Ontological Modeling of Transformation in Heart Defect Diagrams

Venkatesh Viswanath, MS, Tuanjie Tong, MS, Deendayal Dinakarpandian, MD, PhD, Yugyung Lee, PhD
School of Computing & Engineering, University of Missouri-Kansas City, Missouri, MO

The accurate portrayal of a large volume data of variable heart defects is crucial to providing good patient care in pediatric cardiology. Our research aims to span the universe of congenital heart defects by generating illustrative diagrams that enhance data interpretation. To accommodate the range and severity of defects to be represented, we base our diagrams on transformation models applied to a normal heart rather than a static set of defects. These models are based on a domain-specific ontology, clustering, association rule mining and the use of parametric equations specified in a mathematical programming language.

INTRODUCTION

A diagram is a simplified, structured representation that describes the components of a system and their spatial relationships. Mullins diagrams are a standard reference for congenital heart defects [8]. Mullins diagrams are used to document the diagnosis as well as post-operative details in pediatric cardiology. Specifically, when clinical measurements like blood velocity, thickness of septum are encoded with the help of diagrams, they serve as a useful tool to plan, guide, and record appropriate patient care.

In typical practice, one of the 169 Mullins diagrams is chosen and then manually modified to correspond to the exact condition of the patient. This is because there is a large variety of congenital heart defects, many of which are similar but not identical to one or more of the Mullins diagrams. As a result, significant effort and time are expended in manual editing and annotation of Mullins diagram templates based on a patient's particular condition. The availability of automatic tools that support the creation of custom diagrams would be very useful in this regard.

We have previously published work on OntoDiagram (Ontology based Diagram Generation) [8], a project that utilizes clinical and spatial cardiac ontologies to automatically generate a Mullins-like diagram based on patient information in cardiology databases. We expand this work in the present paper to capture the large variability of cardiac defects. In particular, we view heart defects as being the effect of transformation models applied to the normal heart. This is made possible by modularizing the knowledge underlying the heart diagram. We describe the diagrams in terms of sub-models that represent different aspects of the diagram like changes, orientation and spatial relations. We have modeled the diagram as a set of components with specific semantics.

The objectives of the transformation modeling are (1) to generate diagrams based on descriptions given by domain experts and representation formats, (2) to mine implicit domain knowledge stemming from the domain ontology that is a part of the system and (3) to perform dynamic diagram generation of defects using equations. The prototype has been implemented using WEKA [9] for transformation pattern mining, Protégé[1] for creating the ontology representing the transformation model and Maple[2] for generating transformation diagrams.

RELATED WORK

An interesting diagram generation research problem is automating the design of graphical presentations of relational information [4]. The major challenge addressed by the work is determining the graphical representation that best expresses the required information within specified geometric constraints. Intelligent Diagram editor [2] focuses on generating an editor that can parse a diagram as it is being constructed, performing error correction and collecting geometric constraints that capture the relationships between diagram components. The TRIP2 system [5] provides bidirectional translation between abstract and pictorial data to generate common diagrams like transition diagrams and flowcharts. The AVE system [3] generates diagrams by composing basic diagram components that correspond to various data elements; it is particularly useful to generate diagram for heterogeneous data.

Mimos [1] describes the semantics of medical image processing but does not allow the generation of a diagrammatic representation as it lacks the capacity to describe structural aspects of images. Shenkman's work [6] focuses on generating circuit diagrams through an internal mapping between various standard notation and diagram elements. Smith and Linders [7] developed a method for automatic circuit diagram generation by exploiting

[1] http://protege.stanford.edu/
[2] http://www.maplesoft.com/

techniques of partitioning, placement and routing in which geometric constraints and diagram consistency should be considered. We addressed some of these issues in OntoDiagram [8] using a component based approach for diagram generation to increase effects of reusability. In contrast to the use of static contexts in the AVE system [3], we use dynamic contexts to reflect transformational changes in diagrams.

