Research Cycles, Collaboration, and Visualization

VIEW Workshop

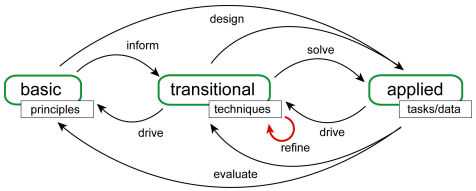
Tamara Munzner, UBC

27 June 2007

Outline

- research cycles and collaborator roles
- value of collaboration: success stories
- difficulty of collaboration: when to walk away

Research Cycles



Johnson, Moorhead, Munzner, Pfister, Rheingans, and Yoo. NIH/NSF Visualization Research Challenges Report. IEEE CS Press, 2006.

- difficult for one person cover all cycles
- collaboration is obvious way to fill in gaps

Collab Roles: Methodology ↔ Vis ↔ Problems

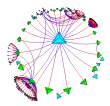
- left: providers of principles/methodologies
  - HCI, cognitive psychology
  - computer graphics
  - math, statistics
- right: providers of driving problems
  - domain experts, target application users
- middle: fellow vis practioners
  - all the usual reasons to work jointly instead of alone
- middle: fellow tool builders, outside of vis
  - often want vis interface for their tools/algs

Outline

- research cycles and collaborator roles
- value of collaboration: success stories
- difficulty of collaboration: when to walk away

Methodology → Informs Vis: Mathematics

- methodology: mathematics of hyperbolic geometry
- collab as way to build expertise
  - webviz: first project, in collab with hyperbolic geometer
  - H3: later project, solo



Munzner and Burchard. Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space. Proc VRML 95

Munzner. H3: Laying Out Large Directed Graphs in 3D Hyperbolic Space. Proc InfoVis 97

Methodology → Informs Vis: Evaluate Techniques

- methodology: evaluate systems/techniques
- collab as way to extend reach of research program
- two projects in collab with HCI researchers

Risden, Czerwinski, Munzner, and Cook. An initial examination of ease of use for 2D and 3D information visualizations of web content. Intl Journal of Human Computer Studies, 53(5):695-714, Nov 2000.

- compare H3 technique to other techniques

Nekrasovski, Bodnar, Guimbretière, McGrenere, and Munzner. An Evaluation of Pan&Zoom and Rubber Sheet Navigation with and without an Overview. Proc CHI 2006, p 11-20.

- investigate distortion-based vs. standard navigation
- investigate Focus+Context assumptions about overview use

Methodology → Informs Vis: Evaluate Perception

- methodology: evaluate low-level perceptual mechanisms
- collab as way to extend reach of research program
- two projects in collab with perceptual psych researcher

Lau, Rensink, and Munzner. Perceptual Invariance of Nonlinear Focus+Context Transformations. Proc APGV 04, p 65-72.

- effect of distortion on visual search
  - zones of no cost, some cost, major cost based on amount

Lam, Rensink, and Munzner. Effects of 2D Geometric Transformations on Visual Memory. Proc APGV 06, p 119-126.

- effect of rigid and nonlinear transforms on visual memory
  - similar cost zones, background grids usually help

User Studies

- designing good user studies is very tricky
  - stats analysis in papers is just tip of iceberg
  - **many** ways to make methodological mistakes in design
  - very strongly recommend collab if your background not HCI/psych: extend reach of research scope
- lab study is only one possibility
  - now carrying out more observational field studies
- user study results are changing my thinking on how people use distortion-based techniques
  - changing direction of my research program
  - research cycles in action

Vis ← Driven By Problems

- three approaches
  - pick one problem domain, immerse yourself in it
    - eventually become near-expert yourself
    - develop close ties to experts in domain
  - pick problem domain where you are already expert
    - you are the target user
  - move between problem domains opportunitistically
    - my usual approach
    - learning enough about the domain to find problem that you think vis can help with
    - critical step: abstracting/generalizing from domain problem to vis problem
    - building shared understanding and vocabulary does require time

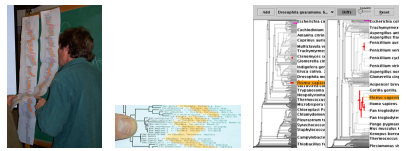
Vis ← Driven By Problems: Constellation

- driven by: needs of computational linguists checking algorithms that traverse semantic networks built from online dictionaries
  - use spatial position to show path structure
  - visual layering to disambiguate edge crossings

Munzner, Guimbretière and Robertson. Constellation: A Visualization Tool For Linguistic Queries from MindNet. Proc InfoVis 99, p 132-135.

