Narrative Visualization

Jessica Hullman

Cave painting [Ennedi Plateau, Chad]
Janet L. Yellen, on the Economy’s Twists and Turns

The Federal Reserve is said to have a “dual mandate”: keeping inflation in check and the unemployment rate low. These measures, which tend to change cyclically and in concert with each other, are charted for every year since the Great Depression.

In speeches and in meetings, Ms. Yellen, the nominee for the next Fed leader, has commented on the Fed’s actions during significant periods, providing a window into her views and priorities.

[Giratikanon and Parlapiano 2013]
Topics

Design space
Theories of interpretation
Construction
  Manual
  Automated
Value of storytelling [Gershon and Page 2001]

A way of structuring information

- Easier to understand than lists
- Uncertainty, conflict, resolution
- Text and visuals can be complementary
Human Development Trends 2005

Interactive presentation of some of the messages in the Human Development Report 2005

English
Dansk
Portuguese
Suomi
Français
Deutsch

Produced in collaboration with:
Gapminder: www.gapminder.org

English translation: Claes Johansson, UNDP

Storytelling techniques

- build in picture
- continuity
- animate events
- increasing attention
- redundancy

[Gershon and Page 2001]
Narrative Visualization [Segel and Heer 2010]

Studied 58 examples, characterized design space
  • News media, blogs, instructional videos, research

Design space dimensions: 1. Genre

[Slide adapted from Segel 2010]
2. Visual narrative tactics

Highlighting

Transition guidance

3. Narrative structure tactics

Ordering
Messaging
Interactivity

Budget Forecasts, Compared With Reality

Just two years ago, airplanes were predicted by 2012. How accurate have past White House budget forecasts been?
Author vs. reader driven stories

**Author driven**
- Prescribed ordering
- Stronger messaging
- Limited interactivity

**Reader driven**
- Multiple orderings
- Less messaging
- More open interactivity

How do we make sense of these examples?

An Organizational Chart of the House Democrat's Health Plan [GOP 2008, Robert Palmer 2008]
Visualization rhetoric [Hullman & Diakopoulos 2011]

Using data and visualization to persuade users to adopt certain interpretations (explicitly or implicitly)

Framing effects: small changes in presentation of an issue result in significant changes in opinion

Method
- 51 professional produced narrative visualizations
- NYT, BBC, Economist, local news, political outlets
- Iterative qualitative coding, seeded scheme with semiotics, persuasion concepts

Taxonomy of rhetorical strategies

Rhetorical Categories
- Information Access
- Provenance
- Mapping
- Procedural
- Linguistic

Editorial Layers
- Data
- Visual Representation
- Annotation
- Interactivity

- Omission (variable selection)
- Metonymy (aggregation, categorization)
Enthymeme (information access)

How the Recession Changed Us [The Atlantic 2009]

Metonymy (information access)

Donald Trump wants massive tax cuts for the rich [Vox 2015]
Taxonomy of rhetorical strategies

Rhetorical Categories
- Information Access
- Provenance
- Mapping
- Procedural
- Linguistic

Editorial Layers
- Data
- Visual Representation
- Annotation
- Interactivity

• Citing sources, methodology
• Exceptions, corrections
• Acknowledging uncertainty

Citing sources / methods (provenance)

[Bloch, Carter, and McLean 2010]
Taxonomy of rhetorical strategies

Rhetorical Categories
- Information Access
- Provenance
- Mapping
- Procedural
- Linguistic

Editorial Layers
- Data
- Visual Representation
- Annotation
- Interactivity

- Visual metaphor
- Semantic encodings (e.g., red:Rep, blue:Dem)

Visual metaphor (mapping rhetoric)
Taxonomy of rhetorical strategies

Rhetorical Categories
- Information Access
- Provenance
- Mapping
- Procedural
- Linguistic

Editorial Layers
- Data
- Visual Representation
- Annotation
- Interactivity

- Rhetorical question, “scare quotes”
- Labeling choices
- Analogy, simile, double entendre

Design process

An editorial process characterized by rhetorical decisions at various points.

Define context/goal

Filter, transform

Visualize

Select

Annotate

Order, interactions

Data
Scope of narrative visualization

- Visualization genres that employ visual and narrative structures to guide attention
- Visualizations that use rhetorical devices to persuade

Discuss with 1 or 2 other people near you:
In your own research or papers you’ve read/cited, is narrative visualization occurring? How do persuasion and rhetoric occur in these contexts?

Theories of Interpretation
Semiotics

"TREE"
Sign

Signified

Viewing codes: the perceptual mechanisms, cognitive mechanisms, conventions and other prior knowledge that influence how we interpret signs.

