Methods for Visualizing Biodiversity & Building Rewarding Collaborations

Tamara Munzner

Department of Computer Science University of British Columbia

UBC Biodiversity Challenge Retreat, Hakai Institute 11 June 2019

www.cs.ubc.ca/~tmm/talks.html#hakail9-methods

DESIGNING for PEOPLE



Data Science Institute



TreeJuxtaposer:Visual tree comparison

- driving problem from UT Austin Hillis Lab in 2001: phylogenetic trees
- algorithm focus on scale, later extended to gene sequences



Ruminant 묘 NNNNNN C R CTACATO Pig 문 L. CTATATO Llama C GTACATO Rhino 뭈 G GTACATO Tapir G GTACATO Horse R ATACATO Cat ATACATO Caniform A P Pangolin G ATACATO **SequenceJuxtaposer**

https://youtu.be/GdaPj8a9QEo **TreeJuxtaposer**

joint work with: Guimbretiere, Li, Zhang, and Zhou



Cerebral: Integrating gene expression w/ interaction network

- automatic network layout by subcellular location, like hand-drawn diagrams
- multiple views with linked highlighting and navigation
- Cytoscape plugin, funded by Agilent



Aaron Barsky



Jenn Gardy (Microbio: Hancock)



Robert Kincaid (Agilent)



MizBee: Comparing genomes between species

- driving problems: Broad Inst. biologists studying fungus (Ma) and stickleback/pufferfish (Grabherr)
- two use phases: first fully validate computational pipeline, then can analyze biological questions
- investigated whole-genome duplication events, refined syntenic block construction algorithm



MizBee <u>https://youtu.be/86p7brwuz2g</u>

kleback/pufferfish (Grabherr) nalyze biological questions construction algorithm

Hanspeter Pfister (Harvard)



Miriah Meyer



Comparative functional genomics

- Pathline: multiple pathways, multiple genes, multiple species over time
 - Broad Institute, Regev Lab
 - curvemap as alternative to heatmap
- MulteeSum: all that + spatial location (cells within fruitfly embryo)
 - Harvard Med School, dePace Lab
 - compare summaries across multiple computational workflows



joint work with: Meyer, Pfister, Wong, Styczynski, dePace

CHEMITES GHOUP	Cu	rvemap
	o teo, kri Kr	
		2

Variant View: Visualizing sequence variants in genetic context

 concise overview supports reasoning about variant type & location -across several levels of biological context (vs extensive navigation w/ genome browsers)



Variant View https://youtu.be/AHDnv_qMXxQ

Joel Ferstay



Cydney Nielsen (BC Cancer)



Aggregated Dendrograms: Visual comparison between many phylogenetic trees

- concisely summarize trees interactively wrt bio meaningful criteria
 - one use case: compare gene trees to species trees



Aggregated Dendrograms https://youtu.be/2SLcz7KNLJw

ny phylogenetic trees aningful criteria

Zipeng Liu



Shing Hei Zhan



Vismon: Fisheries simulation

supporting decision-makers not expert in simulation & stats

- sensitivity analysis, global trade-offs analysis, staged uncertainty



Vismon https://youtu.be/h0kHoS4VYmk

Maryam Booshehrian



Torsten Moeller (SFU)



Integrating visualization & biostats methods

- Human-centered design & qualitative coding
- Epidemiology/health expectations & constraints
- Mixed initiative: automation and manual analysis
- Mixed methods: when to use qual & when to use quant



GEViT: Genomic Epidemiology Visualization Typology https://gevit.net

Anamaria Crisan



Jenn Gardy BCCDC/SPPH



Evidence-Based Design and Evaluation of a Whole Genome Sequencing Clinical Report for the Reference Microbiology Laboratory



A Nested Model

for Visualization Design and Validation

http://www.cs.ubc.ca/labs/imager/tr/2009/NestedModel

A Nested Model for Visualization Design and Validation. Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), 15(6):921-928, 2009.

n	
action	
oding/interaction idiom	
ım	

Nested model: Four levels of visualization design

• domain situation

- -who are the target users?
- abstraction
 - -translate from specifics of domain to vocabulary of visualization
 - what is shown? data abstraction
 - why is the user looking at it? **task** abstraction
- idiom
 - -**how** is it shown?
 - visual encoding idiom: how to draw
 - **interaction** idiom: how to manipulate
- algorithm

[A Multi-Level Typology of Abstract Visualization Tasks Brehmer and Munzner. IEEETVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

– efficient computation



[A Nested Model of Visualization Design and Validation. Munzner. IEEETVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]



Different threats to validity at each level

Domain situation
You misunderstood their needs

Data/task abstraction You're showing them the wrong thing

Wisual encoding/interaction idiom The way you show it doesn't work

Algorithm Your code is too slow



Interdisciplinary: need methods from different fields at each level

• mix of qual and quant approaches (typically)

anthropology/	Domain situation Observe target users using existing tools	qual
ethnography	Data/task abstraction	
design	Visual encoding/interaction idiom Justify design with respect to alternatives	qual
computer science	Algorithm Measure system time/memory Analyze computational complexity	quant
psychology	Analyze results qualitatively Measure human time with lab experiment (<i>lab study</i>	qual) quant
anthropology/	Observe target users after deployment (field study)	qual
ethnography	Measure adoption	quant



technique-driven work

Mismatches: Common problem

L Domain situation

Observe target users using existing tools

Data/task abstraction

Wisual encoding/interaction idiom Justify design with respect to alternatives

Algorithm

Measure system time/memory Analyze computational complexity

Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

Observe target users after deployment (*field study*)

Measure adoption

benchmarks can't confirm design

lab studies can't confirm task abstraction

Problem-driven collaborations

- working with domain scientists
- translating from domain-specific language
 - -how to pull this off?

problem-driven work







Building Rewarding Collaborations



Design Study Methodology

Reflections from the Trenches and from the Stacks

http://www.cs.ubc.ca/labs/imager/tr/2012/dsm/

Design Study Methodology: Reflections from the Trenches and from the Stacks. SedImair, Meyer, Munzner. IEEE Trans. Visualization and Computer Graphics 18(12): 2431-2440, 2012 (Proc. InfoVis 2012).

