

## TWO NETWORK TRANSFORMATIONS

JOHN L. PFALTZ

*Abstract.* In this paper, we study relational networks. They may be as large as social networks or as small as neural networks. We employ the concepts of closure and closure operators to describe their structures, and introduce the idea of functional transformation to model their dynamic qualities.

One transformation,  $\omega$ , reduces a complex network to a much simpler form, yet preserves important properties such as path connectivity and centrality measures. The other transformation,  $\varepsilon$ , expands a network by using grammar-like productions. Both are continuous (with respect to closure) and we show that  $\varepsilon$  is effectively  $\omega^{-1}$  in that  $\omega \cdot \varepsilon \cdot \omega = \omega$ .

It is thought that  $\omega$  may model human memory consolidation and that  $\varepsilon$  may model memory reconstruction.

### 1. INTRODUCTION

A *network*,  $\mathcal{N} = (N, \eta)$ , is a finite collection  $N$  of elements, frequently called *nodes* together with a reflexive relation  $\eta$  on  $N$ . While  $\eta$  is a relation, it is often convenient to treat it as an operator, that is, given  $y \in N$ ,  $\{y\} \cdot \eta = \{z \mid (y, z) \in \eta\}$  which is called the *neighborhood* of  $y$ . (Operators, and later transformations, which are functions that map sets into sets are denoted by Greek characters. We denote their application to specific set arguments using a suffix notation to indicate that they are not ordinary functions.) An operator,  $\eta$ , is said to be *extensible* if

$$Y \cdot \eta = \bigcup_{y \in Y} \{y\} \cdot \eta$$

A relational operator,  $\eta$ , can be arbitrarily complex, but if it is extensible, then it can be visualized as a kind of “adjacency” operator denoting the adjacent neighbors of elements in  $\mathcal{N}$  and we can depict these networks as simple graphs, such as that of Figure 1.<sup>1</sup> We assume extensibility for all neighborhood operators,  $\eta$ , in this paper. Proposition 1.2, which is central to our development, requires it. However, to achieve maximal generality, we will not require the relation  $\eta$  to be symmetric until Section 3.1. Consequently, some of the connections can be directional, denoted by

---

*MSC (2010):* primary 54A05.

*Keywords:* closure operator, continuous map, network centrality, chordless cycle.

<sup>1</sup>With the most general interpretation of  $\eta$ , a network is essentially a “hypergraph”. For instance, one might have  $X \cdot \eta = X \cup \{a\}$ ,  $Y \cdot \eta = Y \cup \{b\}$ , but  $(X \cup Y) \cdot \eta = X \cup Y \cup \{abc\}$ .

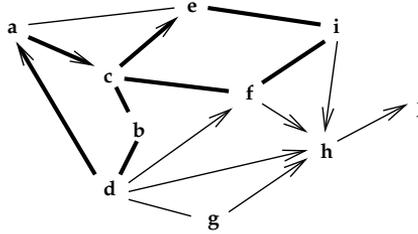


Figure 1. A small network illustrating  $\eta$  on 10 elements.

the arrow heads. Others are symmetric, or bi-directional, to minimize clutter we omit any arrow heads on them. In Figure 1,  $\{d\}.\eta = \{abdfhg\}$ ,  $\{fi\}.\eta = \{cefih\}$ .<sup>2</sup>

It is doubtful if any real networks of interest actually look like Figure 1. We know that real networks, be they neural networks or social networks of individuals, are much larger and much more complex. Moreover, they are subject to constant dynamic change. Nevertheless, graphical depictions such as this figure are useful in visualizing basic concepts.

Since  $\eta$  is reflexive,  $Y \subseteq Y.\eta$ . In some textbooks [3, 15], this is called a *closed* neighborhood of  $Y$ ; and  $Y.\eta - Y$  is called an *open* neighborhood.<sup>3</sup> These are not terms we will use in this paper. We call  $Y.\eta - Y$  the *halo* around  $Y$ . It will be important in Section 3.

