Visualizzazione dell'Informazione

Interrogazioni ed Esplorazioni Visuali

Carlo Combi
Dipartimento di Informatica
Università degli Studi di Verona

(in collaborazione con Luca Chittaro e Giampaolo Trapasso,
Università di Udine)

Rappresentazione di intervalli temporali e di relazioni

Motivations

- A crucial component for turning any temporal reasoning system into a real-world application that can be adopted by a wide base of users is given by its **user interface**.
- A proper consideration of human-computer interaction (HCI) aspects is needed both in:
  – **presenting temporal data** to users (ensuring that it is easy and quick to understand and does not lead to ambiguous interpretations), and
  – **accepting temporal specifications** from users (ensuring an easy formulation of queries or constraints, and minimizing the possibility of errors).

Information Visualization

- MAPPING: How do we visually encode information ?
- PRESENTATION: How do we lay out the visualization on the available screen space ?
- INTERACTIVITY: What tools do we provide to explore and rearrange the visualization ?
- HUMAN FACTORS: Are we taking into account human perception capabilities? Are we taking into account what mental models our users easily develop?
- EVALUATION: How do we test that the visualization is really effective with users on the considered task ?

Taxonomy of IV Data Types

(Shneiderman, 1996)

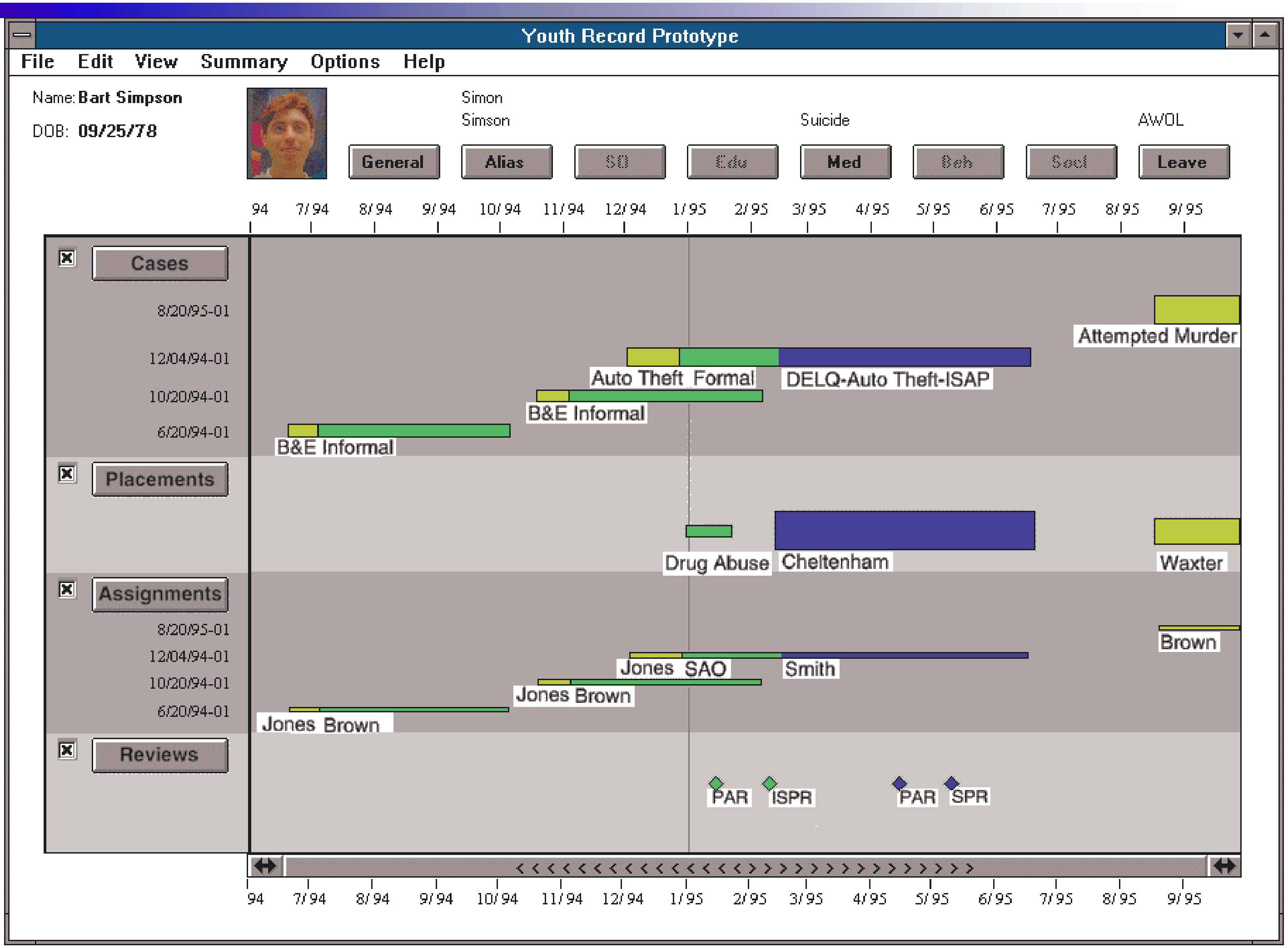
- *1-dimensional* (linear data organized in a sequential manner),
- *2-dimensional* (planar or map data covering some part of an area),
- *3-dimensional* (data with volume and potentially complex relations),

- *temporal* ("data with a start time, finish time, and possible overlaps on a timescale, such as that found in medical records, project management, video editing,…"),

- *multi-dimensional* (data with *n* attributes which becomes points in a *n*-dimensional space),
- *tree* (collections of items linked hierarchically by a tree structure),
- *network* (collections of items linked by a graph structure).

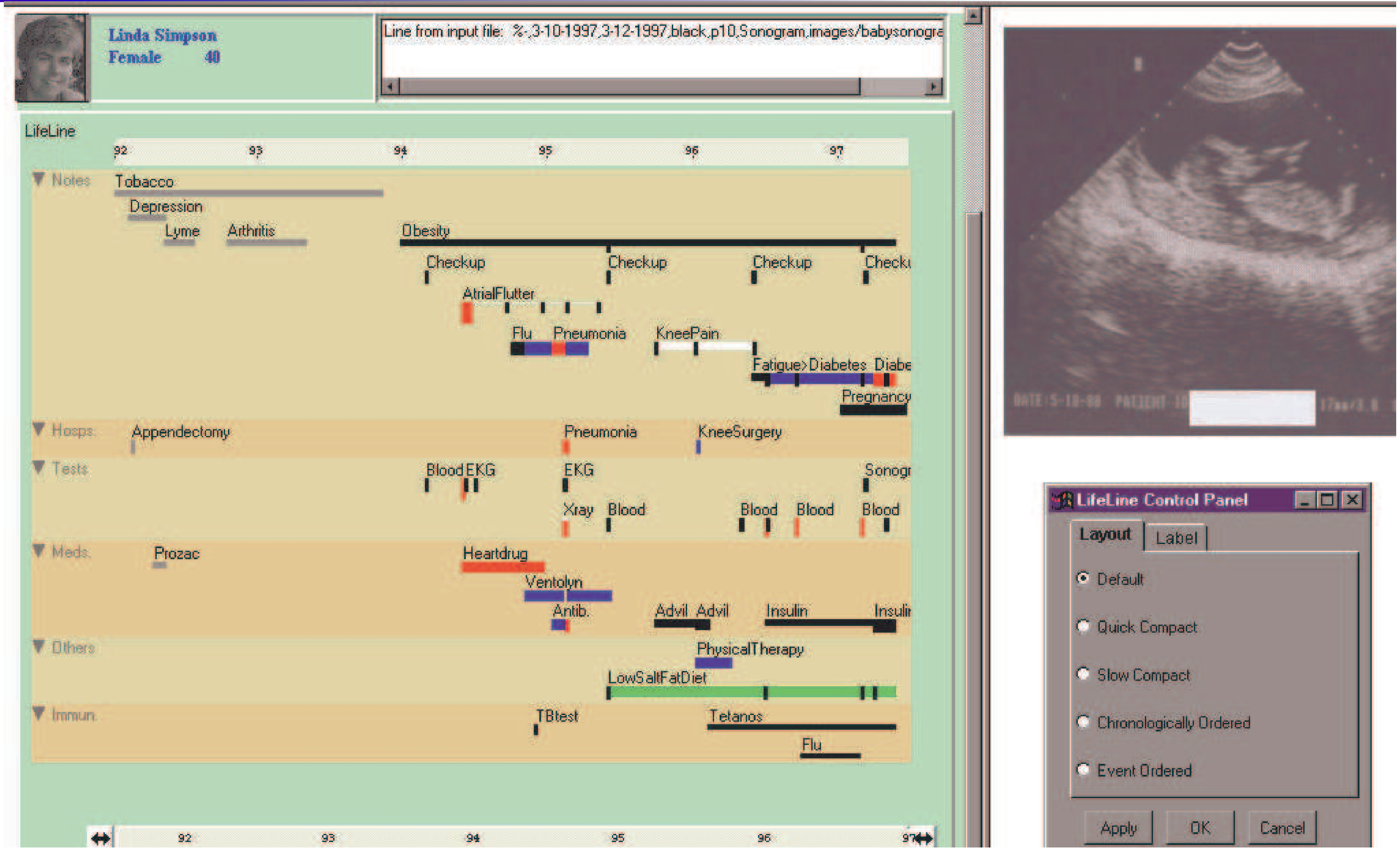
LifeLines

(Plaisant et al., 1996)



