H. Schumann, and C. Tominski. 2007b. Visualizing Time-Oriented Data - A Systematic View. *Computers & Graphics* 31 (3): 401–409.

Aigner, W., S. Miksch, B. Thurnher, and S. Biffl. 2005. PlanningLines: Novel Glyphs for Representing Temporal Uncertainties and their Evaluation. In *Proc. of the 9th Intl. Conf. on Information Visualisation (IV05)*: IEEE Press.

Buono, P., A. Aris, C. Plaisant, A. Khella, and B. Shneiderman. 2005. Interactive Pattern Search in Time Series. In *Proc. of the Conf. on Visualization and Data Analysis (VDA 2005)*, 175–186: SPIE.

Card, S., J. Mackinlay, and B. Shneiderman. 1998. *Readings in Information Visualization: Using Vision to Think*. San Francisco, USA: Morgan Kaufmann Publishers Inc.

Chi, E. H. 2000. A Taxonomy of Visualization Techniques Using the Data State Reference Model. In *Proc. of the IEEE Symp. on Information Vizualization 2000 (Info-Vis'00)*, 69–76: IEEE Computer Society.

dos Santos, S., and K. Brodlie. 2004. Gaining Understanding of Multivariate and Multidimensional Data through Visualization. *Computers & Graphics* 28:311–325.

Fayyad, U., G. G. Grinstein, and A. Wierse. 2001. *Information Visualization in Data Mining and Knowledge Discovery*. San Francisco, USA: Morgan Kaufmann Publishers Inc.

Frank, A. U. 1998. *Spatial and Temporal Reasoning in Geographic Information Systems*, Chapter Different Types of "Times" in GIS, 40–62. New York, USA: Oxford University Press.

Han, J., and M. Kamber. 2005. *Data Mining: Concepts and Techniques*. San Francisco, USA: Morgan Kaufmann Publishers Inc.

Keim, D. A. 2000. Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Trans. on Visualization and Computer Graphics* 06 (1): 59–78.

Laxman, S., and P. Sastry. 2006. A Survey of Temporal Data Mining. *Sadhana* 31:173–198.

Mennis, J. L., D. Peuquet, and L. Qian. 2000. A conceptual framework for incorporating cognitive principles into geographical database representation. *Intl. Journal of Geographical Information Science* 14 (6): 501–520.

Miksch, S., W. Horn, C. Popow, and F. Paky. 1996. Utilizing Temporal Data Abstraction for Data Validation and Therapy Planning for Artificially Ventilated Newborn Infants. *AI in Medicine* 8 (6): 543–576.

Müller, W., and H. Schumann. 2003. Visualization Methods for Time-dependent Data - an Overview. In *Proc. of Winter Simulation 2003*, ed. S. Chick, P. Sanchez, D. Ferrin, and D. Morrice. New Orleans, USA.

Robinson, S. 2004. *Simulation: The Practice of Model Development and Use*. West Sussex, England: John Wiley & Sons, Ltd.

Shneiderman, B. 1996, September 3–6. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proc. of the IEEE Symp. on Visual Languages*, 336–343: IEEE CS Press.

Silva, S. F., and T. Catarci. 2002. Visualization of Linear Time-Oriented Data: a Survey. *Journal of Applied System Studies* 3 (2): 454–478.

Thomas, J. J., and K. A. Cook. 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Science.

Tominski, C., J. Abello, and H. Schumann. 2004. Axes-Based Visualizations with Radial Layouts. In *Proc. of the ACM Symp. on Applied Computing (SAC04)*, 1242–1247: ACM.

Weber, M., M. Alexa, and W. Müller. 2001, October. Visualizing Time-Series on Spirals. In *Proc. of the IEEE Symp. on Information Visualization 2001 (InfoVis01)*, 7–14.

Wright, H. 2007. *Introduction to Scientific Visualization*. Springer.

AUTHOR BIOGRAPHY

WOLFGANG AIGNER is scientific researcher at the Department of Information and Knowledge Engineering, Danube University Krems, Austria and lecturer at Vienna University of Technology. He received his PhD from Vienna University of Technology and his main research interests include Visual Analytics, Information Visualization, Human-Computer Interaction, and User Centered Design. He can be contacted at <wolfgang.aigner@donau–uni.ac.at>.

ALESSIO BERTONE is scientific researcher at the Department of Information and Knowledge Engineering, Danube University Krems, Austria. He is PhD candidate at the Vienna University of Technology and his main research interests include Visual Analytics, Information Visualization, Temporal Data Mining, Multirelational Data Mining and Semantic Web.

SILVIA MIKSCH is University Professor (Univ.-Prof.) and head of the Department of Information and Knowledge Engineering at Danube University Krems, Austria. Since 1998 she is head of the Information and Knowledge Engineering research group, Institute of Software Technology and Interactive Systems, Vienna University of Technology. Her main research interests include Information Visualization, Visual Analytics, Plan Management, and evaluation of Knowledge-Based Systems in real-world environments (health care).

CHRISTIAN TOMINSKI received diploma (MSc) and doctoral degree (Dr.-Ing.) from the University of Rostock in 2002 and 2006 respectively. Christian is working in a postdoctoral position at the same university. His main interests concern visualization of multivariate data in time and space, visualization of graph structures, and visualization on mobile devices.

HEIDRUN SCHUMANN is full professor at the University of Rostock. Since 1992 she is heading the Chair of Computer Graphics at the Institute for Computer Science. Her research profile covers aspects of information visualization and visual data mining, mobile interfaces, rendering, and image presentation. She was the head of the cross-institutional research project "MoVi - Visualization of Multimedia Information on Mobile Computer Systems", supported by the German Research Foundation (DFG).