Generating Coherent Visualization Sequences for Multivariate Data by Causal Graph Traversal

Puripant Ruchikachorn, *Member, IEEE,* Darius Coelho, Jun Wang, *Member, IEEE,* Kristina Striegnitz, and Klaus Mueller, *Fellow, IEEE*

Abstract—Multivariate data contain an abundance of information and many techniques have been proposed to allow humans to navigate this information in an ordered fashion. For this work, we focus on methods that seek to convey multivariate data as a collection of bivariate scatterplots or parallel coordinates plots. Presenting multivariate data in this way requires a regime that determines in what order the bivariate scatterplots are presented or in what order the parallel coordinate axes are arranged. We refer to this order as a *visualization sequence*. Common techniques utilize standard statistical metrics like correlation, similarity or consistency. We expand on the family of statistical metrics by incorporating the rigidity of causal relationships. To capture these relationships, we first derive a causal graph from the data and then allow users to select from several semantic traversal schemes to derive the respective chart sequence. We tested the sequences with a crowd-sourced user study and a user interview to confirm that the causality-informed visualization sequences help viewers to better grasp the relationships that exist in the data, as opposed to sequences derived from correlations or randomization alone.

Index Terms—Causality, Causal Graph, Visualization Sequence, Multivariate Visualization, Parallel Coordinates

1 Introduction

Multivariate data analysis can enable researchers and practitioners to uncover meaningful insights, make informed decisions, and optimize processes. Examples include analyzing healthcare data to identify associations between patient demographics, medical history, and treatment outcomes; investigating consumer behavior data to uncover connections between customer demographics, product preferences, and marketing campaigns; and exploring environmental data to understand the impact of climate change on temperature, precipitation, and species diversity.

A prime goal in multivariate data analysis is to uncover the intricate relationships that exist among the variables, conveying a comprehensive understanding of the complex system that underlies the data. An attractive method for this purpose is to learn a causal model from the data. Causal models can effectively elucidate linear relationships and interdependencies that exist among the variables and are easily visualized as node-link diagrams. While these models can reveal more semantics than simple summary statistics, they do not allow the analyst to detect irregularities, outliers, and unusual data patterns. Conversely, visualization capitalizes on the unsurpassed power of the human visual system to quickly recognize these types of irregularities.

There is a large arsenal of visualization methods, at a wide gamut of complexity. A bivariate scatterplot is the simplest visualization that can give a viewer insight into data relations, but among two variables only. A Parallel Coordinates Plot (PCP) [1] draws each data point as a line across several parallel data axes and as such can visualize more than two variables. Other techniques for multivariate data employ data embeddings, such as MDS [2], t-SNE [3], and UMAP [4], to name the most prominent, but their scatterplot projections primarily focus on the depiction of clusters and neighborhood relations.

Bivariate scatterplots clearly reveal relationships between two continuous variables. Additional attributes can be encoded using depth in 3D scatterplots or retinal variables such as color, shape, and size. But 3D plots often suffer from perceptual issues like occlusion, distortion, and visual clutter [5], while overloading 2D plots with more than one or two retinal encodings can reduce interpretability [6].

To explore multivariate relationships, one alternative is to construct a sequence of bivariate scatterplots. However, with n variables, the number of plots and possible orderings grows rapidly—for example, six variables yield 15 scatterplots and 15! permutations, while PCPs give rise to 6! axis orderings—both resulting in a combinatorial explosion.

Bridging the gap between causal inference and data visualization, we introduce a novel approach that leverages a causal network to create a coherent sequence of bivariate scatterplots or PCP axes. The order in which the plots are presented ensures that each visualization builds upon the previous one, progressively unveiling the underlying causal structure with visual evidence. Viewers follow the sequence from start to finish, thereby grasping the intricate relationships in multivariate data. As a causal graph can be traversed in many ways, multiple causal sequences can be created, each uniquely narrating the same data. We use the term *coherent* to emphasize causal consistency and directional interpretability in these sequences, distinguishing them from arbitrary or correlation-based orderings.

Our paper is organized as follows. Section 2 lists some

- *Puripant Ruchikachorn, Darius Coelho, Jun Wang, and Klaus Mueller are or were with the Visual Analytics and Imaging Laboratory, Department of Computer Science, Stony Brook University, Stony Brook, NY. Emails: {pruchikachor, dcoelho, junwang2, mueller}@cs.stonybrook.edu.*
- *Puripant Ruchikachorn is now with the Chulalongkorn Business School, Chulalongkorn University, Bangkok, Thailand. Email: puripant@cbs.chula.ac.th.*
- *Kristina Striegnitz is with the Computer Science Department, Union College, Schenectady, NY. Email: striegnk@union.edu.*

related work. Sections 3 and 4 present our narrative graph traversal schemes and prototype to create and edit the resulting sequences. Section 5 presents quantitative and qualitative validations. Section 6 and 7 discuss the result, limitations of our work, and some future work.

2 Related Work

We divide the related work into two main areas, namely data visualization with bivariate scatterplots and with PCPs, specifically emphasizing information navigation. Further, we also touch on work in the domain of narrative data visualization to provide a wider context to our proposed data navigation method which we refer to as causal narratives.

2.1 Information Navigation with Scatterplots

A scatterplot can show the relationship between two variables and allow its viewers to immediately understand how strong the relationship is. A shortcoming of this visualization is that it can only display two variables at a time, and so a natural extension is to combine multiple scatterplots into a matrix called Scatterplot Matrix (SPLOM) [7]. It is practical to order the plots by strength of relationship [8], such as correlation or more complex notions of similarity [9], but the order of visual traversal of the plots is still left to the viewer, with no guidance provided.

A more sophisticated method is to use a quality-based visualization matrix [10] whose rows are ordered according to overall dimension quality gauged by e.g. Class Consistency [11]. One may also take advantage of the concept of Scagnostics [12] which parameterizes scatterplots as a vector of appearance attributes, such as skew and stringiness. The ScagExplorer [13] performs clustering in this space and identifies a set of leader and follower plots which are then laid out into various spatial configurations. None of these methods, however, provides a unique sequence of plots that can be "read" from start to finish.

Interaction can significantly improve information navigation. The groundbreaking work by Elmqvist et al. [14] presented the Rolling Dice method that allows users to perform animated transitions among scatterplot matrix tiles that share a common dimension. The motion parallax in the temporary 3D animation fosters a better understanding of the information shared among the two tiles. Users can design multi-tile tours across the scatter matrix (with multiple transitions) and so gradually increase their understanding of the multivariate space. However, there is no analytical guidance to specify the paths of these tours. Our method can complement this paradigm by providing automatically generated causal tours to accelerate data understanding without the need for interaction.

While scatterplot matrices are powerful tools for visualizing pairwise relationships, they are rarely used in mainstream visual storytelling—especially for general audiences. Even highly data-literate outlets like *The New York Times* seldom use scatterplot matrices, including small ones (e.g., 3×3), but frequently employ individual scatterplots to convey complex ideas in intuitive ways. Our goal is to preserve analytical rigor while designing visualizations that are familiar and cognitively approachable, following prior work in visual communication and journalism that emphasizes simplicity, familiarity, and narrative flow as essential for engagement and comprehension [15]. We achieve this by reinterpreting the scatterplot matrix as a causal sequence of bivariate plots, making multivariate data more accessible.

A generalization of the scatterplot is the biplot [16] where data points and dimension axes are projected into the two most significant principal components. Since the projection is linear, the only distortion is due to the lack of contribution of the less significant principal components. The Subspace Voyager [17] allows users to tilt a biplot with a 3D trackball interface, transitioning into adjacent subspaces for a gradual exploration of the multivariate space. This interface is a generalization of the Rolling Dice method and can be supported by our causal network traversal schemes.

Another method to explore multivariate data through a sequence of projections is the Grand Tour [18], which uses projection pursuit [19] to generate dynamic views by optimizing statistical "interestingness"—such as variance, clustering, or correlation. These sequences are mathematically driven and may lack semantic or domain interpretability. In contrast, our method produces a deliberate sequence of standard scatterplots by traversing a causal graph, where each step follows a directed dependency. This results in visualization sequences that are not just statistically informative, but also interpretable in terms of cause–effect mechanisms grounded in domain understanding.

2.2 Information Navigation with Parallel Coordinates

The strength of a PCP is that multivariate data can be read from left to right, forming a narrative structure told by the order of the data dimensions. The information that can be discerned this way strongly depends on the ordering of the axes, and this has been a topic of extensive research, both on automatic and on interactive methods.

Ankerst et al. [9] proposed a similarity metric that compares two dimensions based on the root-mean-square distances of all data points. The dimension ordering is then optimized via an approximate traveling salesman solver. Tatu et al. [20] introduced a similarity-based measure based on Hough transforms. Johansson and Johansson [21] defined a weighted metric that assesses dimensions by their importance in correlation, outlier detection, and subspace cluster significance to select relevant dimensions for a PCP. Similarly, Artero et al. [22] explored the correlation of dimension pairings to optimize the ordering. Zhang et al. [23], [24] allowed users to interactively design an axis ordering on a correlation map of the variables.