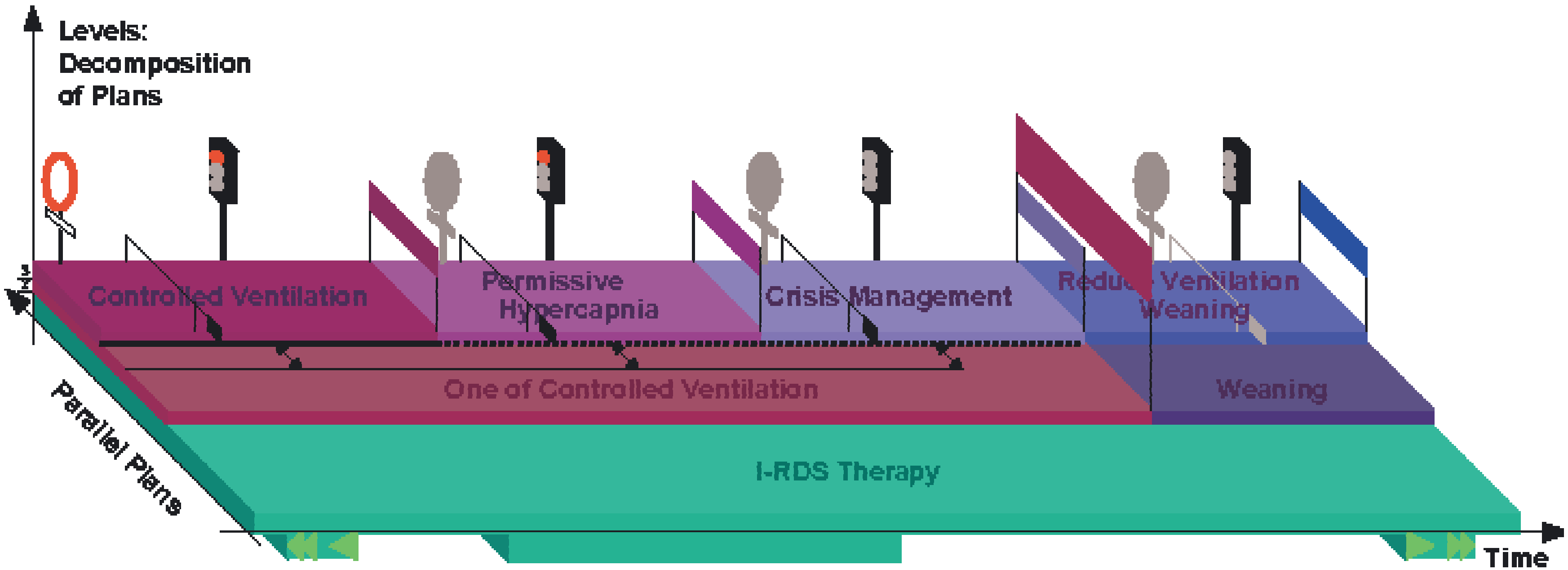
LifeLines: Visualizing Patient Records

(Plaisant et al., 1998)



AsbruView

(Kosara and Miksch, 2000)



Open Issues

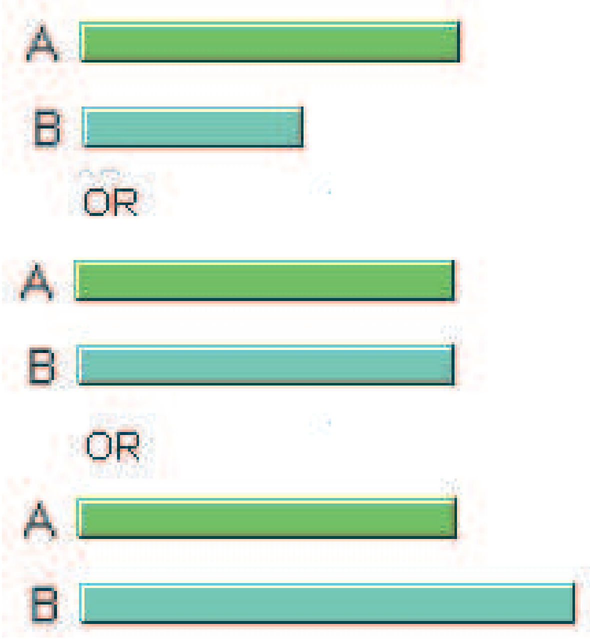
- Designing an effective visual vocabulary for temporal data is not trivial:
  - no disciplined design methodologies and engineering principles for IV have been yet identified,
  - temporal relations are abstract information with no natural and obvious physical representation.
- Lack of user studies and evaluations which follow the rigorous techniques of HCI research.
- Commonly accepted visual representation for intervals and timelines, but much work to be done for the representation of **temporal relations** and complex **temporal patterns**.

Temporal Patterns: Example 1

- ***Case 1***. Intervals A and B start at the same time, but it is irrelevant when they finish, relative to one another.

In terms of Allen's relations: (B *starts* A) $\vee$ (A *starts* B) $\vee$ (A *equal* B)

Using the typical interval visualization:



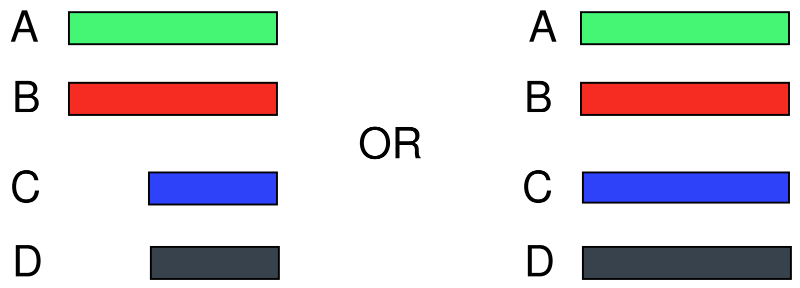
Temporal Patterns: Example 2

- ***Case 2***. Interval A is equal to interval B; interval C is equal to interval D; the four intervals finish at the same time; A and B start before or together with C and D.

A possible representation in terms of Allen's relations:

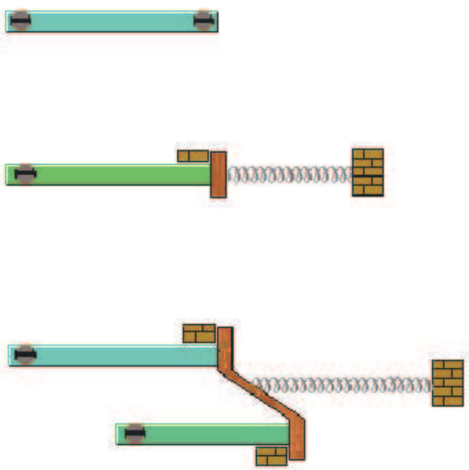
(A *equal* B) $\wedge$ (C *equal* D) $\wedge$ ((C *finishes* B) $\vee$ (A *equal* D))

Using the typical interval visualization:

