A Category of "Undirected Graphs"

A Tribute to Hartmut Ehrig

John L. Pfaltz
Dept. of Computer Science, University of Virginia
jlp@virginia.edu

Abstract

In this paper, a category of undirected graphs is introduced where the morphisms are chosen in the style of mathematical graph theory rather than as algebraic structures as is more usual in the area of graph transformation.

A representative function, ω, within this category is presented. Its inverse, ω^{-1}, is defined in terms of a graph grammar, ε.

1 The abstract category

Hartmut Ehrig was one who helped introduce the graph grammar community (GraGra) to the concept of categories [7]. In this short paper we apply some of his vision to develop a category of *undirected graphs*. A graph (V, E) is undirected if its edge set E consists of sets $\{x, y\}$, not ordered pairs. It is not hard to characterize one version of this category. It consists of *obj* = *UG*, the collection of all finite undirected graphs, together with *hom* = *all functions,* $f : G \to G'$, where $G, G' \in UG$, with composition, that is $f : G \to G', g : G' \to G''$ implies $f{\cdot}g : G \to G'' \in hom$. Let $G = (V, E)$ and $G' = (V', E')$. By $f : G \to G'$ we mean $f : 2^V \to 2^{V'}$ subject to appropriate constraints with respect to the edge sets E and E'.[1] But, without specifying these constraints this kind of category conveys little information.

More interesting is the subcategory whose functions f, g are continuous (see below). Continuity in the familiar continuous manifolds, such as **R** or **C**, is defined in terms of open sets. With discrete, or finite, graphs it can be better defined in terms of closed sets.

Let φ denote an arbitrary closure operator on an arbitrary collection, 2^V, of sets, that is for all subsets $X, Y, \in 2^V$, φ is expansive ($Y \subseteq Y.\varphi$), monotone ($X \subseteq$

[1] The codomain $2^{V'}$ of f need not be 2^V, and its edge set E' need not have the same structure as E. Therefore, elements of the codomain are denoted with a prime.

Y implies $X.\varphi \subseteq Y.\varphi$) and idempotent ($Y.\varphi.\varphi = Y.\varphi$).[2] Such closure systems $(2^V, \varphi)$ are rather well studied, since they include matroids and antimatroids [2, 3, 4, 5, 6, 9]. More importantly, we can now define what we mean by a continuous, discrete, set-valued function f. A function $f : (2^V, \varphi) \to (2^{V'}, \varphi')$ is said to be continuous [12, 13] if for all $Y \subseteq V$,

$$Y.\varphi.f \subseteq Y.f.\varphi'$$

We observe that the closure operator, φ' on V' need not be the same as φ on V. To obtain a category, we must now show that the composition of continuous functions $f \cdot g$ is continous. But, they need not be. The composition $f \cdot g$ of continuous, set-valued functions will be continuous provided f and g are also monotone [13]. To create a subcategory, we need both properties.

Suppose the functions f and g are also "closure preserving", that is the image of any set Y, closed with respect to φ will be closed with respect to φ'. In this case,

$$Y.f.\varphi' \subseteq Y.\varphi.f$$

so $Y.\varphi.f = Y.f.\varphi'$, yielding the categorical diagram

$$\begin{tikzcd}
\mathbf{Y} \arrow[r, "\varphi"] \arrow[d, "\mathbf{f}"'] & \mathbf{Y.\,\varphi} \arrow[d, "\mathbf{f}"] \\
\mathbf{Y'} \arrow[r, "\varphi'"] & \mathbf{Y.f.\varphi'}
\end{tikzcd}$$

Figure 1: A typical categorical diagram

The preceding discussion creates a subcategory of continuous set-valued functions. But as yet, it has nothing to do with undirected graphs!

As before, let obj be the set of all undirected graphs, $G = (V, E)$ where V is a set of vertices, points, or nodes and E is a symmetric binary relation on V, commonly called the edge set. Now, we consider hom to be the collection of all continuous, monotone, set-valued functions mapping subsets of the vertex (point, node) set, V of G into subsets of the vertex set V' of G'. We expect, somehow, that the closure operator on these graphs should reflect their edge structure. Let η be an operator on 2^V such that $y \in \{x\}.\eta$ and $x \in \{y\}.\eta$ if and only if $\{x, y\}$ is an edge in G. It is convenient if η, a *neighbor* operator is reflexive, that is $x \in \{x\}.\eta$. We, now, extend η to subsets $Y \subseteq V$ by $Y.\eta = \cup_{y \in Y}\{y\}.\eta$. Some texts call these "closed neighborhoods".[3] In the case of undirected graphs we prefer to use

[2]We use suffix notation to denote the application of set-valued operators and functions.