Dasgupta and Kosara [25] proposed several metrics based on the visual appearance of polylines to determine optimal dimension orderings. Peng et al. [26] focused on minimizing clutter and outliers between adjacent dimensions while establishing the ordering. Ferdosi and Roerdink [27] ordered the dimensions based on high-dimensional structures identified through subspace clustering. Finally, Yang et al. [28] employed hierarchical dimension filtering in conjunction with dimension clustering using Ankerst's metric, and subsequently enabled users to interactively navigate the hierarchy to create their desired ordering. Regardless of whether the aforementioned works are interactive or not, none utilize causality as their metric for ordering PCP dimensions into a sequence.

Blumenschein et al. [29] investigated how different data patterns, such as correlation, outliers, clusters, skew, and clutter are exposed by the various existing ordering methods. Extending this study, Tyagi et al. [30] developed real-time detection schemes for these patterns and implemented an interactive visual user interface by which users could specify preferences and generate the most appropriate dimension ordering. These approaches focus on interactive PCP axis ordering guided by data patterns while ours aims to automate the ordering process with respect to the user's chosen causal graph traversal strategy.

2.3 Narrative Data Visualization

Our method can be considered an approach for narrative visualization because a visualization sequence chooses which visualization appears first, similar to how a narrative orders scenes. However, the overall design space of narrative visualization is far broader than sequencing. The seminal work of Segel and Heer [15] offered a systematic review of the design space and narrative structures. They described seven fundamental narrative genres that have different characteristics including one without a prescribed ordering [31]. These so-called reader-driven narratives are not in the scope of our work. We focus on a subset of strict author-driven narratives which visually show data in linear sequences directed by causality. Hence we will refer to the narratives we create as *causal* narratives.

3 Methodology

In this section, we describe how causal graphs are used to generate visualization sequences via different traversal strategies, or "narratives." Figure 1 shows an example input graph and resulting sequences for both PCPs and scatterplots. Each narrative corresponds to a specific strategy and is illustrated using examples from multiple datasets, as implemented in our prototype (Section 4).

3.1 Causal Graph Generation

The first step of our method is to generate a causal graph that represents the directional dependencies between a dataset's variables. Following the foundational work by Pearl [32] and Spirtes et al. [33], we represent causal relationships as a directed acyclic graph (DAG), where nodes are variables and edges denote causal influence. Each edge has a polarity (positive or negative) and a strength derived from partial correlation tests.

To construct the causal graph several causal model packages can be used, such as TETRAD, DoWhy, and the causal-learn library. We make use of the interactive visual causality tool by Wang et al. [34] [35], which allows analysts to edit the causal graph when the automatically inferred structure does not fully align with domain knowledge. Without loss of generality, we currently support only numerical data.

What sets causal analysis apart is that it provides stronger explanatory power than correlation alone. While correlation captures statistical associations, causal relationships imply directionality and enable reasoning about counterfactuals—the highest level on Pearl's "causal ladder" [36]. For example, a car's weight affects its fuel economy (MPG), but not vice versa. A causal graph reflects this asymmetry by assigning a directed edge from weight to MPG, enabling sequences that follow meaningful, cause-effect orderings. In contrast, correlation cannot determine direction and may lead to spurious or misleading sequences that obscure the true structure of the data.

However, causal discovery algorithms rely on strong assumptions, such as the absence of hidden confounders and the sufficiency of the measured variables. These assumptions may not hold in practice, and no algorithm can guarantee recovery of the true causal structure without expert input. To address this, the system by Wang et al. [34] allows users to edit and refine the causal graph, enabling domain knowledge to guide the visualization sequences.

3.2 Narrative Graph Generation

With the causal graph of a dataset in place, we can now construct a narrative graph. Since each edge in a causal graph is a causal relation, it represents a logical progression, which is a key ingredient of any author-driven narrative. The causal narratives we are interested in are driven by the presentation of causal relationships that reveal dependencies between different variables. Shown as a directed node-link diagram, a narrative graph traverses a causal graph and orders visualizations into a sequence.

For a bivariate scatterplot, each directed causal edge generates a chart where the cause variable is mapped to the x-axis and the effect variable is mapped to the y-axis. They correspond to the tail and the head of an arrow, respectively. Following the next causal edge, the previous effect variable becomes the cause variable and switches to the x-axis. The new effect variable takes its place on the y-axis. This generates a sequence where adjacent scatterplots are causally connected as a chain.

Likewise, for a PCP, the cause and effect variables of a causal edge are assigned to the left and right axes. Because a PCP can show more than two variables at a time, the next causal edge need not switch axes and can simply add a variable as another axis to the right of the existing ones. A causal chain can be presented as one visualization in PCP.

The main task of generating a narrative graph from a causal graph is to select which causal relationships or chains to include in the narrative and to decide in which order to present them. The selection of causal chains is determined by the type of narrative. Each chain is then turned into a node or a sequence of nodes in the narrative graph. Note that this approach to narrative generation is independent of the domain of the data.

Figure 2 shows narrative graphs constructed from three basic causal relationships: causal chain (mediator), common cause (confounder), and common effect (collider). Different causal graph structures are expected to lead to different narrative graphs. Also, different visualization types can produce different narrative graphs. For instance, a simple chain of causal relationships can be represented by one PCP, while a sequence of bivariate scatterplots (one per causal relation or variable pair) is necessary to convey the same information. Our current prototype implements only two visualization types: scatterplot and PCP.

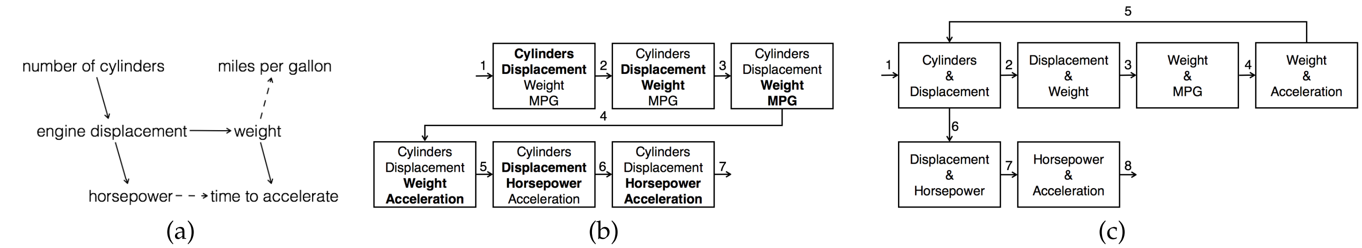


Fig. 1: (a) Causal graph for the *Cars* dataset. Solid and dashed edges denote positive and negative weights, respectively. (b) Causal sequence for the *exhaustive* narrative graph traversal scheme designed to be used with a PCP. The dimension pairs in bold font indicate the focused causal relations in each panel. (c) Causal sequence for the *exhaustive* narrative traversal scheme designed to be used with bivariate scatterplots.

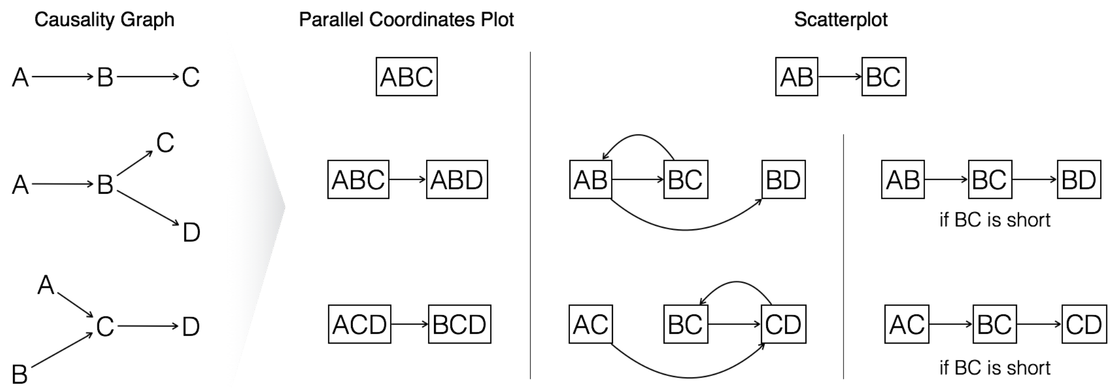


Fig. 2: Different narrative graphs are the result of both the graph topology and visualization type: PCP and scatterplot. From top to bottom, examples here are of three typical patterns: causal chain (mediator), common cause (confounder), and common effect (collider). **A**, **B**, **C**, and **D** denote different data variables that are connected by arrows to form different causal graphs. A rectangle surrounding the letters indicates a visualization of respective variables; a PCP can show many variables at once so a box can contain **ABC**, for instance, while a scatterplot can present at most two variables such as **AB**. A complete traversal in PCP may simply be one plot while the same traversal in scatterplot may need more plots to *flashback* to go through all variables. If a chain, say **BC**, is short, it may not be necessary to traverse back to the previous chain in order to move forward to the next chain.

As shown in Figure 2, each narrative graph node can be revisited as many times as necessary. In some cases, this creates a *flashback* to the sequential flow which we will explain further below.

