

Visualizations for Preference Inspection in Group Decision Making

Jordon Johnson

Department of Computer Science
University of British Columbia

in collaboration with

Emily Hindalong

Giuseppe Carenini

Tamara Munzner

Motivation and background

Group decision making and preference inspection



Group Decision Making is everywhere

- Family choosing where to go for **dinner**
- Software company choosing a **technology stack**
- University department hiring new **faculty members**
- Country selecting sites for new **marine protected areas**



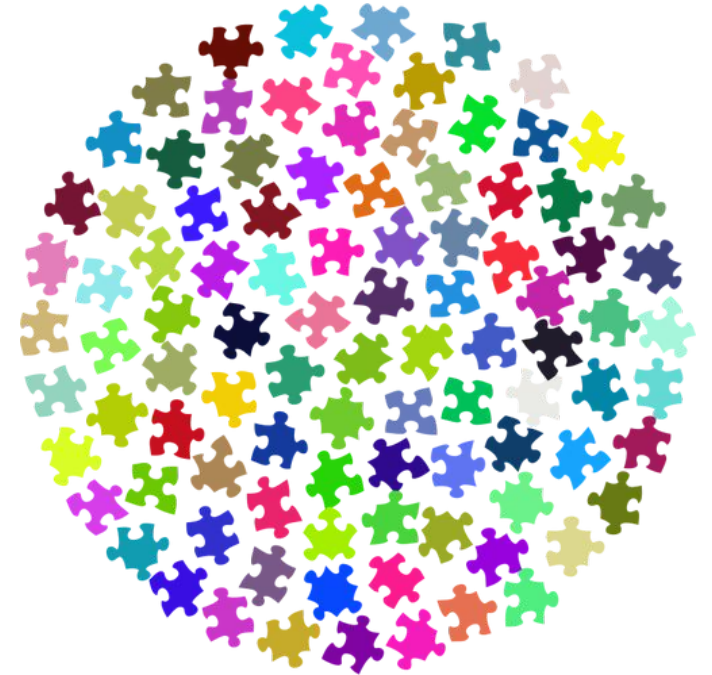
Motivation and background

Group decision making and preference inspection



Group Decision Making is everywhere

- Scenarios vary considerably in **complexity** and **impact**
 - numbers/diversity of **alternatives**
 - numbers and importance of **attributes**
 - numbers/diversity of **stakeholders**
 - **frequency** of the decision
 - ...



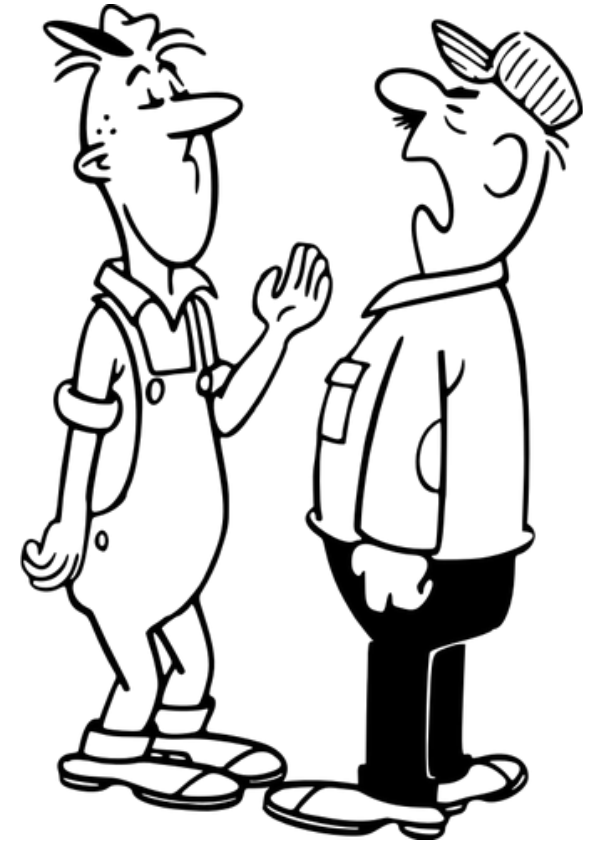
Motivation and background

Group decision making and preference inspection



Interpersonal challenges of Group Decision Making

- **Disagreements**
 - What are disagreements actually about?
 - What are the sources?
 - difference of opinion
 - missing information
- **Misunderstandings**
 - Do decision makers understand others' preferences?
 - Do they truly understand their own preferences?





Questions:

- What is **preference inspection**, and why can it be beneficial to effective group decision making?
- How can **visualizations** support preference inspection?
- Are all visualizations equally **effective**?
- How can we **rigorously design** effective visualizations?

Motivation and background

Group decision making and preference inspection



Eliciting and inspecting preferences

- Have stakeholders explicitly **model** their preferences over the alternatives
 - numerical scales (1-10)
 - PAPRIKA [Hansen and Ombler 2008]
 - ...
- Decision makers can then **examine** those preferences
 - how alternatives **perform**
 - investigate **sources of disagreement**
 - better understand **all points of view**
- This can (and often should be) an **iterative process**



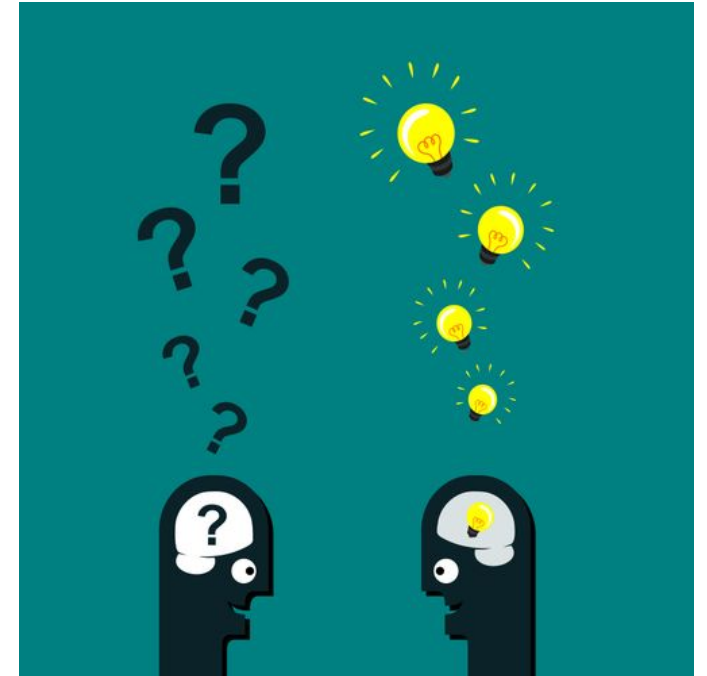
Motivation and background

Benefits of visualization



Ensuring effective preference inspection

- It is vitally important to **gain insights quickly**



Motivation and background

Benefits of visualization



Ensuring effective preference inspection

- It is vitally important to **gain insights quickly**
- This can be difficult with **text**-based formats

	Alice	Bob	Carol	David	Esther
McDonalds	2	6	7	4	9
A&W	3	8	4	6	6
Burger King	5	4	5	5	8
Wendy's	4	2	6	5	3
Dairy Queen	6	8	8	3	2
Kentucky Fried	3	3	3	9	6
Pizza Hut	9	5	2	9	7
Olive Garden	8	9	5	8	7
Arby's	4	1	6	1	5
Applebee's	6	5	9	2	3
Taco Bell	2	3	9	2	1
Chick-Fil-A	4	5	6	4	8
Bob's Burgers	6	2	1	5	6

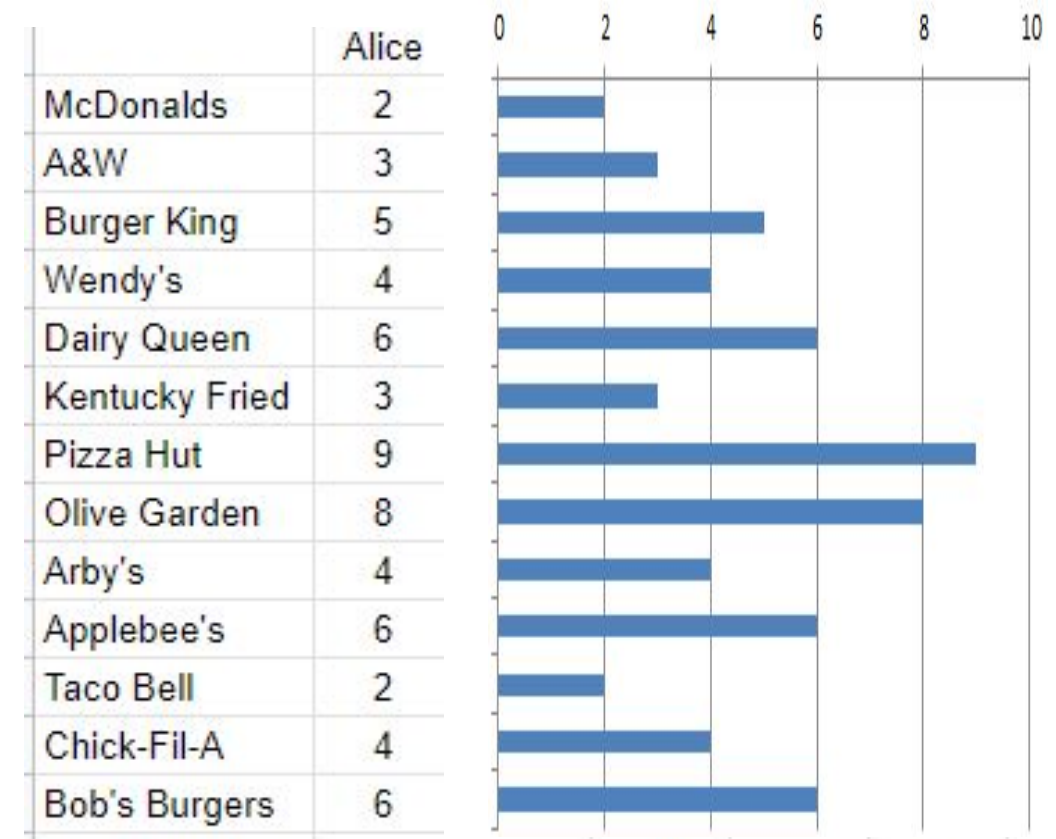
Motivation and background

Benefits of visualization



The advantage of using visualization

- Visualizations leverage the **pattern recognition** and **pre-attentive capabilities** of the human visual system [Munzner 2014]



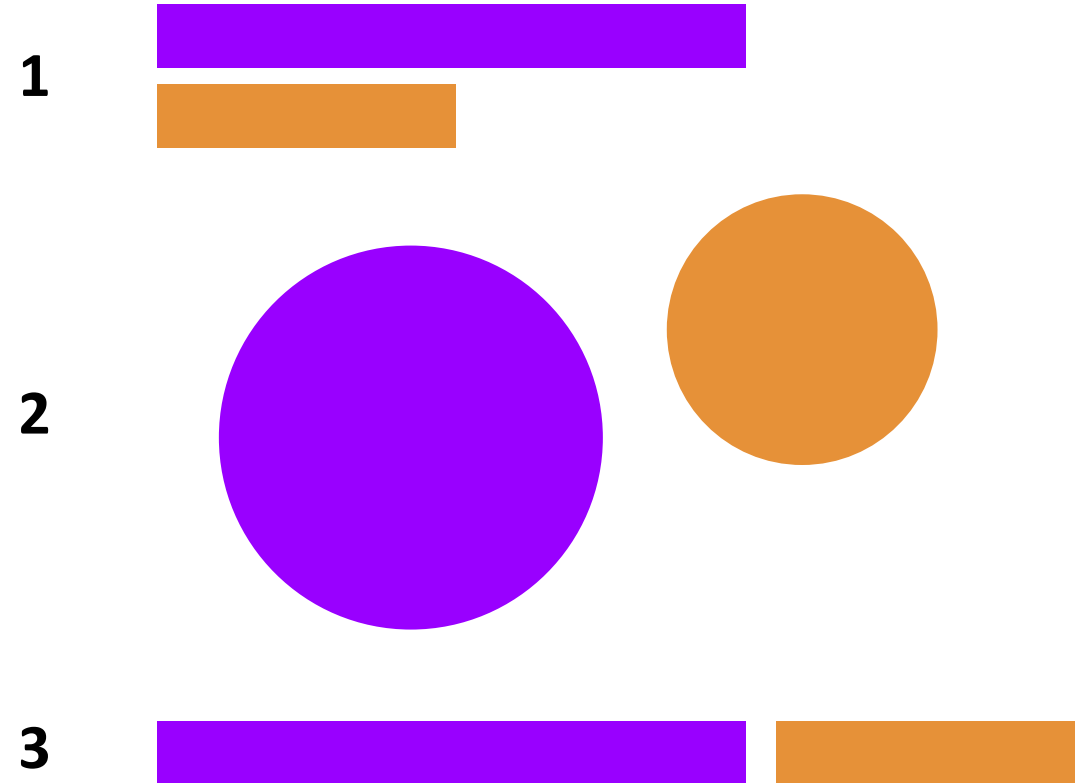
Motivation and background

Challenges in creating effective visualizations



Not all visualizations are equally effective

- We are able to process some visual inputs more precisely than others
- Important to choose effective **idioms**