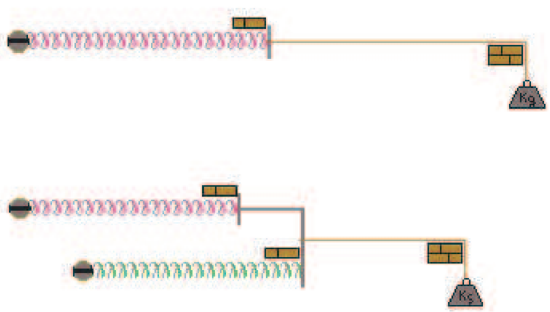


Interval Visualization

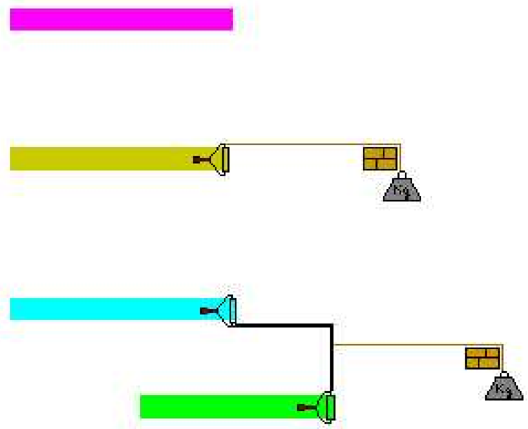
Elastic Bands



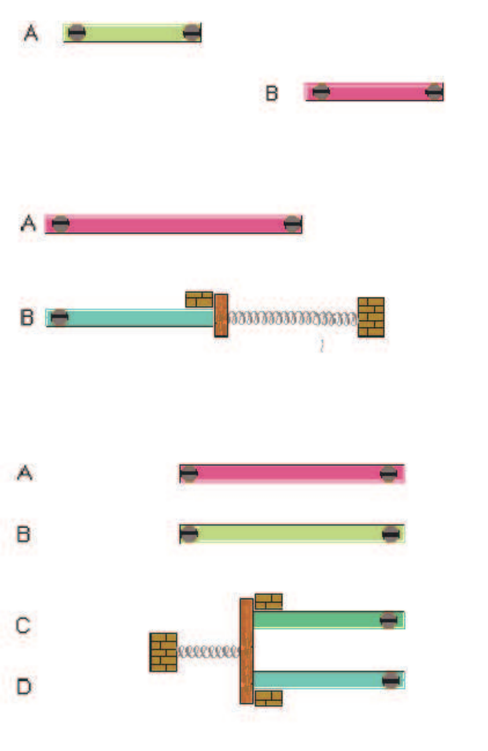
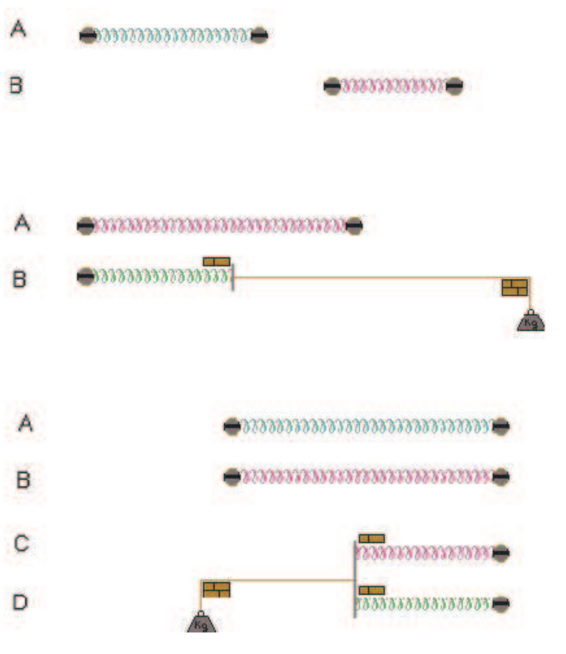
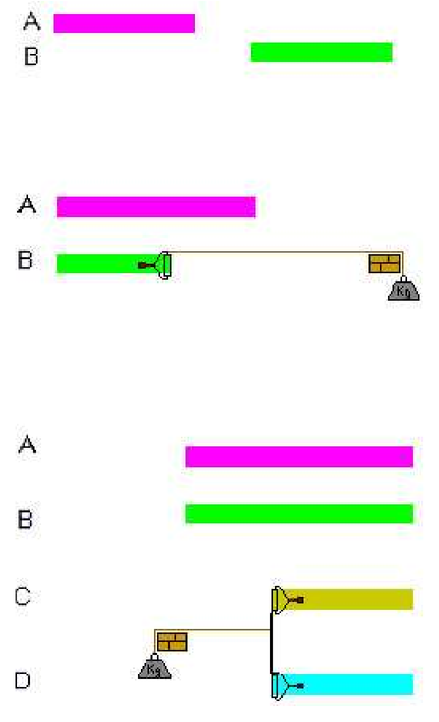
Springs



Paint Strips



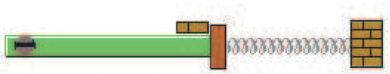
Examples of Interval relations

Elastic Bands	Springs	Paint Strips
		

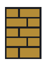
Evaluation: First Experiment

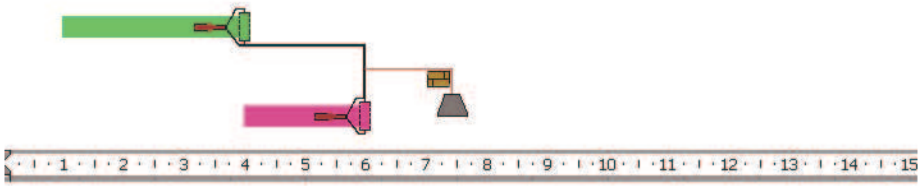
- **Goal:** determining if a visualization is more frequently perceived and understood correctly than the others.
- **Technique:** questionnaire with 4 exercises, for each of the 3 proposals.
- **Subjects:** 30 (13 F and 17 M). Age from 24 to 37, (average: 27), 9 physicians, 20 university students from different fields, 1 subject held a secretarial position.
- **Design:** within-subjects experiment.
- **Counterbalancing:** order of presentation of the 3 proposals.

Questionnaire: sample exercises



Which objects can move among those depicted in the picture above ?

❑	
❑	
❑	
❑	
❑	
❑	Other ______________________________



	Smallest extension		Largest extension	
	From point	To point	From point	To point
Green strip				
Purple strip				