[3]This is a common terminology, but unfortunately such "closed neighborhoods" are not "closed". The intersection of closed sets must be closed, but it easy to show that this is seldom true with "closed neighborhoods".

neighborhood closure φ_η, defined below

$$Y.\varphi_\eta = \{z | \{z\}.\eta \subseteq Y.\eta\} \tag{1}$$

Because η is reflexive, φ_η is expansive; it is monotone by construction; and idempotency is not hard to prove [15, 16].

Now we have the makings of a category, UG, of undirected graphs consisting of obj = the collection of all undirected graphs, and hom = all monotone, set-valued functions $f : 2^V \to 2^{V'}$ that are continuous with respect to φ_η. It is worth observing that this development allows us to continuously enlarge graphs by a function $f : 2^V \to 2^{V'}$ in which $Ø.f = X' \subseteq V'$[4] and to contract graphs with $g : 2^{V'} \to 2^{V''}$ where $Y.g = Ø \subseteq V''$. It is convenient to employ the notation $f : G \to G'$ with the understanding that f is really defined on the power sets of V and V' and that f is continuous with respect to a closure operator φ on the edge set/relation E.

Is UG anything more than an abstract category? Are there really functions in hom?

In the next section we present two graph transformations which define $\omega \in hom$ and $\varepsilon \in hom$. Both have been implemented as algorthmic computer programs.

2 Two functions in $hom(UG)$

Let G be a graph (V, E), with a neighborhood operator η. Suppose $z \in \{y\}.\varphi_\eta$, implying by (1) that $\{z\}.\eta \subseteq \{y\}.\eta$. Since $\{z\}.\varphi_\eta = \{y\}.\varphi_\eta$, the set $\{z\}$ contributes nothing to the closure structure of G; it can be removed from G with little loss of information. We define the transformation $\omega_z : G \to G'$ by $\{z\}.\omega_z = Ø$ where ω_z is the identity map on $V - \{z\}$, $Y \subseteq V$, and $\{u', v'\} \in E'$ if and only if $\{u, v\} \in E$, $u, v \neq z$. We say z has been *subsumed* by y. It is not hard to show that ω_z is both monotone and continuous since $z \in \{y\}.\varphi_\eta$

2.1 Reduction, ω

A computer procedure, *reduce* implements ω. It repeatedly sweeps through all vertices $y \in V$, deleting any vertices $z_i \in \{y\}.\varphi_\eta$, together with all edges incident to z_i, until no such z remain in V.[5] That is, $\omega = \omega_{z_1} \cdot \omega_{z_2} \cdot \ldots \cdot \omega_{z_n}$. Since each ω_{z_i}

[4] We modify the usual definition of monotonicity to read: $X \subseteq Y$ implies $X.f \subseteq Y.f$, provided $X \neq Ø$.

[5] This procedure has been quite effective reducing large graphs $|V| \geq 1,000$, with at worst 6 iterative sweeps of V.

is monotone and continuous, ω is as well, that is $Y.\varphi_\eta.\omega \subseteq Y.\omega.\varphi_\eta{}'$. The process terminates when every singleton subset $\{y\} \subseteq V$ is closed. Such a graph is said to be *irreducible*.

It can be shown that $G' = G.\omega$ is unique (up to isomorphism) regardless of the order in which the vertices $y \in V$ are visited by ω or the order in which vertices $z \in \{y\}.\varphi_\eta$ are deleted [15, 16, 17]. So ω is a well defined function in $hom(UG)$. Because every singleton set (vertex) in G' is closed, ω must also be closure preserving, with $Y.\omega.\varphi_\eta{}' \subseteq Y.\varphi_\eta.\omega$, so the diagram of Figure 1 is applicable when $f = \omega$.