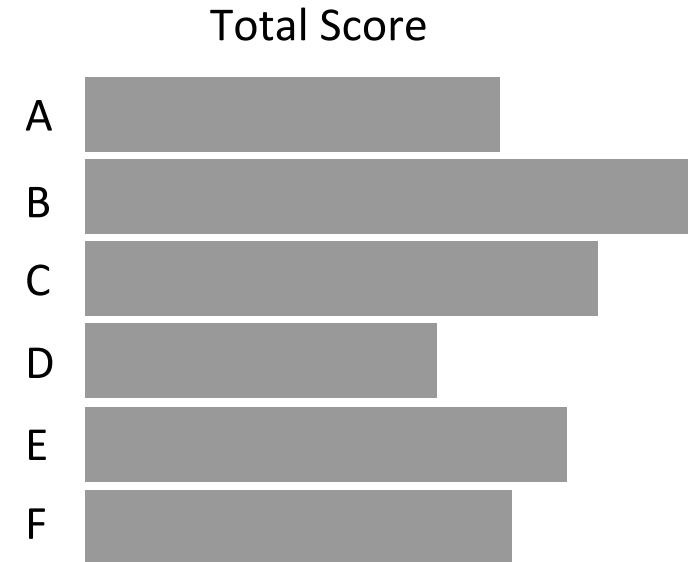
Motivation and background

Challenges in creating effective visualizations



What are we trying to accomplish?

- A visualization may not support all tasks equally well
 - Which alternative has the highest **overall score**?
 - On which alternative is there the **most disagreement**?
 - Are there **outliers**?
 - Are there **polarized groups**?
 - Which alternative does the **majority shareholder** prefer?



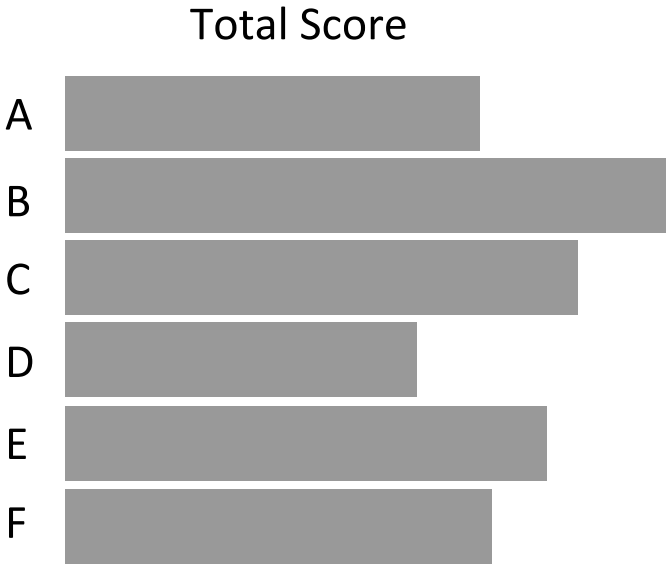
Motivation and background

Challenges in creating effective visualizations



How about lots of targeted visualizations?

- Each could be targeted for a different task
- However, switching between views increases **cognitive load** [Munzner 2014]



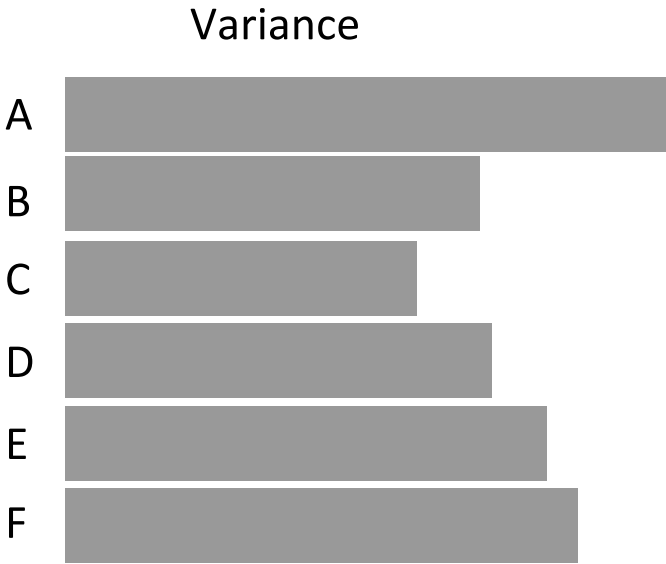
Motivation and background

Challenges in creating effective visualizations



How about lots of targeted visualizations?

- Each could be targeted for a different task
- However, switching between views increases **cognitive load** [Munzner 2014]



Motivation and background

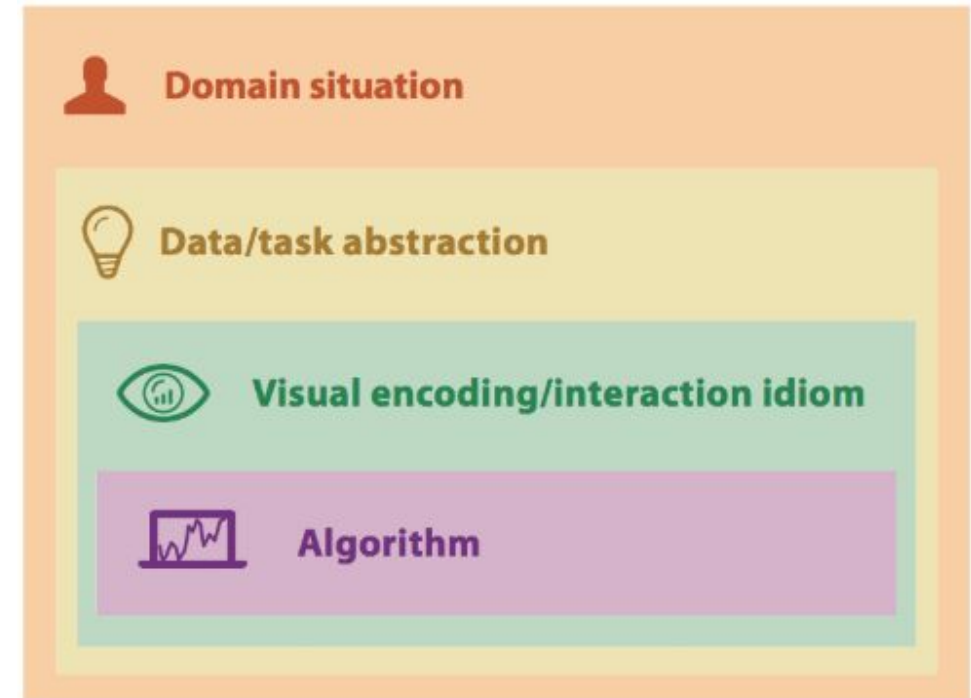
Challenges in creating effective visualizations



The Nested Model of Visualization Design

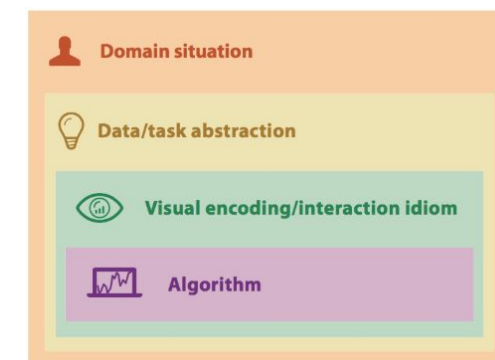
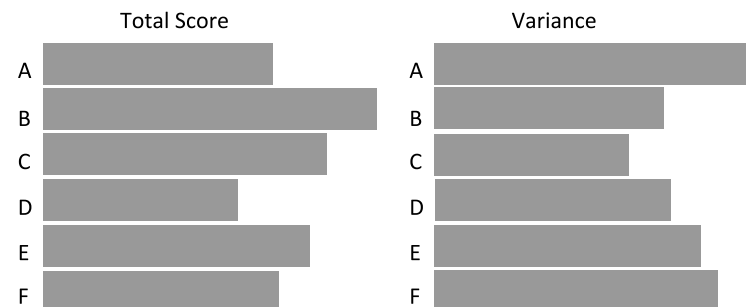
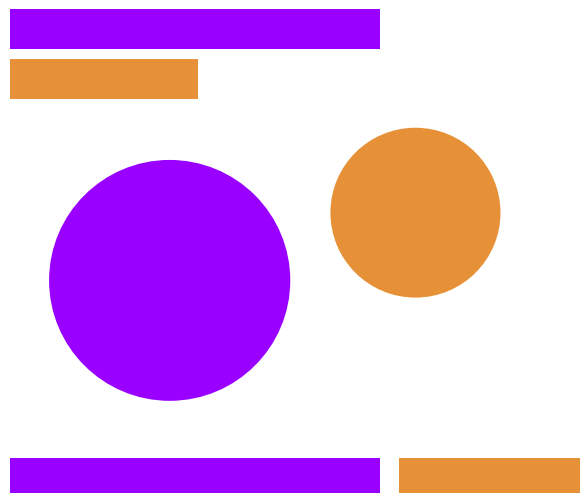
[Munzner 2009]

- A **rigorous** and **domain-independent** framework for designing and validating effective visualizations
- The output of each stage serves as input to the next



Question Break

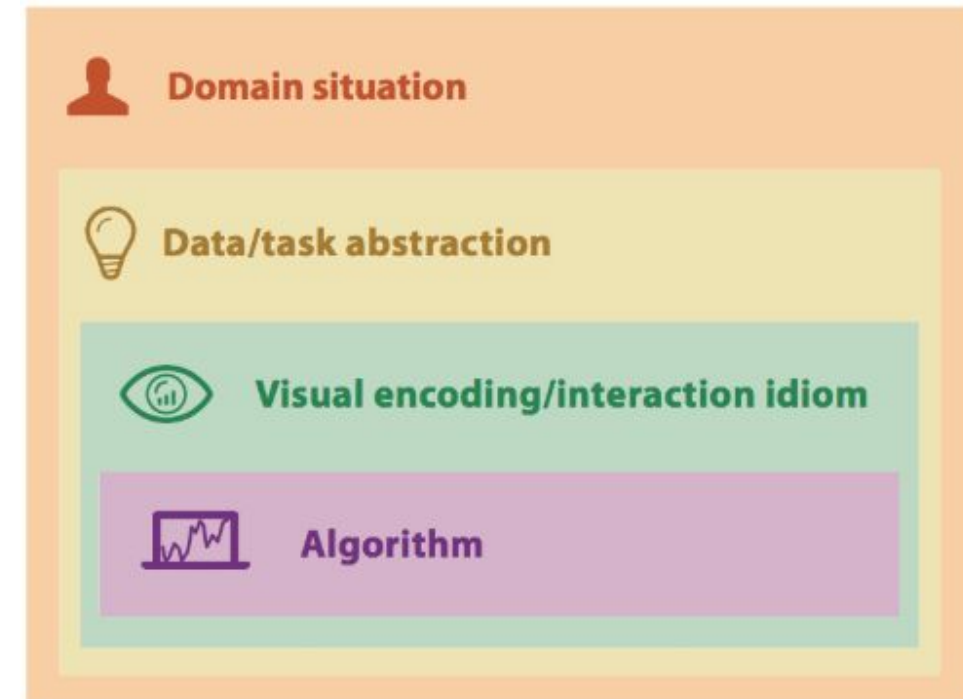
	Alice	Bob	Carol	David	Esther
McDonalds	2	6	7	4	9
A&W	3	8	4	6	6
Burger King	5	4	5	5	8
Wendy's	4	2	6	5	3
Dairy Queen	6	8	8	3	2
Kentucky Fried	3	3	3	9	6
Pizza Hut	9	5	2	9	7
Olive Garden	8	9	5	8	7
Arby's	4	1	6	1	5
Applebee's	6	5	9	2	3
Taco Bell	2	3	9	2	1
Chick-Fil-A	4	5	6	4	8
Bob's Burgers	6	2	1	5	6