TRANSFORMATION MODELING

Since the human heart is a mechanical system, there are several biological factors, hemodynamics and other complex anatomical details that may influence some transformations. As a result, a transformation can induce another transformation in some other component that is not physically related. For example, an instance of *Truncus Arteriosus* implies the presence of *Ventricular Septal defect* (missing component). Sometimes these induced transformations may not be explicitly specified in the domain descriptions. Thus this requires additional instructions to reflect these implicit changes in diagrams. Propagation of transformation addresses changes of components that are anatomically continuous or at the proximity with components getting transformed. However, it does not address changes that might occur due to other complex relationships the components might have despite being physically remote from each other. The Mullins diagrams model complex relationships between components that are beyond boundary relationships or physical closeness.

Mullins diagram domain offers many challenges in terms of transformation complexity. Firstly, some domain conditions are required to be represented differently on different components. For example, *Stenosis* may be used differently to represent different transformations (e.g., *Stenosis* in *Aortic trunk* vs. *Stenosis* in *Aortic valve*). Secondly, propagation effects are complex and do not behave the same way for all transformations. For different transformations different patterns of propagation are observed. Additionally, the propagation pattern is dependent on the component on which the transformation is exerted. Thirdly, implicit domain information is not always specified in the domain description. This implicit information might translate to a new unspecified transformation on components. It is necessary to build a domain model that interprets the effects of changes and ensures consistency of changes with respect to the domain.

The underlying diagram model in the OntoDiagram serves as a reference for generating different variants of the diagram for specific domain descriptions (Table 1). A Transformation is a change applied to the original model that is reflected in its subsequent visualization.

The Transformation model is represented as 3-tuple $\{C_T, T, P\}$, where C_T is the component being transformed, T represents the transformation on the component, and P is a set of parameters for the transformation defined. For example, if T represents scaling then P gives the scaling factors along x and y axes. T can be a geometric or polynomial transformation or other type of custom transformation. In cases of polynomial transformation, the parameter P represents a set of transformation equations and control points that are used as the equation inputs. A transformation model is built based on the transformation effects shown in the heart components defined in the OntoDiagram ontology [8].

Table 1. Examples of Transformation

Defect	Operations	Parameters
Stenosis	Narrowing of tubes	Narrowing width (severity of *stenosis*)
	Narrowing of Valves	Narrowing radius (severity of *stenosis*)
Dilation	Bulges in tubes	Bulge factor (severity of *dilation*)
Bifurcation	Split in tubes	Position of split (location of *bifurcation*)
Coarctation	Stricture of tubes	Location of stricture (which portion of the tube has *coarctation*)
Atresia	Missing Valves	Missing components

Typically, transformations are defined for specific components and direct and indirect types of transformation are defined for comprehensive coverage of cardiac structural variations. Direct transformation explicitly affects the component of the diagram model for which a transformation is defined. For example, a *stenosis* transformation defined on a specific component is said to be direct. When a direct transformation on a component affects the structure of an adjacent component, the latter is said to have undergone indirect transformation. It can be thought of as a change propagated to adjacent or neighboring components.

In order to determine if a component can experience indirect transformations, it is necessary to know how direct transformations propagate across the diagram model. This propagation cannot be trivially determined as it is not only observed on components with anatomical continuity but also on those in the proximity. For example, an indirect transformation on *aortic arch* is observed as an effect of *dilation* on *aortic trunk*.

Typical transformations can be classified as follows: *Abnormal growth* (new components which are not a part of the normal diagram model), *Absence* (missing components), *Transposition* (exchange of positions between two components, e.g., *TGA*), *Deformation* (abnormal shape), and a combination of any of these, e.g., *Tetralogy of Fallot*.

Based on the nature of physical change (e.g., abnormal growth, deformation of component, etc), transformation of components can be classified into geometric and non-geometric (or polynomial transformation) transformations. Geometric transformation includes simple operations like scaling, rotation and translation. Non-geometric transformations are a common occurrence but more complex to represent. Such transformations can be defined as equations to generate the corresponding changes in the diagram. If components are represented as vector images, these equations can be used to generate the transformation of control points or nodes on the vector image.