Michael SedImair



Miriah Meyer





Tamara Munzner @tamaramunzner



Lessons learned from the trenches: 21 between us



Cerebral genomics



MizBee genomics



Pathline genomics



MulteeSum genomics



Vismon fisheries management



MostVis in-car networks



Car-X-Ray in-car networks



ProgSpy2010 in-car networks



RelEx in-car networks



Cardiogram in-car networks



Constellation linguistics



LibVis cultural heritage



Caidants multicast

"(her Took Populatio	 Tanera' Tesk Reputation 	"West-Ch" Task Propulation
1 30 30 300 300 30		
AND DESCRIPTION	1 (A) (A) (A)	Aug. 1
1.000		
1 S-	- E	S
The Real Property lies	Contraction of the second	Contract of the second
- proved	a - provide a state	" The state of the second second
in the second se	() · · · · · · · · · · · · · · · · · ·	
-	and the second division of the second divisio	an or a second second second
1		and the second s
1 1		
		have been a second
 Effetismine 		(1 solu solu
A CONTRACTOR OF A CONTRACTOR OFTA CONTRACTOR O	and the P. P. R. Street and	AND A DESCRIPTION OF THE OWNER.

SessionViewer web log analysis

and a	1.000	1000	-	-			
-	- 100		- 3	har.	Ş.,	120	1
			5		3	10	
-	- Anna		-	1		-	
3	-						
- 2		- 1			-	-	ł
-		_					
		-				_	

LiveRAC server hosting



PowerSetViewer data mining





QuestVis sustainability



WiKeVis in-car networks



AutobahnVis in-car networks



VisTra in-car networks



LastHistory music listening

Methodology for problem-driven work

• definitions

• 9-stage framework

• 32 pitfalls & how to avoid them

comparison to related methodologies







dvance: jumping forward over stages	general
art: insufficient knowledge of vis literature	learn
ommitment: collaboration with wrong people	winnow
available (yet)	winnow
time available from potential collaborators	winnow
visualization: problem can be automated	winnow
xpertise does not match domain problem	winnow
research: engineering vs. research project	winnow
change: existing tools are good enough	winnow

Design study methodology: 32 pitfalls

• and how to avoid them

PF-1	premature advance: jumping forward over stages	general
PF-2	premature start: insufficient knowledge of vis literature	learn
PF-3	premature commitment: collaboration with wrong people	W1NNOW
PF-4	no real data available (yet)	WINNOW
PF-5	insufficient time available from potential collaborators	winnow
PF-6	no need for visualization: problem can be automated	winnow
PF-7	researcher expertise does not match domain problem	winnow
PF-8	no need for research: engineering vs. research project	winnow
PF-9	no need for change: existing tools are good enough	winnow



considerations







... or maybe a fellow tool builder?

METAPHOR Winnowing











initial conversation further meetings requirements analysis

full collaboration





Design study methodology: 32 pitfalls

• and how to avoid them

PF-1	premature advance: jumping forward over stages	general
PF-2	premature start: insufficient knowledge of vis literature	learn
PF-3	premature commitment: collaboration with wrong people	winnow
PF-4	no real data available (yet)	winnow
PF-5	insufficient time available from potential collaborators	winnow
PF-6	no need for visualization: problem can be automated	Winnow
PF-7	researcher expertise does not match domain problem	W1NNOW
PF-8	no need for research: engineering vs. research project	winnow
PF-9	no need for change: existing tools are good enough	winnow

Design study methodology: definitions



More Information

tamaramunzner

- this talk https://www.cs.ubc.ca/~tmm/ talks.html#hakai19-methods
- papers, videos, software, talks, courses http://www.cs.ubc.ca/group/infovis http://www.cs.ubc.ca/~tmm



MEMBERS

RECENT NEWS

Workshop.

9/2018 [PARE]:

8/2018 [evel]:

[paper]





America Cristan





ichael Copermann



using Gaze Data was accepted to ETVIS 2018. [pre-print PDF]

Shannah Fisher







05/2018 [types]:



















- Aggregated dendrograms for visual comparison between many phylogenetic trees.
- by Zipeng Liu, Shing Hei Zhan, and Tamara Munzher was accepted at IEEE Transactions on Visualization and Computer Graphics



- At the BELIV Workshop, Anamaria Crisan and Madison Elliott will present the paper "How to evaluate an evaluation study? Comparing and contrasting practices in vis with those of other disciplines".
- Michael Oppermann will give a talk entitled "Uncovering Spatiotemporal Dynamics from Non-Trajectory Data" at the Urban Data Visualization
- Tamara Kunzher will participate in a panel at the VisGuides Workshop. We're co-hosting the (North) West Coast Party, on Thursday night.
- A systematic method for surveying data visualizations and a resulting genomic epidemiology visualization typology: GEVIT
- by Anamaria Crisan, Jennifer L. Gardy, and Tamara Munzher was published in Oxford Bioinformatics
- Adjutant: an R-based tool to support topic discovery for systematic and Iterature reviews
- by Anamaria Crisan, Tamara Munzner, and Jennifer L. Bardy was published in Oxford Dipinformatics



- GaRSIVis: Improving the Predicting of Self-Interruption during Reading
- by Jan Pilzer, Shareen Mahmud, Vanessa Putnam, and Tamara Munzner.