A *path* in  $\mathcal{N}$  of length  $n$  is a sequence of elements  $\langle y_0, y_1, \dots, y_n \rangle$ ,  $y_i \in N$  such that  $y_{i+1} \in \{y_i\}.\eta$ , for  $0 \leq i \leq n-1$ . If  $y_0 \neq y_1$  and  $y_0 = y_n$ , the path sequence is said to be a *cycle*,  $C$ , of length  $n$ .

A pair of nodes  $y_i, y_k$  in a cycle,  $C$ , is said to constitute a *chord* if  $y_k \in \{y_i\}.\eta$  and  $|k-i| > 1$  and  $i, k \neq 0, n$ . A cycle  $C$  of length  $\geq 4$  is said to be *chordless* if there are no chords in  $C$ . In Figure 1, the sequence of nodes  $C_1 = \langle a, c, f, i, e, a \rangle$  is a cycle.  $\{c\}.\eta = \{ce\}$  is a chord in that cycle.  $C_2 = \langle c, e, i, f, c \rangle$  is a chordless cycle. So is  $C_3 = \langle a, c, b, d, a \rangle$ . Both have be slightly emboldened in Figure 1 The order of the bi-directional links  $(b, c)$  and  $(d, b)$  is reversed in  $C_4 = \langle f, d, b, c, f \rangle$ .

If a collection of subsets  $\{C_i\}$  of elements in  $N$  has the property that no set  $C_i$  is completely contained in another, that is if  $C_i \not\subseteq C_k$  for all  $i \neq k$ , and  $N = \cup_i C_i$ , then  $N$  and the subsets  $C_i$  are said to constitute a Sperner system. Such systems are so called after Emanuel Sperner who first described them [11]. With a little thought, we see that if  $C_i$  and  $C_k$  are chordless cycles, then  $C_i$  cannot contain  $C_k$ , or *vice versa* because the larger cycle must contain a chord. Consider  $C_1 = \{a, c, e, f, i\}$  and  $C_2 = \{c, e, f, i\}$ , now just as sets of elements.  $C_2$  is a subset of  $C_1$  so  $C_1$  cannot be chordless, and indeed it is not, the edge  $(c, e)$  is a chord. Consequently, a system of chordless cycles  $\{C_i\}$  constitutes a Sperner system,  $N$ , with each cycle being a unique set in this system.

Let  $S$  be any set of elements, possibly the set  $N$  of nodes in a network. An operator  $\varphi$  is said to be a *closure* operator if for all subsets  $X, Y$ ,

<sup>2</sup>One might expect to denote the neighborhood of a single node, such as  $d$ , in Figure 1, by  $d.\eta$ . But,  $\eta$  is a mapping on the powerset,  $2^N$ , of  $N$ ; that is  $2^N \xrightarrow{\eta} 2^{2^N}$ . Therefore,  $\{d\}$  denotes a particular singleton set of  $2^N$ . Eliding the braces has caused confusion in earlier papers.

<sup>3</sup>Many authors prefer to use  $Y.\eta \setminus Y$  to denote set difference.

- (C1)  $Y \subseteq Y.\varphi$ , (expansive),  
 (C2)  $X \subseteq Y$  implies  $X.\varphi \subseteq Y.\varphi$ , (monotone),  
 (C3)  $Y.\varphi.\varphi = Y.\varphi$ , (idempotent).

Such a system  $(S, \varphi)$  is called a *closure system*. A set  $Y$  is said to be *closed* if  $Y.\varphi = Y$ . Closure operators have been widely studied, *c.f.* [5,6,10,12,16,20–22,36]. The concept of closure is fundamental to our way of approaching the properties of discrete networks. Since  $\varphi$  is expansive (C1), in any network  $\mathcal{N}$ , the set of all nodes  $N$  must itself be closed.

There are alternative definitions of closure systems and closure operators. For example, let  $\mathcal{C}$  denote an arbitrary collection of sets  $\{Y_i\}$ . This collection,  $\mathcal{C}$ , is a closure system if

- (C4)  $(\forall i) Y_i \in \mathcal{C}$  implies  $\bigcap_i Y_i \in \mathcal{C}$  (closure under all intersections)

Proposition 1.1 is well known, but it is worth repeating here.

