Time Series Data Are Everywhere

Temporal relationships can be highly complex
- temporal ordering is a serious issue
- event may occur in spatially disjoint locations
- what came before what – cause and effect
- what time shifts are acceptable/plausible?

To understand temporal relationships, an analyst:
- might need to reread the paragraph many times
- needs to cognitively make inferences between pieces of information

Visualization is key to externalize these relationships
- put it all out on “paper” and reason with it

Example (from Kuchar et al. IEEE CG&A May/June 2006)

Time is Special

Difficult to handle:
- likely will have to create this sample news story from information embedded in thousands of documents
- say each sentence contains information distilled from 100 documents
- analysts would need to find each of these sentences and then link them together to formulate the story

Why is time difficult?
- what actually is time?
- how can one work with the metaphor of time’s flow?
- what is the proper formalism that captures the time’s special role in reasoning

Time is an important variable
- people consider it as an independent quantity
- our perception is that we have no control over it
- time is an ever-present thread that can help tie events together
**Time as a Reference Frame**

Reference frame:
- standard point of view or perspective for making observations and judgments
- calendars provide such reference frames
- Gregorian calendar, universal time, absolute time

But sometimes time information is incomplete
- “The bank robbery occurred at 2:30 …”
- the date and time zone are unspecified
- what to do then?

Time is also often used in relative terms:
- “today”, “yesterday”, “fortnight”, “before Tuesday”, …
- must normalize different reference systems into a common framework
- it might be unknown what reference system was used individually
- “This robbery is similar to a crime spree that started on the Chinese New Year …” – when is Chinese New Year?
- causes ambiguities, uncertainties, biases, conflicts

**Time: Tasks and Taxonomies**

Often asked questions:
- when was something greatest/least?
- as there a pattern?
- are two series similar?
- does a data element exist at time \( t \), and when?
- how long does a data element exist and how often?
- how fast are data elements changing
- in what order do they appear?
- do data elements exist together?

Different types of time series data:
- discrete vs. interval
- linear vs. cyclic
- ordinal vs. continuous
- ordered vs. branching vs. time with multiple perspectives

**Traditional Time Series Visualizations**

Climate models might be missing an atmospheric ingredient
Traditional Time Series Visualizations

NVIDIA stock vs. NASDAQ (from yahoo! finance)

Fun one… (found in J. Stasko lecture)

Next…

Shall examine a few good visualization metaphors for time
- there are quite a few of them…

ThemeRiver (Havre et al., 2002)

River widens or narrows to depict changes in the collective strength of selected themes in the underlying documents. Individual themes are represented as colored "currents" flowing within the river.

Example shown here: newspaper themes around the Cuban Missile crisis
Stream Graphs

The Ebb and Flow of Movies: Box Office Receipts 1986 — 2008
Summer blockbusters and holiday hits make up the bulk of box office revenue each year, while outliers for the Oscars tend to attract smaller audiences that build over time. Here's a look at how movies have fared at the box office, after adjusting for inflation.

Stacked Area Charts

How Different Groups Spend Their Day
The American Time Use Survey asks thousands of American residents to recall every minute of a day. Here is how people over age 15 spend their time in 2006. Related article

Name Voyager (http://www.babynamewizard.com)

New! Try the NameRapper to see where your favorite names are being used, and Namipedia for full info on every name

Chronological Flow Charts

Plaisant et al. (CHI ’96)
Medical Data

Medical data are often displayed along time
- natural to humans
- progression of disease
- appearance of symptoms
- time course of treatment and outcome
- but also time signals (ECG, blood pressure, etc.)

A popular example is Lifelines and Lifelines2
- Shneiderman and Plaisant et al.

LifeLines: Patient-Centric

Goals:
- bring out temporal categorical patterns across multiple records
- categorical event data such as complaints, diagnoses, treatments
- play important roles in health providers decision making

Features
- allows users to manipulate multiple records simultaneously
- understand relative temporal relationships across records
- 3 operators: align, rank, filter
- temporal summaries allow multiple groups of records to be compared
LifeLines2: Videos
http://www.cs.umd.edu/hcil/lifelines2/

Multi-Scale and Abstractions (Aigner at al., IEEE TVCG, 2008)
Deal with different levels of detail
- illustrative abstraction
- overview + detail
- used here for medical data

Cyclic Patterns
Time data are often cyclic
- spiral displays are good to bring out cyclic patterns
- one period per loop (for example, a year)

Weber et al., 2001

Cyclic Patterns
May have to play around to discover the cycles

from J. Stasko, lecture notes
Combine Space and Time

OculusInfo Geotime application
- events are represented in an X,Y,T coordinate space
- the X,Y plane shows geography
- the vertical T axis represents time
- events animate in time vertically through the 3-D space as the time slider bar is moved.

http://www.oculusinfo.com/SoftwareProducts/GeoTime.html

Geotime

Interaction

As complexity increases, interaction capabilities are key
- show more context of what else was going on at that time
- likely have to abstract some of the information
- allow several different levels of detail at once
- allow drill-down for details
- use dashboard design with many linked information displays

Example: Computer system management
- LiveRAC system (McLachlan et al.)
- next two slides

LiveRAC
Figure 3. LiveRAC shows a full day of system management time-series data using a reorderable matrix of area-aware charts. Over 4000 devices are shown in rows, with 11 columns representing groups of monitored parameters. (a) The user has sorted by the maximum value in the CPU column. The first several dozen rows have been stretched to show sparklines for the devices, with the top 13 enlarged enough to display text labels. The time period of business hours has been selected, showing the increase in the *le_plus* parameter for many devices. (b) The top three rows have been further enlarged to show fully detailed charts in the CPU column and partially detailed ones in Swap and two other columns. The time marker (vertical black line on each chart) indicates the start of anomalous activity in several of *le_plus*’s parameters. Below the labeled rows, we see many blocks at the lowest semantic zoom level, and further below we see a compressed region of highly saturated blocks that aggregate information from many charts.