Vis ← Driven By Problems: TreeJuxtaposer

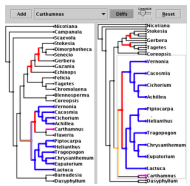
- driven by: needs of evolutionary biologists comparing phylogenetic trees
  - requires both overview and detail
    - stretch and squish navigation maintains orthogonal ordering



Munzner, Guimbretière, Tasiran, Zhang, and Zhou. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. Proc SIGGRAPH 2003, pp 453-462.

Methodology → Informs Vis: TreeJuxtaposer

- collab with computational geometry researchers
  - methodology: algs for computing correspondences between trees
  - although general theoretical area well studied, vis needs led us to new open problem

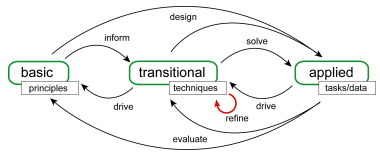


Vis ← Driven By Problems: Cerebral

- driven by: needs of microbiologists studying protein-protein interaction networks
  - protein nodes divided into levels by subcellular location
  - groups determined by biological function

Barsky, Gardy, Hancock, and Munzner. Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Bioinformatics 23(8):1040-1042, 2007.

Vis Technique Refinement



- technique-driven instead of problem-driven
  - goal: algorithm improvement
  - best if original technique justified by real problem
  - close collab with problem domain people not required
- of course, can collab with fellow vis practioners

Vis Technique Refinement: Accordion Drawing

- extend/generalize technique introduced in TreeJuxtaposer

Beerman, Munzner, and Humphreys. Scalable, Robust Visualization of Large Trees. Proc EuroVis 2005, p 37-44.

- scale up to trees of 15M nodes using GPU, 5M nodes on standard hw

Slack, Hildebrand, and Munzner. Partitioned Rendering Infrastructure for Scalable Accordion Drawing (Extended Version). Information Visualization 5(2):137-151, 2006.

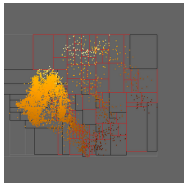
- general rendering framework supporting guaranteed visibility for many data types

Slack and Munzner. Composite Rectilinear Deformation for Stretch and Squish Navigation. Proc Visualization 2006 / TVCG 12(5) Sep 2006.

- stretch and squish navigation algs that handle millions of items

Vis Technique Refinement: MDSteer

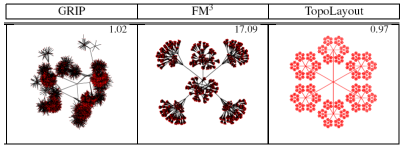
- make dimensionality reduction steerable
  - pure technique refinement
  - well established problem



Williams and Munzner. Steerable, Progressive Multidimensional Scaling. Proc InfoVis 2004.

Vis Technique Refinement: TopoLayout

- multilevel graph drawing
  - break up graph into pieces based on graph connectivity
  - draw each piece with appropriate algorithm



TopoLayout: Multi-Level Graph Layout by Topological Features. Archambault, Munzner, and Auber. IEEE TVCG 13(2):305–317.

Outline

- research cycles and collaborator roles
- value of collaboration: success stories
- difficulty of collaboration: when to walk away

Four Process Questions

- questions to ask before starting collab projects
  - sometimes I asked them early
  - sometimes I wish I'd asked them early
- 1. what is the role of my collaborators?
  - real users with driving problems?
  - fellow tool builders?
  - providers of principles or methodologies?
- 2. is there a real need for my new approach/tool?
- 3. am I addressing a real task?
- 4. does real data exist and can I get it?

Q1. Real Users or Fellow Tool Builders?

- real users
  - target end-users intended to use tool
- fellow tool builders (FTB)
  - non-infovis person, typically from CS domain
  - wants to work with me to build a (better) tool aimed at end-users
- example:
  - data mining FTB wants to add infovis "windshield" to steerable data mining system
  - intended real users are analysts with warehouse of market-basket transaction data

Q1. Real Users or Fellow Tool Builders?

- FTB can be valuable collaborators
- but not a substitute for direct contact with real users
  - even if longstanding project
  - especially if new project
- different situation than user-centered design
  - failure to explicitly distinguish can lead to role confusion

Q2. Real Need?

