

Visualization: Abstractions & Idioms

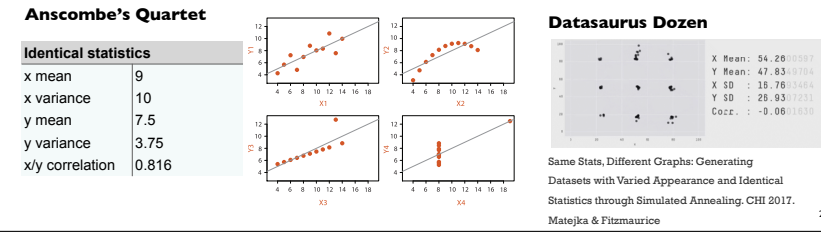
Tamara Munzner
 Department of Computer Science
 University of British Columbia
 @tamaramunzner
 University of Coimbra Guest Lecture
 9 Mar 2022, virtual / Coimbra, Portugal
<http://www.cs.ubc.ca/~tmm/talks.html#coimbra22>



Visualization defined & motivated

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- suitable when human in the loop needs details
 - interplay between human judgement and automatic computation

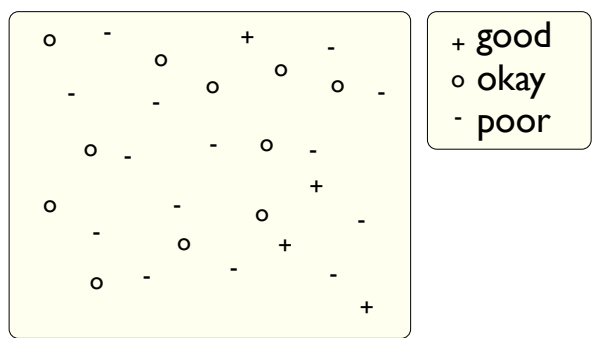


Why focus on tasks and effectiveness?

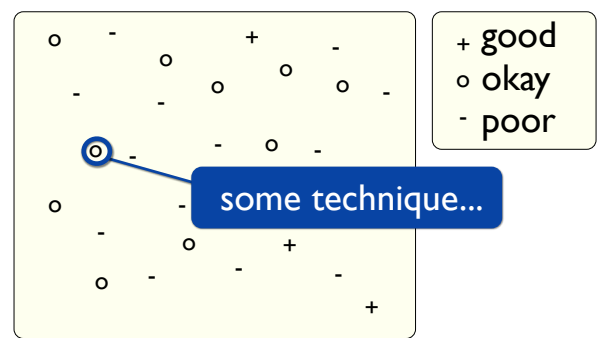
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- effectiveness requires match between data/task and representation
 - set of representations is huge
 - many are ineffective mismatch for specific data/task combo
 - increases chance of finding good solutions if you understand full space of possibilities

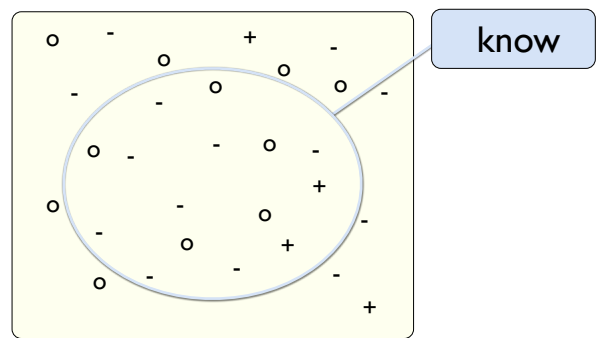
Metaphor: Design space



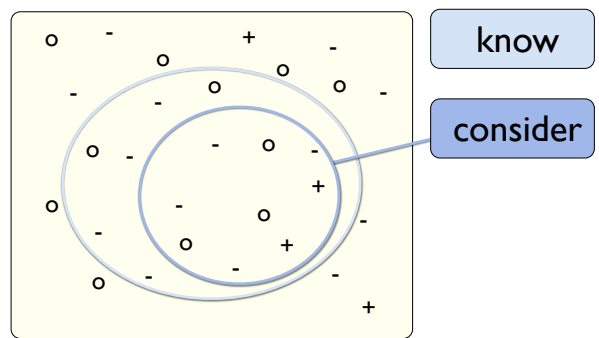
Metaphor: Design space



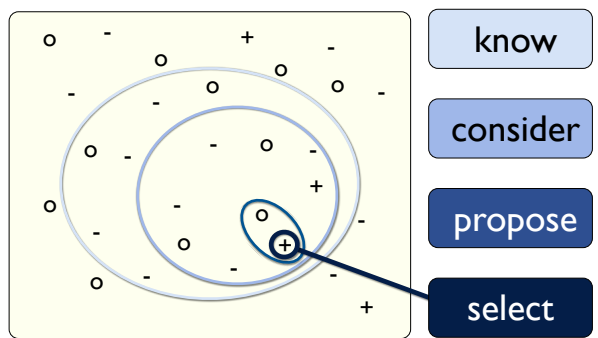
Metaphor: Design space



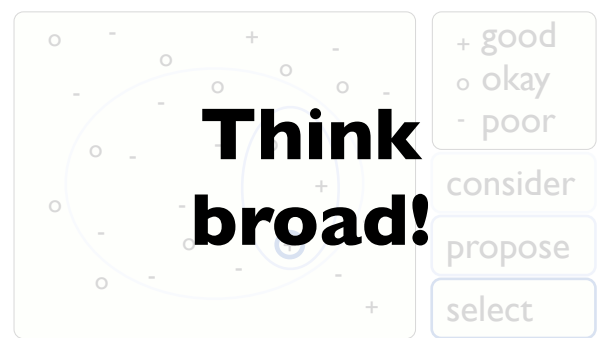
Metaphor: Design space



Metaphor: Design space



Metaphor: Design space

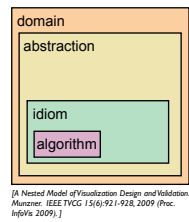


Why focus on tasks and effectiveness?