TRANSFORMATION PATTERN MINING

Though it would be difficult to determine all the possible changes that a diagram or components in the diagram could take, frequently encountered changes in individual components of the diagram may be determined. This will allow the combinatorial coverage of a large number of cardiac anomaly presentations without the need for an exhaustive listing. Studying transformation patterns can be used to determine transformation equations, the effect of transformation on components (i.e., same transformation can have different effect on different components), the effect of transformation on other components (propagation of transformation), parameters of transformation and parametric analysis (how changes in parameters of transformation affect the extent of transformation).

The defect patterns can be thought of as signatures of the diagrams representing cardiac defects. A signature can be represented as a set of <component, transformation> pairs describing the semantics of each diagram or defect. A signature is defined as a set $Di = \{<C_1,T_1>, <C_2,T_2>, \ldots <C_N,T_N> \}$, where $C_1, C_2, \ldots C_N$ are heart components and $T_1, T_2, \ldots T_N$ are the respective transformations on those components. The signature represents the semantics from a diagrammatic perspective. For instance, the signature for right *aortic arch* would contain the set of components that get transformed in right *aortic arch* cases and the transformations on each of them.

In this paper we have analyzed the component changes for each cardiac defect in the Mullins atlas and then determined the component association patterns using data mining approaches. Association rule mining (apriori algorithm [9]) was used to identify association patterns existing among defects, components or effects of transformations based on Mullins atlas. K-Means clustering [9] was used to cluster the defects based on common characteristics of components and their changes. For example, if diagrams having right *aortic arch* form a cluster then each member of the cluster may share similar characteristics. For diagram generation, physician descriptions could be converted into signatures of diagram transformation using the results of the clustering, and the component associations patterns used to complete the signature when incomplete domain descriptions were given.

EQUATIONS FOR BLOOD SUPPLY STYLES AND JOINING PATERNS

Based on analysis of the diagrams in Mullins atlas, we have devised a set of mathematical equations for dynamically representing transformations in diagrams. We have analyzed the blood supply aspect of the Mullins diagrams and modeled blood supply (tubes) patterns and joining patterns. We have modeled six Valve styles: (a) *Valveless style* (b) *Valveless style with orifice* (c) *Atresia of valve style* (d) *Valve leaflet style* (e) *Vented valve leaflet style* and (f) *Stenosis style*.

The blood supply joining patterns are modeled as follows. (a) The *bifurcate* pattern is used to generate cases of bifurcated blood supply. Examples are a *bifurcation* of *pulmonary artery* and *double aortic arch*. (b) The *confluence* pattern is used to generate diagram portions that represent confluence of various blood supply valves. Inferior vena cava is an example. (c) The *merge* pattern is applicable when multiple valves merge at one point instead of being confluent, e.g. *pulmonary artery*. (d) *Merge-into-shape* pattern is useful to generate cases where multiple tubes merge to form a special shape as in the case of *coratrium*. The equations defined for various blood supply elements in conjunction with joining styles have been used to generate several variations of the normal blood supply structure.

RESULTS

We have analyzed 169 different Mullins diagrams and identified 107 defect types, 38 heart components, 51 transformation types, and 432 transformation instances. We now specifically describe the ontology we constructed for representing the transformation model, transformation patterns we discovered using data mining and automatic diagram generation using the patterns.

TRANSFORMATION ONTOLOGY

The Transformation ontology constructed using Protégé[1] contains the 51 transformation concepts shown in Fig. 1. The Transformation ontology is comprised of the transformation model that describes congenital heart defects in terms of effects on different components, and transformation operations for generating diagrams. In our model, a defect is defined as an abnormal condition of heart structure. The condition can be mapped to the changes (effects of transformation) in the heart structure. There could be a single case or a complex case with two or more effects.