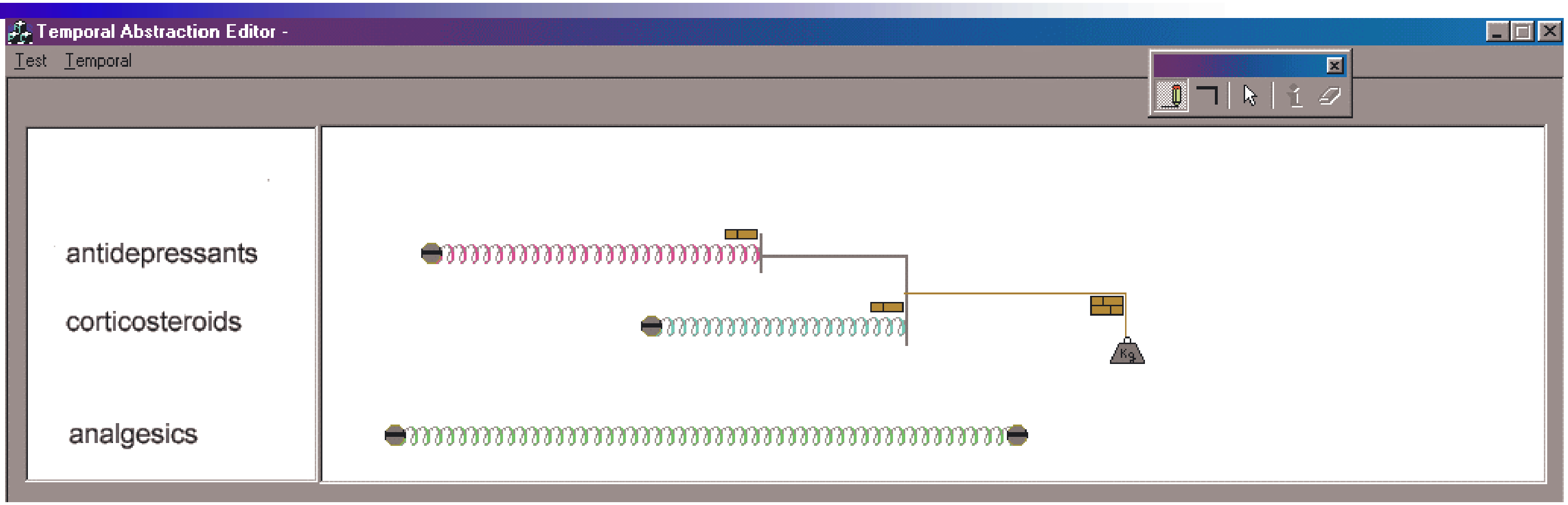
First Experiment: Results

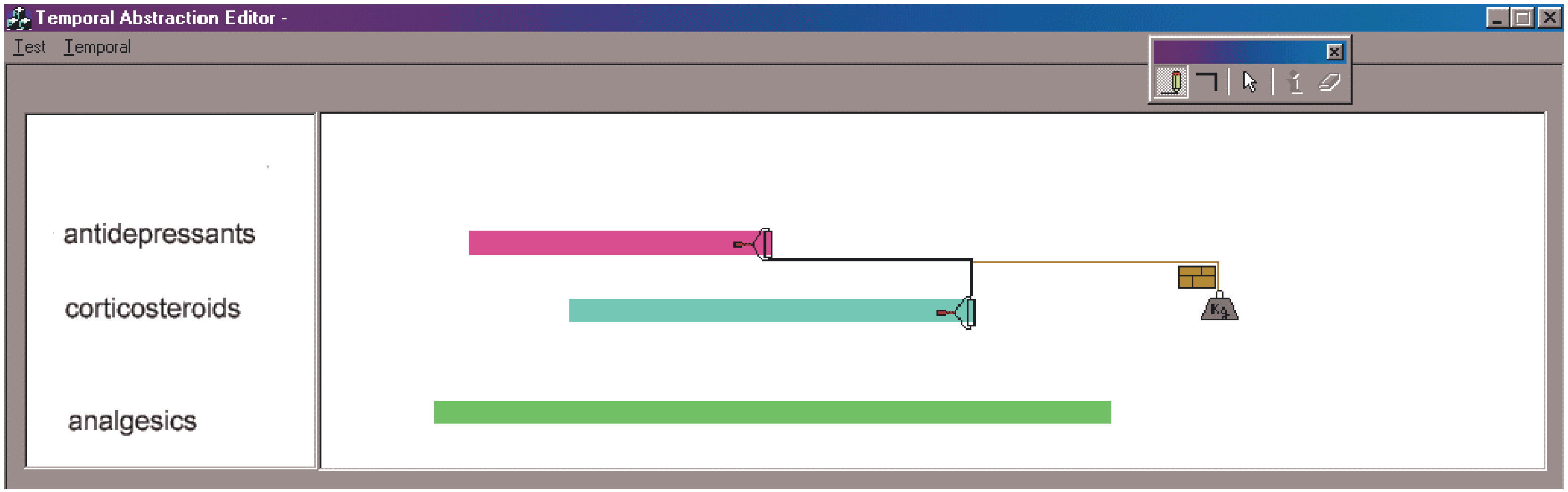
- **Independent Variable:** type of visualization.
- **Dependent Variable:** number of correctly solved exercises.
- **Means:** Paint Strips condition = **3,** Springs Condition = **2.8,** Elastic Bands Condition = **2.2**

Evaluation: Second Experiment

- **Goal:** thoroughly evaluating the two proposals which gave better results, also using fully implemented interfaces.
- **Task:** 4 *interpretation* exercises (on paper), 4 *definition* exercises (using the interface), for each proposal.
- **Subjects:** Experiment performed in a medical clinic with 31 subjects (16 F and 15 M). Age from 23 to 44 (average: 30), 23 physicians, 6 university students in Medicine, 1 psychologist, 1 pharmacologist.

User Interfaces



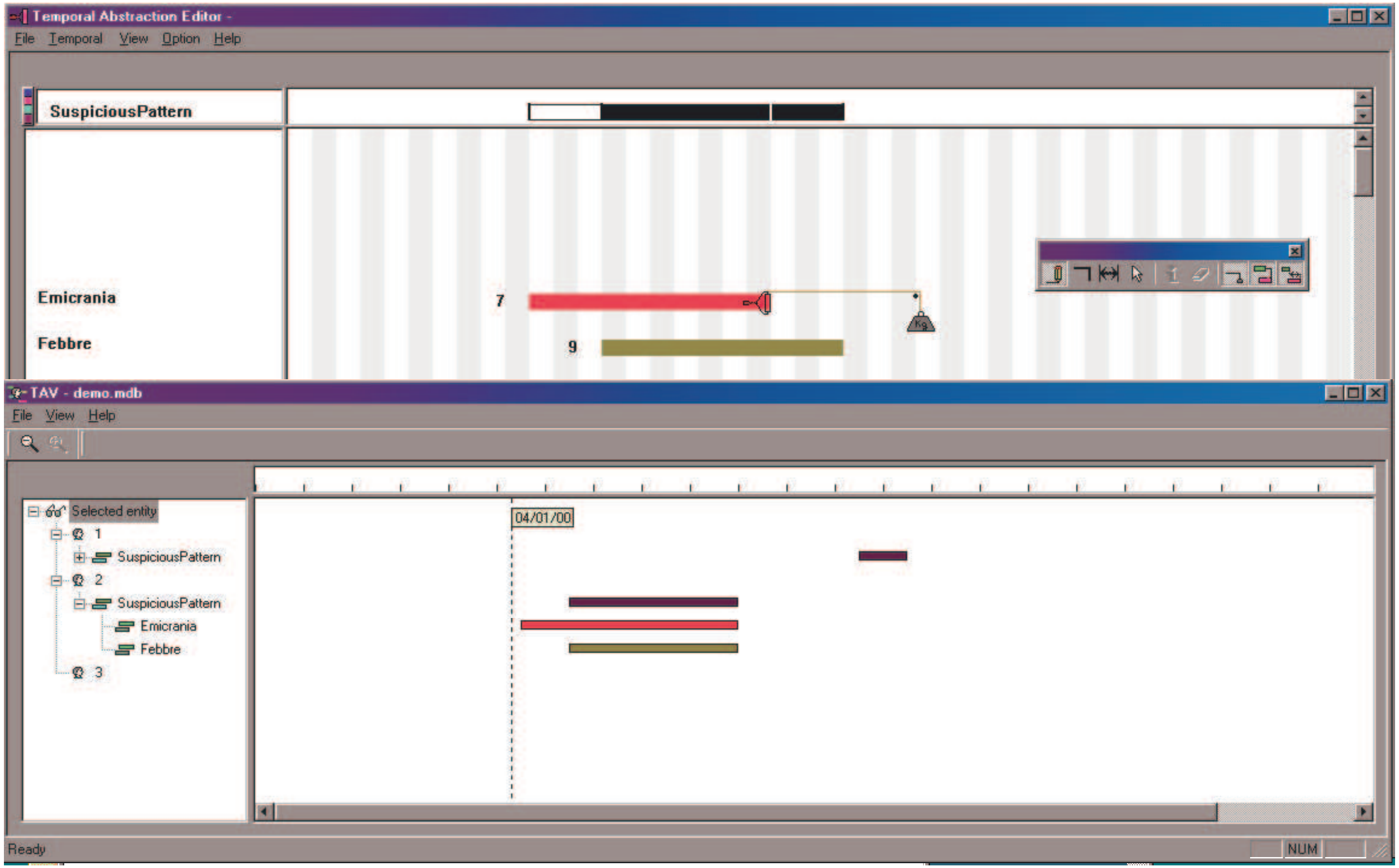


Second Experiment: Results

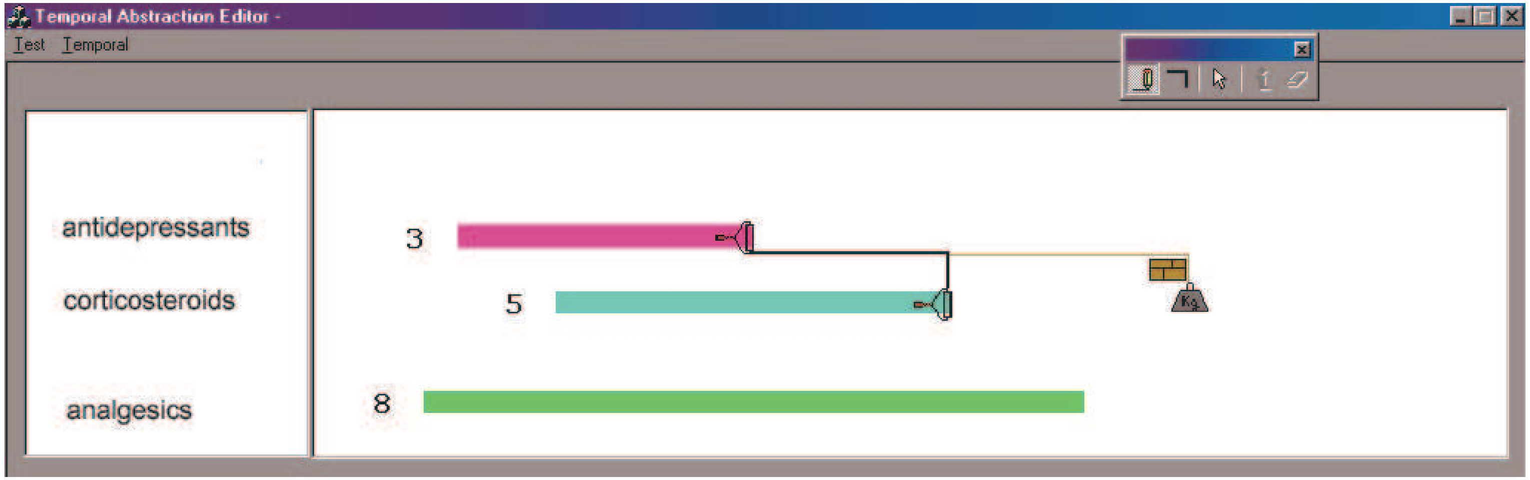
- **CORRECT ANSWERS to EXERCISES:**
  - **Interpretation task:** Paint Strips = **3.52,** Springs = **3.39**
  - **Definition task:** Paint Strips = **2.81,** Springs = **2.55**
- **TIME SPENT:**
  - **Interpretation task:** Paint Strips = **160,** Springs = **189**
  - **Definition task:** Paint Strips = **567,** Springs = **595**
- **QUALITATIVE IMPRESSIONS:**
  - The most important result came from the final question which asked to choose the best of the two interfaces:

	Subjects	Percentage
Paint Strips	22	71%
Springs	9	29%

Querying clinical data

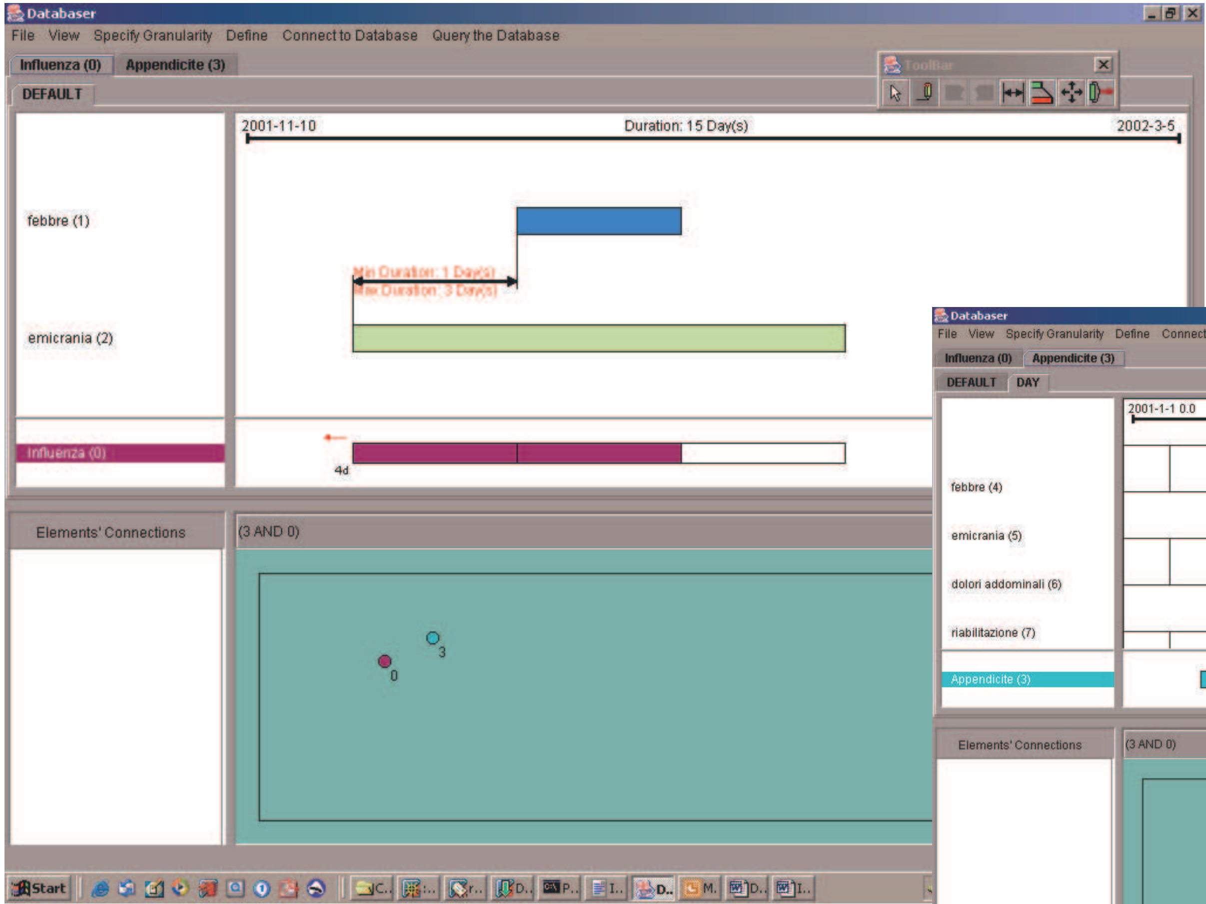


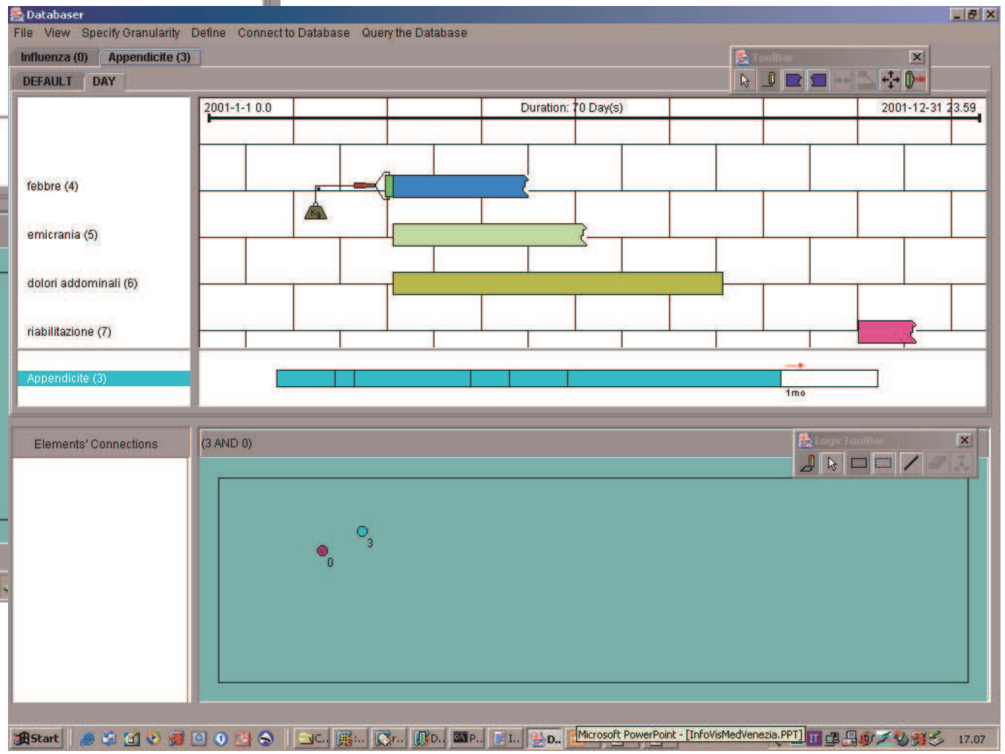
From visual queries to SQL queries




```
SELECT HistoryID
FROM HistoryDB h3, HistoryDB h5, HistoryDB h8
WHERE h3.HistoryID = h5.HistoryID AND h3.HistoryID = h8.HistoryID AND
    h3.FactDescr = "antidepressants" AND
    h5.FactDescr = "corticosteroids" AND
    h8.FactDescr = "analgesics" AND
    h8.FactStart < h3.FactStart AND h3.FactStart < h5.FactStart AND
    h5.FactStart < h8.FactEnd AND h3.FactEnd < h5.FactEnd AND
    h5.FactStart < h3.FactEnd AND h3.FactEnd < h8.FactEnd
```

Extended temporal visual queries





Conclusions and Current Work

- Paint Strips have been adopted in the medical system we built (Chittaro and Combi, AIME 2001), but are a general visual vocabulary that can be used in other temporal applications.
- Our system integrates with a common database by translating visually defined temporal patterns into SQL queries.
- New features have been added to increase the expressiveness of the visualization, e.g. quantitative ranges for distances between interval ends and granularities.
- The introduction of visually defined logical relations is being experimented.

IPBC (*Interactive Parallel Bar Charts*)

Esplorazione di serie temporali da dati di emodialisi

Research Goals

- The capability of interactively mining patient clinical information is an increasingly urgent need in the clinical domain, due to the continuous growth in the number of parameters that can be automatically acquired and in the size of the databases where they accumulate.
- Existing Visual Data Mining prototypes for clinical databases are attempting to exploit familiar visualizations such as *parallel coordinate plots* (Falkman, 2001) or *tables* (Spenke, 2001).
- Our work is aimed at:
  - exploring visualizations based on *interactive bar charts*;
  - applying them to real clinical databases, especially those containing *collections of time-series*.
- Bar charts were preferred to novel approaches to time-series visualization (e.g., spiral structures), because they are very familiar to clinicians.

Hemodialysis

- Hemodialysis is the widely used treatment for patients with acute or chronic end-stage renal failure.
- During an hemodialysis session, blood is circulated through a machine, that removes wastes and excess water. The “cleaned” blood is then returned to the patient’s bloodstream.
- In general, hemodialysis patients are treated *3 times a week* and each session lasts about *4 hours*.
- The number of patients that need hemodialysis is increasing.