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

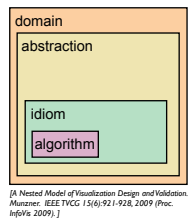
- effectiveness requires match between data/task and representation
 - set of representations is huge
 - many are ineffective mismatch for specific data/task combo
 - increases chance of finding good solutions if you understand full space of possibilities
- what counts as effective?
 - novel: enable entirely new kinds of analysis
 - faster: speed up existing workflows
- how to validate effectiveness
 - many methods, must pick appropriate one for your context

Nested model: Four levels of visualization concerns



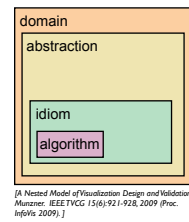
Nested model: Four levels of visualization concerns

- domain situation
 - who are the target users?



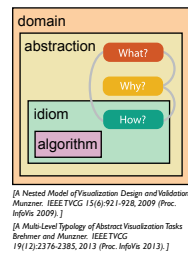
Nested model: Four levels of visualization concerns

- domain situation
 - who are the target users?
- abstraction
 - translate from specifics of domain to vocabulary of vis



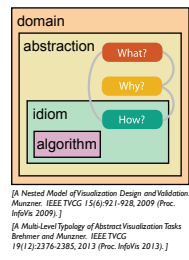
Nested model: Four levels of visualization concerns

- domain situation
 - who are the target users?
- abstraction
 - translate from specifics of domain to vocabulary of vis
 - what is shown? **data abstraction**



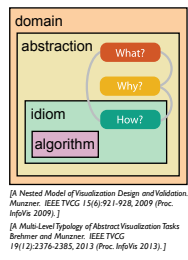
Nested model: Four levels of visualization concerns

- domain situation
 - who are the target users?
- abstraction
 - translate from specifics of domain to vocabulary of vis
 - what is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form



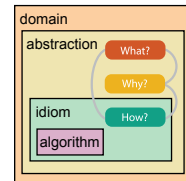
Nested model: Four levels of visualization concerns

- domain situation
 - who are the target users?
- abstraction
 - translate from specifics of domain to vocabulary of vis
 - what is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form
 - why is the user looking at it? **task abstraction**



Nested model: Four levels of visualization concerns

- **domain** situation
 - **who** are the target users?
- **abstraction**
 - translate from specifics of domain to vocabulary of vis
 - **what** is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form
 - **why** is the user looking at it? **task abstraction**
- **idiom**
 - **how** is it shown?

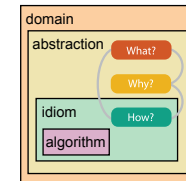


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]
 [A Multi-Level Typology of Abstract Visualization Tasks. Behner and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

17

Nested model: Four levels of visualization concerns

- **domain** situation
 - **who** are the target users?
- **abstraction**
 - translate from specifics of domain to vocabulary of vis
 - **what** is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form
 - **why** is the user looking at it? **task abstraction**
- **idiom**
 - **how** is it shown?
 - **visual encoding idiom:** how to draw

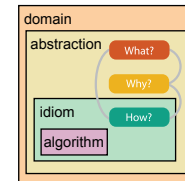


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]
 [A Multi-Level Typology of Abstract Visualization Tasks. Behner and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

18

Nested model: Four levels of visualization concerns

- **domain** situation
 - **who** are the target users?
- **abstraction**
 - translate from specifics of domain to vocabulary of vis
 - **what** is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form
 - **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

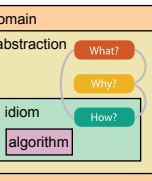


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]
 [A Multi-Level Typology of Abstract Visualization Tasks. Behner and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

19

Nested model: Four levels of visualization concerns

- **domain** situation
 - **who** are the target users?
- **abstraction**
 - translate from specifics of domain to vocabulary of vis
 - **what** is shown? **data abstraction**
 - often don't just draw what you're given: transform to new form
 - **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**
 - efficient computation



[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]
 [A Multi-Level Typology of Abstract Visualization Tasks. Behner and Munzner. IEEE TVCG 19(12):2376-2385, 2013 (Proc. InfoVis 2013).]

20

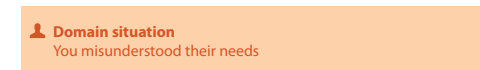
Why is validation difficult?

- different ways to get it wrong at each level

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

Why is validation difficult?

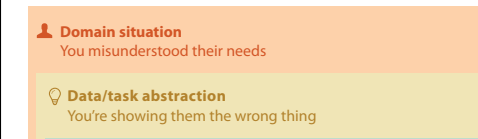
- different ways to get it wrong at each level



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

Why is validation difficult?

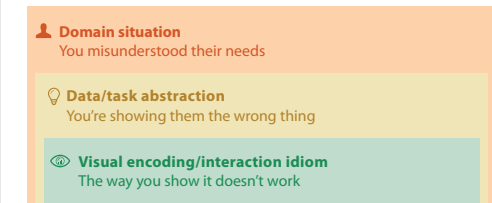
- different ways to get it wrong at each level



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

Why is validation difficult?

- different ways to get it wrong at each level



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

24

Why is validation difficult?

- different ways to get it wrong at each level

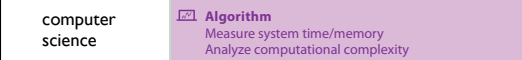


[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: use methods from appropriate fields at each level

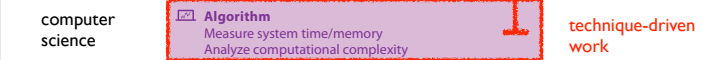
[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: use methods from appropriate fields at each level



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

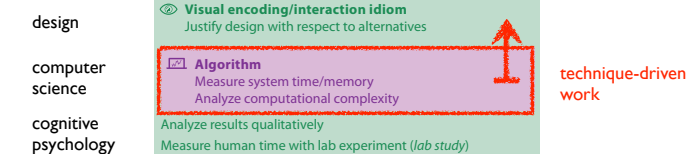
Validation solution: use methods from appropriate fields at each level