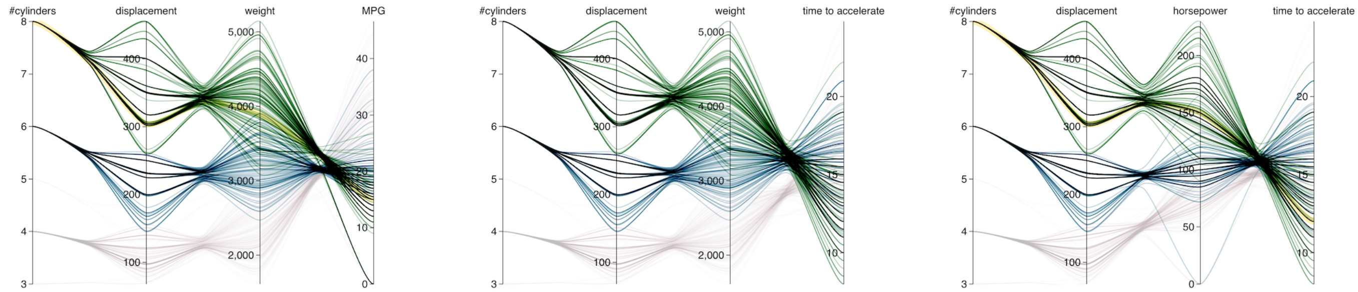
As a causal graph can be traversed in numerous ways, there are many possible narratives. In this paper, we propose four different traversal schemes which lead to four different narratives: (1) exhaustive, (2) detective, (3) dramatic, and (4) journalistic narratives. Some narratives may be more appropriate for certain audiences or purposes. Our goal at this point is not to automatically select the best narrative but to suggest some possible ones to the users of our prototype. The user interface, as described in the supplementary materials allows choosing between these four generated narratives and further edit them.

Exhaustive Narrative

This narrative gives a complete overview of all causal relationships, so it is often the longest and the most detailed of all proposed narratives. For a connected graph structure, common graph traversal algorithms are directly applicable to this task. We chose the depth-first search algorithm as it naturally follows causal chains and backtracks to the closest unvisited variables. Causal graph nodes that have no incoming edges, i.e. they are not an effect of any cause, serve as the root nodes for the search algorithm. The search extracts all chains from those roots to any leaf nodes, which have no outgoing edges or are not a cause of any effect. As a causal graph is a directed acyclic graph, there are always such root and leaf nodes and the algorithm always works.

As mentioned earlier, different visualization types may result in different narrative graphs, given the same causal graph. A PCP can display several variables in one visualization so it can show an entire causal chain. An *exhaustive* narrative graph for PCP can thus simply present all causal chains from the search algorithm. To minimize abrupt visual changes, all chains are sorted by length and then swapped to reduce local edit distances between adjacent paths. The polylines are also bundled and colored for better readability.

Figure 1(b) shows an *exhaustive* narrative graph for a PCP of the *Cars* dataset [37]. There are three causal chains, all starting with the causal relation from '#cylinders' to 'displacement' and then branching off into two relationships of the same cause variable 'displacement' to two different effect variables, 'weight' and 'horsepower'. Two out of three chains end with 'time to accelerate' while the final variable of the third chain is 'MPG'. Figure 3 presents these three

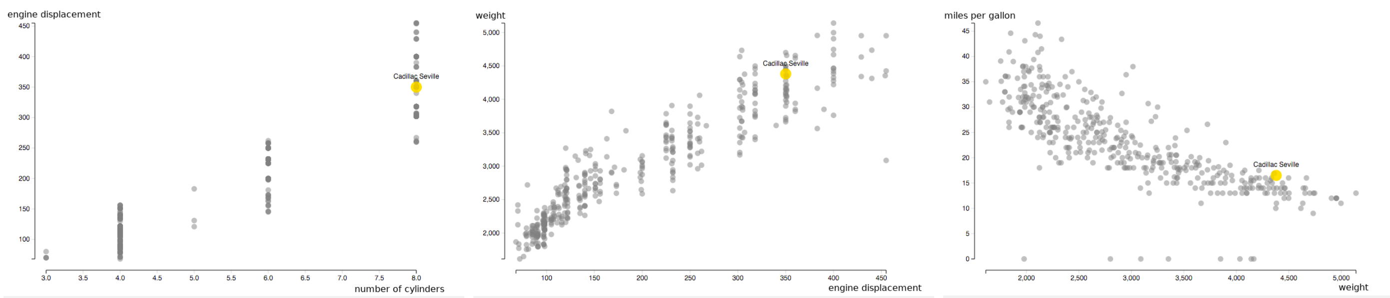


Chain 1. Data in view: #cylinders, displacement, weight and MPG. #cylinders and displacement are positively correlated (r=0.95). Displacement and weight are positively correlated (r=0.93). Weight and MPG are negatively correlated (r=-0.79).

Chain 2. #cylinders, displacement, weight and time to accelerate. #cylinders and displacement are positively correlated (r=0.95). Displacement and weight are positively correlated (r=0.93). Weight and time to accelerate are weakly negatively correlated (r=-0.43).

Chain 3. #cylinders, displacement, horsepower (HP) and time to accelerate. #cylinders and displacement are positively correlated (r=0.95). Displacement and HP are positively correlated (r=0.87). HP and time to accelerate are negatively correlated (r=-0.68)

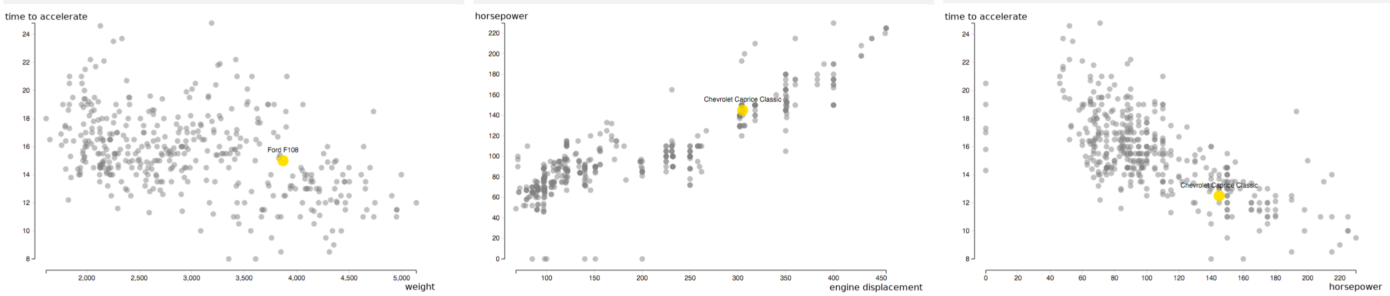
Fig. 3: An automatically generated *exhaustive* sequence for the *Cars* dataset with PCPs.



Frame 1. Data in view: #cylinders (x-axis) and displacement (y-axis). #cylinders and displacement are positively correlated (r=0.95).

Frame 2. Data in view: displacement (x-axis) and weight (y-axis). Displacement and weight are positively correlated (r=0.93).

Frame 3. Data in view: weight (x-axis) and MPG (y-axis). Weight and MPG are negatively correlated (r=-0.79).



Frame 4. Data in view: weight (x-axis) and time to accelerate (y-axis). Weight and time to accelerate are weakly negatively correlated (r=-0.43).

Frame 5. Data in view: displacement (x-axis) and horsepower (y-axis). Displacement and horsepower are positively correlated (r=0.87).

Frame 6. Data in view: horsepower (x-axis) and time to accelerate (y-axis). Horsepower and time to accelerate are negatively correlated (r=-0.68).

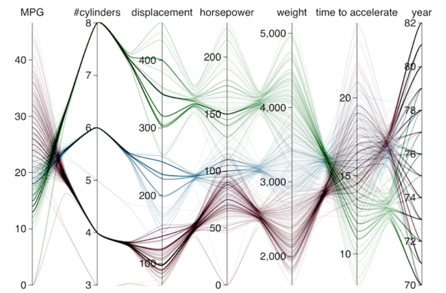
Fig. 4: An automatically generated *exhaustive* sequence for the *Cars* dataset with scatterplots. Not shown is the *flashback* frame after Frame 4, where Frame 1 would be repeated after Frame 4.

chains as a sequence in PCPs.

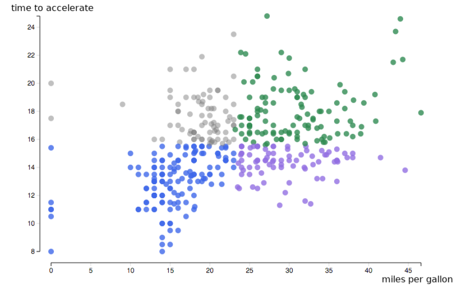
A scatterplot can present only two variables at a time, so each visualization of an *exhaustive* narrative for scatterplot can display only one causal link between two variables. The first visualization starts from one of the root variables. While traversing, the current and the next causal relations share a variable, i.e. the effect of the former and the cause of the latter. When the effect variable is a leaf variable, the search algorithm backtracks to an unexplored branch and the *exhaustive* narrative structure includes a *flashback* to the nearest scatterplot whose effect variable is the cause variable of the new chain's first scatterplot.

