Why is validation difficult?

• different ways to get it wrong at each level

Domain situation
You misunderstood their needs
The way you show it doesn’t work
Algorithm
Your code is too slow
Data/task abstraction
• different ways to get it wrong at each level

Data/task abstraction
– scaffold to help you think systematically about choices
– help you think about the design space
– get an overview of the design space

Why talk about a textbook to a room of experts?

• ways to get it wrong at each level

You misunderstand their needs
The way you show it doesn’t work

End of textbook

How to encode: Arrange space, map channels

• arrange space

– SpaceTree

– Ternary

– SpaceTree

– color

– alpha

– size

• map channels

– single channel

– two channels

– three channels

– four channels

• combination of marks and channels

– points

– lines

– areas

– volumes
How to handle complexity: 3 more strategies

1. **Manipulate Facet Reduce**
   - Change
   - Select
   - Navigate
   - Superimpose
   - Filter
   - Aggregate
   - Embed

2. **Facet Reduce**
   - Change
   - Select
   - Navigate
   - Superimpose
   - Filter
   - Aggregate
   - Embed

3. **Reduce**
   - Change
   - Select
   - Navigate
   - Superimpose
   - Filter
   - Aggregate
   - Embed

**Idiom: bird's-eye maps**
- Encoding: same
- Data: subset shared
- Navigation: shared
- Bidirectional linking
- Why? Benefits: memory
- Lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
- Cons: display area, 2 views side by side each have only half the area of one view

**System: GDMaps**

**Channels:**
- Position on a common scale
- Position on an aligned scale
- Length (1D size)
- Width (2D size)
- Area (2D position)
- Depth (2D position)
- Color
- Closeness
- Value (1D value)
- Volume (3D size)

**Idiom: boxplot**
- Static item aggregation
- Task: find distribution
- Data: table
- Derived data
  - 5 quant attributes
  - Outliers beyond fence cutoffs explicitly shown

**System: Cerebral**

**Coordinate views:**
- Design choice interaction
- Why juxtapose views?
  - Benefits: eyes vs memory
  - Lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
  - Cons: display area, 2 views side by side each have only half the area of one view

**Idiom: Dimensionality reduction for documents**
- Attribute aggregation
  - Derive low-dimensional target space from high-dimensional measured space

**System: EDV**

**Visual Exploration of Large Structured Documents:**

- How to handle complexity: 3 more strategies
- Idiom: Linked highlighting
- System: EDV
- Reduce what is shown within single view

**Channels:**
- Position on a common scale
- Position on an aligned scale
- Length (1D size)
- Width (2D size)
- Area (2D position)
- Depth (2D position)
- Color
- Closeness
- Value (1D value)
- Volume (3D size)
A quick taste of my own work!

• technique-driven work
  - theoretical foundations
  - evaluation

• problem-driven work

---

**Evaluation: Focus+Context**

- Separation vs integrated views
- Distortion impact on search/memory

- Heidi Lam
- Ron Rensink
  (UBC)

- Robert Kincaid
  (Agilent)

- Heidi Lam

---

**Evaluation: Dimensionality reduction**

- Michael Sedlmair
- Melanie Tory

- Points vs landscapes for dimensionally reduced data
- Taxonomy of cluster separation factors

- Melanie Tory
  (UVic)

- Guidance on DR & scatterplot choices

---

**Evaluation: Graph drawing**

**Technique-driven: Graph drawing**

- Papers Process & Pitfalls
  - Visual Encoding Pitfalls
  - Rainbows Just Like In The Sky
  - 2D Good, 3D Better
  - Hammer In Search Of Nail
  - Unjustified Visual Encoding

- http://www.cs.ubc.ca/~tmm/vadbook

---

**Evaluation: Focus+Context**

**Theoretical foundations**

**Evaluation: Graph drawing**

**Technique-driven: Graph drawing**

---

**Evaluation: Dimensionality reduction**

---

**Technique-driven: Many domains**

---

**Problem-driven: Genomics, fisheries**

- Source: Humantestimony
- Pfister
  (Harvard)

- Miriah Meyer
- Aaron Barsky
- Jenn Gardy
  (Microbio)

- Robert Kincaid
  (Agilent)

- Heedful

---

**Problem-driven: Genomics**

- Problem-driven: Many domains
- 38
- 10Mb
- chr1
- chr2
- chr3
- chr4
- chr5
- chr6
- chr7
- chr8
- chr9
- chr10
- chr11
- chr12
- chr13
- chr14
- chr15
- chr16
- chr17
- chr18
- chr19
- chr20
- chr21
- chr22
- chrX
- chrY

---

**Problem-driven: Many domains**

- LiveRAC: systems time-series
- Peter McLachlan
- Stephen North
  (AT&T Research)

- SessionViewer: web log analysis
  (Sud. Zeitung)

- Heidi Lam
- Diane Tang
  (Google)

---

**Theoretical foundations**

- Nested Model
- Papers Process & Pitfalls
- Design Study Methodology

- Michael Sedlmair
- Miriah Meyer
- Aaron Barsky
- Jenn Gardy
  (Microbio)

- Robert Kincaid
  (Agilent)

- Joanna McGrenere

---

**More Information**

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#vad16gi
- book page (including tutorial lecture slides)
  http://www.crcpress.com/product/isbn/9781466508910
- 20% promo code for book+ebook combo: HVX7
- Illustrations: Eamon Maguire

- papers, videos, software, talks, courses
  http://www.cs.ubc.ca/~tmm/talks.html
  http://www.cs.ubc.ca/~tmm