PROPOSAL WHITE PAPER

BAA Number 07-036

Title: An Algebraic Approach to Finding Self-Similar Structures in Graphs

Technical Point of Contact: Prof. D. Paul Benjamin
phone: (212) 346-1012
fax: (212) 346-1863
benjamin@pace.edu

Topic 2: Real-Time Methods for the Analysis of Networks

Overview

We have developed the basis of an algebraic theory that relates the local and global structure of large graphs, implemented it, and successfully applied it to a variety of applications. We propose to extend the theory to include more complex graph structures and to complete the implementation of a set of mathematical tools that can be used to analyze graphs and classify them according to their self-similar structure.

Our approach is based on the mathematics of algebraic linguistics and dynamical systems theory, and focuses on symmetries and invariants of local structures that exist in the graph. These structures often exist on a range of scales in a manner similar to fractals. Each distinct graph possesses a unique such structure. The analytical tools developed by this project will enable the researcher to find this structure and use it to derive properties of the graph. This approach to graph analysis is original to this project.

Identification of the research and issues

Many of the most important military applications are physical, social and biological systems whose complexity creates significant difficulties in the analysis of their structure. Modeling such systems is hindered by problems of accuracy and tractability. In addition, these systems exhibit structure over a large range of space and time scales, ranging from small scales of millimeters and milliseconds to large scales of kilometers and hours. Each of these scales seems to possess a complex structure, and each can potentially constrain any other. One general approach to overcoming these difficulties is to bring powerful mathematical tools to bear, to create tools of great generality that can be applied to entire classes of complex systems.

A general and powerful body of theory is the theory of graphs, which has been used to model a wide variety of systems, including state space graphs, constraint satisfaction graphs, resource-flow graphs and Petri nets. Nodes in the graphs stand for entities or resources, and edges between these nodes can represent actions, flows of materials, or constraints imposed by one node on another. This approach to analysis has yielded a number of useful results; however, this approach has not led to a useful classification of graphs, and in particular to an understanding of how the properties and structure of complex systems arise from the properties and structure of their sub-systems. As a result, current graph-theoretic techniques cannot tractably analyze large systems, especially ones that change over time, and these are precisely the systems that confront the military analysts.

Associated with the theory of graphs is a body of algebraic theory that has been much less explored in these applications. The algebraic approach analyzes graphs in terms of algebras of partial transformations acting on the graphs [5,7,8,10]. Various structures within these algebras correspond to useful properties in the system. For example, a normal subsemigroup within an algebra gives a decomposition for the graph that is analogous to the coset decomposition of a group. This identifies subgraphs (the “orbits” of each coset) whose properties compose to give properties of the entire graph.

The algebraic approach has the advantage that it permits us to talk of families of graphs, and of inheritance of properties within families of graphs [5]. This enables us to handle systems that evolve over time, and to try to classify their structure. A successful classification scheme will enable us to identify anomalous structures within graphs and their corresponding systems, and to identify when an evolving graph changes its classification.

Proposed Technical Approach

Often, research into the analysis of complex systems has been hindered by the belief that complex behavior requires complex mechanisms that explicitly encode complex structure. This has led to much work that constructs elaborate formalisms and simulators for complex systems; the complexity of the analytical tools rivals that of the systems themselves. Recent work in dynamical systems theory suggests a different course of action.

The study of dynamical systems is a field of mathematics with very wide applicability throughout science. Simply formulated dynamical systems can exhibit a wide range of complex behaviors that are not evident from their formulations. Even the simple pendulum turns out to possess chaotic behavior! This relationship between simple systems and complex behaviors poses a difficult hurdle, as sophisticated mathematics may be required to understand the behaviors of even very simply formulated nonlinear systems. However, we are interested in the inverse problem: how to formulate a relatively simple graph representation of a complex system, so this relationship between simple systems and complex behaviors is attractive. The importance of our research project is that it holds the potential to greatly reduce the complexity of analysis by finding methods for building simple hierarchical graph structures that express the complex behaviors of large systems.