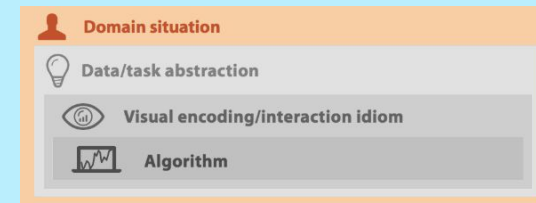
Overview

- Follow the Nested Model
- Begin with simple case: scoring alternatives directly [Hindalong et al. 2020]
- Extend to more complex preference models (ongoing)



Our work

Scoring alternatives directly [Hindalong et al. 2020]
Scenario analysis



Variety of real-world scenarios

Name	Description	Elicitation Method	Alternatives	Evaluators
BP: Best Paper	Researchers choose Best Paper Award recipient	Interview	4 - 15	5
FH: Faculty Hiring	Faculty members choose which candidate to hire	Interview	1 - 4	50 - 100
CR: Campbell River	Stakeholders from diverse interest groups choose a watershed operation strategy	Webinar observation	6	15
VY: Voyager [12]	Scientists choose trajectories for Voyager 1 & 2	Journal Paper	32	10
NC: Nuclear Crisis [27]	Emergency planners choose a strategy in response to a mock nuclear crisis	Journal Paper	6	6
SW: Software	Software company employees choose a technology stack	Interview + In-person observation	2	5
GI: Gift	Lab members choose a gift for a colleague	Interview	3	10

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Scenario analysis

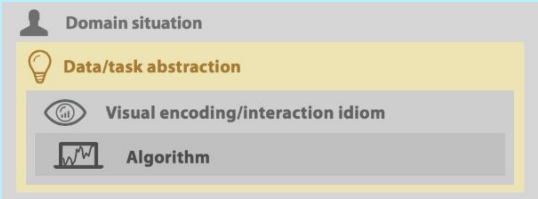


User classes from scenario features

Type	Specialized	Professional	Casual
Scenarios	CR, VY, NC	BP, FH, SW	GI
Stakes	Very High	Medium – High	Low
Work Context	Expert Assistance	Professional	Casual
Timeframe	Days	Hours	Hours
Decision Frequency	Once	Monthly – Annually	Once

Our work

Scoring alternatives directly [Hindalong et al. 2020]
Data and task abstractions



Data abstractions

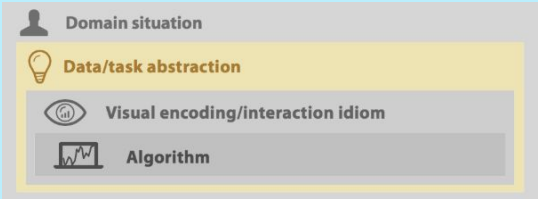
- Represent preferences as **tabular data** [Munzner 2014]
- **Dimensions** (form **keys**)
 - Alternatives
 - Evaluators
- **Measures** (obtain **values**)
 - $AltScore(a,e)$
 - $AltRank(a,e)$
 - **Derived measures:** $TotalScore(a)$, $TotalRank(a)$

	Alice	Bob	Carol	David	Esther
McDonalds	2	6	7	4	9
A&W	3	8	4	6	6
Burger King	5	4	5	5	8
Wendy's	4	2	6	5	3
Dairy Queen	6	8	8	3	2
Kentucky Fried	3	3	3	9	6
Pizza Hut	9	5	2	9	7
Olive Garden	8	9	5	8	7
Arby's	4	1	6	1	5
Applebee's	6	5	9	2	3
Taco Bell	2	3	9	2	1
Chick-Fil-A	4	5	6	4	8
Bob's Burgers	6	2	1	5	6

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Data and task abstractions



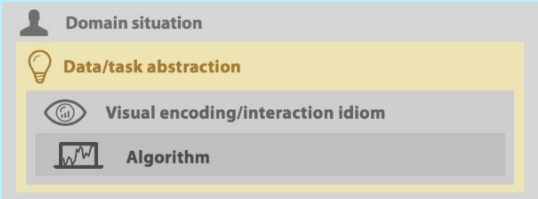
Scenario-independent goals

GENERIC GOAL			SCENARIOS
G1	Discover Viable Alternatives		
	a	Discover high-performing alternatives across evaluators/evaluator groups	<i>all but</i> FH
	b	Discover high-performing alternatives for a single evaluator/evaluator group	SW, GI
G2	Discover Sources of Disagreement (discrepancies across evaluators)		
	a	Discover and explain disagreement about an alternative (across evaluators/evaluator groups)	<i>all but</i> GI
G3	Explain Individual Scores		
	a	Analyze contribution of different evaluators to an alternative’s total score	NC
G4	Validate Model		
	a	Understand sensitivity of preferences to change	NC
	b	Understand sensitivity of total scores to aggregation method	VY
	c	Discover discrepancies between the preferences of one evaluator and the rest	BP
G5	Discover Nuances (not captured by preference models)		<i>all</i>

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Data and task abstractions



High-level preference inspection tasks

- Placed tasks within widely used visualization typology [Brehmer and Munzner 2013]
- **12 tasks identified**

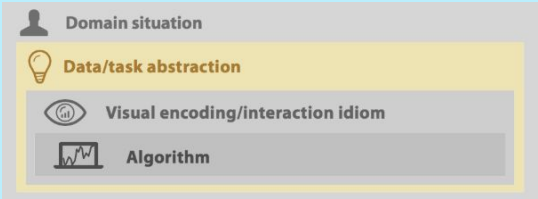
Table 5: Tasks to Support G1: Discover Viable Alternatives.

TASK	
G1a. Discover high-performing alternatives across evaluators	
T1	Discover alternative(s) with best TotalRank/TotalScore
T2	Discover alternatives(s) with low variance in AltRanks/AltScores across evaluators
T3	Discover non-dominated alternatives across evaluators
T4	Discover trade-offs in AltRanks/AltScores between alternatives <i>a</i> and <i>b</i>
T5	Discover pros and cons in AltRanks/AltScores for alternative <i>a</i>
G1b. Discover high-performing alternatives for a single evaluator	
T6	Discover alternative(s) with best AltRank/AltScore for evaluator <i>e</i>

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Data and task abstractions



Auxiliary task functions

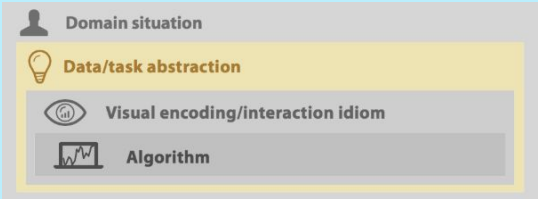
Table 6: Auxiliary Task Functions. The Input column also indicates when there are different cases of the task.

	Action	Input	Output	Supported by
AT1	Query: Identify	single value or distribution	its key-set	
AT2	Query: Compare	pair of values A: one evaluator, two alternatives B: one alternative, two evaluators	difference	
AT3	Query: Compare	pair of same-type distributions A: all evaluators, two alternatives B: all alternatives, two evaluators	tuple of differences	AT2
AT4	Query: Compare	pair of same-type distributions	dominance relation	AT3
AT5	Query: Summarize	single distribution	summary of variance	
AT6	Search: Locate	key-set A: one alternative, one evaluator B: one alternative C: one evaluator	single value or distribution	
AT7	Search: Lookup (in context)	key-set + single value or distribution	single value or distribution	AT6
AT8	Search: Browse	single distribution	outliers	AT2
AT9	Search: Browse	single distribution A: one evaluator B: all data	min/max values	AT2
AT10	Search: Browse	set of distributions	non-dominated distributions	AT4

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Data and task abstractions



Auxiliary task functions

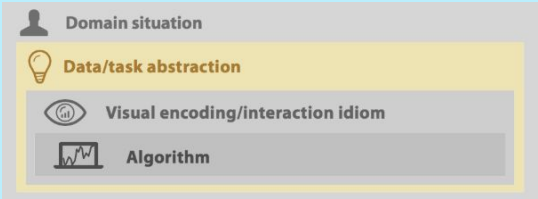
Table 6: Auxiliary Task Functions. The Input column also indicates when there are different cases of the task.

	Action	Input	Output	Supported by
AT1	Query: Identify	single value or distribution	its key-set	
AT2	Query: Compare	pair of values A: one evaluator, two alternatives B: one alternative, two evaluators	difference	
AT3	Query: Compare	pair of same-type distributions A: all evaluators, two alternatives B: all alternatives, two evaluators	tuple of differences	AT2
AT4	Query: Compare	pair of same-type distributions	dominance relation	AT3
AT5	Query: Summarize	single distribution	summary of variance	
AT6	Search: Locate	key-set A: one alternative, one evaluator B: one alternative C: one evaluator	single value or distribution	
AT7	Search: Lookup (in context)	key-set + single value or distribution	single value or distribution	AT6
AT8	Search: Browse	single distribution	outliers	AT2
AT9	Search: Browse	single distribution A: one evaluator B: all data	min/max values	AT2
AT10	Search: Browse	set of distributions	non-dominated distributions	AT4

Our work

Scoring alternatives directly [Hindalong et al. 2020]

Data and task abstractions



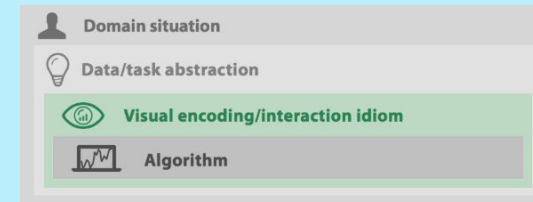
Auxiliary task functions

Table 6: Auxiliary Task Functions. The Input column also indicates when there are different cases of the task.

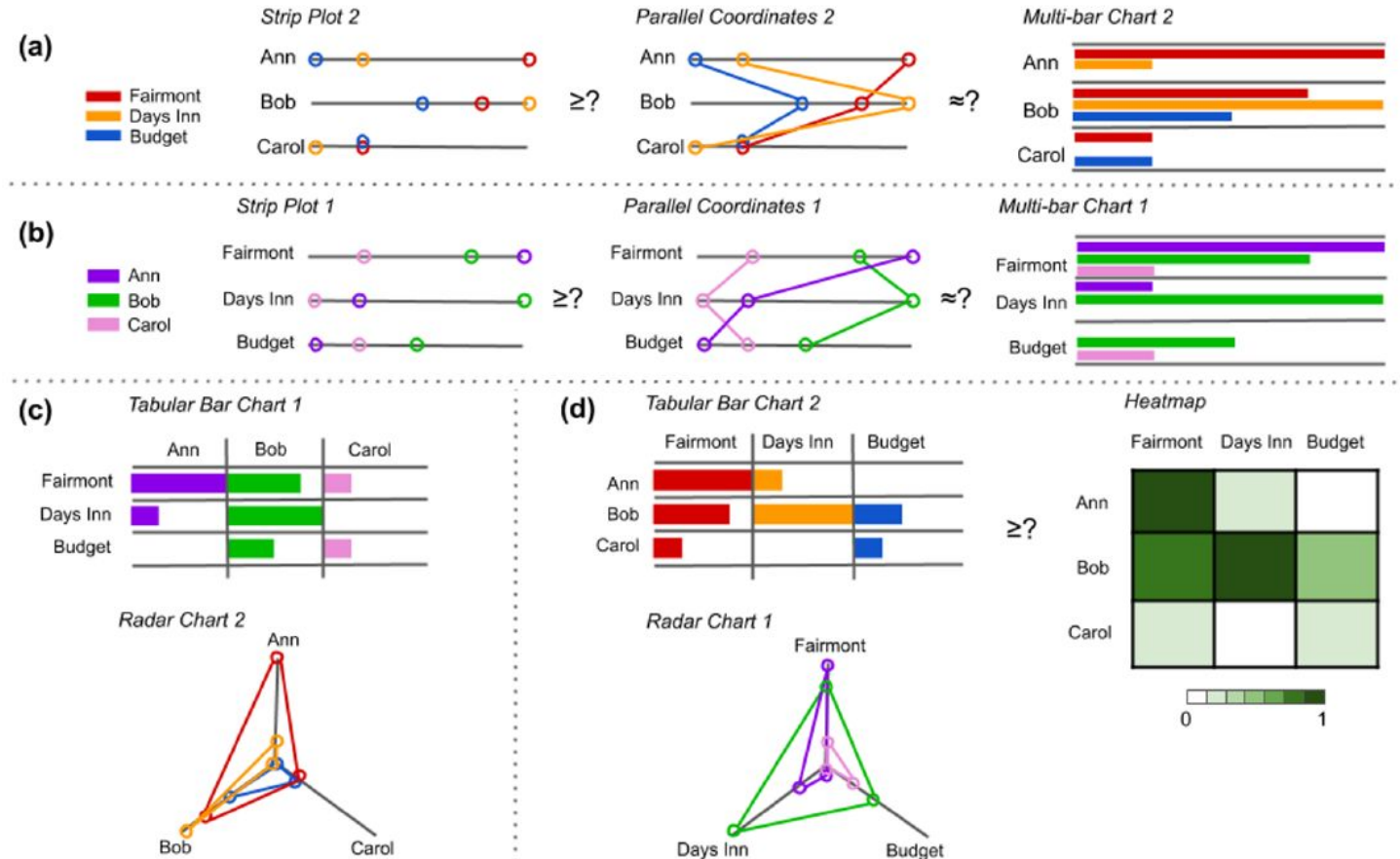
	Action	Input	Output	Supported by
AT1	Query: Identify	single value or distribution	its key-set	
AT2	Query: Compare	pair of values A: one evaluator, two alternatives B: one alternative, two evaluators	difference	
AT3	Query: Compare	pair of same-type distributions A: all evaluators, two alternatives B: all alternatives, two evaluators	tuple of differences	AT2
AT4	Query: Compare	pair of same-type distributions	dominance relation	AT3
AT5	Query: Summarize	single distribution	summary of variance	
AT6	Search: Locate	key-set A: one alternative, one evaluator B: one alternative C: one evaluator	single value or distribution	
AT7	Search: Lookup (in context)	key-set + single value or distribution	single value or distribution	AT6
AT8	Search: Browse	single distribution	outliers	AT2
AT9	Search: Browse	single distribution A: one evaluator B: all data	min/max values	AT2
AT10	Search: Browse	set of distributions	non-dominated distributions	AT4