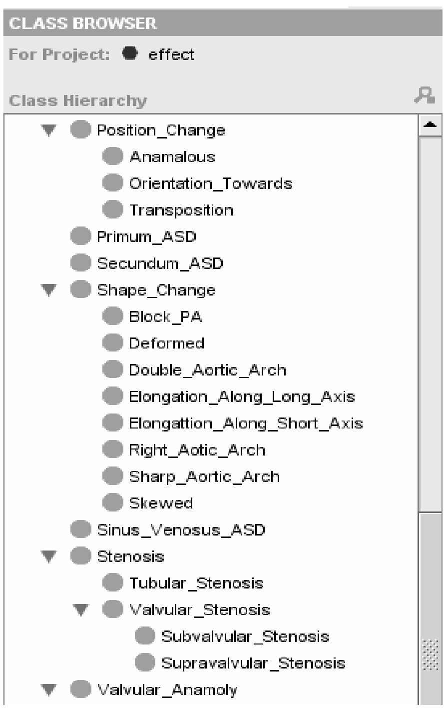


Fig. 1. Transformation Ontology

The following four transformation-specific slots were introduced in the transformation ontology: (a) *Components affected* slot contains information on the components that are influenced by a specific effect. (b) *Severity* slot has information on various severity values that an effect can take. (c) *Association pattern* slot contains information on the association of an effect with different components. (d) *Transformation* slot contains diagram generation information with respect to indirect transformation of components.

MINING TRANSFORMATION PATTERNS

Clustering techniques have been used to organize data in classes based on attribute values. The main objective is to minimize inter-class similarity and maximize intra-class similarity. In our study, K-Means in Weka [9] was used to determine centers to represent *k* cluster (for a given *k*) and assigning objects to the closest center using Euclidian distance similarity measure.

For the clustering, we split the dataset of 432 transformation instances into two portions, 66% for training and the remainder for testing. Table 2 shows the six clusters whose centroid is determined based on the three attributes: defect, component, and transformation.

Table 2. Clusters for Transformation

Centroid	Instance	%
D75, Inter Ventricular septum, Transposition	264	61%
D128, Left Auricle, Drains into	42	10%
D41, Aortic valves, Right aortic arch	39	9%
D72, Inter Atrial septum, Membraneous VSD	56	13%
D131, Left Auricle Drains into	3	1%
D31, Aortic Arch, tubular stenosis	28	6%

We have also applied the association rule mining technique to the transformation dataset and discovered association patterns among defect, component, and transformation. Table 3 shows the association rules generated by Apriori in Weka [9] with minimum support 0.05.

Table 3. Association Rules for Transformation

Rule	Confidence
Secundum ASD ==> Right Ventricle	conf:(1)
Right Aortic Arch==> Aortic Valves	conf:(1)
Membraneous VSD ==> Inter Atrial Septum	conf:(1)
Transposition ==> Aorta	conf:(0.8)

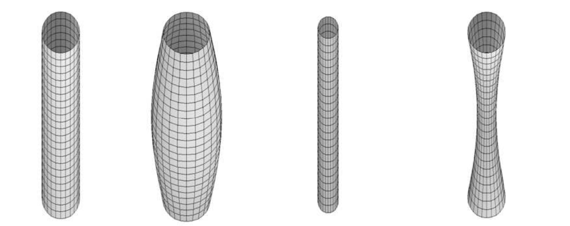
DIAGRAM GENERATION USING MAPLE

We have generated three types of transformation using Maple[2]. The transformation equation table (Table 4) and diagrams (Fig. 2) show the effect of the parameters on three types of transformation applied to (a) Original shape of the tube: (b) *Dilation of Pulmonary*, (c) *Artery Stenosis of Pulmonary* and (d) *Coarcation of Aorta*. The Maple tool allows us to model the various portions of the diagram and visualize the effect of parameter change.