The perspective of this research project is that the basic structure of the graph of a complex system is given by the algebraic structure of the local neighborhoods of its associated semigroup, and the important properties are those properties that hold *locally everywhere* within the neighborhoods. Under certain common conditions, the local neighborhood structure becomes a substitution tiling. In these cases, self-similarity of the neighborhoods permits analysis to be effective at a range of scales, creating a hierarchical structure that is simple and local at each scale, but capable of explaining very complex global structure. Fortunately, such substitution tilings appear to be ubiquitous in the physical world [9,11,12].

Our previous work [2,3,4] and the work of others has shown that this structure exists in tasks ranging from constraint satisfaction problems to the structure of the internet [6] to robot motion planning, computer vision and image analysis [1]. We propose to continue the basic theoretical research by expanding our understanding of the fractal nature of graphs that model large, complex systems. This theory is based on the analysis of local symmetry.

Symmetry is well known to be a basic organizing principle in science. Visual symmetry is a fundamental concept widely used to categorize objects in computer vision. Linguistic symmetry is equally fundamental, e.g. in Galois theory, in which symmetry groups of equations determine whether equations are solvable or not. We focus on the scaling symmetry, in which a pattern is replicated at different scales. Finding and exploiting such patterns is our approach to conquering systems' complexity.

Our previous investigation has shown that global symmetries are not sufficient for this purpose. We need to consider local symmetries and their corresponding local invariants. The mathematics for the study of local symmetries comes from semigroup theory, much of which was developed in the context of algebraic linguistics. Much explanatory detail is omitted here, and can be found in [7,8,10]. Informally, researchers in algebraic linguistics have described a structure in language that is similar to (although weaker than) the structure of a vector space. The words in a language are like the vectors, and it is possible to find local neighborhoods of similar vectors, and then to construct coordinate systems for these neighborhoods. We can then create matrix representations for these coordinate systems, and can manipulate these representations using standard techniques of matrix algebra.

This approach provides a sound basis for analyzing the structure of large graphs and for reformulating these graphs as a hierarchy of simpler graphs. This approach differs fundamentally from existing work, which transforms a graph by partitioning it. Our approach does not produce a graph partition, but rather a sheaf-like composition of subgraphs. This is the well-known approach to relating local and global properties.

Potential Impact on Department of Defense Capabilities

This formalism is extremely general and can be used to analyze virtually any kind of system environment. We have applied this approach to constraint satisfaction problems (such as scheduling) and to robotics, both of which are important military applications.

The ability of this formalism to represent complex systems at differing levels of description would be very useful to military planners, and the fact that this formalism is computable permits the construction of system simulators for evaluation and prediction.

Potential Team and Management Plan

This work is primarily theoretical, consisting of the exploration of new structures, proving necessary results, and programming tools in Mathematica for other researchers' use. The team to carry out this work will consist of Prof. Benjamin and his students at Pace University. This team already exists and has been working for several years. The output of this project will be a number of publications describing the results together with a body of publicly available Mathematica code that implements the analytical methods.

Summary of Estimated Costs

We estimate a total project cost of $275,000 over three years. This breaks down approximately into: $80K for principal investigator support, $90K for graduate student support, $10K for supplies, software and equipment, $5K for travel to conferences, and the rest for benefits and indirect costs.

References

[1] Michael F. Barnsley, Fractal Image Compression, Notices of the American Mathematical Society, Volume 43, Number 6, pp. 657-662, June, 1996.

[2] Benjamin, D. Paul, On the Emergence of Intelligent Global Behaviors from Simple Local Actions, *International Journal of Systems Science: Emergent Properties of Complex Systems*, Vol. 31, No. 7, pp. 861-872, 2000.

[3] Benjamin, D. Paul, A Decomposition Approach to Solving Distributed Constraint Satisfaction Problems, *Proceedings of the IEEE Seventh Annual Dual-use Technologies & Applications Conference*, IEEE Computer Society Press, 1997.

[4] Benjamin, D. Paul, Formulating Patterns in Problem Solving, *Annals of Mathematics and AI*, **10**, pp.1-23, 1994.