Hemodialysis Data: Need for Analysis Tools

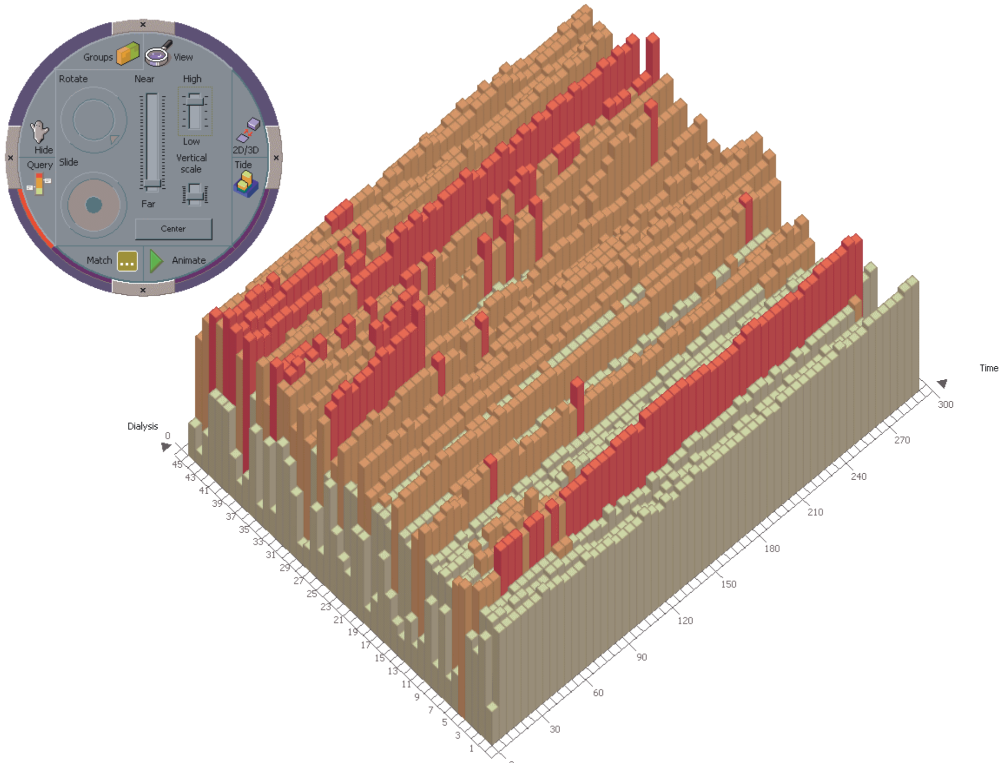
- It is very important to be able to evaluate the quality of hemodialysis, e.g.:
  - for each single hemodialysis session,
  - for all the sessions concerning the same patient, and
  - for sets of sessions concerning a specific hemodialyzer device.
- Medium-size hemodialysis center:
  - up to 60 patients per day (i.e., more than 19’000 sessions per year);
  - up to 50 different parameters acquired from the patient and the process, with a configurable sampling time;
  - Example (25 parameters with a sampling time of 30 seconds): 12’000 values collected in each session, and 228’000’000 values per year.
- Unfortunately, hemodialysis software is more concerned with acquiring and storing data, rather than visualizing and analyzing it.

Hemodialysis and Visual Data Mining

- Clinicians' abilities in recognizing interesting patterns are currently used suboptimally or not used at all. VDM of hemodialytic data would allow clinicians to take decisions affecting different important aspects:
  - *therapy* (personalizing the individual treatment of specific patients);
  - *management* (assessing and improving quality of the hemodialysis centre);
  - *nephrology research* (discovering relations and testing hypothesis).
- Data Mining on hemodialysis databases is (at least, at initial stages) intrinsically vague for clinicians, VDM can be more promising than fully automatic DM, because clinicians can freely explore datasets as they see fit.
- The clinical context is characterized by a need for user interfaces that require minimal technical sophistication and expertise to the users, while supporting a wide variety of information intensive tasks. A proper exploitation of visual aspects and interactive techniques can greatly increase the ease of use and acceptance of the provided solutions.

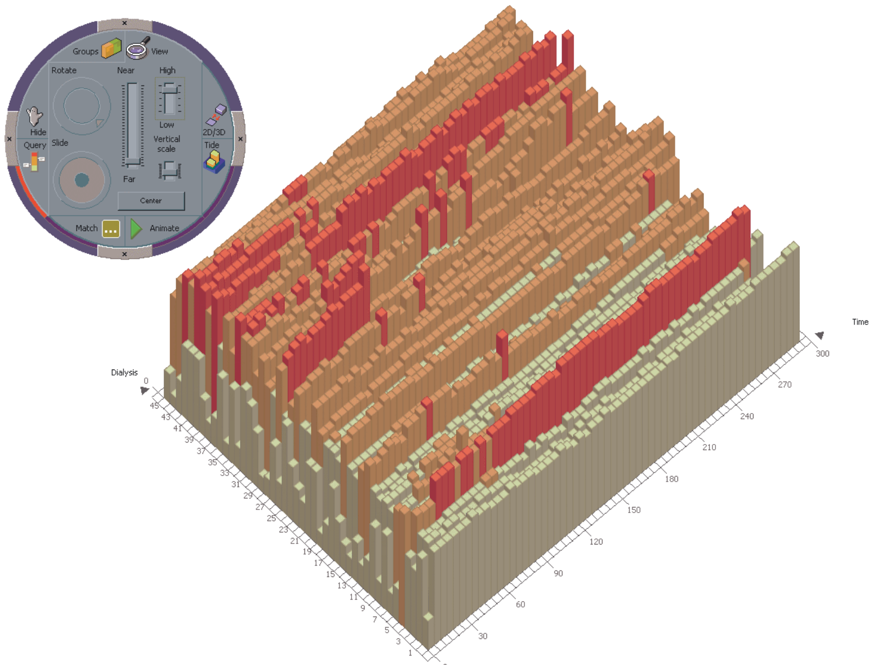
The IPBC System

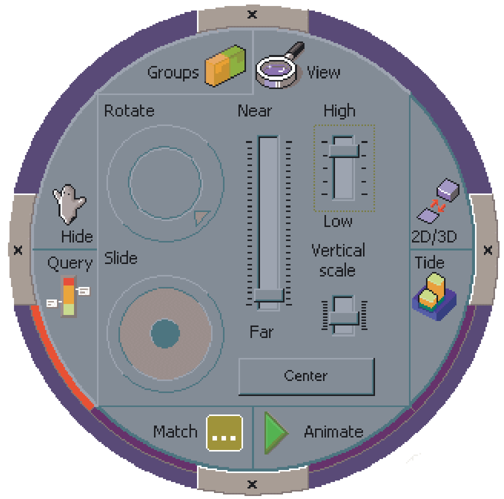
- IPBC (*Interactive Parallel Bar Charts*) connects to the hemodialysis database, produces a visualization that replaces tens of separate screens used in traditional hemodialysis systems, and extends them with a set of interactive tools for exploration.
- **Example:**



The Round Toolbar

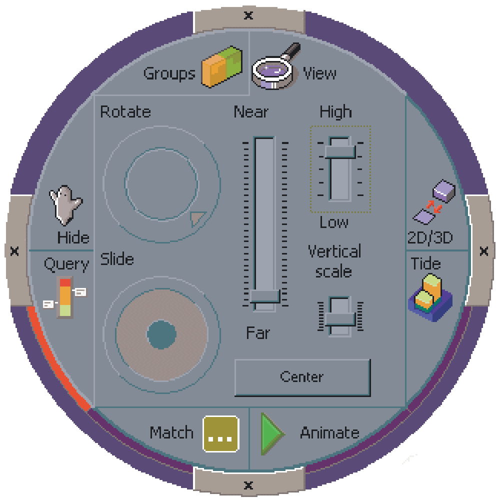
- Every interactive function is activated through a round-shaped pop-up menu that:
  - can exploit unused space in the corners of the screen and
  - minimizes effort needed to locate and activate functions.

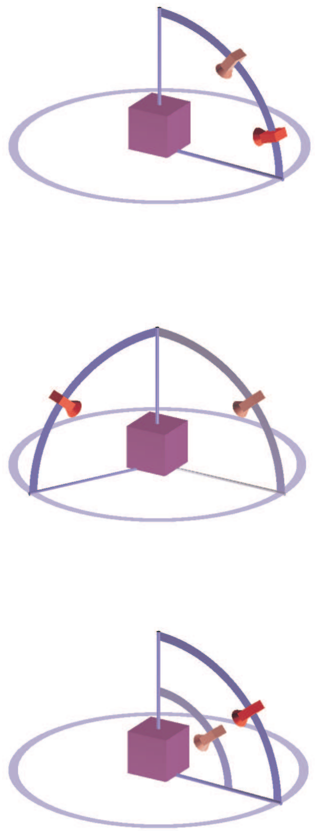




Viewpoint Mode

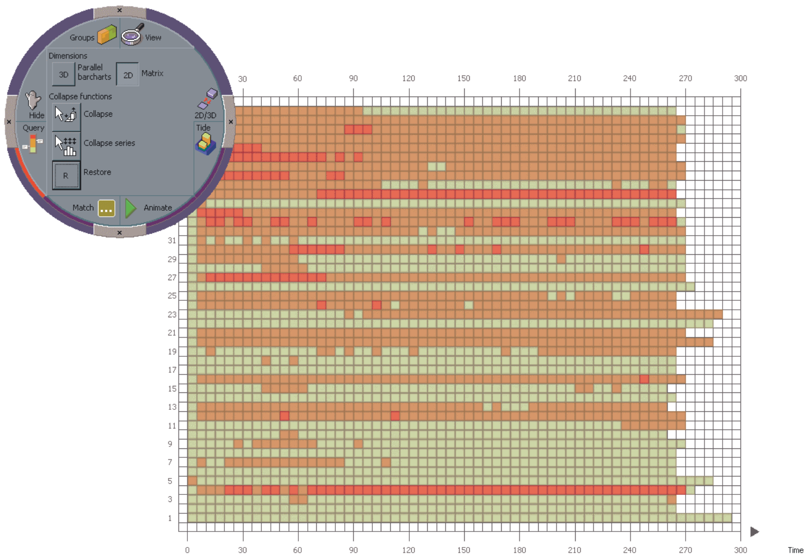
- Viewpoint control is based on predefined trajectories to prevent navigation problems.