In Figure 2, the graph G of 18 vertices is reduced to $G' = G.\omega$ with 10 remaining vertices. In G, the dashed lines encircle the vertices that were subsumed

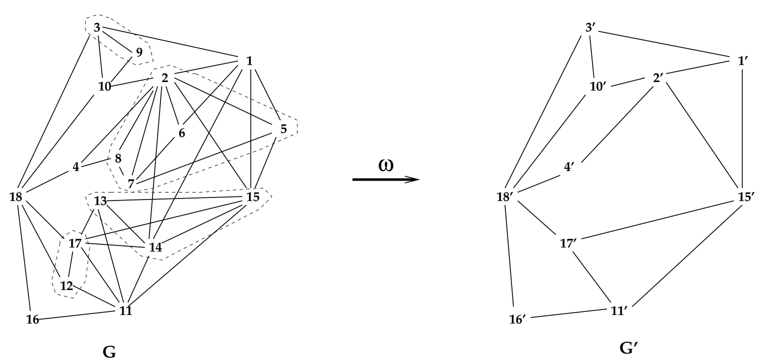

Figure 2: Reduction, ω, of a graph G

by $2', 3', 15'$ and $17'$.

Irreducible graphs, such as G', have a number of interesting properties. It is not hard to show that G' consists of a collection of chordless cycles of length ≥ 4. By a "chordless cycle" we mean a sequence of vertices $< y_1, y_2, \ldots y_n, y_1 >$, where $\{y_i, y_{i+1}\} \in E$, $1 \leq i \leq n-1$, and where $\{y_i, y_{i+k}\} \notin E$ for $k \geq 2$. Of course, we also require $\{y_n, y_1\} \in E$. It's a "pearl necklace" without cross connections. Because there can be no cross connecting edges of the form $\{y_i, y_{i+k}\}, k \geq 2$, each cycle C_α, when considered strictly as a "set" of vertices, is a member of a Sperner set [8]. That is, given a ground set V, for all cycles $C_\alpha, C_\beta \subset V$, $C_\alpha \not\subseteq C_\beta$. Besides the interesting combinatorics associated with Sperner sets, this permits various computer algorithms to process irreducible graphs solely as set systems without regard to individual edges. This reduction, $G.\omega$, of G to an irreducible graph G' has a number of other intriguing properties [17], such as the preservation of paths, of the graph "centers", but this is not relevant to this paper.

2.2 Expansion, ε

It is fairly easy to define the treatment of edges in a function, such as ω, that contracts a graph. If $Y \xrightarrow{f} Ø$, then all edges $\{y, z\}$ such that $y \in Y, z \in Y.\eta$ can be deleted. Expanding a graph, $Ø \xrightarrow{g} Y'$, presents more problems. How is Y' to be embedded in G'? One option is to employ an *expansion grammar* ε, such as explored in [14]. Expansion grammars are quite different from phrase-structured grammars in which a non-terminal symbol A is expanded with a rewrite rule of the form $A \to \sigma$ [20]. The problematic aspect of a phase-structured grammar, explored by Ehrig in [7], is how is the right side σ of the rewrite rule to be embedded in the growing, non-linear structure.

In an expansion grammar, a subset Y of a growing structure is first identified to be the neighborhood of a new element p'. That is $\{p'\}.\eta' = Y \subseteq V$ in the rewritten structure. More precisely, $\varepsilon_i : (V_i, E_i) \to (V_{i+1}, E_{i+1})$ where $V_{i+1} = V_i \cup \{p'_i\}$, $E_{i+1} = E_i \cup \{\{y_k, p'_i\}, y_k \in Y \subset V_i\}$ and $\varepsilon_i : Ø = \{p'_i\}$.

The set-valued procedure, ε can then be defined as a graph grammar with any set of specified rewrite rules, or productions. The following example of an expansion grammar is also given in [14]. Consider the rewrite rule $r1$ below,

$$r1: \quad K_n \quad \xrightarrow{\varepsilon} \quad : \; p' \qquad n \geq 1$$

which specifies that any complete subgraph, K_n, (or clique) of order n in V can serve as the neighborhood of a new point p' provided $n \geq 1$.[6] Every point in K'_n will be adjacent to p' in G'. Call the application of a rewrite rule a step, ε_i, in the process ε. It is a well defined operation in which $Ø.\varepsilon_i = \{p'\}$. The left side of the rewrite rule defines its embedding neighborhood. The right-most part defines any conditions on this neighborhood.

