

# HumMod Browser: An Exploratory Visualization Tool for the Analysis of Whole-Body Physiology Simulation Data

Keqin Wu

Jian Chen\*

University of Maryland, Baltimore County

William A. Pruett

Robert L. Hester

University of Mississippi Medical Center

## ABSTRACT

We present HumMod Browser, a multi-scale exploratory visualization tool that allows physiologists to explore human physiology simulation data with more than 6000 attributes. We first present a tag cloud technique to reveal the significance of time-varying attributes and then study how a chain of tag clouds can form an exploratory visualization that assist multiple dataset comparison and query. One purpose is to reduce the high cognitive workload of understanding complex interactions within the large attribute space. The HumMod Browser produced can give physiologists flexible control over the visualization displayed for quick understanding of complicated simulation results. The visualization is constructed through the metaphorical bubble interface to allow dynamic view controls and the data relationships and context information unfold as physiologists querying groups of connected bubbles within the hierarchical or causal relationships. HumMod Browser contributions to the interaction design and provides multi-scale coordinated interactive exploration for a new type of physiological modeling data. Two case studies have been reported with real datasets containing more than 6000 physiology attributes, which provide supportive evidence on the usefulness of HumMod Browser in supporting effective large-attribute-space exploration.

**Keywords:** — Information visualization, design study, interaction, physiology.

**Index Terms:** H.5.m [Information Systems]: Information Interfaces and Presentation—Miscellaneous; I.3.8 [Computing Methodologies]: Computer Graphics—Applications

## 1 INTRODUCTION

Recent success in real-time modeling and simulation of human physiology has created great challenges for data analysis enabling scientists to understand the intricate functioning of organisms as governed by the laws of physics and chemistry. HumMod [1], possibly today's largest and most comprehensive model of physiological functions, models interactions among the cardiovascular, respiratory, renal, neural, endocrine, skeletal muscle, and metabolic systems. An important use of the model is to understand physiological mechanisms and interactions that are not obvious, allowing one to observe higher-level emergent properties of the complex physiological systems. However, the huge number of attributes prohibits developing a clear and comprehensive view of the attribute interactions. Currently, integrative modeling in HumMod is accomplished by assimilating over 40 years' worth of published physiological models. Good

visualization and interaction techniques are needed to discern non-obvious inter-attribute relationships.

Our HumMod Browser (Figure 1) is a new interactive visualization tool that supports analyzing a large number of attributes with complicated interactions. To help physiologists explore the intricate relationships among attributes, two problems need to be solved: (1) The model has a large attribute space. This poses a cognitive challenge for physiologists to explore their relationships due to the limits of human memory, a situation intensified by the second problem: (2) The model output differs from run to run. It is thus nontrivial for physiologists to perform comparisons manually without the help of a dynamic multi-view environment that supports extracting the most relevant information for different simulation outputs at the same time.

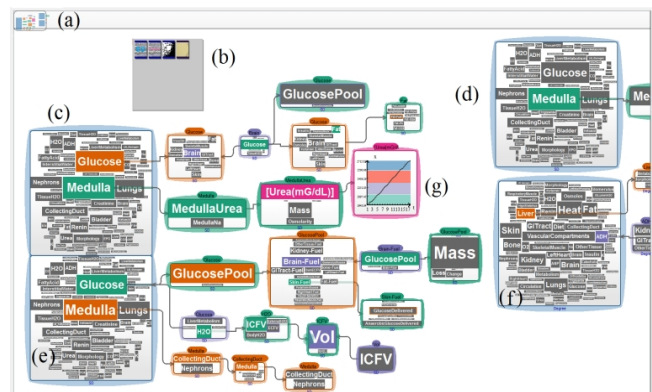


Figure 1: HumMod Browser Interface: an example use letting physiologists query the datasets interactively and rearrange and compare physiological attributes. (a) An overview shown in the panning bar. (b) Control panel. (c) Two grouped bubble sets for comparison. (d) Two individual bubble sets. (e) Tag cloud by standard deviation. (f) Tag cloud by degree. (g) Time series plot.

To address these problems, HumMod Browser is designed as an exploratory visualization tool to let physiologists combine different pieces of information through interactive visualizations in a succinct way. The meteorologists, Dr. Hester and Dr. Pruett, who are also co-authors, provided expert advice and feedback during the development of HumMod Browser. Our tool addresses the first design challenge – the limited human memory for exploring the large attribute space – by providing visible path to record physiologists' data exploration operations and flexible controls over its pace. It also addresses the second challenge – the large workload of performing comparison tasks – by enabling easy access to previous query paths and effective interactions with these paths. This paper contributions to the interaction design, provides multi-scale coordinated interactive exploration for a new type of physiological modeling data, and explores an alternative to the classical graph visualization to assist physiologists in learning causality relationships in mathematical modeling data. This paper has made significant extension to our EuroVis short paper submission [2] added with a concrete task analysis, rich

\* Corresponding author, email: jichen@umbc.edu

interaction designs, an informative case study, and insights from our physiology collaborators.

## 2 DATA CHARACTERIZATION

The many physical- and temporal-scale model attributes in HumMod form a dense jungle of relationships that are of two types: hierarchical and causal.

### 2.1 Attribute Interactions: Hierarchical Relationships

Every attribute in HumMod has an assigned “global” name. For example, *Co2Total* includes the attributes or global names *CO2.CO2Total.Inflow* and *CO2.CO2Total.Outflow*. Each global name has three distinct components, *CO2*, *CO2Total*, and *Inflow*, and *CO2*, *CO2Total*, and *Outflow*. This represents the fact that within the folder describing *CO2* and the specific file describing *CO2Total* are local elements corresponding to both *Inflow* and *Outflow*.

For simplicity we refer to the components of the name of a model attribute as the level-*x* name, where *x* refers to the depth. Therefore, for the attribute *CO2.CO2Total.Inflow*, the level-0 name is *CO2*, the level-1 name is *CO2Total*, and the level-2 name is *Inflow*. In HumMod, there is no inherent restriction on the possible depth of global names. However, currently most global names are composed of a level-0, a level-1, and a level-2 name. Figure 2a illustrates the hierarchical relationships among the local names of these attributes.