Our work

Scoring alternatives directly [Hindalong et al. 2020]

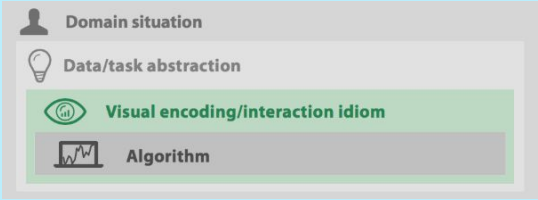


Evaluating core idioms



Our work

Scoring alternatives directly [Hindalong et al. 2020]



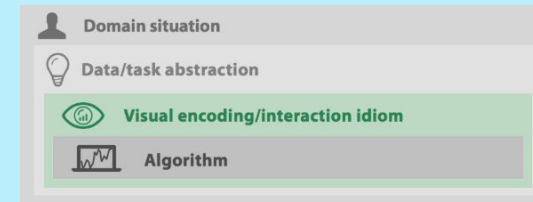
Evaluating core idioms

	AT1	AT2:A	AT2:B	AT3:A	AT3:B	AT4	AT5	AT6:A	AT6:B	AT6:C	AT7	AT8	AT9:A	AT9:B	AT10	Total
Parallel Coords 2	1	2	1	3	2	3	1	1	1	2	n/a	1	3	3	3	27
Parallel Coords 1	1	1	2	2	3	2	2	1	2	0	n/a	2	1	1	2	22
Radar Chart 2	1	1	0	2	0	3	1	1	1	2	1	1	2	2	3	21
Multi-bar 2	2	2	1	1	1	1	1	3	1	2	n/a	1	2	2	1	21
Multi-bar 1	2	1	2	1	1	1	1	3	2	1	n/a	2	1	1	1	20
Tabular Bar 1	3	1	0	2	0	2	0	3	2	3	n/a	0	1	1	2	20
Box Plot 1	1	1	2	1	1	2	3	1	2	0	n/a	3	0	0	2	19
Strip Plot 2	1	3	1	1	1	1	1	1	0	2	n/a	0	3	3	1	19
Box Plot 2	1	2	1	1	1	1	1	1	0	2	n/a	0	3	3	1	18
Heatmap	3	0	0	0	0	1	0	3	2	2	n/a	2	2	2	1	18
Strip Plot 1	1	1	3	1	1	2	2	1	2	0	n/a	2	0	0	2	18
Tabular Bar 2	3	0	1	0	2	0	1	3	3	2	n/a	1	0	0	0	16
Radar Chart 1	1	0	1	0	2	0	2	1	2	2	n/a	2	1	1	0	15
Stacked Bar Chart	1	0	0	0	0	0	0	1	1	2	3	0	0	3	0	11

Figure 4: Support for each auxiliary task function (see Table 6 above for full descriptions) by encoding. 3=best, 2=strongly effective, 1=weakly effective, 0=ineffective. *n/a* indicates the encoding is not applicable. The *Total* column contains the row totals. The rows are sorted by Total.

Our work

Scoring alternatives directly [Hindalong et al. 2020]

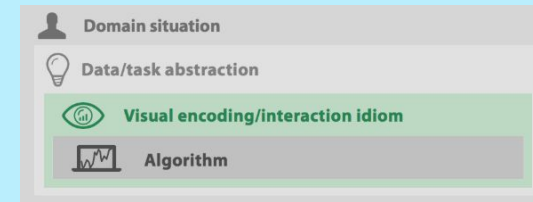


Recommendations for casual users

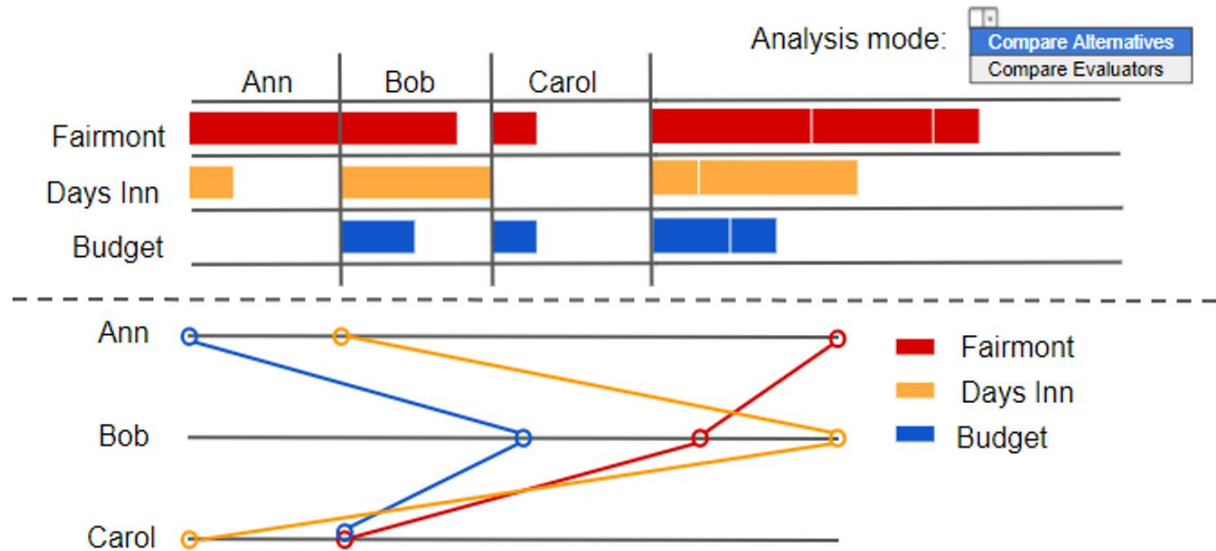
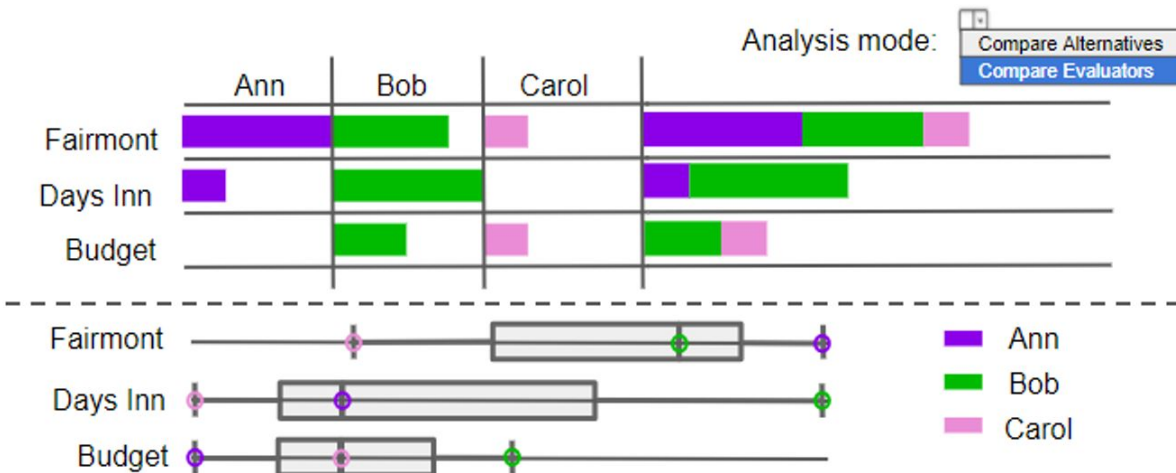


Our work

Scoring alternatives directly [Hindalong et al. 2020]