- do users need a new tool/technique/approach?
  - are existing tools good enough to do the job?
  - even if not perfect from infovis research standpoint
  - some users do have infovis needs without knowing it
- is problem on the table best solved with infovis?
  - or other methods?
  - some users who ask for infovis, don't have real need
- are users willing to try new tool?
  - success is hard enough with enthusiastic end users
  - not worth uphill struggle to deal with reluctant users

Example: Power Grid Control Room Vis

- FTB collaborator conjecture: control room operators had specific problem during crisis use that infovis would solve
  - new project, just funded
  - FTB connection with real users allowed control room visit
- investigation led me to disagree
  - existing tools satisfied users, were adequate for normal use
  - plus, in midst of upgrade to new systems
    - unclear if user buyin or available data
- outcome: walked away early, before engaging in earnest

Example: Accordion Drawing For Cancer Research

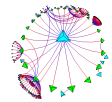
- cancer researchers looking at sequence-registered data
  - used accordion drawing infrastructure to quickly make prototype for their dataset
  - careful observation and interviews showed found that AD capabilities of showing multiple noncontiguous regions simultaneously not what they really needed
- outcome: walked away after several months, giving recommendations on specific engineering improvements to make on their existing tool
- principle: hard for domain scientists to make judgement about vis tools unless they see their own data
  - do not assume they can generalize from other data, show them theirs

Q3: Real Task - Showing the Right Structure?

- is the structure I'm showing really what they need to see?
  - or am I just showing data that's easy to gather?
  - or am I just addressing need of FTB, but not real users?
- example: showing fine-grained structure of search space
  - if user's main task is finding information, does user need to construct and maintain mental model of search space?
  - or does that add cognitive overhead, rather than reduce it?!

Examples: Showing Information Spaces

- visualize hyperlink structure of web for browsing users
  - my entry into infovis (common story!)
  - assertion of lost-in-hyperspace, without real use case
  - outcome: VRML 95 paper



- later, H3 use case was for webmasters instead of browsers
- later, semantic network vis inquiries
  - outcome: walk away very early, after initial discussion

Q3: Real Task - Will Their Need Persist?

- do they do chosen task seldom or occasionally or always?
- will they keep doing it?
- example: Constellation project
  - by the time system done, their needs had shifted
  - outcome: InfoVis 99 paper was careful design study, but could not say users had adopted

- later, with TreeJuxtaposer, pick stable task: biologists have been inspecting trees for centuries!

Q3: Real Task - Does It Exist?

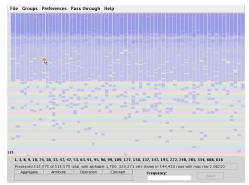
- real users, real data... but no clear questions
  - "maybe there's something interesting lurking in there"
  - hard to know if you solved problem
  - hard to learn new things about infovis
- examples: networking, security
  - outcome: nascent collaboration possibilities not pursued

Q4: Real Data - Can I Have It?

- is data proprietary?
  - many reasons for data producer to not release it
  - expose intellectual property, embarass organization
- example: data mining dashboard
  - never occurred to me to ask if real data available
    - ...because collaborator approached me
    - did not explicitly consider FTB vs. RU roles!
  - discovered DM cultural norm of synthetic data for benchmarks, only after many months into project!
  - conjecture: we're not seeing something useful because nothing to see in fake data, will change when get real data
    - continued with major effort to extend datamining server, refine and scale up nifty technique for infovis client

Q4: Real Data - Can I Have It?

- example: data mining dashboard, cont.
  - reality: could not get real data
    - eventually scrounged quasi-real data
    - alas, nifty scalable technique still didn't show anything useful
    - realized approach didn't match task 2 years into project
  - outcome: tech report



Characteristics I Look For In Collaborators

- people with driving problems
  - big data
  - clear questions
  - enthusiasm/respect for vis possibilities
- all collaborators
  - has sufficient time for the project
  - research meetings are fun
  - project has funding (ideally)

Collaboration Conclusions

- simple story misses some complexity
  - go forth and collaborate
  - three cheers for interdisciplinary research
- nuanced message: collaboration is a challenging dance
  - learning each others' language
  - finding the right people
  - finding the right problems
- big picture: often very rewarding and worthwhile
  - but keep checking that needs on both sides are being met