### 2.2 Attribute Interactions: Causal Relationships

In addition to containing local names, the value of an attribute may have a dependence on the value of other attributes. This means that a change in the latter can cause the former to change. For example, there is a dependence of the value of *CO2.CO2Total.Inflow* on the value of attributes corresponding to the outflow of *CO2* from other body organs, as well as the dependence of the value of *CO2.CO2Total.Outflow* on the values of *Lung.LungCO2.Expired* and *CO2.CO2Tools.LitersToMols*. These relationships constitute the functional relationships among different global attributes. When the “direction” of dependence is taken into account, the global names and interlaced causal relationships form the directed graph shown in Figure 2.

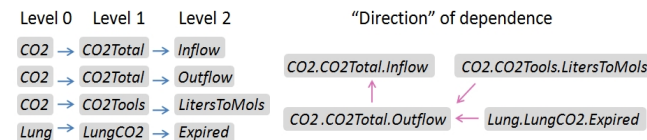


Figure 2: Attribute interactions: hierarchical relationships (left) and causal relationships (right).

HumMod generates time varying values for its attributes which forms interlaced feedback loops. Similar to meteorology, causality cannot be computed analytically but through ensemble simulations. As the attribute space grows, there is no robust computational approach in classical control theory to measure causality.

## 3 RELATED WORK

Several effective tools provide visualization for relationships. Cytoscape [3] and VisANT [4] are mainly designed for biomolecular network visualization. MizBee [5] and Circos [6] use circular layouts and are created for comparative genomics. Classical algorithms such as tree ring [7] intuitively encode both hierarchical and network relationships as layers and edges but use large display spaces to avoid clutter when datasets become large. More space-saving methods such as tree maps [8] express

hierarchical relationship in static images and thus do not provide enough support for physiologists’ dynamic data exploration need.

Causal relationship visualization is a relatively new area, and most causality visualizations use some form of graph layout [9], but scaling the graph display to a large network remains problematic. Some techniques demand a large display space, such as tree-rings [9] [10]; others avoid this by removing edges (e.g., adjacency matrix [11] and its extensions of Matrix Zoom [12] and [13]) or by dynamic data exploration [14]. Empirical studies suggest that an animated graph is better than static graphs for enhancing users’ speed and accuracy in recalling causal relationships [15]. Our HumMod design follows the design implications from these studies to focus on interactive data exploration.

Our design faces workflow issues such as high repetition and limited human memory because the visual knowledge discovery process involves frequent reading, searching, and comparison of data in various hierarchies or locations of the graph. The workflow tool VisTrails [16] developed a pipelined approach to automating repetitive visualization tasks by providing detailed history information. Several narrative visualization and visual storytelling methods [17], [18], [19], [20], [21] have inspired our work. Visual storytelling has the following properties: the pace at which a story unfolds, the context information provided to users to understand their data, information access through interaction, and the specific goals of story-miners such as data exploration and sharing findings with others [17]. VisBubbles addresses a unified environment supporting programming, visualization, and interaction for bat flight analysis that gives users flexible control over the context of the bat flight data [20]. Ferreira et al. present BirdVis, a visualization tool for understanding bird populations [21], which allows visualizations to be communicated through interactive multi-views of visualizations as users explore through space and time and across different bird species. All these methods support effective knowledge learning and information discovery. Our method builds on this trend but addresses the memory and workload issues while lessening visual clutter by careful design of the interface for user to control both context of the visualization and pace of their exploration process.

## 4 HUMMOD BROWSER

We first elicit the main activities and tasks of physiologists and the challenges they face when exploring HumMod data; then, we discuss the visualization techniques and interaction designs employed to achieve effective data exploration with HumMod Browser.

### 4.1 Design Analysis

Our collaboration with physiologists has revealed three main data exploration activities with HumMod simulation data. Our previous work [2] focused on high-level task characterization in model validation. In the paper, we expand the task space based on physiologist data exploration activities.

**T1: Relationship exploration** — Physiologists need to examine the relationship among attributes to learn how these attributes interact with each other. For example, different age- and sex-matched individuals can have different bifurcating responses to 750 ml haemorrhage: some have circulatory collapse while others exhibit no physiological symptoms of the blood loss. Careful examination of why these differences occur is studied in ensemble simulations. Example questions to answer through relationship exploration include: Do a set of attributes change concurrently? Which attributes cause other attributes to change or vice versa?

**T2: Importance analysis** — Attribute importance involves measurement metrics, such as statistical significance. During the

relationship exploration above, physiologists need to carry out analyses of time varying values of selected attributes. A question to ask is: is the output time series of an attribute linear, non-linear, or cyclic?

**T3: Comparison** — It is important to search correlations and discern differences between model attributes, in particular, comparing between different ensemble runs. For instance, how similar are two simulation outputs? Do their attributes show similar trends? Or, how different are two simulation outputs? Are there any attributes that vary significantly between different ensemble runs?

To support physiologists' data exploration activities, our design needs to meet following goals.

**G1: Memory enhancement** — the limited human memory for exploring the large attribute space is one major challenge faced by physiologists. The path of their exploration should be retrievable. In addition, showing dynamic visualization views instead of static enhances users' memory.

**G2: Replicability** — the large workload of performing comparison tasks is another main challenge as physiologists often need to identify and compare various ensemble runs with simultaneous or time-dependent variations. Those activities are highly repetitive. Accordingly, efficient interaction is needed to reduce workload by replicating past query and synchronizing the concurrent queries of different ensembles.

**G3: Importance indicator** — Physiologists need to reason through their analysis with a tremendous amount of data. Visualization with intuitive visual cues indicating important attributes is needed to facilitate this process.