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

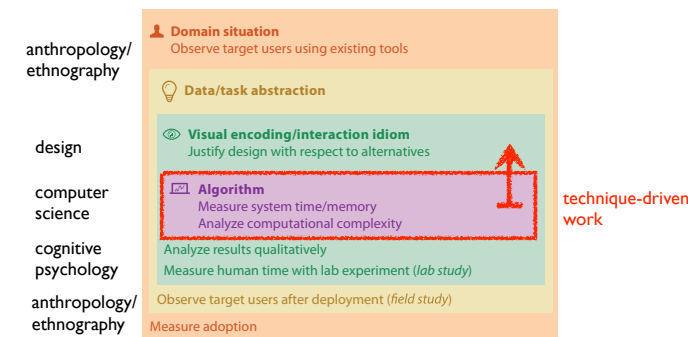
28

Validation solution: use methods from appropriate fields at each level



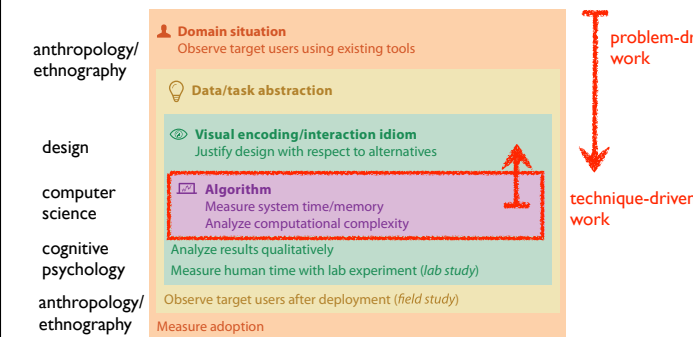
[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

Validation solution: use methods from appropriate fields at each level



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

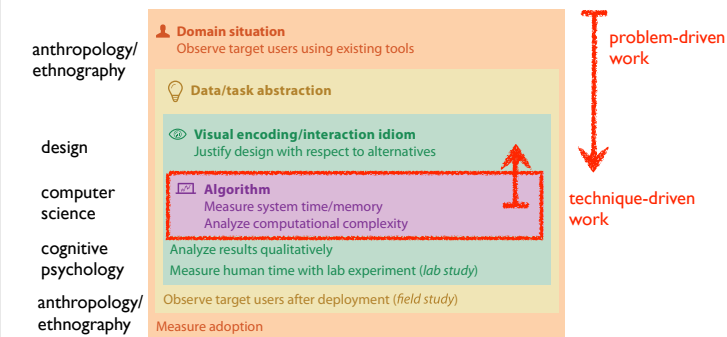
Validation solution: use methods from appropriate fields at each level



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

Validation solution: use methods from appropriate fields at each level

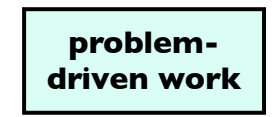
- avoid mismatches between level and validation



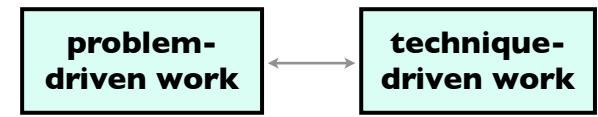
[A Nested Model of Visualization Design and Validation. Munzner. IEEE TVCG 15(6):921-928, 2009 (Proc. InfoVis 2009).]

32

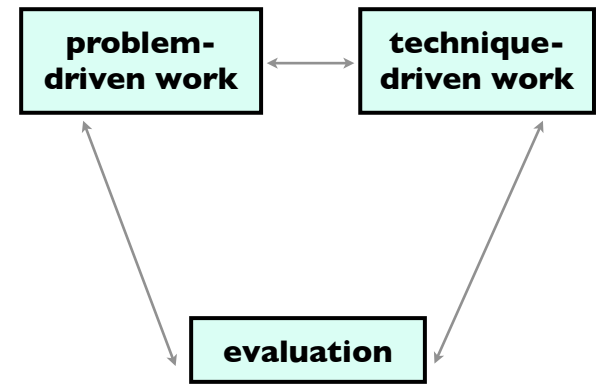
Visualization: Angles of attack



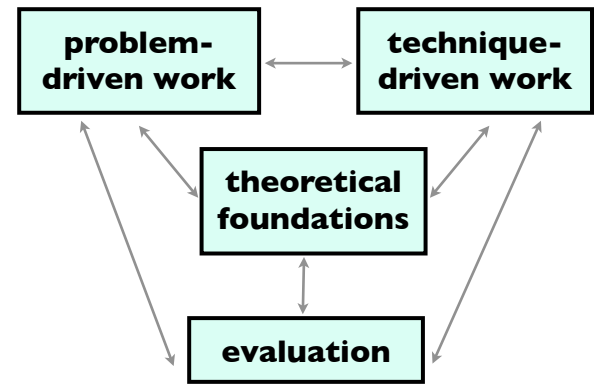
Visualization: Angles of attack



Visualization: Angles of attack



Visualization: Angles of attack



Three case studies: Abstractions & idioms

- e-commerce
- facilities management
- biology

e-commerce by shaashank singh from the Noun Project Biology by lezar tartiblar from the Noun Project Business by Colourcreastype from the Noun Project

Three case studies: Abstractions & idioms

- e-commerce
- facilities management
- biology

e-commerce by shaashank singh from the Noun Project Biology by lezar tartiblar from the Noun Project Business by Colourcreastype from the Noun Project

Kim Dextras-Romagnino

Segmentifier

Interactive Refinement of Clickstream Data

<http://www.cs.ubc.ca/labs/imager/tr/2019/segmentifier>

Segmentifier: Interactive Refinement of Clickstream Data.
Dextras-Romagnino and Munzner. Computer Graphics Forum (Proc. EuroVis 2019) 38(3):623–634 2019

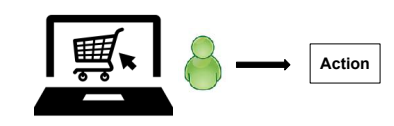
E-commerce: mobile apps for large companies

What are the Data and Task Abstractions for Clickstream Data Analysis?

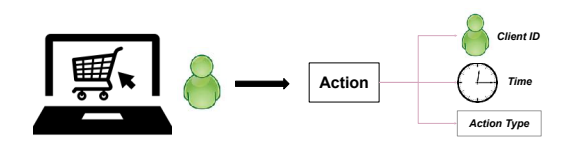
- Clickstream Data
- Clickstream Analysis Tasks
- Segmentifier Analysis Model

What is Clickstream Data?