[5] Cohn, Paul Moritz, Universal Algebra. Dordrecht, Netherlands: D.Reidel Publishing, 1981.

[6] Dill, Stephen, et al., “Self-similarity in the Web”, 27th International Conference on Very Large Databases (VLDB 2001), 2001.

[7] Howie, J.M., *An Introduction to Semigroup Theory*, Academic Press, 1976.

[8] Lallement, Gerard, *Semigroups and Combinatorial Applications*, Wiley & Sons, 1979.

[9] Peitgen, H., H. Jurgens, D. Saupe, Chaos and Fractals: New Frontiers in Science, Springer-Verlag, 1992.

[10] Petrich, Mario, *Inverse Semigroups*, (John Wiley & Sons, Inc., New York, 1984).

[11] Senechal, M., *Quasicrystals and Geometry*, Cambridge University Press, 1995.

[12] Steinhardt, Paul J., New Perspectives on Forbidden Symmetries, Quasicrystals, and Penrose Tilings, Proceedings of the National Academy of Science USA, Vol. 93, pp. 14267-14270, December 1996.

Resume of D. Paul Benjamin

School of Computer Science and Information Systems
Pace University
1 Pace Plaza
New York, NY 10038

benjamin@pace.edu
http://csis.pace.edu/~benjamin
http://csis.pace.edu/robotlab

Education:

Ph.D.	Computer Science - Courant Institute, New York University, 1985
M.S.	Computer Science - Courant Institute, New York University
B.F.A.	Music - Carnegie-Mellon University
M.S.	Mathematics - Carnegie-Mellon University
B.S.	Mathematics - Carnegie-Mellon University

Research Interests:

· Design of cognitive architectures in robotics, focusing on the relationship between perception, problem solving and language
· Application of semigroup and dynamical systems theory to knowledge representation and reformulation, problem solving and learning
· Reformulation and solution of distributed constraint satisfaction problems

Professional Experience:

Pace University, School of Computer Science and Information Systems, 1997 – Present
Professor of Computer Science
Department Chair, 2000 - 2003

Rome Laboratory, Air Force Office of Scientific Research
Visiting Research Professor, 1996 - 1998

Syracuse University, Dept. of Electrical Engineering and Computer Science
Visiting Assistant Professor, 1995 - 1997

Rome Laboratory, Griffiss Air Force Base, Rome, NY
Air Force Office of Scientific Research Summer Research Fellow, 1995

Rome Laboratory, Griffiss Air Force Base, R\ome, NY
Air Force Office of Scientific Research Summer Research Fellow, 1994

Oklahoma State University, Department of Computer Science
Assistant Professor, 1992 - 1995

St. Joseph's University (Philadelphia, Pa.),
Department of Mathematics and Computer Science
Visiting Assistant Professor, 1991 - 1992

Philips Research Laboratory, Briarcliff Manor, NY
Senior Member of Research Staff, 1984 - 1990
Project Leader in Machine Learning, 1985 - 1987

Sponsored Research:

Sponsor: DARPA, “Cognitive Support for Survivability Against Sophisticated Attacks”, contract N00178-07-C-2003, $1.5 million with BBN Technologies, Inc., December, 2006 - June, 2008.

Sponsor: Department of Energy, “Integrating Perception and Action Through Local Symmetries and Invariants”, award number ER 45903, $296,079, September 2001 - August, 2005.

Sponsor: National Science Foundation, “Integrating Formal Methods Tools into the Undergraduate Curriculum”, PI: Prof. Sotorios Skevoulis, Co-PIs: Prof. P. Benjamin and Prof. D. Anderson, Award No. DUE-0126991, $73,423, September, 2002 - August, 2005.

Sponsor: Hudson Valley Center for Emerging Technologies, “Hudson Valley Intelligent Agents Laboratory”, $10,000, 2001.

Sponsor: Air Force Office of Scientific Research / Rome Laboratory, “Pragmatic Approaches to Composition and Verification of Assured Software”, $80,500, with Dr. Shiu-Kai Chin of Syracuse University and Dr. Susan Older of Syracuse University, January, 1998 - December, 1998.