Table 4. Maple Equations and Parameters for Transformation

Transformation	Maple Equation and Parameters
Dilation	*Animate3d(Pi+t/8*sin(Pi*y/(8*Pi)), x=0..2*Pi, y=0..8*Pi, t=0..4*Pi, coords=cylindrical,)*

Artery Stenosis of Pulmonary	*Animate3d(Pi-t/8, x=0..2*Pi, y=0..8*Pi, t=0..4*Pi, coords=cylindrical);*
Coarcation of Aorta	*animate3d(Pi-t/8*sin(Pi*y/(8*Pi)), x=0..2*Pi, y=0..8*Pi, t=0..4*Pi, coords=cylindrical);*



(a) Original Shape (b) Dilation of Pulmonary (c) Artery Stenosis of Pulmonary (d) Artery Coarcation

Fig. 2. Diagrams generated using Maple

DISCUSSION

Mullins diagrams are an invaluable resource in Pediatric Cardiology. However, the set of diagrams is best considered as being representative rather than an exhaustive enumeration of the spectrum of congenital heart defects. We have addressed this issue by providing a scheme for automatic diagram generation by using transformation models that combine explicit instructions together with implicit ones based on a multi-faceted ontology.

Automatic diagram generation is made possible due to the interpretation of complex transformations through our ontological modeling. The degree of change is represented by a set of parameters. The parameterized transformation facilitates geometric or polynomial transformation. This equation-based transformation generation increases the diagram generation capability in terms of the variety of structures that can be generated. Since the required transformation can be dynamically generated, static diagrams are no longer required in the repository.

A current limitation of our framework is the fact that we have manually analyzed the Mullins diagrams to collect the transformation data. Ideally, we would like to utilize more domain knowledge from diverse domain resources such as UMLS[3]. Also, we need to validate the proposed model through in depth assessment of domain experts.

ACKNOWLEDGEMENTS

This research was funded in part by grants from the Heartlab Inc., University of Missouri Research Board (UMRB) and UMKC Faculty Research Grant (FRG).

CONCLUSION

We have designed a transformation model for congenital heart defects extracted from the Mullins diagrams. Three important approaches have been explored: (1) ontology to encode domain knowledge and implicit knowledge discovered from data mining process (2) mining of association and clustering patterns to find relationships between defects and resulting effects (3) automatic transformation generation using arithmetic equations. Consequently, this allows us to generate an automatic diagram for a given domain description of congenital heart defects. A prototype system has been implemented.

References

1. F. Aubry, A. Todd-Pokropeka, Mimos: A Description Framework for Exchanging Medical Image Processing Results, MEDINFO 2001. 891–895.
2. S. Chok and K. Marriott, Automatic Generation of Intelligent Diagram Editors, ACM Transactions on Computer-Human Interaction, 10(3):244 - 276, 2003.
3. G. Golovchinsky, T. Kamps, K. Reichenberger, Subverting Structure: Data-driven Diagram generation, IEEE Conference on Visualization, 217 – 223, 1995.
4. J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics, 5(2): 110 - 141, 1986.
5. S. Matsuoka, S. Takahashi, T. Kamada, A. Honezawa, A Generational Framework for Bidirectional Translation between Abstract and Pictorial Data, ACM Transactions on Information Systems, 10(4):408 - 437, 1992.
6. S. Shenkman, Circuit Diagram Generation Via Functional Logic, Proceedings of ACM IEEE Design Automation Conference on Design automation, 267 - 273, 1973
7. J. Smith and J. Linders, Automatic Generation of Logic Diagrams, Proceedings of the ACM IEEE Design Automation Conference 377 - 391, 1976.
8. K. Vishwanath, V. Viswanath, W. Drake, Y. Lee, OntoDiagram: Automatic Diagram Generation for Congenital Heart Defects in Pediatric Cardiology, Proceeding of American Medical Informatics Association (AMIA), 2005.
9. I.H. Witten, E. Frank, L. Trigg, M. Hall, G. Holmes, and S. J. Cunningham, Weka: Practical Machine Learning Tools and Techniques with Java Implementations.

[3] http://www.nlm.nih.gov/research/umls/