Application of $r1$ is illustrated in Figure 3. Each expanded neighborhood (in this case clique) has been made bold; and the expansion point, p', circled. The dashed edges indicate those links which define the clique as the neighborhood of the expansion point p'. It is not hard to see that any graph generated in this fashion must be chordal.[7]

A more relaxed version of the rewrite rule $r1$ above, will allow Y, the new neighborhood of p', to be any subset of the neighborhood of an existing vertex $y \in V_i$. Specified as a rewrite rule $r2$ it is,

$$r2: \quad Y \quad \xrightarrow{\varepsilon} \quad : \; p' \qquad \exists y \in V_i, Y \subseteq \{y\}.\eta$$

Figure 4 shows one possible application of this expansion grammar ε to the graph G' of Figure 2. Here, the rewrite rule $r2$ has been used 8 times, to create $a, b, \ldots h$. The vertex d is generated by $r2$ using the neighborhood $\{17\}.\eta = \{15, 17, 18\} =$

[6] A graph, K_n is **complete** if all n nodes are mutually connected by an edge.

[7] Because extreme points are simplicial (neighborhood is a clique), and because every chordal graph must have at least two extreme points [9, 10], every chordal graph can be so generated.

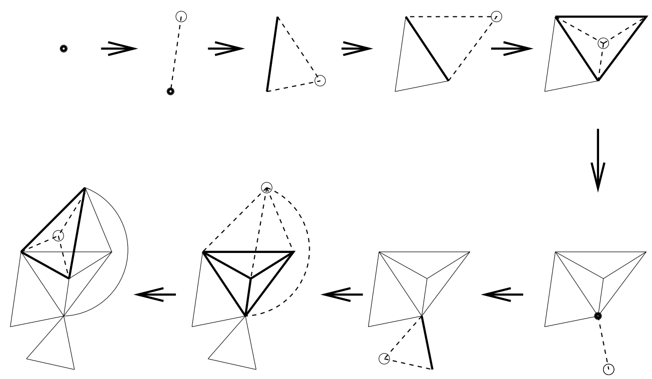

Figure 3: A sequence of neighborhood expansions generating chordal graphs

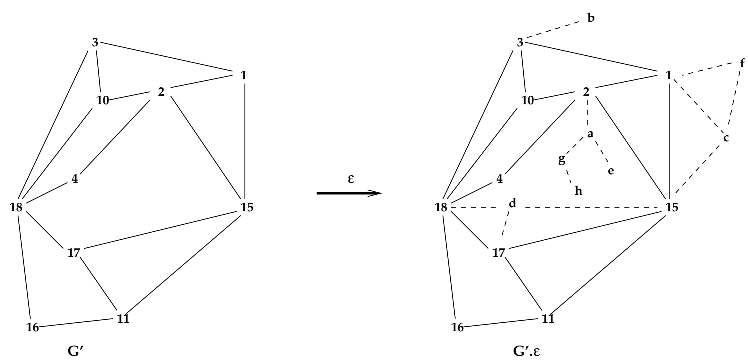

Figure 4: A member of $G'.\varepsilon$ where $G' = G.\omega$ in Figure 2.

$\{d\}.\eta'$. The new vertex c was attached to $\{1, 15\} \subset \{1, 2, 3, 15\} = \{1\}.\eta$; and f was later attached to $\{1, c\} \subset \{1\}.\eta$.

2.3 The Inverse Set, ω^{-1}

The two procedures ω and ε are intertwined. The requirement in the second rewrite rule $r2$ that $\{p'\}.\eta = Y \subseteq \{y\}.\eta$ ensures that if ω is applied to $G'.\varepsilon$, p' will at some iteration be subsumed by y. Thus, if G' is irreducible, $G'.\varepsilon.\omega = G'$ This characteristic is evident in Figure 4 where b will be subsumed by 3, *etc*. It is also true for the graph $G'.\varepsilon$ of Figure 5 as well. Consequently, ω is a right-inverse of ε over the subspace of irreducible undirected graphs. The inverse of ω, that is $G.\omega.\omega^{-1}$ is the collection of all undirected graphs $\{G_k\}$ such that $G_k.\omega = G' = G.\omega$. Each invocation of the non-deterministic procedure ε is single-valued; but ε is not a function. The execution of ε will yield a graph, $G_k \in G.\omega.\omega^{-1}$.

In the rewrite rule $r2$ the choice of $y \in V_i$ and the choice of $Y \subseteq \{y_i\}.\eta$ are completely arbitrary. Given different choices for y and Y yields Figure 5 which seems to be a far more interesting graph. Both Figures 4 and 5 were generated by

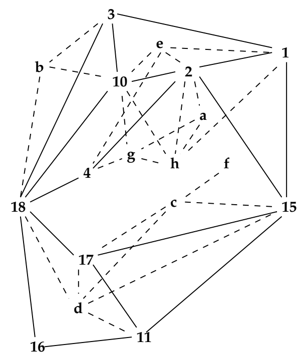

Figure 5: Another graph $G'.\varepsilon$ in $G'.\omega^{-1}$.

a computer version of ε using a random number generator.