There is a fire!

Interpretant

Object
(Signified)

Representamen
(Sign)
Denotation vs. connotation

(Make explicit)

The Democrats proposed health-plan is disorganized and inefficient.

(Imply)

Narrative theory

Story grammars:
Models of narrative cognition based on systematic studies of what impacts peoples’ ability to recall parts of a story

Reader mentally indexes events by time, space, protagonist, causality, intention [Zwaan 1995]

Remembrance of Things Parsed [Mandler and Johnson 1977]
Narrative theory applied

Figure 3. Graph comic example illustrating historical events preceding WWI targeted to a general audience.

Graph Comics [Bach et al. 2016]

C1. Element identity—Characters in traditional comics are recognizable via their distinct visual features. In graph visualization, nodes and edges are usually represented by circles, sometimes using labels for identifications. Variations in shape and color reflect attributes in the data.

C2. Depicting change—In comics, readers can infer changes even if not shown explicitly. For example, a baby being shown implies it has been born. In graphs, a (new) node being shown does not mean it has not been there before.

C3. Order of events—Many changes exhibit a natural order (e.g., humans grow old, objects fall down, day turns into night). Changes in graphs do not have a natural order nor duration; a cluster can grow or shrink, divide or merge.

C4. Spatial context—In comics, the spatial context of a scene is usually indicated by the panel background: the same background means the same place and often the same scene.

C5. Number of elements—Comics usually show a small set of characters in every panel. Changes in graphs may involve many more elements than it is possible to show in a panel.

Graph Comics [Bach et al. 2016]
Narrative theory applied

Presentation sequence → recall, interpretation
[Black & Bower 1979, Pennington & Hastie 1982]

How to design an effective sequence?

[Hullman et al. 2013]

Design tools

Can we automatically identify sequences to recommend to a human designer?

Data

- filter, transform
- visualize
- select
- annotate
- order, interactions

define context / goal
Discovery order ≠ presentation order

Interactive dynamics in analysis: Filter, derive, etc. 
[Shneiderman 1996, Heer & Shneiderman 2012]

Provenance: How did I get to this result? 
[Heer et al. 2008, Bavoli et al. 2005, Scheidegger et al. 2007]

What makes an effective sequence?

What do designers and journalists do?

Examined 400+ transitions in explicitly ordered visualizations
• Interactive slideshows
• Scrollable visualizations
• Animated videos
• Analysis presentations
Dialogue (e.g., question/answer)
Causal
Temporal (e.g., chronology, future)
Hierarchy (general-to-specific)
Comparative Data (Dimension walk, Measure walk)
Spatial (e.g., cardinal directions)

Dimension walk - change I.V.

At Top Colleges, an Admissions Gap for Minorities

Few black students at elite colleges

Roughly 15 percent of public high school graduates are black, but despite the widespread use of affirmative action at elite colleges, only one college with a graduation rate of more than 70 percent has that many black students in its freshman class.

By JOHN KILLOR
Source: National Center for Education Statistics, 2011
Some observations about transitions

Some types (time, hierarchy, comparative) more common
Defined by a single change to data schema
• Implies a goal of minimizing cognitive cost

Examining the cost of transitions

Crowd workers decided between pairs of transitions and explained choices
Lower cost transitions are preferred.
Examining the cost of transitions

Type preferences among equal cost transitions preferred over others:

Temporal > (Dimension | Measure ) > Hierarchy

Some observations about transitions

Graph-based approach
Minimize cost between adjacent visualizations
Cost model should account for preferences between types of transitions and changes to data schema

Goal:
Automatic detection of effective sequences
Micro-vs. macrostructure

Fictional stories are described in terms of how larger events are organized (macro-), each of which is comprised of sub-events (micro-).

Remembrance of Things Parsed [Mandler and Johnson 1977]

Multiple transition patterns
Parallel structure

Are groupings preferred despite cost?

How do authors balance high-level and low-level structuring strategies?
Narrative structure in visualization: model

Top-down: parallel structure, semantic groupings

Bottom-up: cognitive cost of transitions

Narrative patterns to support EDA

Voyager [Wongsuphasawat et al. 2015]

Suggest paths through recommended visualizations
Theories: Narrative as constructive

Presentation sequence influences the stories we tell ourselves (Wilensky 1997)

 Sender (designer)  
  
  
  Receiver (user)  

Theories: Narrative as constructive

Theories of Data Analysis: From Magical Thinking Through Classical Statistics

Persi Diaconis
Stanford University

Exploratory data analysis (EDA) seeks to reveal structure, or simple descriptions, in data. We look at numbers or graphs and try to find patterns. We pursue leads suggested by background information, imagination, patterns perceived, and experience with other data analyses.
A fix for base-rate neglect?