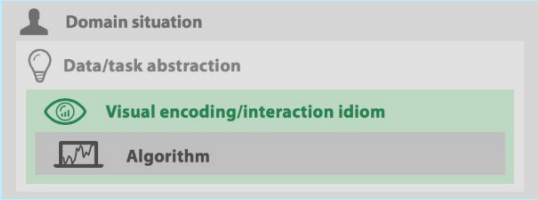


Recommendations for advanced users

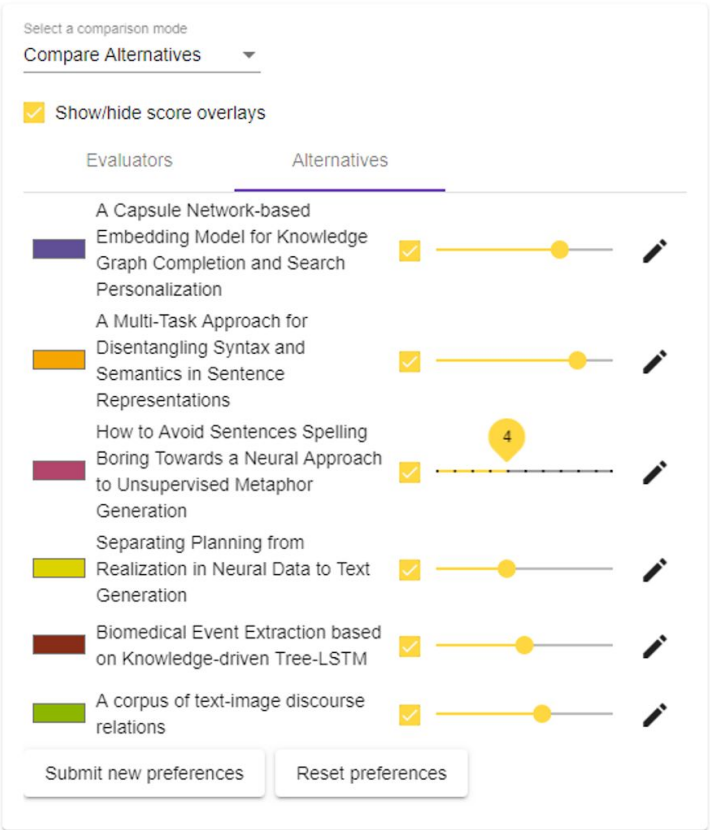
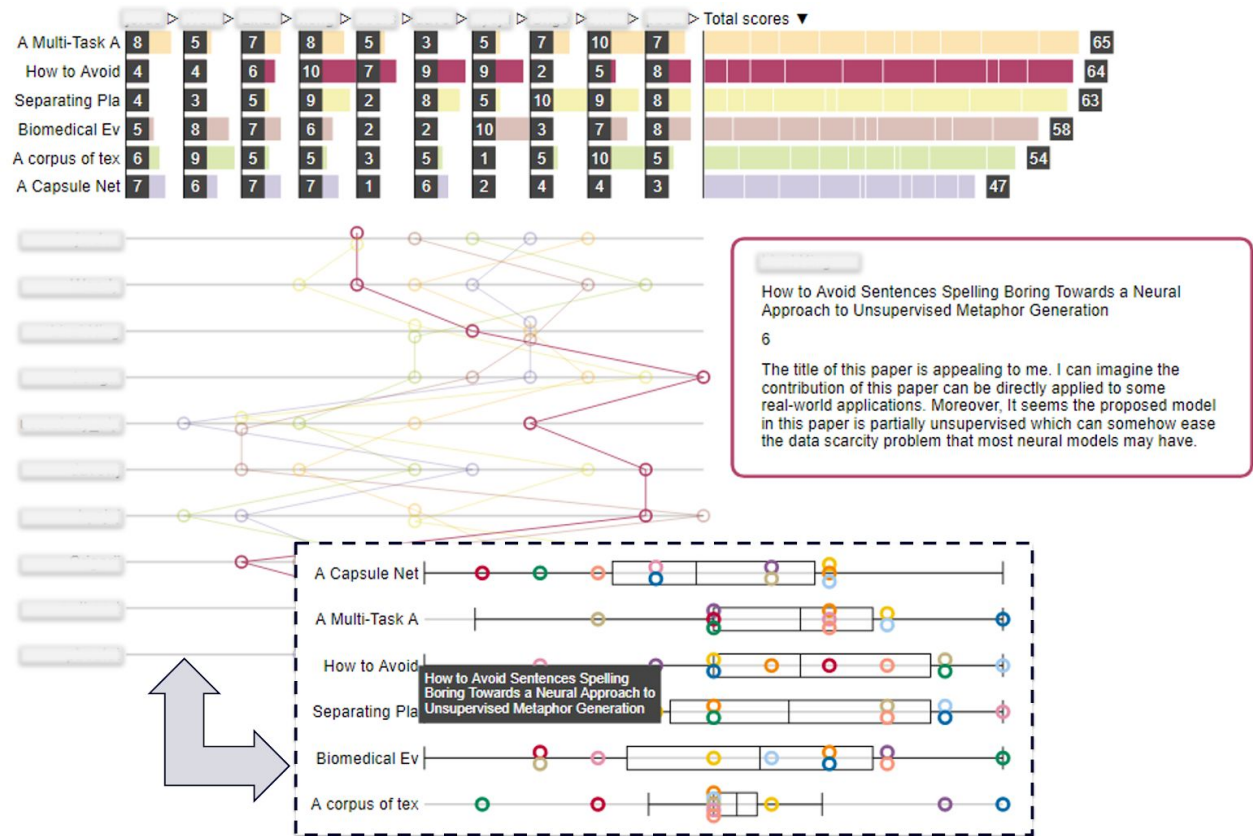


Our work

Scoring alternatives directly [Hindalong et al. 2020]

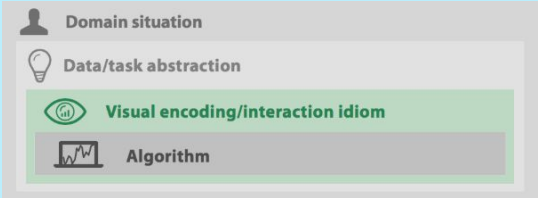


Prototype

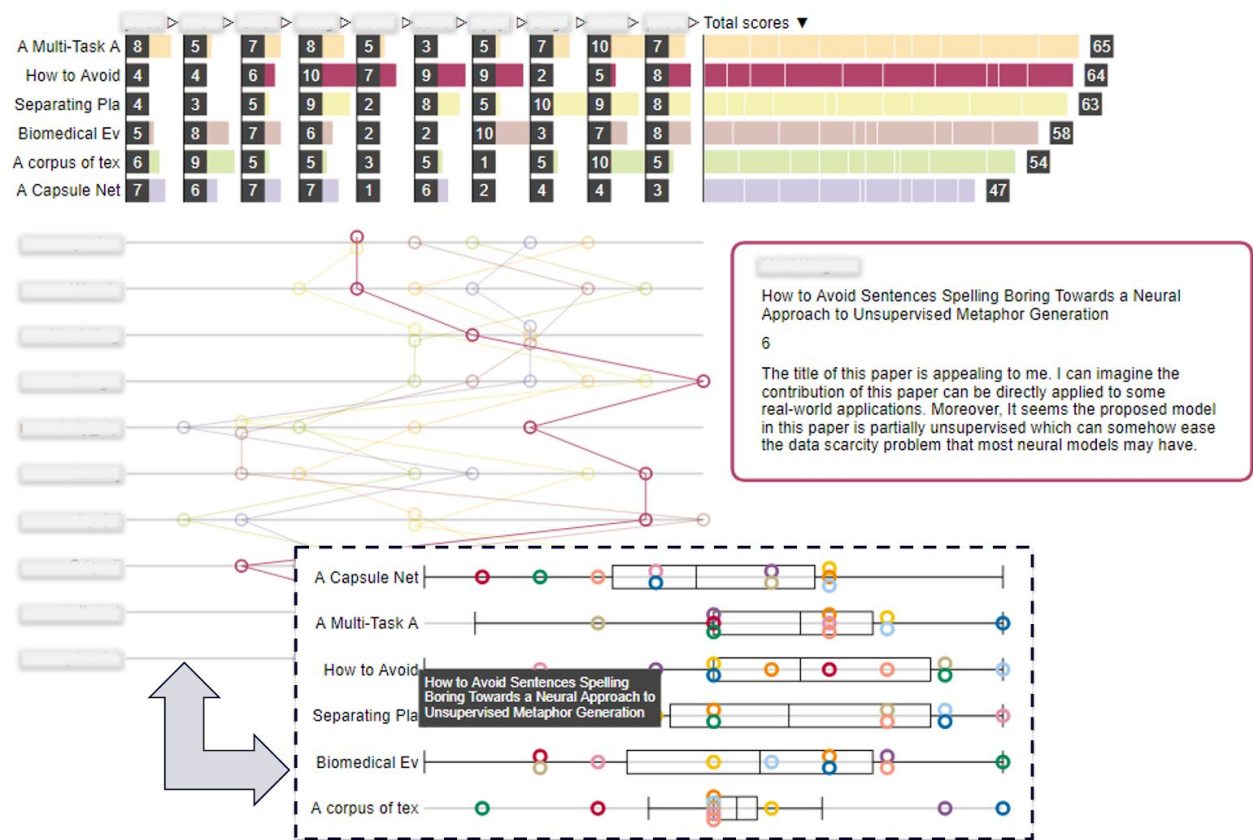


Our work

Scoring alternatives directly [Hindalong et al. 2020]



Prototype



Select a comparison mode

Compare Alternatives

☒ Show/hide score overlays

Alternatives

A Capsule Network-based

Embedding Model for Knowledge Graph Completion and Search Personalization

A Multi-Task Approach for Disentangling Syntax and Semantics in Sentence Representations

How to Avoid Sentences Spelling Boring Towards a Neural Approach to Unsupervised Metaphor Generation

Separating Planning from Realization in Neural Data to Text Generation

Biomedical Event Extraction based on Knowledge-driven Tree-LSTM

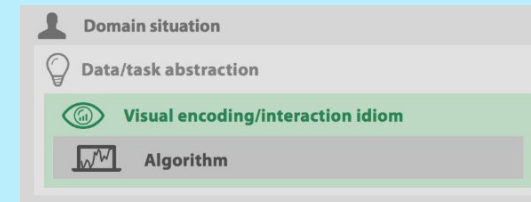
A corpus of text-image discourse relations

Submit new preferences

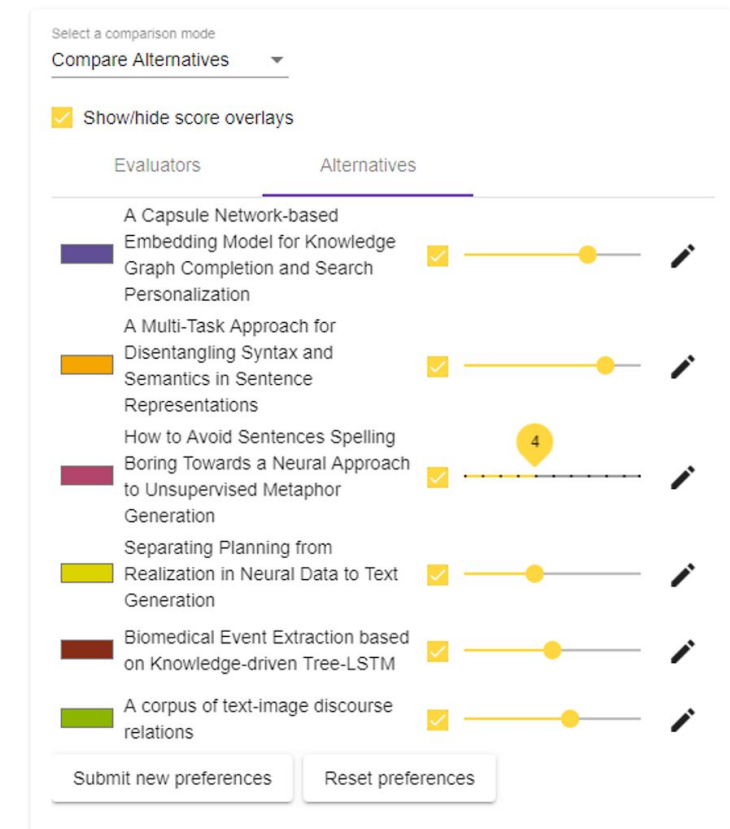
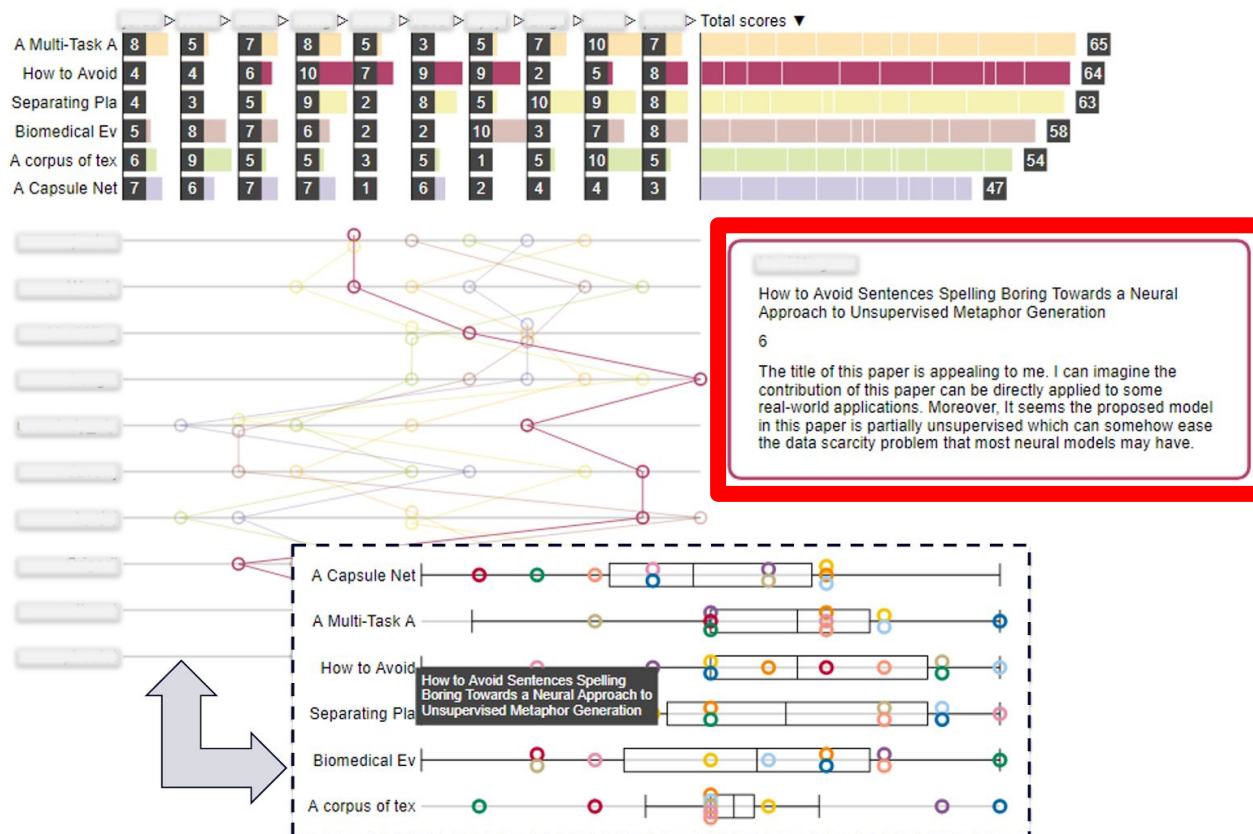
Reset preferences