**G4: Scalability** — The scalability issue in context of HumMod data analysis is primarily concerned with the intricate hierarchical and causal relationships in the large attribute space. There is a need to reduce the complexity of the visualization and solve the conflict between the size of visualization, screen space constraint, and limited human memory.

To address these requirements, we develop a dynamic multi-view visualization which support an uncluttered interactive exploration based on tag-cloud visualization and a set of intuitive interactions in a metaphorical bubbles interface [20], [22] with some significant extensions. Our designs support user's information query in four ways, among which (1) and (4) were migrated to our HumMod Browser whilst (2) and (3) are new, compare to those in [20] and [22]:

(1) *Dynamic view control.* Partly inspired by the pace control of visual storytelling approaches [19], [17], we show a piece of information with similar level of detail in each bubble (Figure 1 c and d) and allow viewers to decide when to go to the next level by generating a new bubble. The new bubble reveals further information about the current bubble. This gives users a better control of the progression of their exploration operations.

(2) *Contextual information on demand.* The fragment of contextual information is encapsulated in each bubble, revealing more detail of an attribute as the user's query process proceeds. Physiologists primarily wish to know which attributes change together. We thus consider attributes with higher standard deviation or more connections to other attributes to be more significant than others. To allow attributes with higher significances to stand out for an efficient information discovery, we use a tag-cloud technique to indicate the significances of different attributes (Figure 1 e and f). Users are also interested in exploring the time series of the value of certain attributes, e. g. they may want to see how two attributes with a causal relationship change overtime. Therefore, we provide time series representations [23] (Figure 1g).

(3) *Rich interactions.* We adopt the rich interactions in bubbles interface [20], [22] such as grouping, dragging, context editing of bubbles, copy and paste, to support efficient information access. Conventional interactions like zoom in, zoom out, and dragging bubbles to different locations are also provided. In addition, HumMod Browser lets users apply a query instance to different attribute or datasets, query the visualization history, and comparatively visualize different results.

(4) *Pannable and scalable workspace.* To further support scalability of the visualization, we adopt a panning bar (Figure 1a) which shows an overview map of the entire workspace, and location indicators showing the sizes and locations of bubbles [22]. Users can pan and zoom in or zoom out the screen by panning and moving the mouse wheel down or up. This provides them a continuous virtual space where users do not need to toggle among the virtual desktops in the classical discrete space [24]. This, along with (1), (2), and (3), allow users to scale up or down the complexity and size of the visualizations. Tag Clouds with Random Space-Filling

We develop tag clouds to support importance analysis task (T1). One way to facilitate physiologists' exploration of the large attribute space is to provide visual cues to differentiate the significant attributes from less significant ones. Tag clouds provide a schematic overview about relative magnitude of a large number of tags. Other visualizations such as heat-map [17] and line plots [21] are less capable of conveying a large amount of textual information in limited space. The importance of each attribute is measured by the standard deviation of its time-varying value or the degree, indegree, or outdegree of the attribute in the causal relationship graph. The higher standard deviation an attribute has, the higher variability it may have. The higher degree, indegree, or outdegree an attribute has in the graph, the greater influence it may have on other attributes. We adopt a tag cloud in which each attribute name is scaled according to its significance [25], [26] since size (length, area, volume) is considered to be the visual property best suited, next to positions, for quantitative purposes among Bertin's eight visual attributes [11].

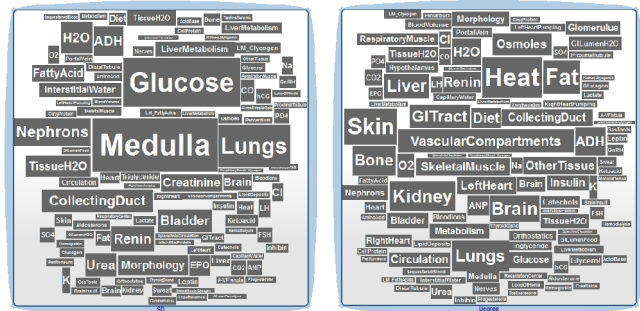


Figure 3: Tag clouds for posing spatial queries. Tag clouds encode the average standard deviation (left) or the degree (right) of attributes with level 0 names. New bubbles can be brought out as users query the attributes that interest them.

A tag cloud is an effective way to plot large numbers of attributes while representing their relative importance [27]. Mundad, et al. [28] and Ferreira et al. [21] suggest sorting the tags in alphabetical order. However, to make the tag clouds more spatially efficient and to prevent mistakenly linking importance to the alphabetical order of the attribute names, we adopt a random space-filling method [29] to randomly locate the tags. We utilize

a Poisson random-number generator and place tags in an order such that the larger tags have a greater chance of being placed near the center of the bubble while the smaller tags are more likely to be placed near the boundary, since the inner area may already be occupied by the larger ones placed earlier. To increase both the legibility of a tag and the accuracy of size reading, we scale a gray rectangle proportionally to a tag's importance and show it on the tag's background. As shown in Figure 3, tag clouds are effective in helping viewers quickly perceive the relative importance of attributes in a bubble.

The tag cloud is implemented by adapting space-filling methods with random fractal non-overlapping simple shapes [29]. Figure 4 shows the pseudo code used to generate tag cloud. The local names of each attribute are the input tags. The Poisson random-number generator is that in the Standard (Std) C++ Library. We set the mean distribution to be the plane center.

```

Assign the sizes of n tags according to their importance;
Sort tags with decreasing sizes;
Initiate Poisson random number generator;
for (i=0; i<n; i++) {
    do {
        Choose a random position for tag i in the region of the plane;
        for (j=0; j<i; j++) {
            if (the new tag does not intersect with tag j )
                place the tag at this position;
        }
    } while(the tag is not placed);
}

```

Figure 4: Pseudo code for generating tag clouds

## 4.2 Interaction Designs

The interface provides various user interactions to support physiologists' interactive data exploration. Users can show or hide a bubble and edit the visual encoding inside it. Users can also drag a bubble to any location. Dragging one bubble into another bubble groups the two bubbles to show the same exploration operations. The exploration operations of an attribute recorded though the bubble chain connected with it can be applied to another attribute by a simple copy-and-paste operation.