1% of the women had breast cancer at the time of the screening.

Of those with breast cancer, 80% received a positive result on the mammogram.

Of those without breast cancer, 15% received a positive result on the mammogram.

All others received a negative result.

Suppose a woman gets a positive result during a routine mammogram screening. Without knowing any other symptoms, what are the chances she has breast cancer?

[Krynski and Tenenbaum 2007]
Causal models matter

Figure 5. Histogram of responses in Experiment 2. The correct answer was 51%. Responses were classified as correct (51%), base-rate neglect (>65%), and other. A significant difference was found between false positive and benign cyst scenarios (Fisher’s exact test, $p < .05$). Error bars represent the standard error of the normal approximation to the binomial distribution.

Construction
Manual design processes

“require skills like those familiar to movie directors, beyond a technical expert’s knowledge of computer science and engineering....”

Gershon and Page 2001

Lyra (Satyanarayan and Heer 2014)
Ellipsis (Satyanarayan and Heer 2014)

**Forecasts worse**
The forecast for the next decade is somewhat worse than it was a year ago. The revised economic outlook from 2011 to 2020 now predicts a total deficit of $6.5 trillion is expected.

**Tableau Story Points**

*Austin's Teacher Turnover Problem*

Many Austin schools are still struggling to hire and keep teachers, despite a multi-million-dollar effort, called Reach, to stem turnover in troubled schools since 2010.

- Austin's teacher turnover is above the national average.
- Teacher turnover is a city-wide problem.
- The problem is particularly acute at some schools.
- And city-wide, the problem isn't getting better.

*Austin's teacher turnover rate by year. Turnover was above the national average for most years.*

- Turnover rates:
  - 2006: 15.3%
  - 2007: 17.7%
  - 2008: 19.6%
  - 2009: 20.0%
  - 2010: 17.6%
  - 2011: 15.8%
  - 2012: 16.1%
  - 2013: 17.4%

*Source: 2014*
## Automated Construction

### Requirements

A model of effective visual narrative that can be operationalized algorithmically

**Examples:**
- Automated support for chart reading operations
- Automated generation of annotated news visualizations
- Personalized text stories and graphics
Brainstorming

Think about annotations (text or visual) that are useful in presenting or analyzing visualizations.
Are any of these amenable to automated generation?

With 1 or 2 people near you, talk about what types of annotations you would support and how your algorithm would generate them.

Facilitating chart reading

Elementary processes in graphical perception

[Simkin and Hastie 1987]
Facilitating chart reading

Graphical Overlays [Kong and Agrawala 2012]

AI-generated stories

Artificial intelligence based on human narrative cognition
- Conceptual dependency theory (Schank 1969)
- Case based reasoning (Schank 1982)
Annotated visualizations with news

Can we produce an automated solution?

Annotations tell the story
Can we identify criteria designers use to add context?

Examined annotation content, placement (function) in 136 news visualizations

Add relevant info

Observe salient values

[Hullman et al. 2013]
Contextifier system

Input: news article about a company

Yahoo Decides to Friend Facebook

(visualize)

Generate features, Select annotations

(Article Corpus)

Article

Corpus

(process text)

(obtain stock data)

Relevancy feature

Kullback-Leibler (KL) divergence

\[ D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \]
Article volume

Prioritize dates with unusually high numbers of company articles.

Visual saliency

Visual saliency proxy via simple analyses on stock series.

$max(t), min(t)$

$t_i - t_{i-1}$
Combine terms in an objective function

Personalization

The Best and Worst Places to Grow Up: How Your Area Compares

Children who grow up in some places gain or lose much more than they would if they grew up elsewhere. (via) RELATED ARTICLE

Washtenaw County is very bad for income mobility for children in poor families. It is better than only about 8 percent of counties.

Location matters. If you’re poor and live in the Detroit area, it’s better to be in Washtenaw than in Wayne County or Genesee County. Not only that, the proxies you use when you move to Washtenaw, let alone to Beaumont, don’t help you escape an advantage. Children who are at another age or live in a much better area do become rich parents, more likely to go to college and more likely to pass on that privilege.

[Aisch, Bloch, Cox, and Quealy 2015]
Summary

Narrative visualizations blend communication / exploratory techniques

Messaging, metaphor, sequencing, and other suggestive strategies have a powerful impact on interpretation

Semiotics, narrative theory, causal reasoning are critical

Automated systems possible by formalizing features