Our work

Scoring alternatives directly [Hindalong et al. 2020]

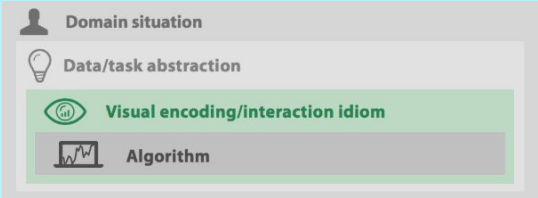


Prototype

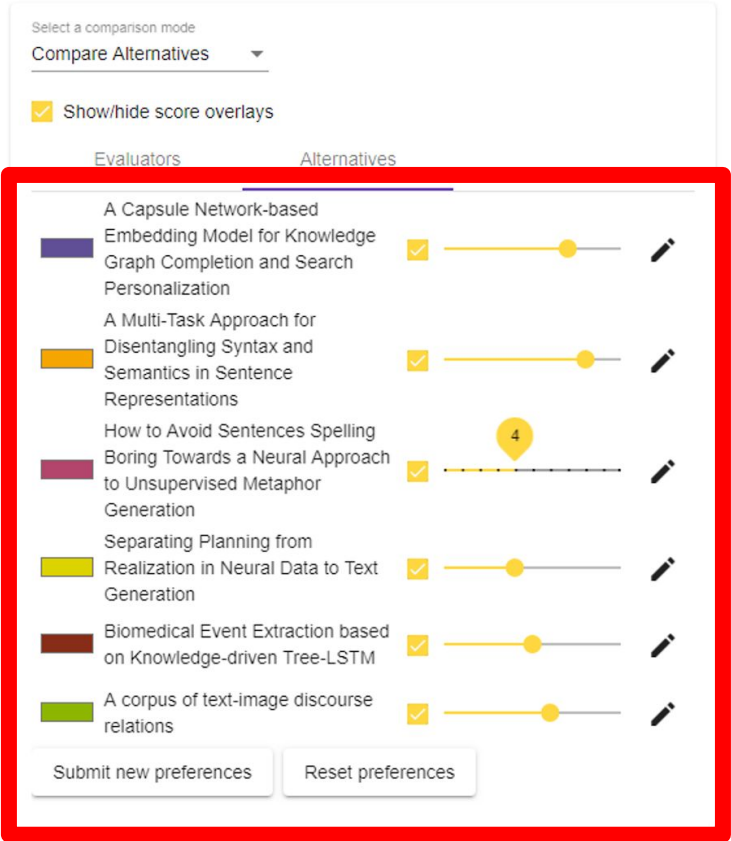
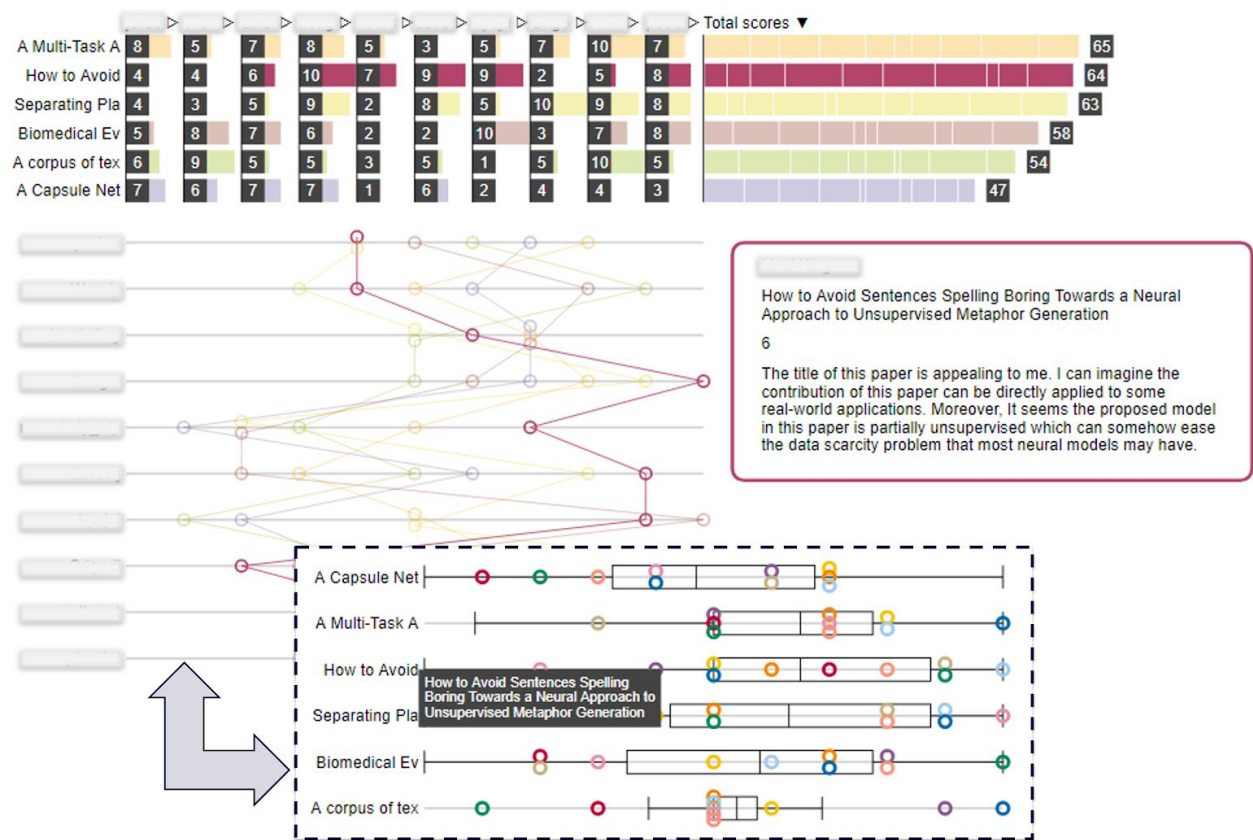


Our work

Scoring alternatives directly [Hindalong et al. 2020]



Prototype



Our work

Scoring alternatives directly [Hindalong et al. 2020]



User studies

Table 7: Mean responses to questions on a 5-point Likert scale (1 = “strongly disagree”, 5 = “strongly agree” unless otherwise specified).

	MER	NLP
Prior experience: (<i>1 = none, 5 = extensive</i>)		
Bar chart	4.83	4.12
Stacked bar chart	4.17	3.00
Box plot	4.00	3.25
Parallel coordinates plot	1.83	1.75
Microsoft Excel	4.50	3.62
Tableau	1.83	2.00
Visualizations in general	3.67	3.12
Usefulness: (<i>1 = not useful, 5 = very useful</i>)		
Bar chart	4.33	4.25
Stacked bar chart	4.67	4.62
Box plot	4.50	3.12
Parallel coordinates plot	2.33	3.12
Legend	4.67	4.38
Mean of Visualization Elements	4.10	3.90
Showing/hiding alternatives	4.00	4.00
Showing/hiding evaluators	4.00	4.12
Sorting by evaluator/total score	4.50	4.88
Switching between comparison modes	3.00	3.62
Showing/hiding the score overlay	3.33	4.12
Linked highlighting	4.50	3.88
Mean of Interactive Techniques	3.89	4.10

	MER	NLP
I modified my own preferences during the session (<i>1 = not at all, 5 = significantly</i>)	4.33	2.12
This tool helps make our discussions more participatory	4.33	4.12
The tool helps identify agreements and disagreements	4.67	4.88
The tool helps make informed decisions based on everyone’s preferences	3.67	4.00
I would be happy if the alternative with the highest aggregate score were chosen	3.67	3.62
Please rate the tool’s potential to affect group interaction (<i>1 = worse, 5 = better</i>)	3.83	4.00
Please rate the tool’s potential to affect information exchange among participants (<i>1 = less exchange, 5 = more exchange</i>)	4.33	3.75
This tool was suitable for the complexity of the decision being made	3.67	3.88
I would like to use this tool for making similar collaborative decisions in the future	4.00	3.62
I would like to use this tool for making more complex collaborative decisions in the future	3.83	3.75

Our work

Scoring alternatives directly [Hindalong et al. 2020]



User studies

Table 7: Mean responses to questions on a 5-point Likert scale (1 = “strongly disagree”, 5 = “strongly agree” unless otherwise specified).

	MER	NLP
Prior experience: (<i>1 = none, 5 = extensive</i>)		
Bar chart	4.83	4.12
Stacked bar chart	4.17	3.00
Parallel coordinates plot	1.83	1.75
Microsoft Excel	4.50	3.02
Tableau	1.83	2.00
Visualizations in general	3.67	3.12
Usefulness: (<i>1 = not useful, 5 = very useful</i>)		
Bar chart	4.33	4.25
Stacked bar chart	4.67	4.62
Parallel coordinates plot	2.33	3.12
Legend	4.67	4.50
Mean of Visualization Elements	4.10	3.90
Showing/hiding alternatives	4.00	4.00
Showing/hiding evaluators	4.00	4.12
Sorting by evaluator/total score	4.50	4.88
Switching between comparison modes	3.00	3.62
Showing/hiding the score overlay	3.50	4.12
Linked highlighting	4.50	3.88
Mean of Interactive Techniques	3.89	4.10

	MER	NLP
I modified my own preferences during the session (<i>1 = not at all, 5 = significantly</i>)	4.33	2.12
This tool helps make our discussions more participatory	4.33	4.12
The tool helps identify agreements and disagreements	4.67	4.88
The tool helps make informed decisions based on everyone’s preferences	3.67	4.00
I would be happy if the alternative with the highest aggregate score were chosen	3.67	3.62
Please rate the tool’s potential to affect group interaction (<i>1 = worse, 5 = better</i>)	3.83	4.00
Please rate the tool’s potential to affect information exchange among participants (<i>1 = less exchange, 5 = more exchange</i>)	4.33	3.75
This tool was suitable for the complexity of the decision being made	3.67	3.88
I would like to use this tool for making similar collaborative decisions in the future	4.00	3.62
I would like to use this tool for making more complex collaborative decisions in the future	3.83	3.75

Our work

Scoring alternatives directly [Hindalong et al. 2020]



User studies

Table 7: Mean responses to questions on a 5-point Likert scale (1 = “strongly disagree”, 5 = “strongly agree” unless otherwise specified).

	MER	NLP		MER	NLP
Prior experience: (<i>1 = none, 5 = extensive</i>)			I modified my own preferences during the session (<i>1 = not at all, 5 = significantly</i>)	4.33	2.12
Bar chart	4.83	4.12	This tool helps make our discussions more participatory	4.33	4.12
Stacked bar chart	4.17	3.00	The tool helps identify agreements and disagreements	4.67	4.88
Box plot	4.00	3.25	The tool helps make informed decisions based on everyone’s preferences	3.67	4.00
Parallel coordinates plot	1.83	1.75	I would be happy if the alternative with the highest aggregate score were chosen	3.67	3.62
Microsoft Excel	4.50	3.62	Please rate the tool’s potential to affect group interaction (<i>1 = worse, 5 = better</i>)	3.83	4.00
Tableau	1.83	2.00	Please rate the tool’s potential to affect information exchange among participants (<i>1 = less exchange, 5 = more exchange</i>)	4.33	3.75
Visualizations in general	3.67	3.12	This tool was suitable for the complexity of the decision being made	3.67	3.88
Usefulness: (<i>1 = not useful, 5 = very useful</i>)			I would like to use this tool for making similar collaborative decisions in the future	4.00	3.62
Bar chart	4.33	4.25	I would like to use this tool for making more complex collaborative decisions in the future	3.83	3.75
Stacked bar chart	4.67	4.62			
Box plot	4.50	3.12			
Parallel coordinates plot	2.33	3.12			
Legend	4.67	4.38			
Mean of Visualization Elements	4.10	3.90			
Showing/hiding alternatives	4.00	4.00			
Showing/hiding evaluators	4.00	4.12			
Sorting by evaluator/total score	4.50	4.88			
Switching between comparison modes	3.00	3.62			
Showing/hiding the score overlay	3.33	4.12			
Linked highlighting	4.50	3.88			
Mean of Interactive Techniques	3.89	4.10			

Our work

Scoring alternatives directly [Hindalong et al. 2020]



User studies

Table 7: Mean responses to questions on a 5-point Likert scale (1 = “strongly disagree”, 5 = “strongly agree” unless otherwise specified).