**Proposition 1.1.** *Let  $(S, \varphi)$  be a finite system whose operator  $\varphi$  satisfies the axioms (C1), (C2), and (C3), then (C4) is true for all subsets  $Y \subseteq S$  such that  $Y.\varphi = Y$ , (a closed subset). Conversely, if  $\mathcal{C}$  is a collection of sets such that (C4) is true, then an operator  $\varphi$  defined  $Y.\varphi = \bigcap_{Z \in \mathcal{C}} \{Y \subseteq Z\}$  will satisfy (C1), (C2) and (C3).*

Given the importance of the neighborhood concept  $\eta$ , we define the *neighborhood closure operator*<sup>4</sup>  $\varphi_\eta$  to be

$$Y.\varphi_\eta = \{x | \{x\}.\eta \subseteq Y.\eta\}. \quad (1.1)$$

Effectively,  $Y.\varphi_\eta$  consists of all elements of  $Y$  itself, plus those adjacent elements  $x$  with no larger a relational horizon than that of  $Y$ . In a social network, this is indeed a “closed” group! While there can be many closure operators, such as transitive closure, on a set of elements, we are solely concerned with  $\varphi_\eta$  in this paper. Readily, for all  $Y$ ,

$$Y \subseteq Y.\varphi_\eta \subseteq Y.\eta. \quad (1.2)$$

**Proposition 1.2.** *If  $\eta$  is an extensible operator, then  $\varphi_\eta$  is a closure operator.*

*Proof.* (C1) Readily,  $Y \subseteq Y.\varphi_\eta$  by definition.

(C2) Let  $X \subseteq Y$  and let  $z \in X.\varphi_\eta$ . By (1.1)  $\{z\}.\eta \subseteq X.\eta \subseteq Y.\eta$ , hence,  $z \in Y.\varphi_\eta$ .

(C3) Let  $z \in Y.\varphi_\eta.\varphi_\eta$ . Then,  $\{z\}.\eta \subseteq Y.\varphi_\eta.\eta = \bigcup_{x \in Y.\varphi_\eta} \{x\}.\eta \subseteq Y.\eta$ , hence,  $z \in Y.\varphi_\eta$ .  $\square$

To see why extensibility is required to establish (C3), let  $N = \{a, b, c\}$  and let  $\{a\}.\eta = \{ab\}$ ,  $\{b\}.\eta = \{ab\}$ ,  $\{c\}.\eta = \{c\}$ , but  $\{ab\}.\eta = \{abc\}$ .  $\eta$  is not extensible and  $\{a\}.\varphi_\eta.\varphi_\eta = \{ab\}.\varphi_\eta = \{abc\} \neq \{ab\} = \{a\}.\varphi_\eta$ , so  $\varphi_\eta$  as defined by (1.1) is not idempotent. (We often elide the commas when denoting a set by enumeration.)

**Proposition 1.3.** *If  $X.\eta \subseteq Y.\eta$  then  $X.\varphi_\eta \subseteq Y.\varphi_\eta$ .*

*Proof.* Let  $X.\eta \subseteq Y.\eta$ . Let  $z \in X.\varphi_\eta$  then  $\{z\}.\eta \subseteq X.\eta \subseteq Y.\eta$ . Hence  $z \in Y.\varphi_\eta$ .  $\square$

---

<sup>4</sup>Neighborhood closure has been called “experiential” and “relational” closure in other papers [27, 28].

One might expect that every point in a discrete network must be closed, *e.g.*  $\{x\}.\varphi_\eta = \{x\}$ . But, this need not be true, as shown in Figure 1. Here,  $\{d\}.\eta = \{abdfgh\}$  and  $\{g\}.\eta = \{dgh\}$ , so  $\{d\}.\varphi_\eta = \{dg\}$ . Similarly,  $j \in \{h\}.\varphi_\eta$ .