### 4.2.1 Editable Bubble Context

The context of a bubble can be (1) a tab cloud showing all the level-0 names (Figure 3), (2) all the level-n ( $n=0, 1, 2$ ) names sharing the same hierarchical or causal relationship with another attribute that is labelled on the top of the bubble, and (3) a time-series plot of an attribute whose highest level (level-2) name appears on the top of the bubble.

Bubbles are allowed to change the visual encoding inside it through menu operations. Currently, a bubble that contains tag clouds may switch its importance measurements within standard deviations and its degree in the causal relationship graph to encode the tag size. A bubble showing time series may switch from the traditional time series plot to a quantile representation [23]. The examples of showing the tag clouds of a same set of attributes with different importance measurements or showing different time series plots for a same attribute are illustrated in Figure 3 and 5. By switching and comparing between the different visualizations of the same data, users obtain a more comprehensive view of it.

Time series plots can be shown in bubbles as users query an attribute. Currently, each attribute contains a time series with 18 steps. We employ two types of time-series plots [23]: the naive time-series plot and quantile representation.

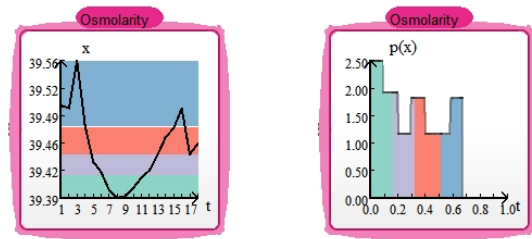


Figure 5: A time series (left) is split into four quantiles (colored shading) (right).

### 4.2.2 Dynamic Bubble Chains

The dynamic bubble chains are developed to support relationship exploration task (T2). A user's exploration operations are dynamic and are enabled by bringing out a new bubble through a typical menu operation involving one right-click and one left-click on the corresponding menu item on a pie menu. Edges are connected between the bubble with the queried attribute name and the newly generated bubble. The bubbles containing level-0 attribute names are the first that users drag out from the control panel for a selected ensemble dataset. Users can always bring up a new bubble by clicking on one attribute name inside the existing bubble. Once a new bubble is generated, its corresponding tag is highlighted with the color of the boundary of the new bubble. An edge is used to connect the selected tag in the current and the newly generated bubbles. The bubble clicked is considered the parent bubble of the newly created bubble. If a new bubble contains attribute names dependent on its parent bubble, the edge points to the parent bubble. In other cases, the edges always point to the newly generated bubbles.

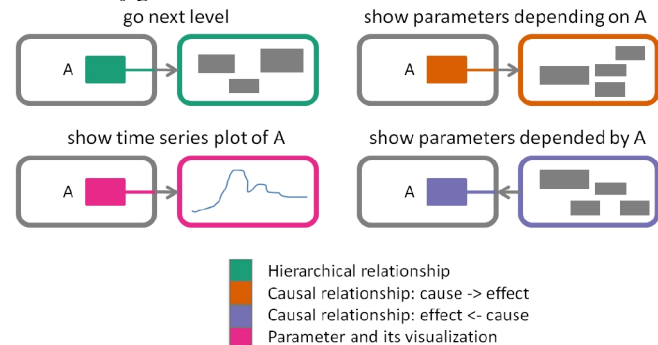


Figure 6: The color scheme and directional edges used to encode the relation between a pair of connected bubbles.

Along a bubble chain, the color and direction of the edge connecting a pair of bubbles indicate the different relationships between the selected tag in the parent bubble and the context of the child bubble.

(1) Exploring hierarchical relationships. Users interested in knowing the level- $n+1$  names of a level- $n$  name inside a bubble can show its corresponding level- $n+1$  names in a new bubble by clicking on the tag and then making the corresponding choice on a popup pie menu.

(2) Exploring causal relationships. Users interested in seeing all the attributes depending on a level- $n$  name inside a bubble can choose to show it in a new bubble. In the same way, a bubble containing all the attributes dependent on the selected name inside an existing bubble can be shown. Those names in the new bubble are also in the same level  $n$  and are shown through tag clouds.

(3) Showing a time-series plot of an attribute. Showing a visualization of the time-series data of a selected attribute is

allowed for a highest level (level-2) name through the same click-and-menu-selection operation.

The color scheme is chosen among the qualitative color sequences provided by ColorBrewer [30]. Figure 6 and 7 illustrate our color scheme and edge direction usage. The corresponding menu items in the menu are also colored accordingly.

Show different level parameter names which depend on A



Show different level parameter names which are depended by A

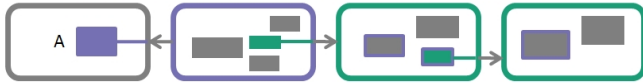


Figure 7: Dependence of attributes in different levels. Users often want to query which attribute names on the higher levels  $m$  ( $m > n$ ) depend on or are depended on by the selected attribute in current level  $n$ . The rectangular boundary of the tags is highlighted in orange or purple to imply that they correspond to the selected attribute in level  $n$ .

#### 4.2.3 Replicable Exploration Operations

Since any previous exploration operations are presented and made visible by the linkages of bubbles on the screen, users can copy any part of an existing exploration operation and apply it to query another attribute through a simple copy-and-paste operation (Figure 8). This function simplifies repeating the same exploration operation on different attributes.

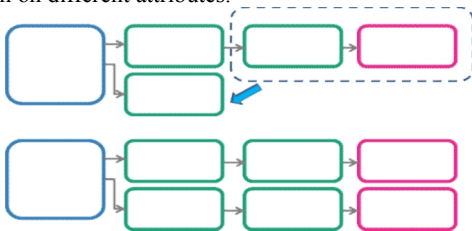


Figure 8: Copy-and-paste bubble sequence. Top: initial configuration. Bottom: after a user copies and pastes part of a bubble chain at another bubble position, the same exploration operation is applied to the second attribute.