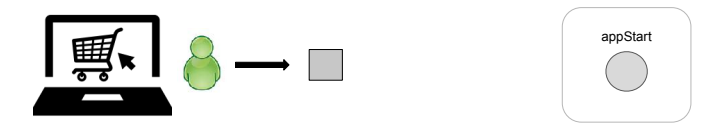
Data: Actions



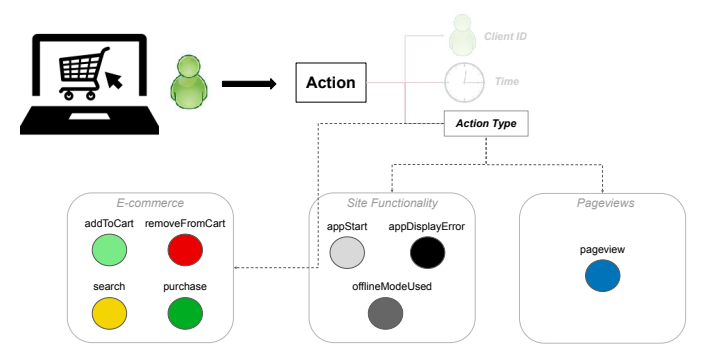
Data: Action Attributes



Data: Sequences



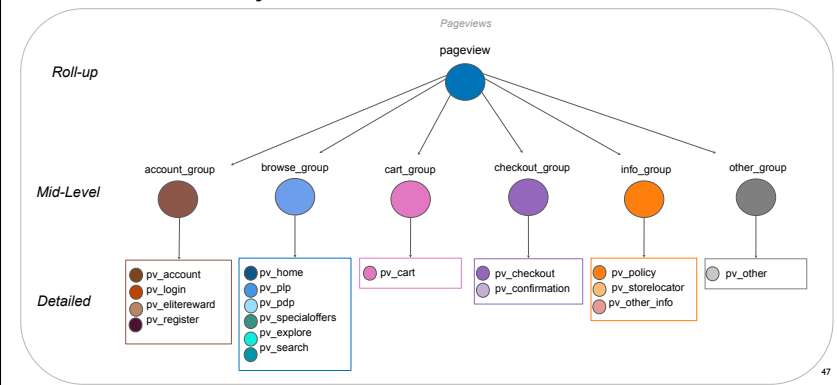
Data: Action Types



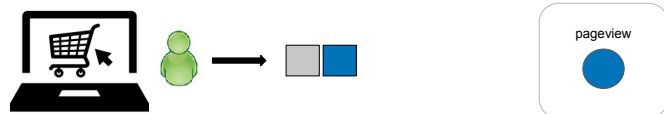
Action Hierarchy



Action Hierarchy

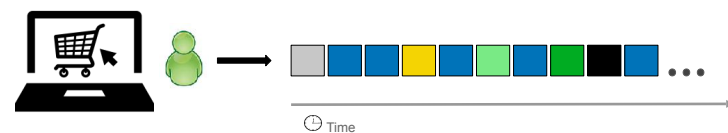


Data: Sequences



49

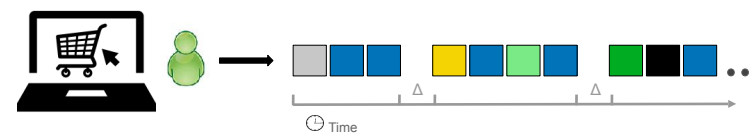
Data: Client Sequences



Client Sequences: all actions performed by a single user

50

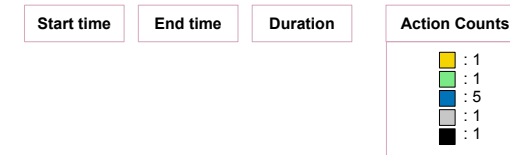
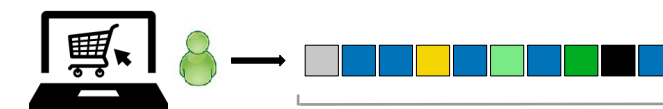
Data: Session Sequences



Session Sequences: all actions performed by a single user within a defined amount of time (Δ) from each other. Δ is usually 30 min.