Sponsor: Air Force Office of Scientific Research, “Formal Approaches to Software Design, Planning, and the Design of Secure Systems”, $105,915, Grant No. F49620-93-C-0063, October 1, 1997 - September 30, 1998.

Sponsor: National Science Foundation, “Reformulation of System Theories in Robotics”, $49,678, SGER grant 9696060, August 1, 1995 - March 31, 1998.

Sponsor: Air Force Office of Scientific Research / Rome Laboratory, “Embedding Process and Specification Descriptions within a Categorical Framework for Refinement and Composition”, $76,000, with Dr. Shiu-Kai Chin of Syracuse University, January, 1997 - September, 1997.

Sponsor: Air Force Office of Scientific Research, “Application of Decomposition and Reformulation to Transportation Scheduling”, $103,831, Grant No. F49620-93-C-0063, October 1, 1996 - September 30, 1997.

Sponsor: Air Force Office of Scientific Research / Rome Laboratory, “Application of Process Algebra and Logic'”, $82,613, with Dr. Shiu-Kai Chin of Syracuse University, May, 1996 - December, 1996.

Sponsor: Air Force Office of Scientific Research, “Reformulating Domain Theories to Improve Their Computational Usefulness”, $24,807, SREP grant, January, 1996 - December, 1996.

Sponsor: Air Force Office of Scientific Research, “Transformational Software Design by Decomposition using the Kestrel Interactive Development System”, $24,970, SREP grant, January, 1995 - December, 1995.

Patents Held:

"System for Intrusion Detection and Vulnerability Assessment in a Computer Network using Simulation and Machine Learning", (pending) U.S. provisional patent application Serial No. 60/654,415, 2005, with Pace University.

"Semantic Encoding and Compression of Database Tables", U.S. Patent 6,691,132, February, 2004, in partnership with Adrian Walker of Reengineering, LLC.

Relevant Publications:

"A Cognitive Robotics Approach to Comprehending Human Language and Behaviors", with Deryle Lonsdale and Damian Lyons, Proceedings of the Human-Robot Interaction Conference 2007 (HRI2007), Washington, D.C., March, 2007.

"ADAPT: A Cognitive Architecture for Robotics", with Damian Lyons and Deryle Lonsdale, Proceedings of the International Conference on Cognitive Modeling, (ICCM-2004), Pittsburgh, Pa., July, 2004.

"On the Emergence of Intelligent Global Behaviors from Simple Local Actions, International Journal of Systems Science, special issue: Emergent Properties of Complex Systems, Vol. 31, No. 7, pp. 861-872, 2000.

"Connecting Perception and Action by Associating Symmetries in Vision and Language", International Journal of Artificial Intelligence Tools, Vol. 8, No. 3, 1999.

"Decomposing Robotic Representations by Identifying Local Symmetries and Invariants", Proceedings of the NSF Design and Manufacturing Grantees Conference, Mexico, 1998.

"A Decomposition Approach to Solving Distributed Constraint Satisfaction Problems", Proceedings of the IEEE Seventh Annual Dual-use Technologies & Applications Conference, IEEE Computer Society Press, 1997.

"Transforming System Formulations in Robotics for Efficient Perception and Planning", Proceedings of the IEEE International Symposia on Intelligence and Systems, Washington, D.C., IEEE Computer Society Press, 1996.

"Behavior-preserving Transformations of System Formulations", Proceedings of the AAAI Spring Symposium on Learning Dynamical Systems, Stanford University, March, 1996.

"Formulating Patterns in Problem Solving", Annals of Mathematics and AI, Vol. 10, pp.1-23, 1994.

"Reformulating Path Planning Problems by Task-preserving Abstraction", Journal of Robotics and Autonomous Systems, 9, pp. 1-9, 1992.

"An Algebraic Approach to Abstraction and Representation Change", by D. Paul Benjamin, Leo Dorst, Indur Mandhyan, and Madeleine Rosar, in Proceedings of the AAAI-90 Workshop on Automatic Generation of Approximations and Abstractions, Boston, July, 1990.