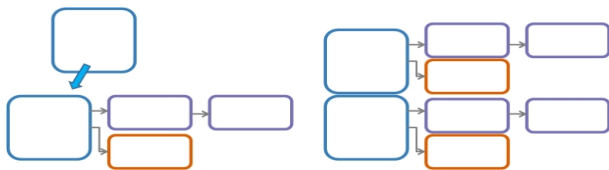


Figure 9: Grouping bubbles. Left: initial configuration. Right: after a user has dragged a root bubble toward another root bubble, the two bubbles are grouped and automatically share the same exploration operations thenceforth.

#### 4.2.4 Synchronizable Exploration Operations

We enable synchronizable exploration operations to support comparison tasks (T3). Users can always group two sets of connected bubbles to form a flexible multiple-view environment by dragging the second bubble of the set to a location overlapping the level-0 bubble of the first. Once grouped, data are automatically linked to support interactive operations in such a way that interacting in one view activates the same interactions in all other grouped views (Figure 9). This function is especially

suitable in comparing different ensemble runs. Whenever users want to compare different simulation ensembles, they can drag out a new bubble for the new ensemble dataset from the control panel and drop it at the location overlapped by an existing level-0 bubble: the two bubbles are immediately linked and share the same query process, with the queried information shown in a side-by-side view for easy comparison.

#### 4.2.5 Implementation

HumMod Browser was implemented in Visual Studio C++ and open source version of Qt 4.8. Qt is used for GUI design. Classes of different bubbles inherit both QObject and QGraphicsItem and the communications between bubbles are implemented through Qt's signal/slots mechanism. Data structure and tag cloud algorithm are implemented in C++ while the visualization is drawn through Qt QPainter. All the interactions are performed in real-time. The executable software is available to download at <https://sites.google.com/a/umbc.edu/pathbubbles/download>. The demonstration video is available to download at <https://sites.google.com/a/umbc.edu/pathbubbles/videos>.

### 5 CASE STUDIES

Here we present two case studies and illustrate the usefulness of HumMod Browser for analysis of time-varying physiology simulation data. The first case shows how HumMod Browser enabled physiologists to explore both hierarchical and causal relationships within physiological subsystems. The second case study shows how tag clouds and synchronizable exploration operations were used to compare between different ensembles.

#### 5.1 Understanding Attribute Interactions within the Lungs

Lungs are important physiological organs and have correspondingly large standard deviation (Figure 10) as well as large degree in the causal relation graph (Figure 1). We conducted a case study to reveal attribute interactions with the lungs. An interactive visualization was constructed as we conducted a series of operations to produce the display shown in Figure 10. The number indexes the order of the bubbles being created. The HumMod attributes sharing the level-0 name *Lungs* are all the attributes hierarchically under the lung's physiological system. When a dataset is selected by dragging out of a bubble from the control panel, the bubble only shows level-0 attribute names (Figure 10(1)). Among these, the attribute *Lungs* has high average standard deviations. Showing its level-1 attributes reveals that the attribute *Breathing* contributes the most to its large standard deviation. Bringing up the level-2 attributes in a bubble reveals the level-2 names sharing the same level-0 name *Lungs* and level-1 name *Breathing*. For instance, as shown in Figure 10(1), a possible attribute *Lungs.Breathing.TotalVentilation* causing the attributes of *Lungs* to vary is found by a glance through the sequence of bubbles indicated by the edges connecting them (Figure 10(1), (2), and (3)).

The exploration operations continue as we further explore the attributes functionally related to *Lungs*. First, we show all the level-0 attribute names depending on *Lungs* in the bubble with an orange boundary; then we show the level-0 attribute names dependent on *Lungs* in the bubble with a purple boundary. Side-by-side comparison of these two bubbles reveals that the level-0 name *CO2* appears in both bubbles. This means that *CO2* both causes attributes under *Lung* to change and is made to change by

those attributes. We then bring up bubbles showing the next level names under *CO2*. The boundary of the specific attributes under the level-0 name *CO2* and functionally connected to *Lung* is highlighted. In this case (Figure 10) they are *CO2.CO2Total.Outflow* and *CO2.CO2Veins.Pressure*. Finally, we look at the time-series plots of these two attributes and see that their values changed with similar trends. This case study thus reveals the interaction between attributes that may play important roles in the simulated lung physiological system.

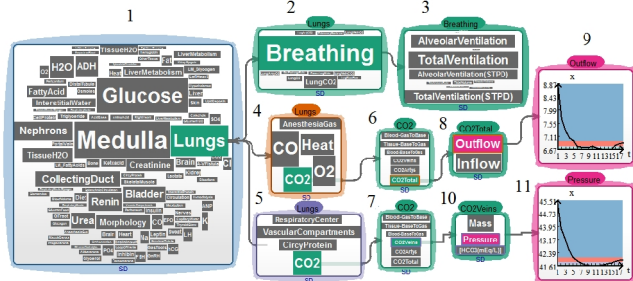


Figure 10: Exploration operations for discovering relationships of the selected attribute "Lungs".

## 5.2 Analyzing the Difference between Normal and Abnormal Cases

In the context of physiology modeling, the primary objectives of simulation studies are to predict how the human body will respond to certain inputs and to enhance our current knowledge of the living system affected by such input. HumMod generates abnormal outputs under pathological conditions. Comparison between the abnormal outputs with the normal outputs would provide certain insights into possible attributes that relate to abnormal outputs.

In this case study, we performed synchronized operations on two ensemble runs, one generated using baseline values from in-vivo experiment and one generated for a haemorrhage condition. As shown in Figure 11, the two big bubbles on the left contain the level-0 attributes of two ensemble runs. The haemorrhage case is shown on the top while the baseline case is shown on the bottom. The physiologists expected attributes such as skeletal muscle blood flow, cardiac output, heart rate, as well as some related to the nervous activities to be mostly different between the haemorrhage and the baseline cases. We grouped two bubbles together to allow them share same data exploration operations.