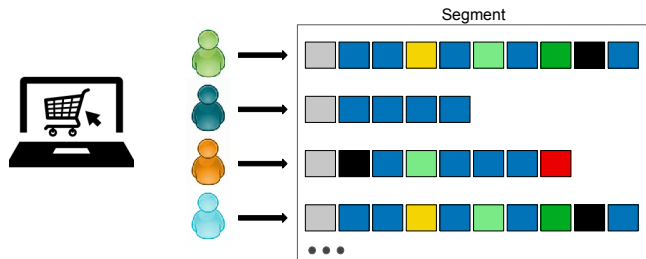
51

Data: Sequence Attributes



52

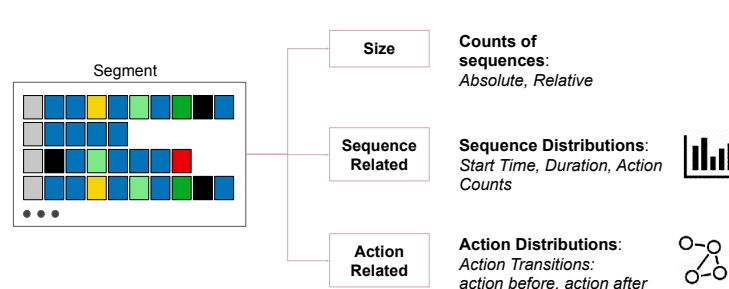
Data: Segments



Segment: any set of sequences

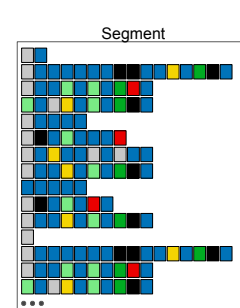
53

Data: Segment Attributes



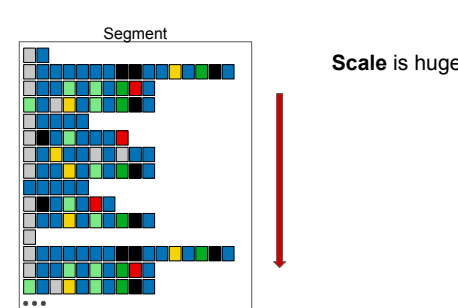
54

Real-world Clickstream Data



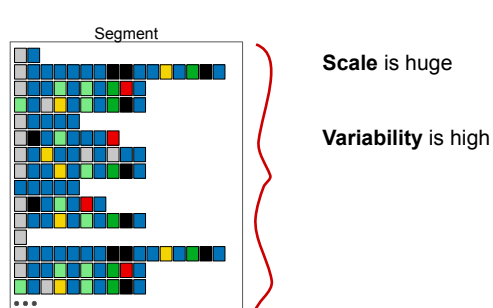
55

Real-world Clickstream Data



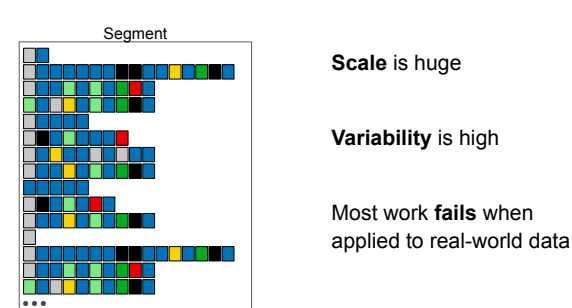
56

Real-world Clickstream Data



57

Real-world Clickstream Data



58

What are
Clickstream Data Analysis Tasks?

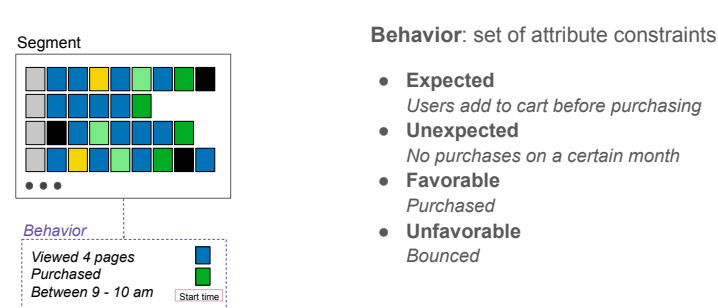
59

Tasks: Segment Behavior



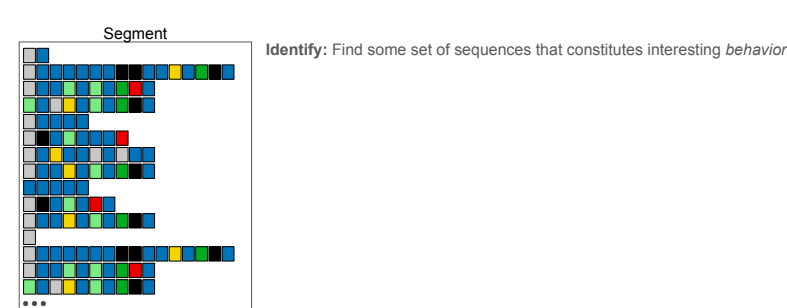
60

Tasks: Segment Behavior



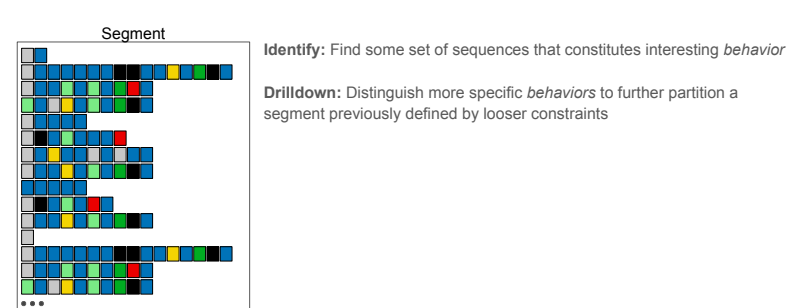
61

Tasks: Task Abstraction



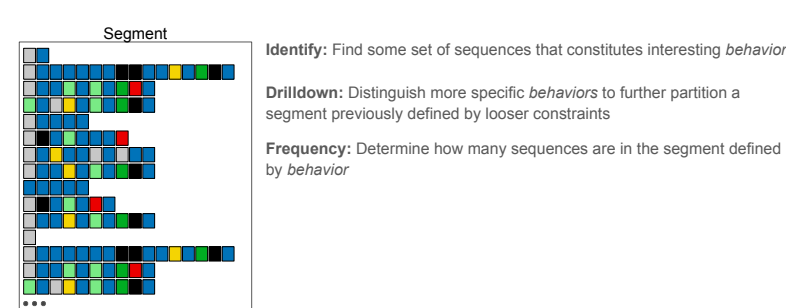
62

Tasks: Task Abstraction



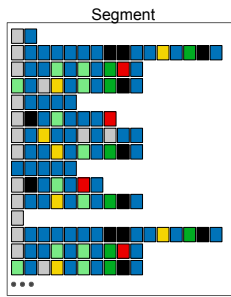
63

Tasks: Task Abstraction



64

Tasks: Task Abstraction



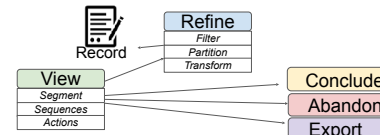
- Identify:** Find some set of sequences that constitutes interesting *behavior*
- Drilldown:** Distinguish more specific *behaviors* to further partition a segment previously defined by looser constraints
- Frequency:** Determine how many sequences are in the segment defined by *behavior*
- Ordering** within sequence: Match if one action subsequence occurs before (or after) another action subsequence in a sequence

65

High-Level Segmentifier Analysis Model

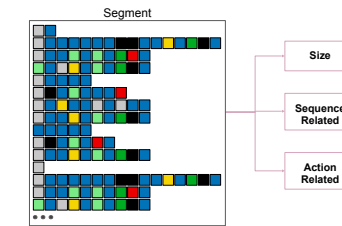
High-Level Segmentifier Analysis Model

- Abstraction above task/data level to provide design rationale
- Take a *giant, noisy dataset* and refine it into *small, clean segments* for
 - actionable insights
 - downstream analysis
- Bridge the gap between *real-world data* and other techniques



66

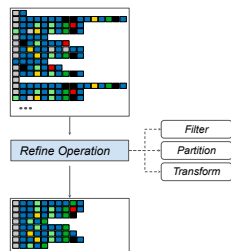
High-Level Segmentifier Analysis Model



- Gives Insight into underlying data of segment
 - o Action Attributes
 - o Sequence Attributes
 - o Segment Attributes
- Leads to:
 - o Insights
 - o New ways on how to *refine*
 - o Whether segment should be *abandoned*
 - o Whether segment should be *exported*