This is not the only category of undirected graphs, but it is a promising one [18]. Unfortunately, undirected graphs, and mappings between such graphs, have little of the regular structure seen in the different abstract algebras that gave rise to the categorical approach of [1, 11, 19], or that of [2] which was applied to general closure operators. Yet, the rudiments are there, as this short treatise shows. In the early 70's, Hartmut Ehrig urged us to view graph grammars and graph manipulation through a categorical lens. He was ahead of his time.

References

[1] Michael Arbib and Ernest Manes. *Arrows,Structures, and Functors: The Categorical Imperative*. Academic Press, New York, 1975.

[2] Gabriele Castellini. *Categorical Closure Operators*. Birkhauser, Boston, 2003.

[3] Vašek Chvátal. Antimatroids, Betweenness, Convexity . In William Clark, László, and Jens Vygen, editors, *Research Trends in Combinatorial Optimization*, pages 57–64. Springer, 2009.

[4] Brenda L. Dietrich. Matroids and Antimatroids — A Survey. *Discrete Mathematics*, 78:223–237, 1989.

[5] Paul H. Edelman. Abstract Convexity and Meet-distributive Lattices. In *Combinatorics and Ordered Sets*, pages 127–150, Arcata, CA, 1986.

[6] Paul H. Edelman and Robert E. Jamison. The Theory of Convex Geometries. *Geometriae Dedicata*, 19(3):247–270, Dec. 1985.

[7] H. Ehrig, M. Pfender, and H. J. Schneider. Graph Grammars: An Algebraic Approach. *IEEE Conf. SWAT*, 1973.

[8] Konrad Engle. Sperner Theory . In Michiel Hazewinkle, editor, *Encyclopedia of Mathematics*. Springer, Berln, 2001.

[9] Martin Farber and Robert E. Jamison. Convexity in Graphs and Hypergraphs. *SIAM J. Algebra and Discrete Methods*, 7(3):433–444, July 1986.

[10] Martin Charles Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Academic Press, New York, 1980.

[11] Saunders MacLane. *Categories for the Working Mathematician*. Springer Verlag, New York, second edition, 1998.

[12] Oystein Ore. Mappings of Closure Relations. *Annals of Math.*, 47(1):56–72, Jan. 1946.

[13] John Pfaltz and Josef Šlapal. Transformations of discrete closure systems. *Acta Math. Hungar.*, 138(4):386–405, 2013.

[14] John L. Pfaltz. Neighborhood Expansion Grammars. In *Intern'l Workshop on Theory and Application of Graph Transformation, TAGT'98*, Lecture Notes in Computer Science, #1764, pages 30–44, Paderborn, Germany, Nov. 1998. Springer-Verlag.

[15] John L. Pfaltz. Finding the Mule in the Network. In Reda Alhajj and Bob Werner, editors, *Intern. Conf on Advances in Social Network Analysis and Mining, ASONAM 2012*, pages 667–672, Istanbul, Turkey, August 2012.

[16] John L. Pfaltz. Mathematical Continuity in Dynamic Social Networks. *Social Network Analysis and Mining (SNAM)*, 3(4):863–872, Dec. 2013.

[17] John L. Pfaltz. The Irreducible Spine(s) of Discrete Networks. In Xuemin Li, Yannis Manolopoulos, Divesh Srivastava, and Guangyan Huang, editors, *Web Information Systems Engineering - WISE 2013*, volume LNCS # 6984, Part 2, pages 104–117), Nanjing, PRC, Oct. 2013.

[18] John L. Pfaltz. Computational Processes that Appear to Model Human Memory. In Daniel Figueiredo, Carlos Martín-Vide, Diogo Gratas, and Miguel A.Vega-Rodríguez, editors, *Proc. 4th Intern. Conf. on Algorithms for Computational Biology (AlCoB 2017)*, pages 85–99, Aveiro Portugal, 2017.

[19] Benjamin C. Pierce. *Basic Category Theory for Computer Scientists*. MIT Press, Cambridge, MA, 1991.

[20] Gregorz Rozenberg, editor. *The Handbook of Graph Grammars*. World Scientific, 1997.