Figure 1(c) shows an *exhaustive* narrative graph for scatterplots of the *Cars* dataset and Figure 4 presents its corresponding sequence. Due to space limitation, the figure does not show the *flashback* visualization that is the scatterplot of the number of cylinders and engine displacement which is the first cause variable of the new chain.

To accentuate the causal flow in a sequence, certain data samples are selected to give examples to the narrative. They are highlighted as a yellow mark in a scatterplot and a yellow polyline in a PCP. For positively correlated variables, the system selects one of the data samples whose values are high in both variables. Conversely, data samples whose values are either high and low or low and high are good examples for negatively correlated variable pairs. Picking a data sample of both types is done via rejection sampling which randomly chooses a sample until all criteria are met. All adjacent variable pairs in a PCP are simultaneously considered. The same method to select data samples as examples also applies to other narratives.

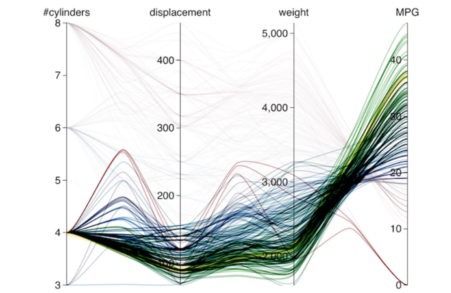


Intro. Data dimensions: MPG, #cylinders, displacement, horsepower, weight, time to accelerate and year



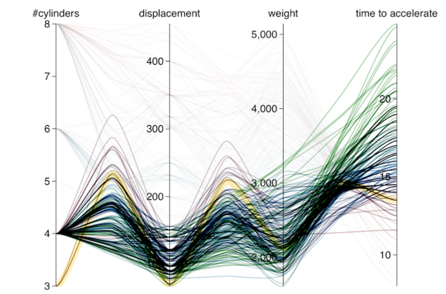
Goal 1: high MPG, **Goal 2:** low time to accelerate

Coloring indicates if samples meet both (purple), one (blue/green), or neither (grey) goal polarities. Colors differ from those used for clustering in the PCPs.



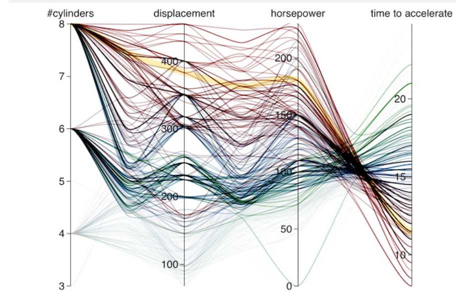
Optimize MPG by (chain 1): decreasing weight, decreasing displacement, and decreasing #cylinders.

The highlighted example has: low #cylinders, low displacement, low weight, and high MPG.



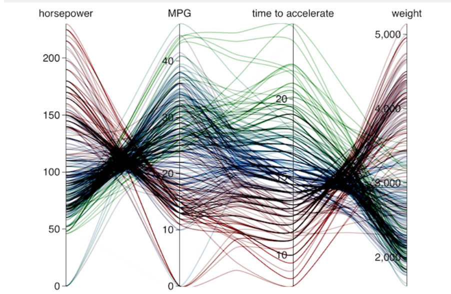
Optimize time to accelerate by (chain 2): decreasing weight, decreasing displacement, and decreasing #cylinders.

The highlighted example has: low #cylinders, low displacement, low weight, and low time to accelerate.



Optimize time to accelerate by (chain 3): increasing horsepower (HP), increasing displacement, and increasing #cylinders.

The highlighted example has: high #cylinders, high displacement, high HP, and low time to accelerate.



Suggestions: increase horsepower and decrease weight.

Fig. 5: Selected panels from an automatically generated *detective* sequence for the *Cars* dataset. Some visualizations are developed incrementally over the course of several panels. PCPs can be built up incrementally adding one dimension at a time. Here, we only show the final views. The data examples are highlighted in yellow

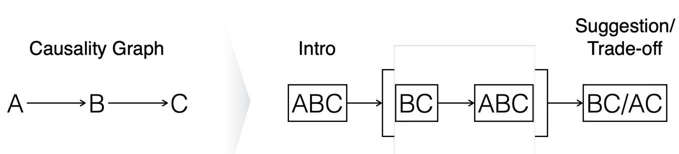


Fig. 6: A detective or journalistic PCP sequence from a simple causal chain starts with an introduction of all variables then builds up the chain from the goal or effect variable. The two narratives lead to differing conclusions. A detective sequence aims to show how to optimize the goal variables, while a journalistic sequence points out how the goal variables cannot be optimized at the same time.

Detective Narrative

An *exhaustive* narrative and sequence may seem uneventful as it traverses all causal relations from cause to effect. On the other hand, a *detective* narrative reverses the traversal, unraveling from an effect or ultimate goal to immediate causes and finally unveiling the primary causes. In the same fashion as a *whodunit* story [38], this narrative functions like a puzzle for a reader to deduce answers or cause variables from clues or effect variables.

All leaf variables without outgoing edges are legitimate goal variables and can be traced back through their causal chains to explain themselves. The goal variables and their desired polarities are user-defined. The *Cars* dataset has two leaf variables: 'MPG' and 'time to accelerate'. The former should be high while the latter should be low.

To effectively use all supported visualization types in a *detective* narrative, we mix scatterplot and PCP together and use them for showing exactly two variables and more than two variables, respectively. A simple causal chain turns into a visualization sequence as illustrated in Figure 6 for PCP.

Figure 5 shows the visualization sequence from a *detective* narrative for the *Cars* dataset. The sequence begins with an introductory PCP that presents all data variables. The goal variables are shown in a scatterplot because there are two goal variables. Then each chain is presented from its goal variable; a new axis is incrementally added to the left of the PCP. After presenting the whole chain, a data sample that optimizes its goal is highlighted as an example. For instance, the third chart of Figure 5 shows the causal chain of '#cylinders', 'displacement', 'weight', and 'MPG' with a highlighted data sample that has low '#cylinders', low 'displacement', and light 'weight' and leads to the desired goal of high 'MPG'. We use the same rejection sampling as explained in the *exhaustive* narrative.

In the end of the *detective* narrative a conclusion summarizes what variables should be increased or decreased to achieve the optimal values of the goal variables. This is particularly useful for conflicting causal relations. For example, the two causal chains to optimize 'time to accelerate' in the *Cars* dataset are irreconcilable. One chain needs low '#cylinders' and 'displacement' while the other requires the opposite. Our system recognizes these conflicting variables

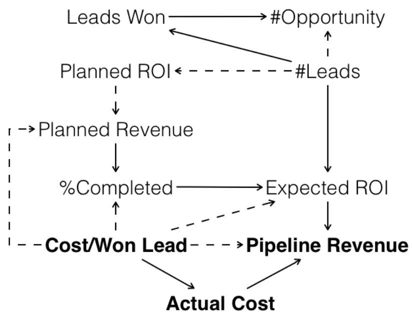


Fig. 7: Causal graph for the *Sales* dataset. Bold nodes involve in a causal and narrative conflict.

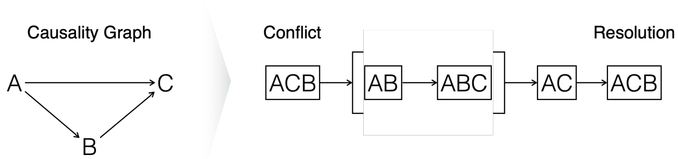


Fig. 8: A dramatic PCP sequence from a simple causal conflict starts with an overview of the conflict, shows the two contrasting causal paths, and concludes with a resolution.

further down the causal chains for all goals and excludes them from the suggestions in the summary.

In the last chart of Figure 5, the suggestions for the *Cars* dataset involve only the immediate causes 'horsepower' and 'weight'. The goal variables in the suggestions are shown in the middle of a PCP with the suggested immediate causes to the left and right. This chart succinctly reveals an engineering challenge exposed by the causal network, i.e. the need to increase horsepower while simultaneously keeping the weight low. The polyline coloring effectively visualizes this. The maroon lines (high 'horsepower') lead to low 'MPG' but achieve low 'time to accelerate', while the green lines (low 'weight') lead to high 'time to accelerate' but achieve high 'MPG'. A solution to this challenge may be incorporating more lightweight material that reduces weight and does not affect acceleration time.

Dramatic Narrative

The conflicting variables in the *detective* narrative and Freytag's pyramid of dramatic structure [39] inspired the *dramatic* narrative. The dramatic structure of five stages — exposition, rising action, climax, falling action, and resolution — appear in many stories and should engage a reader's attention as it revolves around a *conflict*. We propose that a *causal conflict* can serve as a dramatic conflict and drive a dramatic narrative.

To create such a narrative, a conflict must be formally established. We use the definition of a causal collider or a variable with more than one cause to establish a conflict in a causal graph. A causal conflict occurs when a collider has at least a pair of causes that are directly linked and their two different paths are inconsistent.

The *Cars* dataset does not possess such a conflict, but the causal graph for the *Sales* dataset [23] shown in Figure 7 contains one. In this graph, we observe that 'Cost/Won Lead' positively affects 'Actual Cost' which also positively influences 'Pipeline Revenue', while 'Cost/Won Lead' directly causes 'Pipeline Revenue' in a negative way.