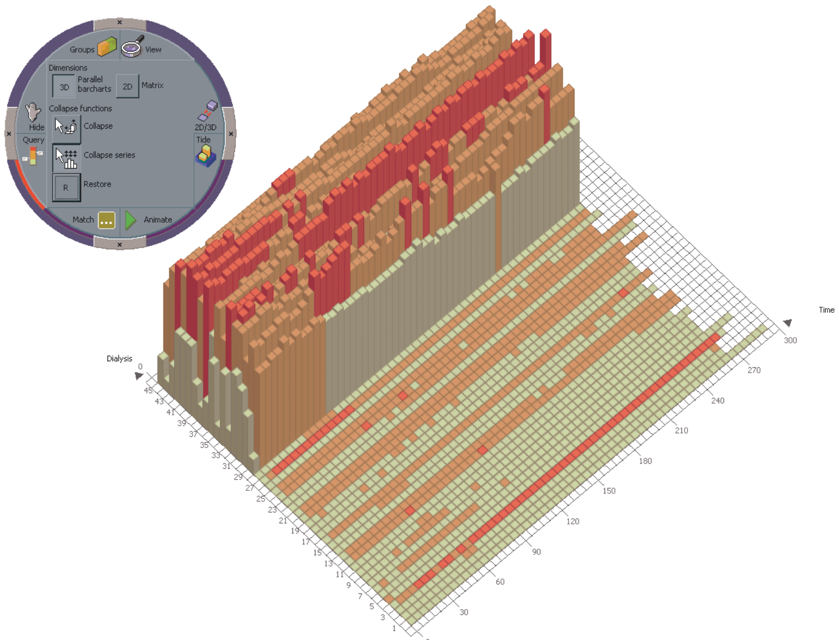




Managing Occlusions

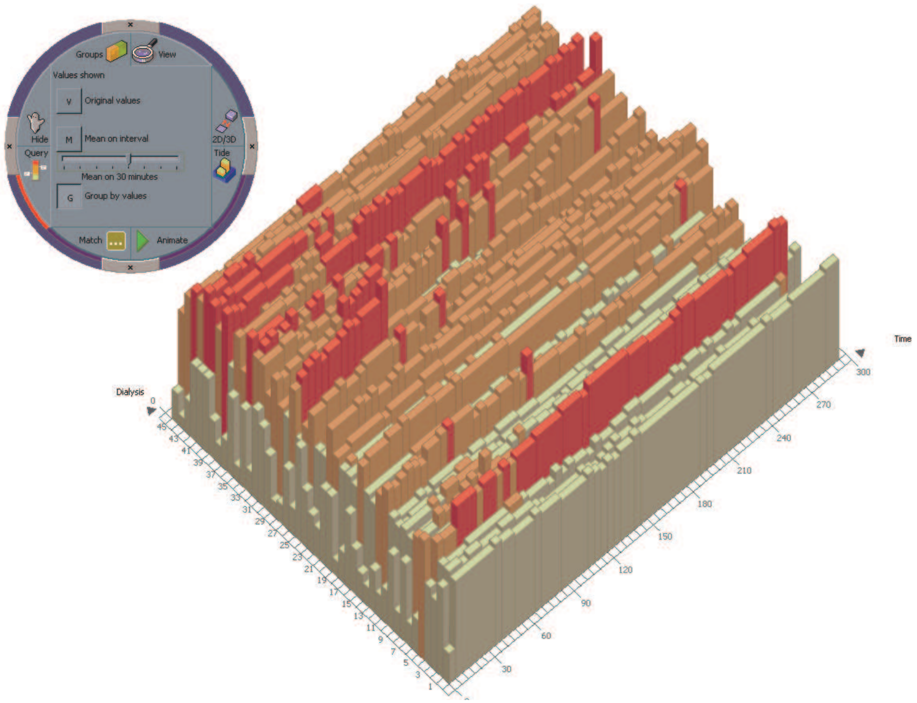
- Besides viewpoint control, the typical occlusion problem of 3D visualizations is handled through:
  - alternative 2D matrix visualization;
  - interactive functions for occlusion elimination.
- Examples:

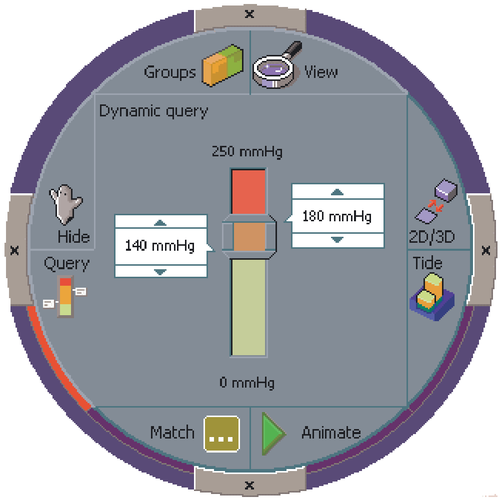




Visual Dynamic Queries

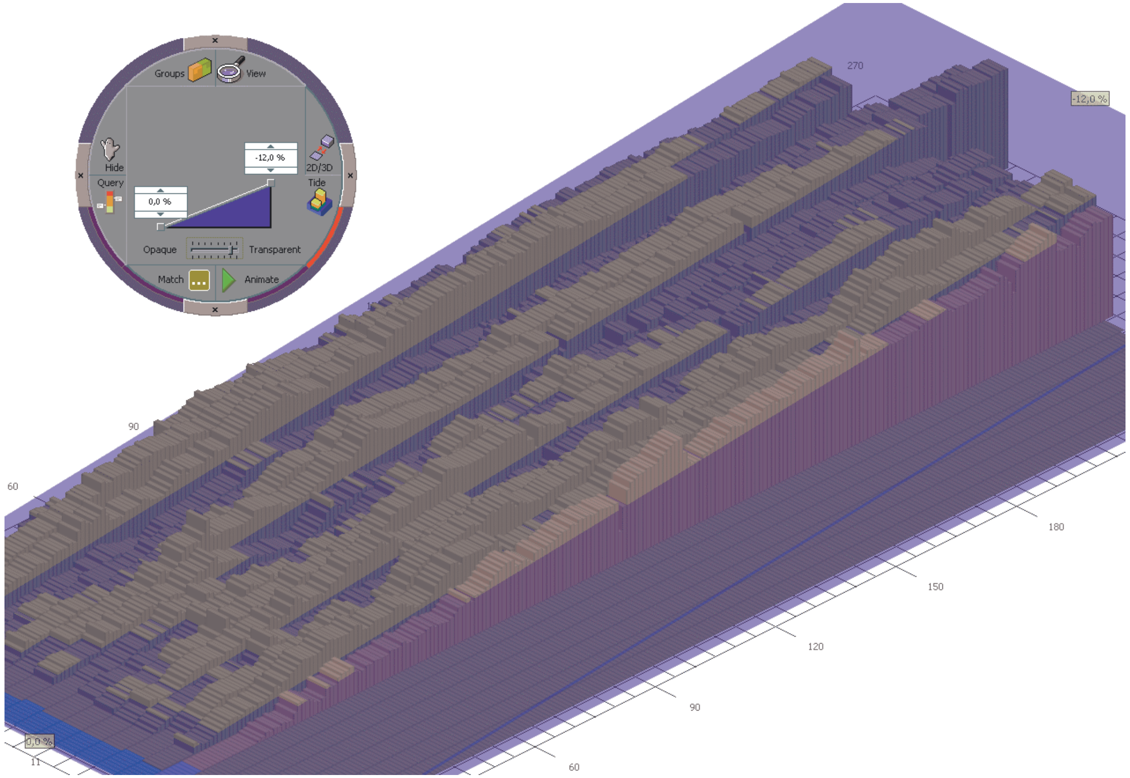
- Color classifies time-series values into different ranges.
- Ranges can be interactively modified by moving sliders in the toolbar.
- Slider elements can be moved *independently* or *together.*