	MER	NLP
Prior experience: (<i>1 = none, 5 = extensive</i>)		
Bar chart	4.83	4.12
Stacked bar chart	4.17	3.00
Box plot	4.00	3.25
Parallel coordinates plot	1.83	1.75
Microsoft Excel	4.50	3.62
Tableau	1.83	2.00
Visualizations in general	3.67	3.12
Usefulness: (<i>1 = not useful, 5 = very useful</i>)		
Bar chart	4.33	4.25
Stacked bar chart	4.67	4.62
Box plot	4.50	3.12
Parallel coordinates plot	2.33	3.12
Legend	4.67	4.38
Mean of Visualization Elements	4.10	3.90
Showing/hiding alternatives	4.00	4.00
Showing/hiding evaluators	4.00	4.12
Sorting by evaluator/total score	4.50	4.88
Switching between comparison modes	3.00	3.62
Showing/hiding the score overlay	3.33	4.12
Linked highlighting	4.50	3.88
Mean of Interactive Techniques	3.89	4.10

	MER	NLP
I modified my own preferences during the session (<i>1 = not at all, 5 = significantly</i>)	4.33	2.12
This tool helps make our discussions more participatory	4.33	4.12
The tool helps identify agreements and disagreements	4.67	4.88
The tool helps make informed decisions based on everyone’s preferences	3.67	4.00
I would be happy if the alternative with the highest aggregate score were chosen	3.67	3.62
Please rate the tool’s potential to affect group interaction (<i>1 = worse, 5 = better</i>)	3.83	4.00
Please rate the tool’s potential to affect information exchange among participants (<i>1 = less exchange, 5 = more exchange</i>)	4.33	3.75
This tool was suitable for the complexity of the decision being made	3.67	3.88
I would like to use this tool for making similar collaborative decisions in the future	4.00	3.62
I would like to use this tool for making more complex collaborative decisions in the future	3.83	3.75

Our work

Scoring alternatives directly [Hindalong et al. 2020]



User studies

Table 7: Mean responses to questions on a 5-point Likert scale (1 = “strongly disagree”, 5 = “strongly agree” unless otherwise specified).

	MER	NLP
Prior experience: (<i>1 = none, 5 = extensive</i>)		
Bar chart	4.83	4.12
Stacked bar chart	4.17	3.00
Box plot	4.00	3.25
Parallel coordinates plot	1.83	1.75
Microsoft Excel	4.50	3.62
Tableau	1.83	2.00
Visualizations in general	3.67	3.12
Usefulness: (<i>1 = not useful, 5 = very useful</i>)		
Bar chart	4.33	4.25
Stacked bar chart	4.67	4.62
Box plot	4.50	3.12
Parallel coordinates plot	2.33	3.12
Legend	4.67	4.38
Mean of Visualization Elements	4.10	3.90
Showing/hiding alternatives	4.00	4.00
Showing/hiding evaluators	4.00	4.12
Sorting by evaluator/total score	4.50	4.88
Switching between comparison modes	3.00	3.62
Showing/hiding the score overlay	3.33	4.12
Linked highlighting	4.50	3.88
Mean of Interactive Techniques	3.89	4.10

	MER	NLP
I modified my own preferences during the session (<i>1 = not at all, 5 = significantly</i>)	4.33	2.12
This tool helps make our discussions more participatory	4.33	4.12
The tool helps identify agreements and disagreements	4.67	4.88
The tool helps make informed decisions based on everyone’s preferences	3.67	4.00
I would be happy if the alternative with the highest aggregate score were chosen	3.67	3.62
Please rate the tool’s potential to affect group interaction (<i>1 = worse, 5 = better</i>)	3.83	4.00
Please rate the tool’s potential to affect information exchange among participants (<i>1 = less exchange, 5 = more exchange</i>)	4.33	3.75
This tool was suitable for the complexity of the decision being made	3.67	3.88
I would like to use this tool for making similar collaborative decisions in the future	4.00	3.62
I would like to use this tool for making more complex collaborative decisions in the future	3.83	3.75

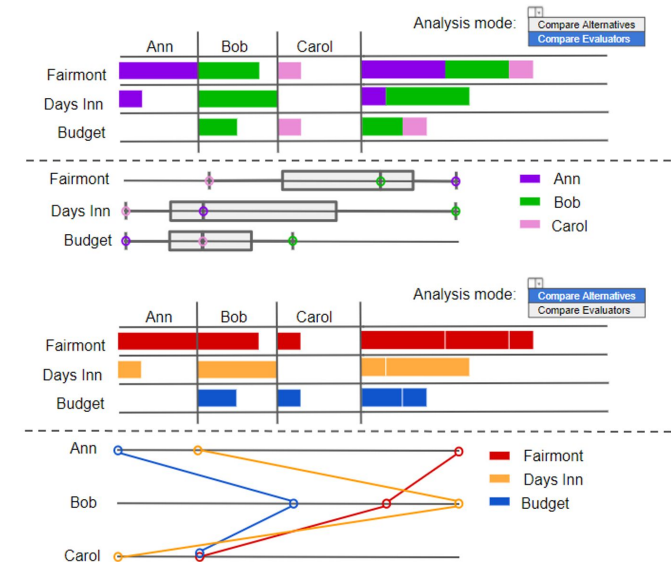


Question Break

Type	Specialized	Professional	Casual
Scenarios	CR, VY, NC	BP, FH, SW	GI
Stakes	Very High	Medium – High	Low
Work Context	Expert Assistance	Professional	Casual
Timeframe	Days	Hours	Hours
Decision Frequency	Once	Monthly – Annually	Once

	AT1	AT2-A	AT2-B	AT3-A	AT3-B	AT4	AT5	AT6-A	AT6-B	AT6-C	AT7	AT8	AT9-A	AT9-B	AT10	Total
Parallel Coords 2	1	2	1	3	2	3	1	1	1	2	n/a	1	3	3	3	27
Parallel Coords 1	1	1	2	2	3	2	2	1	2	0	n/a	2	1	1	2	22
Radar Chart 2	1	1	0	2	0	3	1	1	1	2	1	1	2	2	3	21
Multi-bar 2	2	2	1	1	1	1	1	3	1	2	n/a	1	2	2	1	21
Multi-bar 1	2	1	2	1	1	1	1	3	2	1	n/a	2	1	1	1	20
Tabular Bar 1	3	1	0	2	0	2	0	3	2	3	n/a	0	1	1	2	20
Box Plot 1	1	1	2	1	1	2	3	1	2	0	n/a	3	0	0	2	19
Strip Plot 2	1	3	1	1	1	1	1	1	0	2	n/a	0	3	3	1	19
Box Plot 2	1	2	1	1	1	1	1	1	0	2	n/a	0	3	3	1	18
Heatmap	3	0	0	0	0	1	0	3	2	2	n/a	2	2	2	1	18
Strip Plot 1	1	1	3	1	1	2	2	1	2	0	n/a	2	0	0	2	18
Tabular Bar 2	3	0	1	0	2	0	1	3	3	2	n/a	1	0	0	0	16
Radar Chart 1	1	0	1	0	2	0	2	1	2	2	n/a	2	1	1	0	15
Stacked Bar Chart	1	0	0	0	0	0	0	1	1	2	3	0	0	3	0	11

Figure 4: Support for each auxiliary task function (see Table 6 above for full descriptions) by encoding. 3=best, 2=strongly effective, 1=weakly effective, 0=ineffective. n/a indicates the encoding is not applicable. The Total column contains the row totals. The rows are sorted by Total.





DEMO

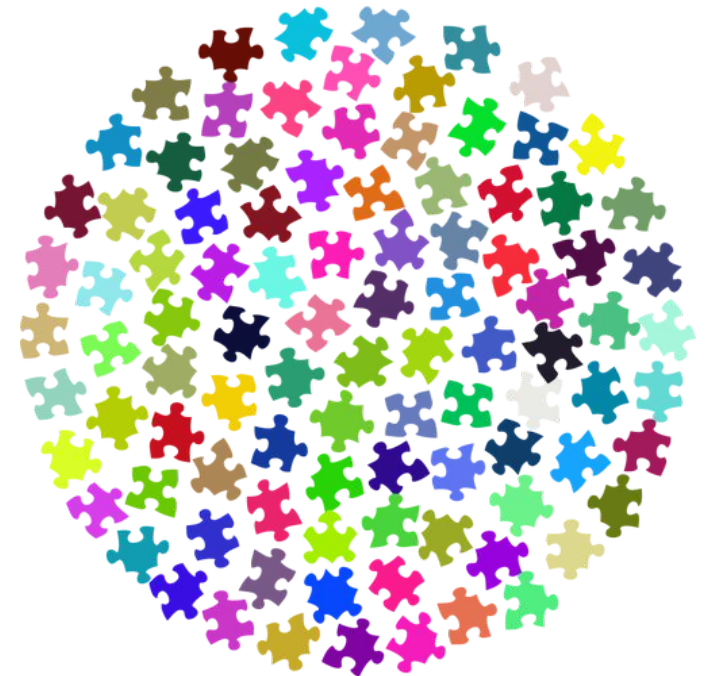
Our work

Extensions to more complex preference models



Scoring alternatives directly isn't always enough

- Choosing a **hotel**:
 - amenities
 - wifi?
 - pool?
 - restaurant?
 - customer service



Our work

Extensions to more complex preference models [submitted for publication]



Extended scenario analysis

- Considers additional **structural features** as well as **characterizations of attributes**

Name	Numbers of Entities			Structural Features		
	Alternatives	Evaluators	Attributes	Evaluator Weights	Evaluator Groups	Attribute Tree
BP	4 - 15	5	N/A			
FH	1 - 4	50 - 100	6 (s)	*	*	
CR	6	15	15 (o)			✓
VY	32	10	N/A	✓		
NC	6	6	7 (m)	✓		✓
SW	2	5	8 (m)	*	*	
GI	3	10	5 (m)	*	*	

Name	No Attributes		Subjective Attributes		Objective Attributes	
	Rank	Score	Score	Score+ Weight	Score+ Weight	Weight only
BP	✓					
FH			✓			
CR	✓	✓				✓
VY	✓	✓				
NC				✓		✓
SW				✓	✓	
GI				✓	✓	

Our work

Extensions to more complex preference models [submitted for publication]



Extended data abstractions

- Take **attributes** into account as well as relative **importance** of different attributes and evaluators

AltAtr Score	Pool			Location		
	Sheraton	Days Inn	Four Seasons	Sheraton	Days Inn	Four Seasons
Ann	0.5	0.5	0	0	0.5	0
Bob	0.2	0.2	0	0.8	0	0.8
Carol	0	0	0	1	0	1



Dimension	Type		Base
EVALUATORS (E)	C		ALT
ALTERNATIVES (A)	C		ALT
ATTRIBUTES (R)	C		ATR
OUTCOMES (O)	{C,Q}		OUT

Measure	Key		
TotalRank	A	$C \rightarrow Q$	-
TotalScore	A	$C \rightarrow Q$	-
EvaluatorWeight	E	$C \rightarrow Q$	-
AltRank	(A,E)	$(C,C) \rightarrow Q$	ALT
AltScore	(A,E)	$(C,C) \rightarrow Q$	ALT
UnweightedAltScore	(A,E)	$(C,C) \rightarrow Q$	-
AtrWeight	(E,R)	$(C,C) \rightarrow Q$	ATR
AltAtrRank	(A,E,R)	$(C,C,C) \rightarrow Q$	ATR
AltAtrScore	(A,E,R)	$(C,C,C) \rightarrow Q$	ATR
UnweightedAltAtrScore	(A,E,R)	$(C,C,C) \rightarrow Q$	ATR
Outcome	(A,R)	$(C,C) \rightarrow \{C, Q\}$	OUT
OutRank	(E,R,O)	$(C,C,\{C,Q\}) \rightarrow Q$	OUT
OutScore	(E,R,O)	$(C,C,\{C, Q\}) \rightarrow Q$	OUT
UnweightedOutScore	(E,R,O)	$(C,C,\{C, Q\}) \rightarrow Q$	OUT
ExternalOutScore	(R,O)	$(C,\{C, Q\}) \rightarrow Q$	w