To construct a *dramatic* narrative graph, all colliders are identified and each pair of causes of the same collider is analyzed. First, both causes and the common effect are displayed in the same visualization to emphasize their confounding relations. Then a step-by-step walkthrough of direct influence (the first cause to the collider) and indirect influence (the first cause to the second cause to the collider) is presented. In the end, a statistical analysis shows the dominant path and cause. Data samples are highlighted to visually aid causal inspection. The causal conflict and its resulting visualization sequence are illustrated in Figure 8.

Figure 9 shows the crucial parts of the *dramatic* narrative generated based on the *Sales* dataset. The narrative evolves around whether 'Cost/Won Lead' or 'Actual Cost' is more influential and it concludes that it is 'Cost/Won Lead' that actually affects 'Pipeline Revenue'. The conflict resolution should invite a reader to take a closer look at the data; the polylines for 'Cost/Won Lead' look far more coherent than those for 'Actual Cost' which have a more random appearance suggesting a weaker causality.

Note that only single-level causal inconsistencies are currently taken into consideration. Multilevel conflicts such as the paths of displacement-weight-acceleration and displacement-horsepower-acceleration in the *Cars* dataset are possible but may result in a complicated narrative.

Journalistic Narrative

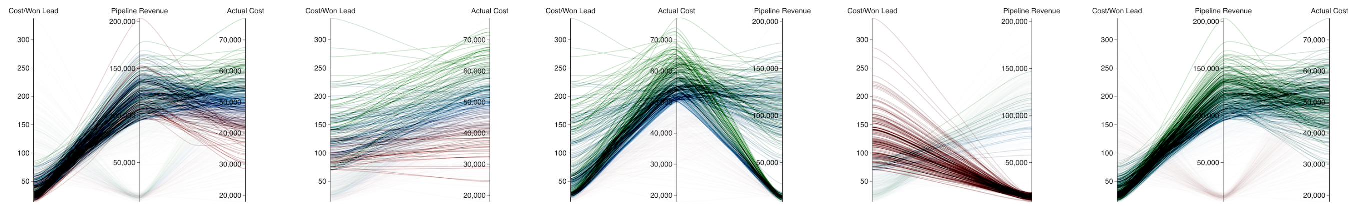
Another way to show conflicting variables is to present them as a trade-off. Because this narrative is similar to how a journalist would write a balanced story [40], we call it a *journalistic* narrative. The process to find a conflict is the same as a *dramatic* narrative. Similar to a *detective* narrative, a *journalistic* narrative begins with an overview of all variables. Then it shows user-defined goals and explains the causes of each goal in a progressively built-up causal chain in PCP, as shown in Figure 6.

Instead of a summary with suggestions as in a *detective* narrative, a *journalistic* narrative points out the variables that optimize only a subset of all goals. Such causes can be found in a network among the lowest common ancestors of the goals. The data polarities of all goal variables — whether high or low values are preferable — are propagated through the causal network. A trade-off happens when there is a variable whose propagated polarities are opposing i.e. its values cannot be adjusted to optimize two incompatible requirements at the same time.

In Figure 10, engine displacement in the *Cars* dataset is the variable that trades off 'MPG' and 'time to accelerate'. The system highlights two extreme cases and the PCP in the summary tells the reader to choose between them.

4 Implementation

We implemented a prototype for creating a chart sequence with the help of several libraries such as Simple Statistics for basic data statistics, D3.js for visualizations, and Syntagmatic's Parallel Coordinates for PCP bundling. All animations and interactions were responsive on any modern



Conflict: Cost/Won Lead and Actual Cost explain Pipeline Revenue in conflicting ways.

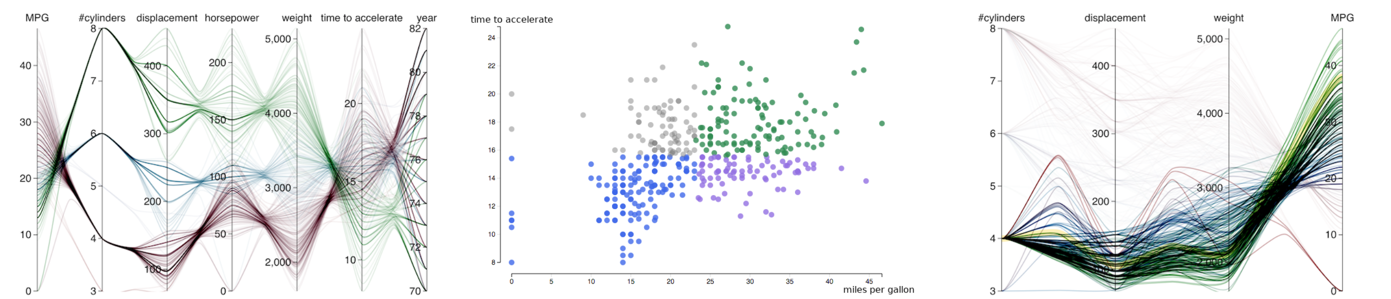
Protagonist's view 1: High Cost/Won Lead causes high Actual Cost.

Antagonist's view: High Actual Cost causes high Pipeline Revenue; some exceptional low Cost/Won Lead earn higher Pipeline Revenue (at 75 C/WL, the broken up blue line bundle).

Protagonist's view 2: However, high Cost/Won Lead causes low Pipeline Revenue, with the exceptions at 75 C/WL (the blue line bundle).

Resolution: A statistical analysis concludes that Cost/Won Lead dominates the effect of Pipeline Revenue, confirming the observed exceptions.

Fig. 9: An excerpt of a *dramatic* sequence for the *Sales* dataset. This set of frames describes the conflict that arises from the relationships between Cost/Won Lead, Actual Cost, and Pipeline Revenue. See Figure 7 for the causal graph.



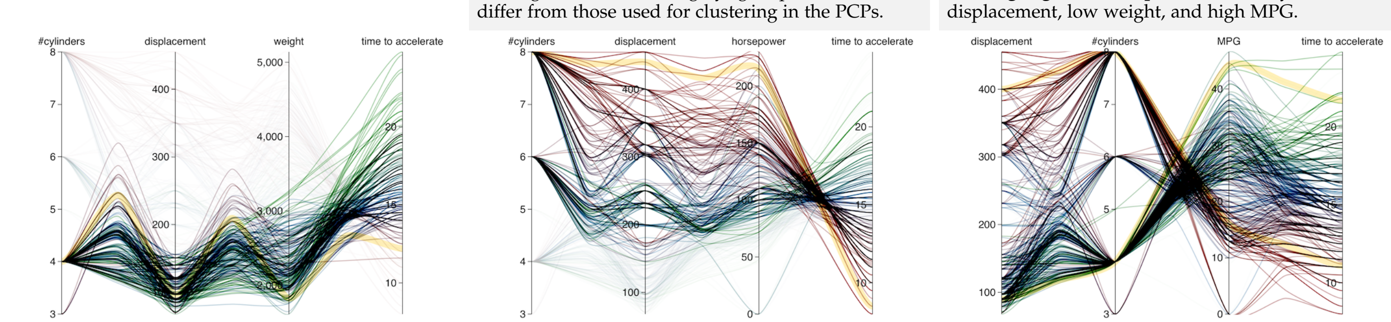
Intro. Data dimensions: MPG, #cylinders, displacement, horsepower, weight, time to accelerate and year

Goal 1: high MPG, **Goal 2:** low time to accelerate

Coloring indicates if samples meet both (purple), one (blue/green), or neither (grey) goal polarities. Colors differ from those used for clustering in the PCPs.

Optimize MPG by (chain 1): decreasing weight, decreasing displacement, and decreasing #cylinders.

The highlighted example has: low #cylinders, low displacement, low weight, and high MPG.



Optimize time to accelerate by (chain 2): decreasing weight, decreasing displacement, and decreasing #cylinders.

The highlighted example has: low #cylinders, low displacement, low weight, and low time to accelerate.

Optimize time to accelerate by (chain 3): increasing horsepower (HP), increasing displacement, and increasing #cylinders.

The highlighted example has: high #cylinders, high displacement, high HP, and low time to accelerate.

Trade-off: MPG and time to accelerate.

The two cars highlighted separately in previous panels are now shown together as incompatible choices. Yet, the considerable fluctuations along both goal variables suggest that there might be compromises.

Fig. 10: Selected panels from an automatically generated *journalistic* sequence for the *Cars* dataset. As in Figure 5, we only show the final views of visualizations that are developed over multiple panels.

browser. As described in Section 3, we used a previously implemented tool to compute the causal graphs.

The datasets in all previous and upcoming examples have 7–10 dimensions and up to 100s of data samples. They were primarily chosen because their domains are suitable for a general audience. The only input necessary from a storyteller were a few additional properties to complement a dataset such as data polarity. For example, in the *College* dataset [41], we required user input to specify that a desirable college should have a high ranking and low tuition.