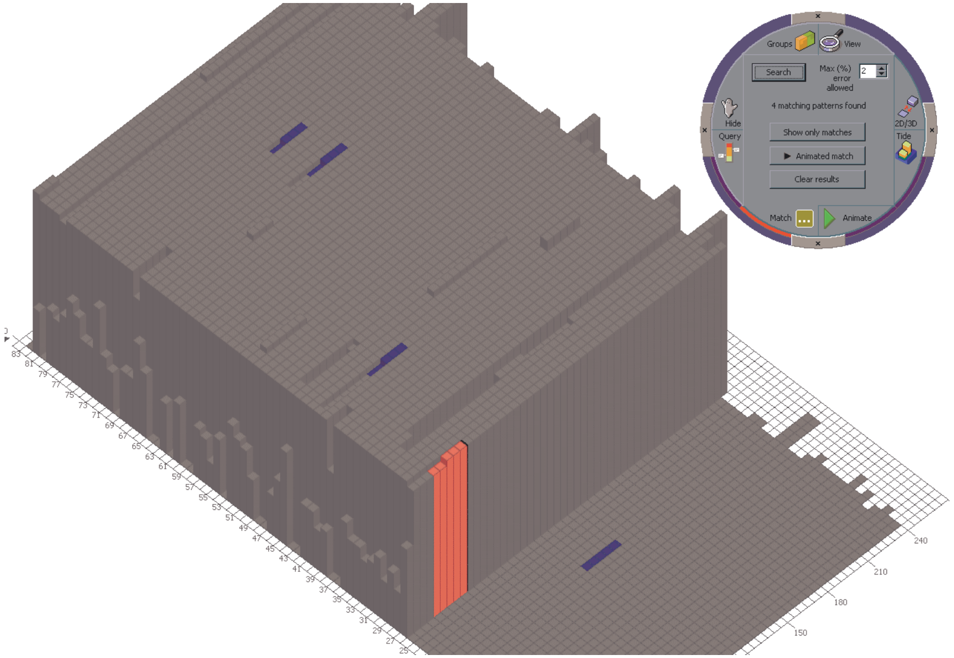
Comparing data with (time-varying) thresholds

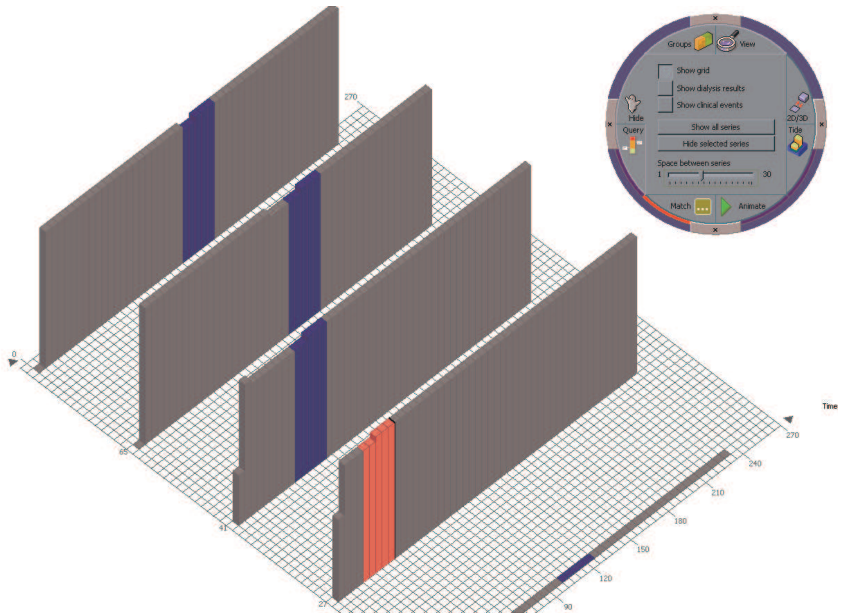
- A frequent need is to quickly perceive how many and which values are below or above a given (time-varying) threshold.
- For this need, IPBC offers a visual function based on a tide metaphor.
- Slope and height are interactively set through the Round Toolbar.



Automatic Pattern Matching

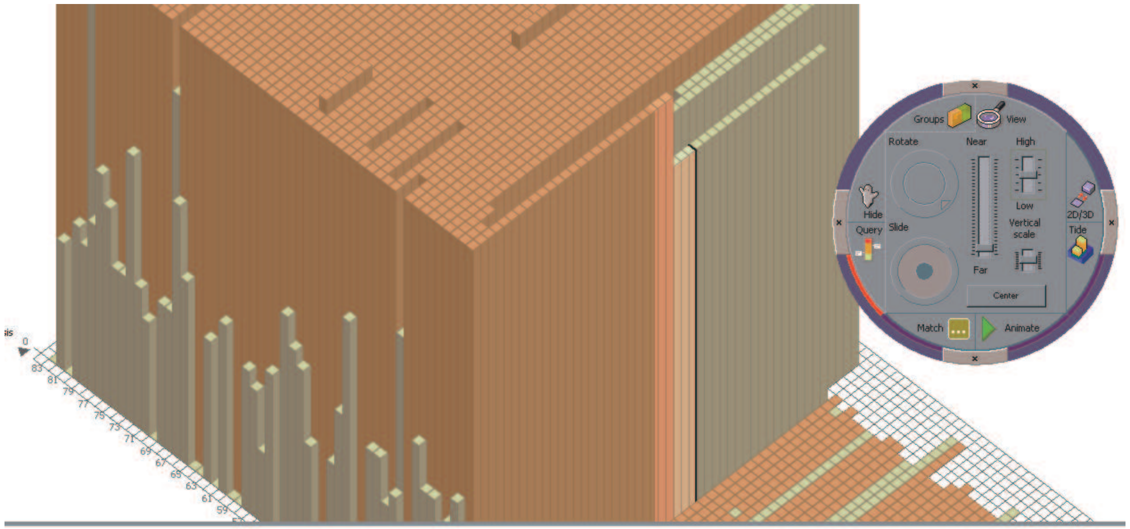
- When the user notices an interesting sequence of values in one of the time-series, IPBC offers her the opportunity to automatically search for and highlight occurrences of a similar pattern.
- The user selects the sequence by simply dragging the mouse over it.

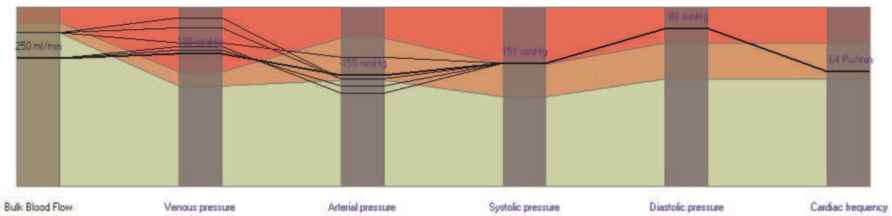




Integration of parallel bar charts with PCP

- Parallel Coordinate Plots: a trajectory in a plot connects the *quality indexes* (typically, 5-7 values) of a session, and this high-level perspective is linked to the much more detailed perspective of the parallel bar chart.





Conclusions and Current Work

- We carried out a field evaluation of IPBC with real users (at the Hemodialysis Center of the Hospital of Mede, PV, Italy).
- Major advantage: the visualization and interactive features are very quickly learned and remembered by clinicians.
- A need (considered relevant by clinicians) is dealing with time-series at different abstraction levels, allowing for both a fine exploration of time-series (e.g., to detect specific unusual values) and their coarse exploration (to focus on more abstract, derived information).

Main References

- CHITTARO, L., COMBI, C., TRAPASSO, G. Clinical Data Mining: a Visual Approach and its Application to Hemodialysis. Journal of Visual Languages and Computing, vol. 14, n.6, p. 591-620, 2003
- CHITTARO, L., COMBI, C. Visualizing Queries on Databases of Temporal Histories: New Metaphors and their Evaluation. Data & Knowledge Engineering, vol. 44, n. 2, p. 239-264, 2003
- CHITTARO, L., COMBI, C., TRAPASSO, G. Visual Data Mining of Clinical Databases: an Application to the Hemodialytic Treatment based on 3D Interactive Bar Charts. Proceedings of VDM 2002: 2nd International Workshop on Visual Data Mining, Helsinki, Finland, 2002.
- CHITTARO, L., COMBI, C. Visual Definition of Temporal Clinical Abstractions: A User Interface based on Novel Metaphors. In: QUAGLINI, S., BARAHONA, P., ANDREASSEN, S. (eds.) Proceedings AIME '01 (Artificial Intelligence in Medicine in Europe). Berlin, Springer, LNCS, p. 227-230, 2001
- CHITTARO, L., COMBI, C. Representation of Temporal Intervals and Relations: Information Visualization Aspects and their Evaluation. In: BETTINI, C., MONTANARI, A. (eds.) Proceedings eighth international symposium on temporal representation and reasoning (TIME 2001). Los Alamitos, IEEE Computer Society Press, p. 13-20, 2001
- COMBI, C., PORTONI, L., PINCIROLI, F. Visualizing Temporal Clinical Data on the WWW. In: HORN, W., SHAHAR, Y., LINDBERG, G., ANDREASSEN, S., WYATT, J. (eds.), Proceedings Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99) Berlin, Springer, LNAI 1620, p. 301-311, 1999