We first found that the tags of level-0 attribute names in two level-0 bubbles are very different. Successively, we queried the attributes mentioned above to find the ones significant different between the haemorrhage and baseline cases. For example, we queried level-0 attribute name *SkeletalMuscle* (highlighted in green), and found in the haemorrhage case that its level-1 attribute *SkeletalMuscle-Flow* has the highest standard deviation as well as the highest difference from that of *SkeletalMuscle-Flow* in the baseline case. Further exploration on the level-2 attributes under *SkeletalMuscle-Flow* revealed that the standard deviation of level-2 attribute *BloodFlow* in haemorrhage case are very low and much lower than that of *BloodFlow* in baseline case, confirming physiologist's previous assumption. The bubbles explored for *SkeletalMuscle* are the two green bubbles shown on the top and the two green bubbles shown on the bottom in Figure 11. When we finished the exploration of certain attributes, we cleaned out some bubbles from the screen and started new exploration operations.

Level-0 attribute *LeftHeart* are among those related to haemorrhage cases as suggested by the physiologists. We

performed exploration operations on it and found a notable difference between haemorrhage case and baseline case in its both level-1 and level-2 attributes (shown in the four green bubbles in the middle of Figure 11). Interested in knowing the possible attributes causing this difference, we explored the attributes depended by it through bringing out two purple bubbles (shown in middle of Figure 11 with the top one for the haemorrhage case while the bottom one for the baseline case). Both purple bubbles show the same list of attributes depended by *LeftHeart*. Among those attributes, *O2* shows the highest standard deviations, indicating its possible correlation with *LeftHeart* and hence a possible role in causing attributes under *LeftHeart* to change. Except for the attributes mentioned above, we noticed that level-0 attribute *RespiratoryCenter* (highlighted in purple) shows the highest standard deviation among all the level-0 attributes in the baseline case while has much smaller tag size in the haemorrhage case. Being curious about the reason, we explored the attributes depended by it. Since *Nerves* (shown in the two purple bubbles in the middle of Figure 11) is the only attribute which may cause *RespiratoryCenter* to change, we further explored the attributes depended by *Nerves*. Among the attributes shown in the new purple bubbles connected to *Nerves*, we found *O2* again. This suggests that *O2* may be an attributes playing an important role in causing both *LeftHeart* related attributes and *RespiratoryCenter* related attributes to change. Therefore, we further explored its dependent attributes in two orange bubbles and attributes depended by it in two purple bubbles. After we viewed the attributes contained in both the purple and orange bubbles (in the bottom-right or the top-right of Figure 11), we found that they share most attribute names such as *Skin*, *Fat*, *kidney*, and *Brain*, indicating possible loops in causal relationships.

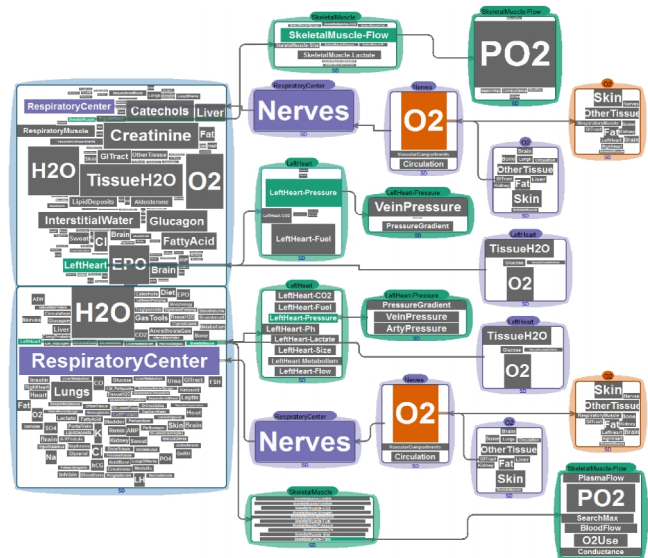


Figure 11: Comparison of two different simulation outputs. The two ensemble data, normal and abnormal, are shown through the two sets of bubbles.

This case study demonstrates that HumMod Browser helps a fast identification of the significant difference between normal and abnormal outputs, suggests possible causal relationships among the attributes, and also a possible link between certain attributes and a given pathological condition.

## 6 DISCUSSION

HumMod Browser is designed as an exploratory visualization tool for analyzing whole-body physiology simulation data characterized with large attribute space and complicated

relationships. We collected comments from two physiologists, who are also our co-authors, on the use of HumMod Browser, providing some evidence that our visualization and interaction design, grounded in a characterization of the domain requirements, support an efficient exploration in large-attribute-space. We discuss the limitations, solutions, and possible extensions of our design in the section as well.

### 6.1 Comments from Physiology Experts

Physiologists considered the importance-driven tag cloud visualization useful in locating the important attributes, and they also suggested that a tag cloud showing the differences between ensemble runs can be a new metric. They mentioned that the order of attribute names were not consistent in different tag clouds making the searching for specific attributes inefficient. This can be addressed by either adding a search function or by the brush-and-link technique to highlight the same attribute in the exploration operations. The dynamic linkages between bubbles were considered helpful in exploring the hierarchical and casual relationship between attributes. They also liked the simultaneous views and flexibility to edit bubble context, expand and hide bubbles. The left-to-right sequence view was designed by following the order of their exploration and used a space save algorithm. Another method could be to layout the newly created bubble closest to the bubble of interest. Our physiologist collaborators do not have preferences one way or the other. It will also be interesting to compare this design with our previous tree-ring layout [10] for data analysis where a single view is used. The bubbles replication and synchronized query by grouping were found useful. They confirmed that memory requirement when exploring HumMod simulation data was lowered with visible exploration operations shown in the screen and they did not need to memorize the specific attribute names and relationships queried in previous steps. One physiologist suggested an extension of the tool to detect repeated attributes (loops) along a causal chain. Eventually, physiologists want to answer more challenging questions such as which attributes relate to a specific outcome and when these attributes are irrevocably linked to that outcome. In addition, more importance measurements, such as the difference between same attributes in different ensemble runs, deviation between eventual values and baseline values from in-vivo experiment might be considered, except for the currently used standard deviation and degree of an attribute in the causal relationship graph.