Our approach does not depend on any specific data domain. To demonstrate this, we present two additional sequences of the *journalistic* and *detective* narratives. Based on the *College* dataset, Figure 11 shows a sequence that exhibits a trade-off between school rankings and affordability. From the *PM10* dataset [37], another sequence in Figure 12 investigates the cause of air pollution and describes that the concentration of particles in the air is lower in the morning (Hour is low) with fewer cars (low #Cars/hr) and when there is a breeze (Wind Speed is higher).

5 Validation

We conducted two studies to validate the usefulness of causal sequences and the tool to control them. The first study was crowd-sourced to measure whether our causal sequences could help viewers grasp the relationships embedded in multivariate data better than an ordinary sequence uninformed by causal relations. Besides quantitative

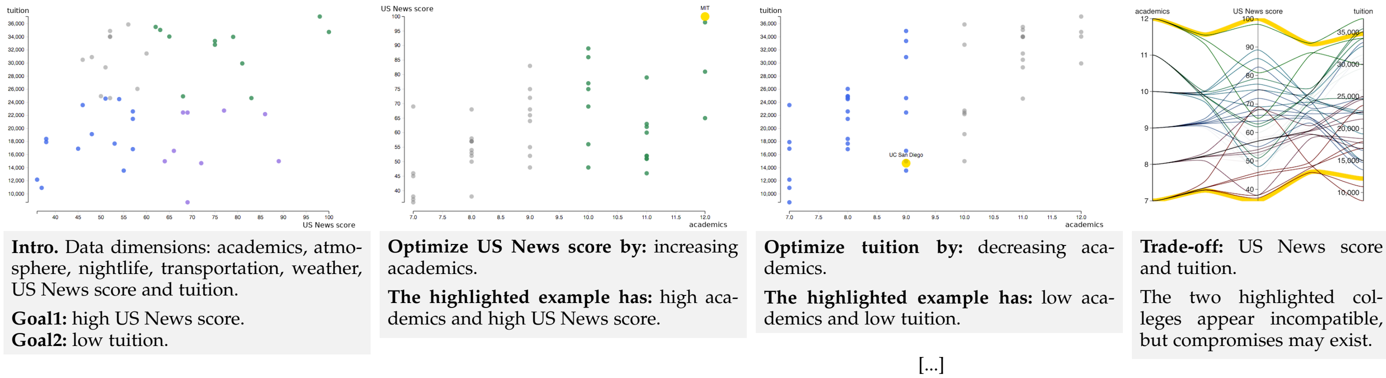


Fig. 11: An abridged version of an automatically generated *journalistic* sequence for the *College* dataset explaining the trade-off between college ranking and tuition. [...] indicates where the sequence has been shortened. As before, only the final views are shown when visualization and text are developed over the course of multiple panels.

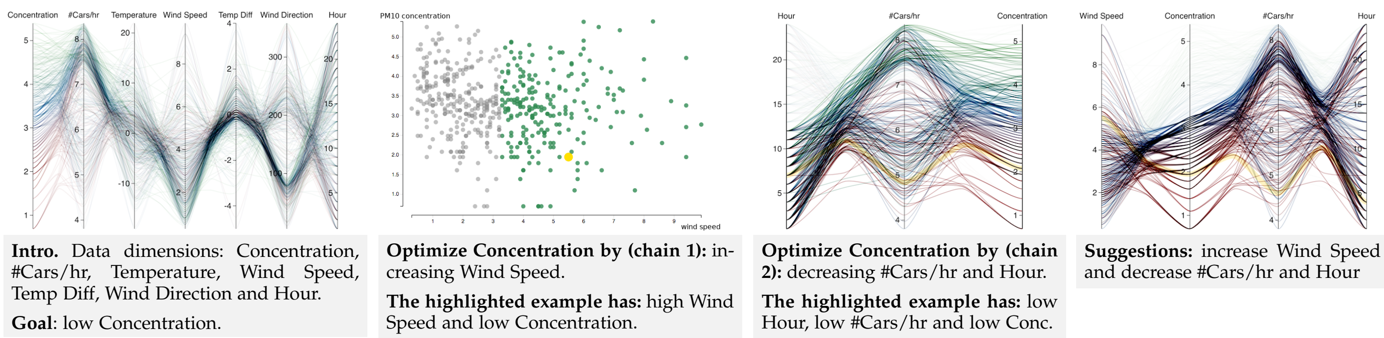


Fig. 12: An abridged version of an automatically generated *detective* sequence for the *PM10* dataset. as before, only the final views of multi-panel developments are shown.

measures, we interviewed potential users of such sequences to comment on causal sequences and our prototype. The second study was qualitative in nature. Both studies were granted an IRB exempt review as it presented minimal risks and collected no identifiable information of any participant.

5.1 Quantitative Validation

We designed an experiment to compare how different scatterplot and PCP sequences support causal data understanding. The sequence types were limited to *exhaustive* and *detective* narratives because they require no causal conflicts and exist in all datasets. 50 participants were recruited via Amazon Mechanical Turk and each was compensated $5 for participating in the study.

5.1.1 Experiment Design

The study was conducted as a web-based questionnaire consisting of an introduction, a tutorial on one of two visualization types (scatterplot or PCP), a practice task, and a series of four tasks featuring different chart sequences. For each task, participants were asked to write a brief narrative describing the relationships between dataset attributes that they could infer from the provided chart sequences.

Before beginning the main task, participants received a brief tutorial introducing the visualization format (scatterplot or PCP) and the concept of a sequence. The tutorial used a few simple examples—such as showing that as temperature increases, ice cream sales also increase—to illustrate how one variable can relate to another within a single view. No causal language was used. Participants then answered a multiple-choice test question based on a similar example to confirm their understanding of how to read the plots. This ensured consistency across conditions while minimizing interpretive bias. The full tutorial text and all example sequences are provided in the supplement.

Each task was based on a different dataset — *Cars*, *Sales*, *College*, and *PM10* — to minimize learning effects. The chart sequences varied across four types: *exhaustive*, *detective*, spurious, and random. *Exhaustive* and *detective* sequences corresponded to the narratives described in Section 3.2. In random sequences, variables were grouped into pairs at random. For spurious sequences, variables that were correlated were grouped together and presented in a random order. Each participant was shown 4 out of the 16 possible dataset-sequence combinations.

Responses from 9 out of 50 participants were excluded due to invalid submissions for at least one task.[1] For each visualization type, we collected approximately 5 valid responses per dataset-sequence combination. These responses were analyzed to determine whether different chart sequences influenced participants' ability to extract causal relationships between dataset attributes.

1. Some responses consisted of random or irrelevant text, such as single characters, random numbers, or a list of attribute names. Others provided explanations of dataset concepts, as generated with ChatGPT, without addressing the relationships shown in the sequences.

TABLE 1: The averages of quantitative metrics — precision (P), recall (R), and F1 score in percentage — of all sequences in scatterplot and PCP.

Sequence	Scatterplot			PCP		
	P	R	F1	P	R	F1
Exhaustive	82.38	39.51	53.41	70.81	44.11	54.36
Detective	74.05	59.79	66.16	73.75	90.74	81.37
Spurious	48.50	23.52	31.68	32.78	18.18	23.39
Random	15.04	4.48	6.90	15.88	7.26	9.96

5.1.2 Clause Analysis

We divided each response into clauses (relations) that implied causality. Two of the co-authors independently annotated these clauses, and any non-causal clauses were removed. After discussion and reconciliation, each clause was assigned a score based on its accuracy: 0 for clauses that identified an incorrect causal variable pair, 0.5 for clauses that identified a correct pair but with the wrong causal direction, and 1 for clauses that correctly identified both the variable pair and the causal direction.

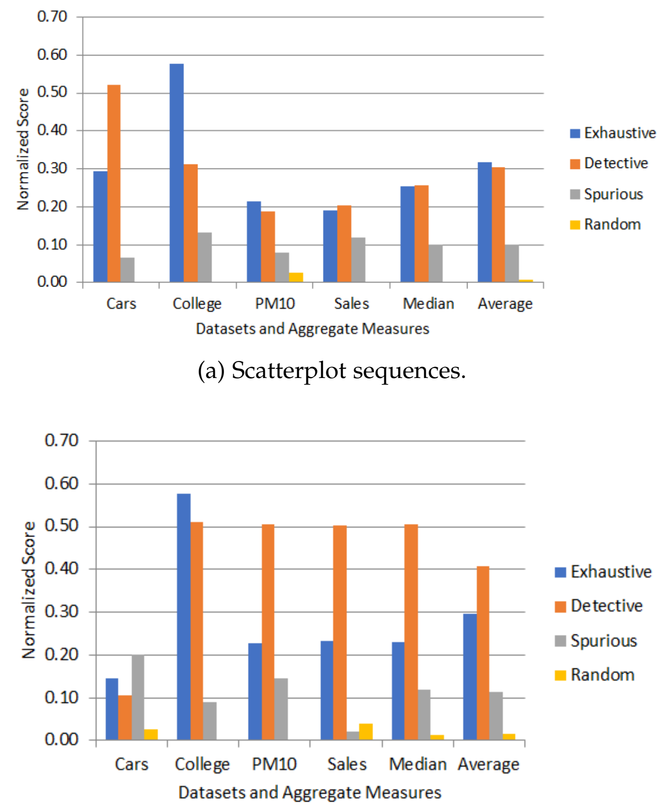
As the visualization sequences varied in length, we computed the total score for each response as well as two normalized metrics. The ratio of the total score to the number of clauses written was treated as the precision of the response, while the ratio of the total score to the number of variable pairs shown in the visualization was treated as the recall. Following common practice in classification tasks, we also calculated the F1 score as the harmonic mean of precision and recall.