Our work

Extensions to more complex preference models [submitted for publication]



Extended data and task abstractions

- Preference inspection tasks increased to **34** (up from **12**)
- *No new auxiliary tasks needed*

TASK	
G1a. Discover high-performing alternatives across evaluators	
T1	Discover alternative(s) with best TotalRank/TotalScore
T2	Discover alternative(s) with low variance in AltRanks/AltScores across evaluators
T3	Discover dominated alternatives across evaluators
T4	Discover trade-offs in AltRanks/AltScores between alternatives a and b
T5	Discover pros and cons in AltRanks/AltScores for alternative a
G1b. Discover high-performing alternatives across attributes	
T6	Discover alternative(s) with low variance in AltAtrRanks/AltAtrScores across attributes (aggregated over evaluators)
T7	Discover dominated alternatives across attributes (aggregated over evaluators)
T8	Discover trade-offs in AltAtrRanks/AltAtrScores between alternatives a and b (aggregated over evaluators)
T9	Discover pros and cons of alternative a (aggregated over evaluators)
G1c. Discover high-performing alternatives for a single evaluator	
T10	Discover alternative(s) with best AltRank/AltScore for evaluator e
T11	Discover alternative(s) with low variance in AltAtrRank/AltAtrScore across attributes for evaluator e
T12	Discover dominated alternatives across attributes for evaluator e
T13	Discover trade-offs in AltAtrRanks/AltAtrScores between alternatives a and b for evaluator e
T14	Discover pros and cons of alternative a for evaluator e

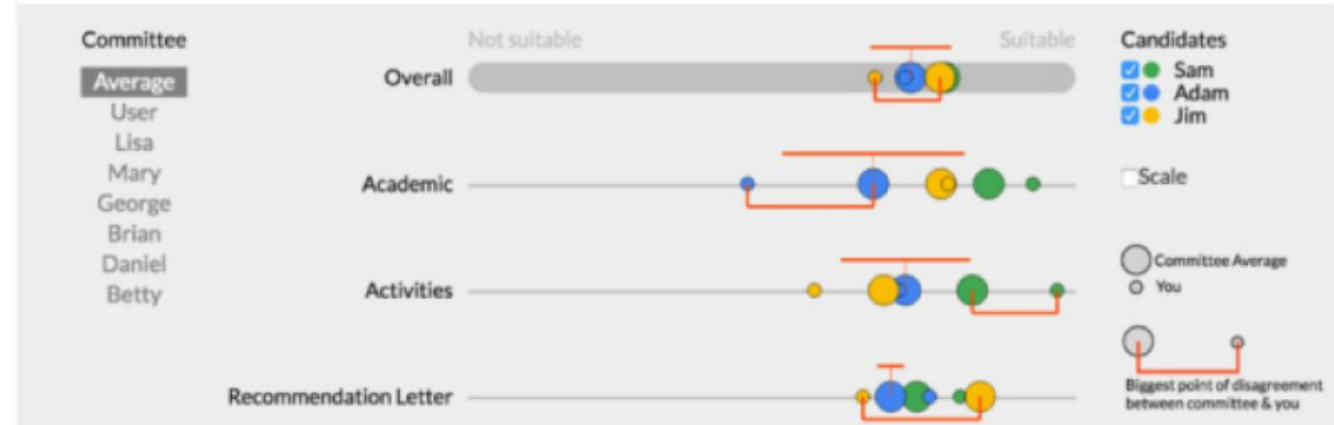
Our work

Extensions to more complex preference models [submitted for publication]



Assessment of existing tools

ConsensUs [Liu et al. 2018]



WebValueCharts
[Mishkin and Hindalong 2018]



Our work

Extensions to more complex preference models [submitted for publication]



Assessment of existing tools

- All tools strongly supported **some** of the visualization tasks
 - None supported **all** of them
- Seem to focus on **one kind** of preference model

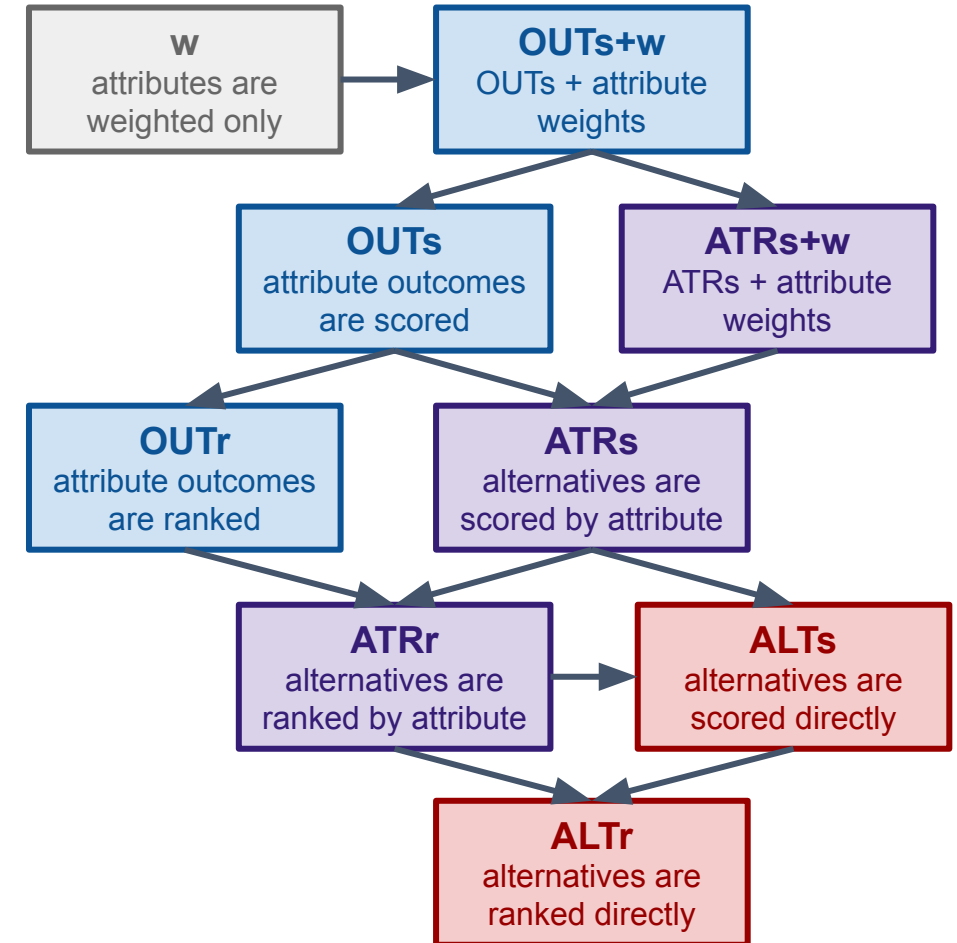
Our work

Extensions to more complex preference models [submitted for publication]



Preference Model Taxonomy

- Classifies preference models
 - **What** is being assessed?
 - **How** is it being assessed?
 - Are attributes being **weighted**?
- Data transformations allow **transitions** between preference models



Our work

Extensions to more complex preference models



Design updates

- Tabular bar chart nested with stacked bars for attributes
 - toggle to tabular bar chart of attributes



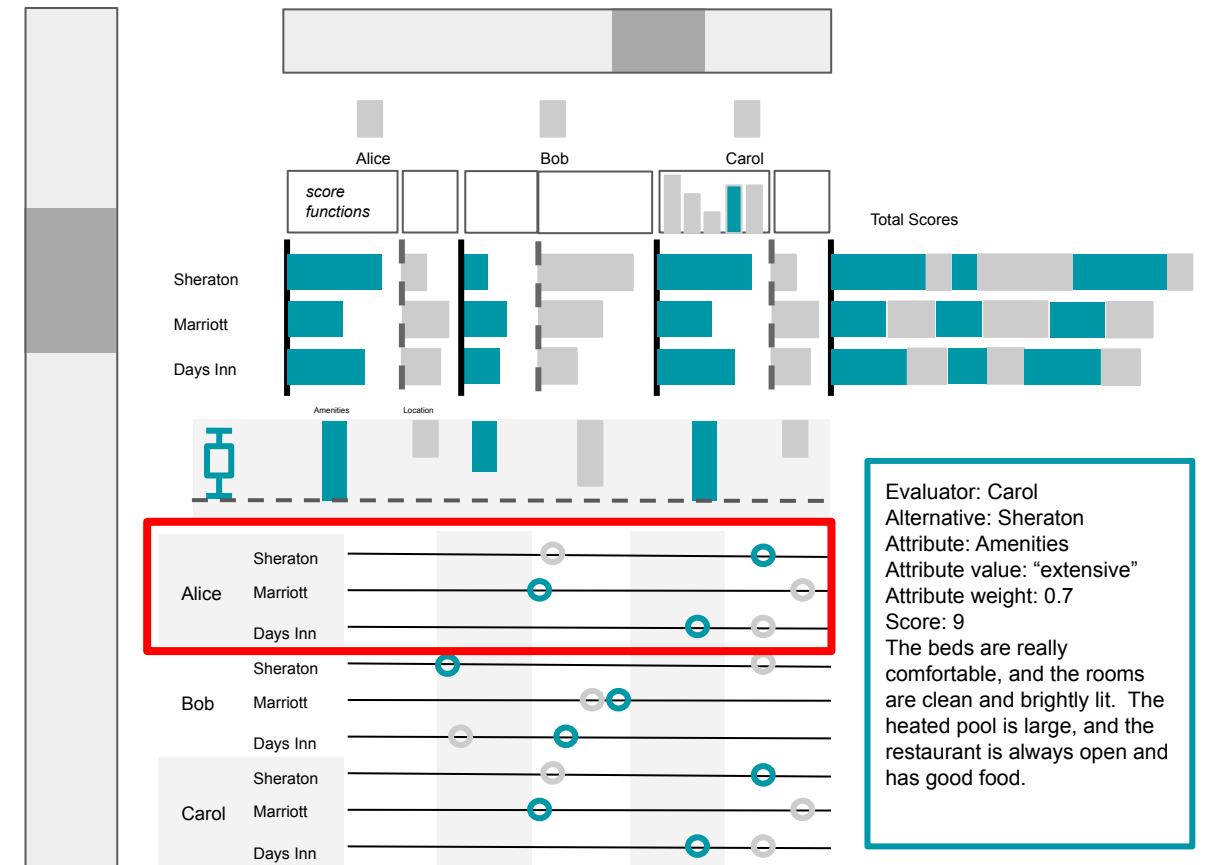
Our work

Extensions to more complex preference models



Design updates

- More flexible strip plots
 - aggregation/splitting options
 - boxes/parallel coordinates on demand



Our work

Extensions to more complex preference models



Design updates

- Optional charts
 - attribute weights
 - evaluator weights
 - attribute outcome scores





- Integrate preference visualizations with data/spatial contexts
- Update our data/task models
 - examine additional group decision scenarios
 - incorporate additional models/theories of group decision making
- Closer look at goal ***G5: Discover nuances not captured by preference models***

References



- Brehmer, Matthew, and Tamara Munzner. "A multi-level typology of abstract visualization tasks." *IEEE transactions on visualization and computer graphics* 19.12 (2013): 2376-2385.
- Hansen, Paul, and Franz Ombler. "A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives." *Journal of Multi-Criteria Decision Analysis* 15.3-4 (2008): 87-107.
- **Hindalong, Emily, et al. "Towards Rigorously Designed Preference Visualizations for Group Decision Making." 2020 *IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, 2020.**
- Hindalong, Emily, et al. *Ongoing work under submission for publication. Expected 2021.*
- Liu, Weichen, et al. "ConsensUs: Supporting multi-criteria group decisions by visualizing points of disagreement." *ACM Transactions on Social Computing* 1.1 (2018): 1-26.
- Mishkin, Aaron and Emily Hindalong. 2018. *Web ValueCharts*. <http://www.cs.ubc.ca/group/iui/VALUECHARTS/>
- Munzner, Tamara. "A nested model for visualization design and validation." *IEEE transactions on visualization and computer graphics* 15.6 (2009): 921-928.
- Munzner, Tamara. *Visualization analysis and design*. CRC press, 2014.

All clipart was obtained from <https://publicdomainvectors.org>



Thank you for your time

You may contact me at jordon@cs.ubc.ca

QUESTIONS?