67

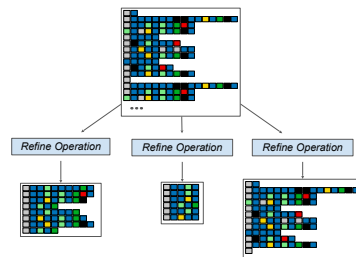
High-Level Segmentifier Analysis Model



- Apply operation to create new segments
- Type of Refinements
 - o Filter
 - o Partition
 - o Transform

68

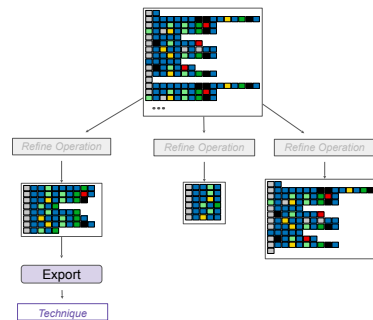
High-Level Segmentifier Analysis Model



- Record all refinement steps automatically
- Keep track of questions asked and hypotheses tested
- Ability to create and view multiple segments from the same segment

69

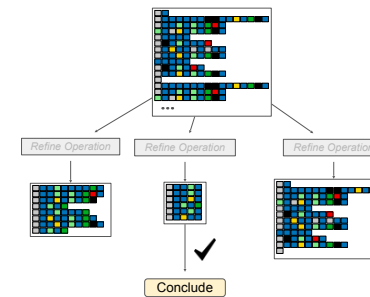
High-Level Segmentifier Analysis Model



- Export refined segments for further downstream analysis, to more specific tools:
 - o Pattern mining
 - o Clustering

70

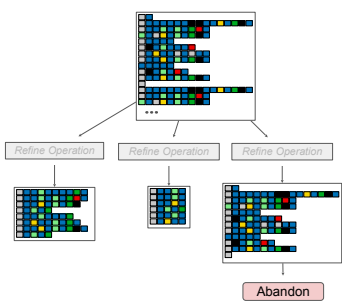
High-Level Segmentifier Analysis Model



- Discover actionable insight by *viewing* segment

71

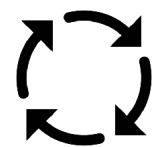
High-Level Segmentifier Analysis Model



- By *viewing* the segment, analyst *abandons* if:
 - o No actionable insights
 - o No further ways to *refine*
 - o Not suitable for *export*

72

Why Visual Analytics?



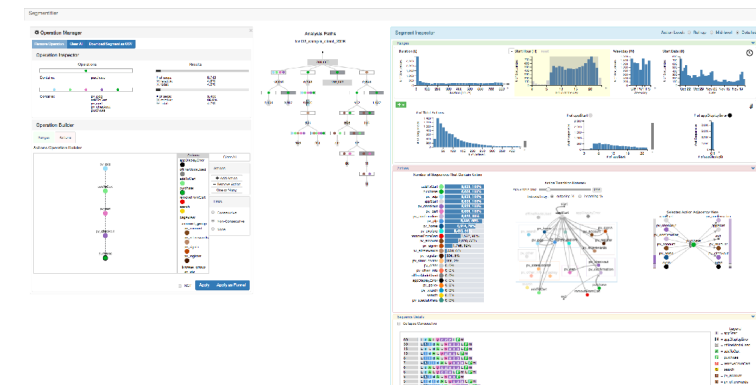
- Automation would be nice...
 - o Put data in, actionable results appear
 - o ... but it is not realistic
 - o Many possible questions, data-driven interplay between finding answers and generating new questions
- Human-in-the-loop visual data analysis
 - o Integrate computing power of machine with intuition of domain experts

73

Solution

74

The Segmentifier Interface



75

Video

Segmentifier: Interactively Refining Clickstream Data into Actionable Segments

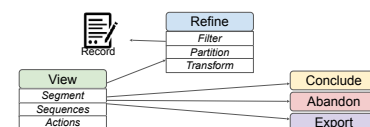


<https://www.youtube.com/watch?v=TobYDFelSOg>

76

Segmentifier Contributions

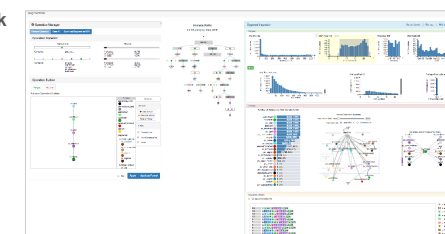
- Thorough characterization of task and data abstraction for clickstream data analysis



77

Segmentifier Contributions

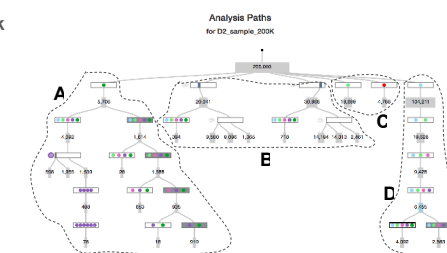
- Thorough characterization of task and data abstraction for clickstream data analysis
- **Segmentifier: novel analytics interface** for refining data segments and viewing characteristics before downstream fine-grained analysis



78

Segmentifier Contributions

- Thorough characterization of task and data abstraction for clickstream data analysis
- **Segmentifier: novel analytics interface** for refining data segments and viewing characteristics before downstream fine-grained analysis
- Preliminary evidence of utility



79

Three case studies of problem-driven work

• e-commerce



• facilities management



• biology



© commerce by shashank singh from the Noun Project Biology by Izzar Tantular from the Noun Project Business by Colourcreastype from the Noun Project



Ocupado

Visualizing Location-Based Counts Over Time Across Buildings

<http://www.cs.ubc.ca/labs/imager/tr/2020/ocupado/>

Ocupado: Visualizing Location-Based Counts Over Time Across Buildings.
Oppermann and Munzner. Computer Graphics Forum (Proc. EuroVis 2020) 39(3):127-138 2020.