According to the average F1 scores shown in Table 1, the *exhaustive* and *detective* sequences clearly outperformed the spurious and random sequences. For the scatterplot condition, a one-way Analysis of Variance (ANOVA) confirmed that there was a statistically significant difference among sequence types ($F(3, 76) = 24.85, p < 0.01$). A post-hoc Tukey's Honestly Significant Difference (HSD) test found that the average F1 score was significantly different in all sequence pairs ($p < 0.05$) except the pair of *exhaustive* and *detective* sequences ($p = 0.67$).

There was a statistically significant difference among sequence types shown in PCP as well ($F(3, 80) = 21.47, p < 0.01$). In a HSD test, the average F1 score was significantly different in all sequence pairs ($p < 0.01$) except the pair of causal sequences (*exhaustive* and *detective*, $p = 0.15$) and the pair of control sequences (spurious and random, $p = 0.38$).

5.1.3 Causal Structure Analysis

To better understand which sequence strategies might support a more comprehensive causal understanding, we analyzed the degree of connectivity among individual clauses. A node (attribute) in a causal graph can act as a mediator (with both incoming and outgoing edges), a collider (with only incoming edges), a confounder (with only outgoing edges), or a combination of these. We consider a user to have acquired a more comprehensive causal understanding if their narrative contains valid clauses that reference a node multiple times, regardless of its role as a mediator, collider, or confounder.



(a) Scatterplot sequences.

(b) PCP sequences.

Fig. 13: Average normalized causal structure scores for all (a) scatterplot and (b) PCP sequences and the 4 datasets. Each score averages 5 participants. A score of 1.0 would denote that for each node with more than one in- or outgoing edge all edges were identified. The last two sets of bars are the Median and Average scores, respectively, across all datasets.

For this analysis, we counted how often each node was referenced by a user. Because the number of edges per node varied, we normalized the counts by dividing by the maximum possible mentions for each node. We visualize both the total and normalized mention counts as heatmaps for each dataset, included in the supplementary material.

Figure 13 provides a summary of the observed patterns. Overall, we observe a clear trend suggesting an advantage of both the *detective* and *exhaustive* narratives over the *random* and *spurious* ones, although the distinction between *detective* and *exhaustive* is more subtle – an ANOVA test failed to establish a statistically significant difference. For PCP narratives, *detective* appears to outperform *exhaustive* in both median and average scores, showing convincingly better results for the PM10 and Sales datasets, with a near tie for Cars and College. For scatterplot narratives, *exhaustive* appears to perform better in the College dataset, whereas *detective* shows an advantage in the Cars dataset, potentially influenced by the larger number of plots forming the sequence (12 vs. 8, see supplement). The differences across the other datasets are relatively minor. Taken together, these exploratory results suggest that if the goal is to identify causal structures, *detective* may be the safer choice for PCP, while for scatterplots, the choice between *detective* and *exhaustive* may depend more on user preference.

5.2 Qualitative Validation

To complement our controlled study, we conducted an exploratory qualitative evaluation with five domain experts: four data storytellers (S1–S4) and one journalism professor (P1). All had experience in descriptive data analysis and crafting narratives to communicate complex insights—particularly on social issues—in a way that is accessible and engaging to general audiences. The goal was to understand how users with communication expertise engage with different narrative styles generated by our system and gather early insights into their interpretability, storytelling potential, and support for multivariate reasoning.

Each 20–30 minute session included a walkthrough of narrative sequences—exhaustive, detective, dramatic, and journalistic—based on the College, Cars, and Sales datasets. Participants were asked to reflect on the clarity, structure, and applicability of these sequences, and encouraged to give open-ended feedback on usability and alignment with professional storytelling practices.

Participants generally found the sequences understandable and aligned with their communication goals. They appreciated the use of example data points beneath each view, which helped ground the narratives and mirrors storytelling approaches they use in practice. Notably, while four of the five participants had not encountered parallel coordinate plots before, they were able to make sense of the visual encodings within the context of the narrative.

Participants differentiated between narrative types and found some more compelling than others. The journalistic sequence for the College dataset (Figure 11) drew particular attention. P1 remarked that it resembled the structure of an investigative story, presenting trade-offs and directional relationships in a way that mirrored how real-world reporting builds tension and supports argumentation. S4 commented that the sequence "raises the kinds of questions a journalist or policy analyst would want to explore further." This suggests that the journalistic narrative, in particular, may naturally align with professional storytelling conventions.

Participants also offered constructive feedback for future extensions. P1 and S4 emphasized the need for more traditional journalistic patterns such as grouping by geography or time, surfacing extremes (e.g., top 10 or bottom 10), and highlighting contrast. S1 and S3 requested support for additional chart types, while S2 proposed the integration of multiple datasets within a single narrative. Several participants suggested the ability to inject prior domain knowledge into the causal sequence (e.g., asserting a known link), or to adjust the narrative length and complexity based on audience needs.

Overall, these sessions suggest that causal narrative styles generated by our system are interpretable and resonate with domain communicators. They offer structure and entry points that align with familiar storytelling patterns and participants were able to distinguish among narrative types in terms of their communicative potential. While exploratory, this feedback provides promising evidence that the system supports meaningful engagement with complex multivariate data and lays the foundation for more systematic comparative studies of narrative style in future work.

6 Discussion and Limitations

Our quantitative evaluation showed that causal sequences improve understanding compared to random and correlation-based (possibly spurious) sequences, underscoring the importance of directionality in distinguishing cause from effect. In the scatterplot condition, the exhaustive sequence yielded higher average precision but lower recall than the detective sequence. In contrast, for the PCP condition, the detective sequence outperformed the exhaustive one across all metrics—most notably in recall, likely due to PCPs' ability to show entire causal chains in a single view. This suggests that PCPs may be inherently better suited than scatterplots for visualizing causal structures. Finally, the use of edge bundling may have influenced participants' perception; a cleaner rendering style [42], [43] could help clarify data relationships in PCPs.

To better understand the depth of causal reasoning, we also analyzed participants' narratives at the level of causal structure. By measuring how often participants referenced nodes with multiple causal connections, we observed that detective and exhaustive sequences supported more comprehensive causal understanding than random and spurious ones. While these trends were consistent across datasets, the small number of participants per dataset-strategy combination (approximately five) means that the results should be viewed as exploratory.

Nonetheless, the convergence of results across both clause-level and structure-level analyses increases our confidence in the general patterns observed. In future work, we plan to replicate and extend these findings with larger and more diverse participant samples to further validate the effectiveness of different sequencing strategies.

Note that our quantitative results relied on free-form text answers and their causal words and structures. For instance, "A causes B" could be naturally phrased as "B increases as A increases" or "the more A is, the more B becomes." Some clauses were debatable, but the co-authors, as two independent coders, tried our best to give each clause the benefit of the doubt. Furthermore, although individual causal pair recall may not fully represent an understanding of the entire network, theories in Mental Models [44] and Constructivist Learning [45] suggest that knowing individual relationships aids in building a comprehensive mental model of a system. We plan to implement more direct methods like interviews and scenario analyses in future research to more effectively gauge overall network comprehension.

The qualitative validation confirmed the usefulness of causal sequences—a feature unavailable in current tools—according to the group of selected professionals. As suggested by the participants, this system could be a tool in a journalistic writing process. Journalists would be in the loop as some computer-generated recommendations might seem unconventional. For example, in Figure 10, the system suggests higher wind speed and early hours to reduce PM10 concentration. Journalists can interpret the suggestions as possible real-life solutions including removing wind barriers or reducing highway fees in the early morning.

In our study we also experimented with the structure of exhaustive sequences, particularly whether to include full flashbacks to the root node in every path. For shorter chains, we opted to skip the root-to-intermediate node recap to reduce cognitive load and avoid redundancy, assuming viewers could retain key steps in working memory. However, in longer or more branching sequences—such as in the Sales dataset—we included full flashbacks to preserve clarity (see Figure 2 middle row for the two options; the supplement provides a visualization of these individual chains for all four datasets). Interestingly, the more verbose Sales sequence received slightly lower comprehension scores, suggesting that shorter, more memory-efficient sequences may be equally or more effective. This tradeoff warrants further investigation in future work.

Our current implementation has a few technical limitations. First, scalability has not yet been tested beyond small datasets—those used in our examples have fewer than a dozen dimensions and a few hundred data points. In practice, standard feature selection or dimensionality reduction techniques can be applied to keep larger datasets manageable without significantly reducing interpretive value. Second, the system is not yet integrated with a live causal inference engine; all causal relationships are precomputed. While on-the-fly computation is technically feasible, it would introduce execution delays that may affect interactivity.

7 Conclusion

We presented a framework that generates coherent visualization sequences by traversing a causal graph derived from a dataset with minimal metadata. Our evaluation shows that these sequences outperform correlation-based and random alternatives in supporting causal understanding.