### 6.2 Limitation, Solution, and Extensions

Intuitively conveying relative importance was the main motivation for us to use tag clouds. One known issue of tag clouds is that the lengths of the tags can influence their perceived sizes [21]. This is alleviated by placing bigger tags closer to the center so that bigger (more important) but shorter tags may be noticed before those smaller (less important) and longer tags placed further from the center. The inconsistent order of attribute names in different tag clouds can be addressed by adding a search or brush-and-link function to allow easier search of same attributes across different tag clouds, as mentioned in section 6.1.

Although we did not use traditional diagram such as compound graphs [31] and time-series analysis [32] for relation exploration, our interaction design which supports memory enhancement and replicability is compatible with these diagrams. Examples of integrating our interaction design with compound graphs and time-series plots are shown in Figure 12. The rich interactions, including dynamic view control, pannable and scalable workspace, grouping, copy and paste, etc. would become available for users to generate highly interactive and less cluttered graph visualizations (Figure 12 (a)) and look into detailed

segments of a time-series dataset and comparing them side-by-side (Figure 12 (b)). In addition, our design can be applied to study other models and data sets such as gene pathway [33] and flowmap dataset [7] whose variables have hierarchical and/or networking relationships as well.

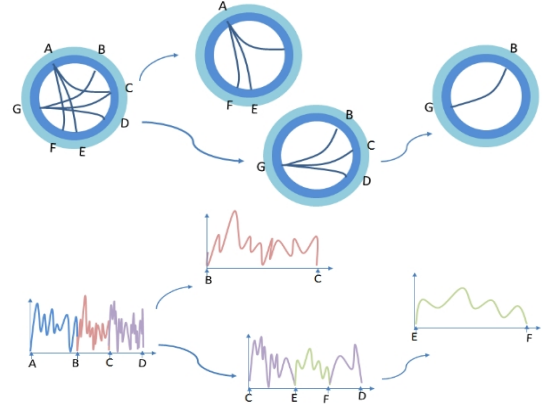


Figure 12: Examples of extending our design to other visualization forms and data types. Top: Exploration of attribute space with compound graphs. Bottom: Hierarchical exploration of time-series data.

Although we believe that our tool help physiologist gain insight over attributes' relationships, importance, and differences between model runs. An insight-based evaluation may be needed to further evaluate the effectiveness of our design for providing insights [34]. In addition, one interesting direction we would like to take in the future is to compare our virtual workplace with the large displays, e. g. tilted displays in terms of the scalability of space uses [35] though user studies.

## 7 CONCLUSION

We have presented HumMod Browser, a new interface to assist physiologists in analyzing time-varying human modeling data. We have demonstrated the use of HumMod Browser in two case studies, providing some evidence that our interaction designs make possible a fluent exploration operations that could potentially overcome the limitations of human memory. We have summarized a list of physiologist activities and introduced four design goals that are crucial for creating such a system. Our results indicate that we can design new exploratory visualization tools using a new metaphorical design in a new domain of uses visualizing physiological causality. Our important measurement metrics can be reused by other physiology causality analysis toolkits. The rich metaphorical interactions exemplified in current HumMod Browser can provide beneficial alternatives to the classical graph visualization to assist physiologists in learning causality relationships which is important in validating mathematical modeling data.

## ACKNOWLEDGMENTS

This work is partially supported by the National Science Foundation under grant numbers DBI-1260795, IIS-1341254, EPS-0903234, and HL-51971.

## REFERENCES

- [1] R.L. Hester, A.J. Brown, L. Husband, R. Iliescu, D. Pruett, and R. Summers, "HumMod: A modeling environment for the simulation of integrative human physiology. *Frontiers in physiology*, 2011, doi: 10.3389/fphys.2011.00012.
- [2] J. Chen, K. Wu, W.A. Pruett, and R.L. Hester, "HumMod Browser: An Exploratory Visualization Tool for Model Validation of Whole-