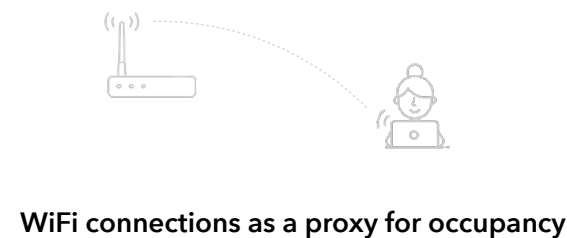
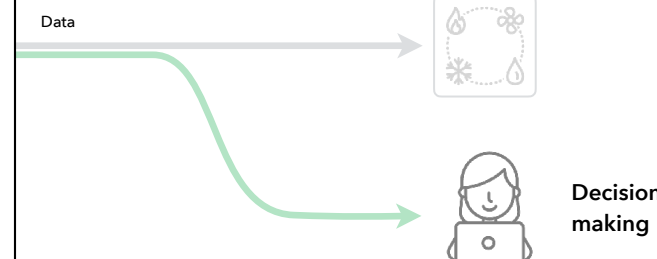
Video



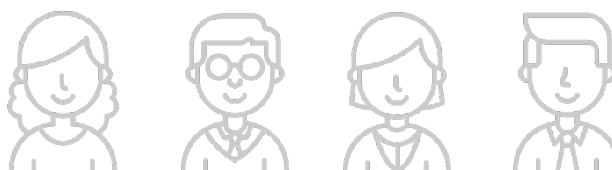
<https://www.youtube.com/watch?v=KcwIVK8eUdw>

Location-Based Counts

- › Regular intervals (e.g., every 5 minutes)
- › Spatial hierarchy (Zone → Floor → Building → Campus)
- › No trajectories or device identifiers are recorded
- › Intrinsic privacy advantages



Interviews with potential stakeholders



Focus Domains

- › Space planning
- › Building management
- › Custodial services
- › Classroom management
- › Data quality control

Focus Domains

- › Space planning
- › Building management
- › Custodial services
- › Classroom management
- › Data quality control



Tasks

- ✓ Confirm assumptions or previous observations.
Do students occupy room x in evenings or on weekends?



Tasks

- ✓ Confirm assumptions or previous observations.
- ✓ Monitor the current/recent utilization rate.
Which rooms are empty/busy?



Tasks

- ✓ Confirm assumptions or previous observations.
- ✓ Monitor the current/recent utilization rate.
- ✓ Communicate space usage and justify decisions.
Space usage improved after renovation.



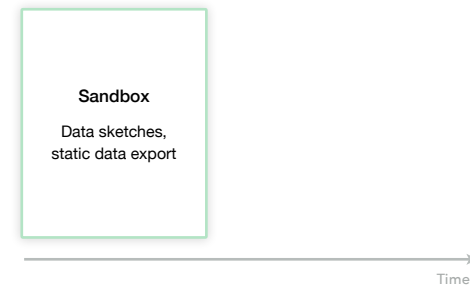
Tasks

- ✓ Confirm assumptions or previous observations.
- ✓ Monitor the current/recent utilization rate.
- ✓ Communicate space usage and justify decisions.
- ✓ Validate the data (quality control).
Check minimum size of a room that can be captured.

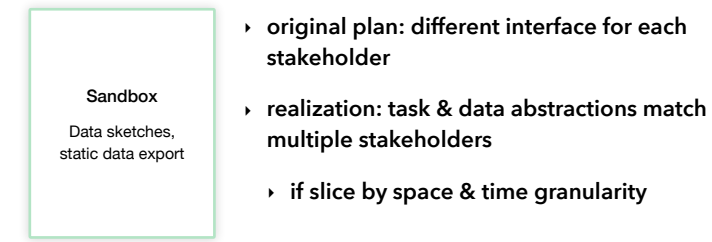


Spatial and Temporal Data Granularities

Visualization Prototypes



Visualization Prototypes

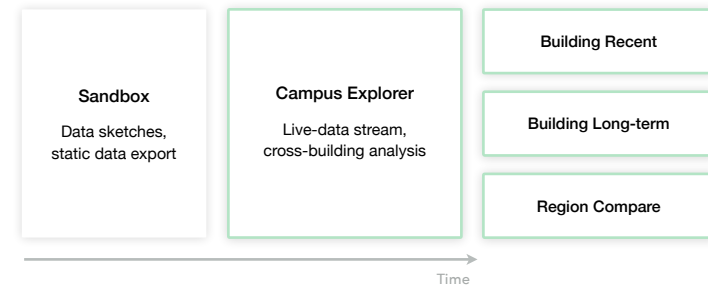


Spatial and Temporal Data Granularities

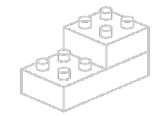


Spatial and Temporal Data Granularities

Visualization Prototypes



Reusable Visualization Components



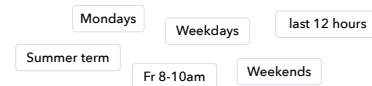
Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)

Regions of interest



Periods of interest



Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)
	Box-plot-bars	Juxtaposition	Repeating patterns, trends, outliers (non-contiguous)

Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)
	Box-plot-bars	Juxtaposition	Repeating patterns, trends, outliers (non-contiguous)
	Confidence band line chart	Aggregation	Typical utilization profiles

Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)
	Box-plot-bars	Juxtaposition	Repeating patterns, trends, outliers (non-contiguous)
	Confidence band line chart	Aggregation	Typical utilization profiles
	Superimposed line chart	Superposition	Within-session patterns, outliers

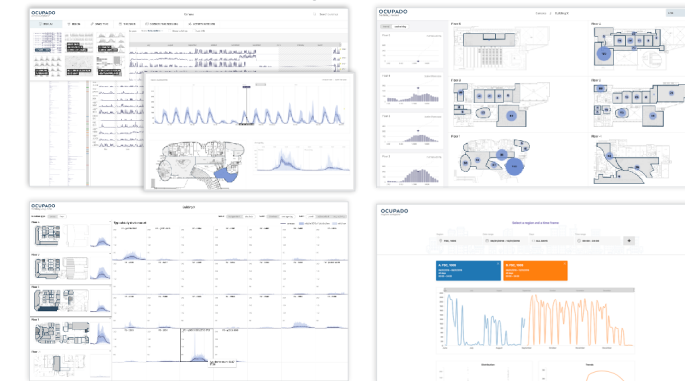
Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)
	Box-plot-bars	Juxtaposition	Repeating patterns, trends, outliers (non-contiguous)
	Confidence band line chart	Aggregation	Typical utilization profiles
	Superimposed line chart	Superposition	Within-session patterns, outliers