In future work we would like to extend the approach to additional traversal strategies, visualization types, and data modalities such as time-varying and categorical variables, as suggested by participants. We also plan to incorporate natural language descriptions, generated by large language models, to complement the visuals and enhance comprehension of causal relationships.

Finally, while our studies confirm the effectiveness of causal narrative sequences, they did not explore why certain styles resonate more or how preferences vary by user expertise or context. Future work will investigate these factors to better tailor narrative designs to different audiences.

Acknowledgements

Partial funding was provided by NSF grant IIS 1527112.

References

[1] A. Inselberg and B. Dimsdale, "Parallel Coordinates: a Tool for Visualizing Multi-Dimensional Geometry," in *IEEE Visualization*, 1990, pp. 361–378.

[2] A. Mead, "Review of the Development of Multidimensional Scaling Methods," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 41, no. 1, pp. 27–39, 1992.

[3] L. Van der Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. 11, 2008.

[4] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv preprint arXiv:1802.03426*, 2018.

[5] T. Munzner, "Visualization analysis and design," in *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Courses*, 2025, pp. 1–2.

[6] J. Bertin, *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press, 1983.

[7] J. Hartigan, "Printer Graphics for Clustering," *Journal of Statistical Computation and Simulation*, vol. 4, no. 3, pp. 187–213, 1975.

[8] ArccGISPro, "Scatter Plot Matrix," https://pro.arcgis.com/en/pro-app/3.0/help/analysis/geoprocessing/charts/scatter-plot-matrix.htm, [Online; accessed 19-July-2023].

[9] M. Ankerst, S. Berchtold, and D. Keim, "Similarity Clustering of Dimensions for an Enhanced Visualization of Multidimensional Data," in *IEEE Information Visualization*. IEEE, 1998, pp. 52–60.

[10] G. Albuquerque, M. Eisemann, D. Lehmann, H. Theisel, and M. A. Magnor, "Quality-based Visualization Matrices," in *VMV*, 2009, pp. 341–350.

[11] M. Sips, B. Neubert, J. Lewis, and P. Hanrahan, "Selecting Good Views of High-Dimensional Data using Class Consistency," in *Computer Graphics Forum*, vol. 28, no. 3, 2009, pp. 831–838.

[12] L. Wilkinson, A. Anand, and R. Grossman, "Graph-Theoretic Scagnostics," in *IEEE Info Vis*, 2005, pp. 21–21.

[13] T. Dang and L. Wilkinson, "Scagexplorer: Exploring Scatterplots by their Scagnostics," in *IEEE Pacific Vis*, 2014, pp. 73–80.

[14] N. Elmqvist, P. Dragicevic, and J.-D. Fekete, "Rolling the Dice: Multidimensional Visual Exploration Using Scatterplot Matrix Navigation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1539–1148, 2008.

[15] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1139–48, 2010.

[16] J. Gower, S. Lubbe, and N. Le Roux, *Understanding Biplots*. John Wiley & Sons, 2011.

[17] B. Wang and K. Mueller, "The Subspace Voyager: Exploring High-Dimensional Data Along a Continuum of Salient 3D Subspaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 2, pp. 1204–1222, 2017.

[18] D. Asimov, "The grand tour: a tool for viewing multidimensional data," *SIAM Journal on Scientific and Statistical Computing*, vol. 6, no. 1, pp. 128–143, 1985.

[19] J. Friedman and J. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Transactions on Computers*, vol. 100, no. 9, pp. 881–890, 1974.

[20] A. Tatu *et al.*, "Combining Automated Analysis and Visualization Techniques for Effective Exploration of High-Dimensional Data," in *IEEE VAST*, 2009, pp. 59–66.

[21] S. Johansson and J. Johansson, "Interactive Dimensionality Reduction through User-Defined Combinations of Quality Metrics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 993–1000, 2009.

[22] A. Artero, M. de Oliveira, and H. Levkowitz, "Enhanced High Dimensional Data Visualization through Dimension Reduction and Attribute Arrangement," in *IEEE InfoVis*, 2006, pp. 707–712.

[23] Z. Zhang, K. T. McDonnell, and K. Mueller, "A Network-based Interface for the Exploration of High-Dimensional Data Spaces," in *IEEE Pacific Vis*, 2012, pp. 17–24.

[24] Z. Zhang, K. T. McDonnell, E. Zadok, and K. Mueller, "Visual Correlation Analysis of Numerical and Categorical Data on the Correlation Map," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 2, pp. 289–303, 2015.

[25] A. Dasgupta and R. Kosara, "Pargnostics: Screen-Space Metrics for Parallel Coordinates," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1017–1026, 2010.

[26] W. Peng, M. Ward, and E. Rundensteiner, "Clutter Reduction in Multi-Dimensional Data Visualization using Dimension Reordering," in *IEEE InfoVis*, 2004, pp. 89–96.

[27] B. Ferdosi and J. Roerdink, "Visualizing High-Dimensional Structures by Dimension Ordering and Filtering using Subspace Analysis," in *Computer Graphics Forum*, vol. 30, no. 3, 2011, pp. 1121–1130.

[28] J. Yang, W. Peng, M. Ward, and E. Rundensteiner, "Interactive Hierarchical Dimension Ordering, Spacing and Filtering for Exploration of High Dimensional Datasets," in *IEEE InfoVis*, 2003, pp. 105–112.

[29] M. Blumenschein, X. Zhang, D. Pomerenke, D. Keim, and J. Fuchs, "Evaluating Reordering Strategies for Cluster Identification in Parallel Coordinates," in *Computer Graphics Forum*, vol. 39, no. 3, 2020, pp. 537–549.

[30] A. Tyagi, T. Estro, G. Kuenning, E. Zadok, and K. Mueller, "PC-Expo: A Metrics-Based Interactive Axes Reordering Method for Parallel Coordinate Displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 1, pp. 712–722, 2023.

[31] C. Tong, R. Roberts *et al.*, "Storytelling and Visualization: An Extended Survey," *Information*, vol. 9, no. 3, p. 65, 2018.
[32] J. Pearl and Judea, *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2000.
[33] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*. MIT Press, 2001.
[34] J. Wang and K. Mueller, "The Visual Causality Analyst: An Interactive Interface for Causal Reasoning." *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 230–9, 2016.
[35] ——, "Visual causality analysis made practical," in *2017 IEEE Conference on visual analytics science and technology (VAST)*. IEEE, 2017, pp. 151–161.
[36] J. Pearl and D. Mackenzie, *The Book of Why: the New Science of Cause and Effect*. Basic books, 2018.
[37] "StatLib: Datasets Archive," http://lib.stat.cmu.edu/datasets/.
[38] D. Herman, M. Jahn, and M.-L. Ryan, *Routledge Encyclopedia of Narrative Theory*. Routledge, 2010.
[39] G. Freytag, *Technique of the Drama: an Exposition of Dramatic Composition and Art*. S. C. Griggs, 1894.
[40] C. Cotter, *News Talk: Investigating the Language of Journalism*. Cambridge University Press, 2010.
[41] J. Nam and K. Mueller, "TripAdvisorN-D: A Tourism-Inspired High-Dimensional Space Exploration Framework with Overview and Detail." *IEEE Transactions on Visualization and Computer Graphics*, 2012.
[42] G. Palmas, M. Bachynskyi, A. Oulasvirta, H. Seidel, and T. Weinkauf, "An Edge-Bundling Layout for Interactive Parallel Coordinates," in *2014 IEEE Pacific Vis*, 2014, pp. 57–64.
[43] K. McDonnell and K. Mueller, "Illustrative Parallel Coordinates," *Computer Graphics Forum*, vol. 27, no. 3, pp. 1031–1038, 2008.
[44] P. Johnson-Laird, *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press, 1983.
[45] L. Vygotsky and M. Cole, *Mind in society: Development of higher psychological processes*. Harvard University Press, 1978.

Puripant Ruchikachorn received his PhD from Stony Brook University, New York, as a recipient of the International Fulbright Science and Technology Award. He now teaches at Chulalongkorn University, Bangkok, Thailand. His research interests include information visualization, data analytics, and human-computer interaction.

Darius Coelho holds a PhD in Computer Science from Stony Brook University. His research interests include information visualization, visual analytics, and human-computer interaction. He has received two best paper awards for his work and currently works as software architect at Akai Kaeru LLC.

Jun Wang earned a PhD in computer science from Stony Brook University and a BS and MS in computer science from Shandong University, China. He is now employed at Google in a R&D position. Jun's research interests are visual analytics, information visualization, and data mining. For more information, see https://junwang23.github.io/.

Kristina Striegnitz received a joint PhD from Saarland University in Saarbrücken, Germany, and Henri Poincaré University in Nancy, France. She is currently an Associate Professor of computer science at Union College, NY. Her research interests are natural language generation and dialog systems. For more information see http://cs.union.edu/~striegnk.

Klaus Mueller received a PhD in computer science from The Ohio State University. He is currently a professor of computer science at Stony Brook University. His research interests include visual analytics, explainable AI, data science, and medical imaging. For more information see http://www.cs.stonybrook.edu/~mueller.