- Body Physiology Simulation," Eurographics Conference on Visualization (EuroVis) (short paper) 2013.
- [3] P. Shannon, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, pp. 2498-2504, 2003, doi: 10.1101/gr.1239303.
  - [4] Z. Hu, "VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology," *Nucleic Acids Research*, vol. 37, no. (web server issue), pp. W115-W121, 2009, doi: 10.1093/nar/gkp406.
  - [5] M. Meyer, T. Munzner, and H. Pfister, "MizBee: A Multiscale Synteny Browser," *IEEE Transactions on Visualization and Computer Graphics* vol. 15, no. 6, pp. 897-904, 2009, doi: 10.1109/tvcg.2009.167.
  - [6] M.I. Krzywinski, J.E. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S.J. Jones, and M.A. Marra, "Circos: An information aesthetic for comparative genomics," *Genome Research*, vol. 9, no. 9, pp. 1639-1645, 2009.
  - [7] D. Holten, "Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 741-748, 2006, doi: 10.1109/tvcg.2006.147.
  - [8] B. Johnson and B. Shneiderman, "Tree-maps: a space-filling approach to the visualization of hierarchical information structures," *Proceedings of IEEE Visualization Conference*, pp. 284-291, 1991, doi: 10.1109/visual.1991.175815.
  - [9] N. Elmqvist and P. Tsigas, "Causality visualization using animated growing polygons," *Proceedings of IEEE Symposium on Information Visualization*, 2003, pp. 189-196, 2003, doi: 10.1109/infvis.2003.1249025.
  - [10] L. Xu, J. Lyle, Y. Wu, Z. Pan, M. Zhang, D.H. Laidlaw, R.L. Hester, and J. Chen, "HumMod explorer: a multi-scale time-varying human modeling navigator," *SIGGRAPH Asia Posters*, 2011.
  - [11] J. Bertin, "Semiology of Graphics," WI: The University of Wisconsin Press, 1983.
  - [12] J. Abello and F. van Ham, "Matrix Zoom: A Visual Interface to Semi-External Graphs," *Proceedings of IEEE Symposium on Information Visualization* pp. 183-190, 2004, doi: 10.1109/infvis.2004.46.
  - [13] N. Henry and J. Fekete, "MatrixExplorer: a Dual-Representation System to Explore Social Networks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 677-684, 2006, doi: 10.1109/tvcg.2006.160.
  - [14] J. Stasko, "An evaluation of space-filling information visualizations for depicting hierarchical structures," *International Journal of Human-Computer Studies*, vol. 53, no. 5, pp. 663-694, 2000, doi: 10.1006/ijhc.2000.0420.
  - [15] N.R. Kadaba, P.P. Irani, and J. Leboe, "Visualizing Causal Semantics Using Animations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1254-1261, 2007, doi: 10.1109/tvcg.2007.70528.
  - [16] S.P. Callahan, J. Freire, E. Santos, C.E. Scheidegger, C.T. Silva, and H.T. Vo, "VisTrails: visualization meets data management," *Proceedings of ACM SIGMOD international conference on Management of data*, pp. 745-747, 2006, doi: 10.1145/1142473.1142574.
  - [17] M. Kwan-Liu, I. Liao, J. Frazier, H. Hauser, and H.N. Kostis, "Scientific Storytelling Using Visualization," *Computer Graphics and Applications, IEEE*, vol. 32, no. 1, pp. 12-19, 2012, doi: 10.1109/mcg.2012.24.
  - [18] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1139-1148, 2010, doi: 10.1109/tvcg.2010.179.
  - [19] Y. Tanahashi and M. Kwan-Liu, "Design Considerations for Optimizing Storyline Visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2679-2688, 2012, doi: 10.1109/tvcg.2012.212.
  - [20] G. Li, A.C. Bragdon, Z. Pan, M. Zhang, S.M. Swartz, D.H. Laidlaw, C. Zhang, H. Liu, and J. Chen, "VisBubbles: a workflow-driven framework for scientific data analysis of time-varying biological datasets," *SIGGRAPH Asia Posters*, 2011.
  - [21] N. Ferreira, L. Lins, D. Fink, S. Kelling, C. Wood, J. Freire, and C. Silva, "BirdVis: Visualizing and Understanding Bird Populations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2374-2383, 2011, doi: 10.1109/tvcg.2011.176.
  - [22] A. Bragdon, S.P. Reiss, R. Zeleznik, S. Karumuri, W. Cheung, J. Kaplan, C. Coleman, F. Adeptra, and J.J. LaViola, "Code bubbles: rethinking the user interface paradigm of integrated development environments," *Proceedings of ACM/IEEE 32nd International Conference on Software Engineering*, pp. 455-464, 2010, doi: 10.1145/1806799.1806866.
  - [23] A. Campanharo, M. Simer, R. Malmgren, F. Ramos, and L. Amaral, "Duality between Time Series and Networks," *PLoS ONE* vol. 6, no. 8, 2011, doi: doi:10.1371/journal.pone.0023378.
  - [24] S. van den Elzen and J.J. van Wijk, "Small Multiples, Large Singles: A New Approach for Visual Data Exploration," *Computer Graphics Forum*, vol. 32, no. 3, pp. 191-200, 2013, doi: 10.1111/cgf.12106.
  - [25] S. Bateman, C. Gutwin, and M. Nacenta, "Seeing things in the clouds: the effect of visual features on tag cloud selections," *Proceedings of the 19th ACM conference on Hypertext and hypermedia*, pp. 193-202, 2008, doi: 10.1145/1379092.1379130.
  - [26] A.W. Rivadeneira, D.M. Gruen, M.J. Muller, and D.R. Millen, "Getting our head in the clouds: toward evaluation studies of tagclouds," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 995-998, 2007, doi: 10.1145/1240624.1240775.
  - [27] A. Slingsby, J. Dykes, J. Wood, and K. Clarke, "Interactive Tag Maps and Tag Clouds for the Multiscale Exploration of Large Spatio-temporal Datasets," *Proceedings of 11th International Conference Information Visualization*, pp. 497-504, 2007, doi: 10.1109/iv.2007.71.
  - [28] P. Mundada and A. Ghotkar, "An approach to second generation tag cloud for assessment of business search," *Proceedings of IEEE International Conference on Technology Enhanced Education (ICTEE)*, pp. 1-3, 2012, doi: 10.1109/ictee.2012.6208610.
  - [29] J. Shier, "Filling Space with Random Fractal Non-Overlapping Simple Shapes," in *Hyperseeing*, summer issue. ISAMA (International Society of the Arts, Mathematics, and Architecture). pp. 131-140, 2011,
  - [30] M. Harrower and C.A. Brewer, "ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps," *The Map Reader*, John Wiley & Sons, Ltd. pp. 261-268, 2011.
  - [31] D. Holten, "Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 5, pp. 741-748, 2006, doi: 10.1109/tvcg.2006.147.
  - [32] R. Dahlhaus and M. Eichler, eds. "Causality and graphical models for time series," *Highly structured stochastic systems*, ed. N.H. P. Green, and S. Richardson, University Press: Oxford.
  - [33] U. Dogrusoz, E. Giral, A. Cetintas, A. Civil, and E. Demir, "A Compound Graph Layout Algorithm for Biological Pathways," *Graph Drawing*, J. Pach, Springer Berlin Heidelberg. pp. 442-447, 2005.
  - [34] P. Saraiya, C. North, and K. Duca, "An Insight-Based Methodology for Evaluating Bioinformatics Visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 4, pp. 443-456, 2005, doi: 10.1109/tvcg.2005.53.
  - [35] C. Andrews, A. Endert, B. Yost, and C. North, "Information visualization on large, high-resolution displays: Issues, challenges, and opportunities," *Information Visualization*, vol. 10, no. 4, pp. 341-355, 2011, doi: 10.1177/1473871611415997.