Reusable Visualization Components

Layout	Visual Encoding	Facet	Comparisons
	Sparkline	Juxtaposition	Repeating patterns, trends, outliers (contiguous)
	Box-plot-bars	Juxtaposition	Repeating patterns, trends, outliers (non-contiguous)
	Confidence band line chart	Aggregation	Typical utilization profiles
	Superimposed line chart	Superposition	Within-session patterns, outliers
	Floor plan with symbols	Superposition	Within local spatial neighborhood
	Spatial heatmap	Containment (nested)	Across distributed regions

Ocupado Interfaces



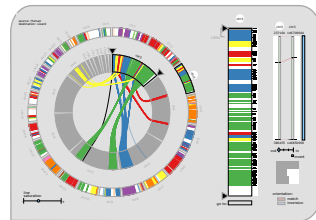
Ocupado Contributions

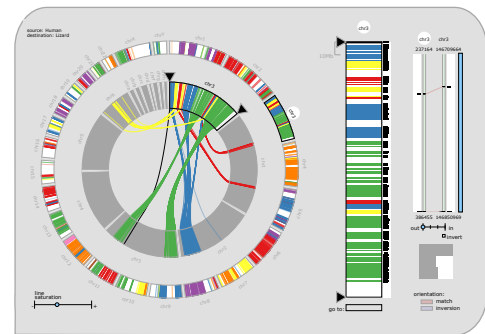
- Analysis and abstraction of data and tasks for studying space utilization
- Ocupado, a set of visual decision support tools
- Generalizable design choices for visualizing non-trajectory spatiotemporal data relating to large-scale indoor environments

MizBee

A Multiscale Synteny Browser

joint work with:
Miriah Meyer, Hanspeter Pfister
<http://www.cs.utah.edu/~miriah/mizbee>



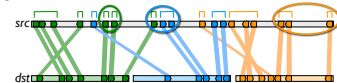


<https://www.youtube.com/watch?v=8Sp7brwuz2g>

113

What: Data abstraction

- data: multiscale lists
 - features: hundreds of thousands
 - ordered attribute: position in chromosome sequence coordinates
 - categorical attributes: orientation, chromosome of matching feature
 - quantitative attributes: length, similarity score
 - syntenic blocks: thousands
 - contiguous sets of features on same chromosome
 - combine thresholded features if
 - destination chromosome and orientation match
 - close together
 - chromosomes: dozens
 - genomes: two



114

Why: Tasks in domain language

- analyze conservation (similarity) relationships between genomic features
 - high-level biology questions
 - evolution
 - how long ago did two species share common ancestor?
 - function
 - which segment of the genome is responsible for specific function in the cell?
 - ...
 - low-level data-centric questions
 - algorithm refinement
 - are paired features within a block contiguous?
 - which chromosomes share conserved blocks?
 - are similarity scores alike within block?
 - ...

115

Why: Tasks abstraction

relationship scale

- relationship types: proximity, size, orientation, similarity
- data scales: genome, chromosome, block, feature
- topics: algorithm in/out, block reliability, high-level science

Which chromosomes share conserved blocks?		x				x			
For one chromosome, how many other chromosomes does it share blocks with?	x	x				x			
What is the density of coverage and where are the gaps on chromosomes? blocks?	x	x	x			x			
Where are the blocks on chromosomes? around a specific location on a chromosome?	x	x				x			
What are the sizes and locations of other genomic features near a block?		x				x		x	
How large are the blocks?		x				x		x	
Do neighboring blocks go to the same chromosomes? relative location on a chromosome?	x	x				x			
Are the orientations matched or inverted for block pairs? feature pairs?		x	x			x			x
Do the orientations match for pairs of neighboring blocks? features within a block?		x	x			x			x
Are similarity scores alike: with respect to neighboring blocks? within a block?		x	x			x			x
Are the paired features within a block contiguous?						x		x	
How large is a feature relative to other genes within a block?						x		x	
What are the sizes, locations, and names of features within a block?						x		x	
What are the differences between individual nucleotides of feature pairs?							x		x

116

How: Idiom design choices

- encode match relationships between chromosome segments with both
 - color
 - src
 - dst
 - connection marks
 - src
 - dst

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape
- Marks As Links
- Containment
- Connection

117

How: Arrange space

- design space of arrangements

Mauve [Darling04], Cinteny, MizBee, Apollo [Lewis02]

118

How: Idiom design choices

- juxtapose linked views
 - multiform overview-detail
 - three views: genome, chromosome, block
 - different visual encoding in each

Facet

Juxtapose

Encoding	Same	All	Subset	None
	Different	Redundant	Overview/Detail	Small Multiples
		Multiform	Multiform, Overview/Detail	No Linkage

119

How: Idiom design choices

- axis orientation
 - radial: genome
 - rectilinear: chromosome, block
 - aligned position more accurate than angle

Arrange

Axis Orientation

- Rectilinear
- Parallel
- Radial

Magnitude Channels: Ordered Attributes

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle

120

How: Idiom design choices

- filter

Reduce

Filter

121

How: Idiom design choices

- outer ring: summarize relationships with color
 - select one chromosome from set of source chromosomes
- inner ring:
 - destination chromosomes around copy of selected source chromosome
 - show relationship details with connection marks as well as color

Manipulate

- Select
- Actions
- Query
- Identify
- Compare
- Summarise

122

MizBee contributions

- first syntenic browser with side-by-side linked views
 - across the range of scales
 - encoding all four conservation relationship types
 - proximity, size, orientation, similarity
- open source
 - <http://www.cs.utah.edu/~miriah/mizbee>

123

Visualization: Abstractions & idioms

- levels of design
 - identify abstractions
 - crucial & difficult, iterative process
 - select appropriate idioms
 - or create new ones if necessary
- three examples
 - different domains
 - different abstractions
 - different idioms

124

More information

- theoretical foundations: book (+ tutorial/course lecture slides)
 - <http://www.cs.ubc.ca/~tmm/vadbook>
- papers, videos, software, talks, courses
 - <http://www.cs.ubc.ca/group/infovis>
 - <http://www.cs.ubc.ca/~tmm>
- this talk
 - <http://www.cs.ubc.ca/~tmm/talks.html#coimbra22>

Visualization Analysis and Design. Munzner. AK Peters Visualization Series. CRC Press, 2014.

@tamaramunzner

UBC UBC

DESIGNING for PEOPLE

CAIDA

125