Equation (1.2) suggests an effective computer algorithm to calculate the closure  $Y.\varphi_\eta$  of any set  $Y$ . Initially, let  $Y.\varphi_\eta = Y$ ; then since  $Y.\varphi_\eta \subseteq Y.\eta$  examine only the points  $z$  in the halo,  $Y.\eta - Y$ , of  $Y$ . If  $\{z\}.\eta \subseteq Y.\eta$ , add  $z$  to  $Y.\varphi_\eta$ .

We say  $\eta_1 \subseteq \eta_2$  if for all  $Y, Y.\eta_1 \subseteq Y.\eta_2$ . (If one thinks in terms of simple graphs, then  $E_1 \subseteq E_2$ .) Given Proposition 1.3, one would expect that we could prove a proposition of the form  $\eta_1 \subseteq \eta_2$  implies  $\varphi_{\eta_1} \subseteq \varphi_{\eta_2}$ . But neighborhood closure is more complex than that. Consider two neighborhood operators  $\eta_1, \eta_2$  defined on a set  $N = \{a, b, c\}$  as shown in Table 1.

**Table 1.** Two neighborhood operators  $\eta_1$  and  $\eta_2$  on  $N$  and their closure operators  $\varphi_{\eta_1}, \varphi_{\eta_2}$ .

$N$	$\eta_1$	$\eta_2$	$\varphi_{\eta_1}$	$\varphi_{\eta_2}$
$\{a\}$	$\{ab\} \subseteq$	$\{abc\}$	$\{ab\} \subseteq$	$\{abc\}$
$\{b\}$	$\{ab\} =$	$\{ab\}$	$\{ab\} \supseteq$	$\{b\}$
$\{c\}$	$\{c\} \subseteq$	$\{ac\}$	$\{c\} =$	$\{c\}$

Readily,  $\eta_1 \subseteq \eta_2$  in that  $\{x\}.\eta_1 \subseteq \{x\}.\eta_2$  for all  $x \in N$ , but no consistent pattern exists between  $\varphi_{\eta_1}$  and  $\varphi_{\eta_2}$ . ( $a \notin \{b\}.\varphi_{\eta_2}$  because although  $a \in \eta_2$ ,  $\{a\}.\eta_2 = \{abc\} \not\subseteq \{b\}.\eta_2$ .)

We can, however, prove the weaker assertion,

**Proposition 1.4.** *Let  $\eta_1 \subseteq \eta_2$ . For any  $Y \subseteq S$ , if  $Y.\eta_2 = Y.\eta_1$  then  $Y.\varphi_{\eta_2} \subseteq Y.\varphi_{\eta_1}$ .*

*Proof.* Let  $Y$  be such that  $Y.\eta_2 = Y.\eta_1$ . Then,

$$\begin{aligned}
 Y.\varphi_{\eta_2} &= \cup_{z \in Y.\eta_2} \{\{z\}.\eta_2 \subseteq Y.\eta_2\} \\
 &\subseteq \cup_{z \in Y.\eta_2} \{\{z\}.\eta_1 \subseteq Y.\eta_2\} && \text{since } \{z\}.\eta_1 \subseteq \{z\}.\eta_2 \\
 &= \cup_{z \in Y.\eta_1} \{\{z\}.\eta_1 \subseteq Y.\eta_1\} && \text{since } Y.\eta_1 = Y.\eta_2 \\
 &= Y.\varphi_{\eta_1}.
 \end{aligned}$$

□

The second line of Table 1 illustrates this result.

## 2. NETWORK TRANSFORMATIONS

By a *transformation*  $\mathcal{N} \xrightarrow{\tau} \mathcal{N}'$  of a network  $\mathcal{N} = (N, \eta)$ , we mean a mapping between the powersets of nodes  $N, N'$  of two networks  $\mathcal{N}, \mathcal{N}'$  a function  $\tau$  mapping the subsets of  $N$  onto subsets of  $N'$ , that is  $2^N \xrightarrow{\tau} 2^{N'}$ . The concepts of *operator* and *transformation* are indistinguishable, except that the domain and codomain of an operator are the same set, whereas with a transformation they may be distinct. The important property of transformations is that they map sets onto sets, therefore,  $Y.\tau = \emptyset$  is meaningful and provides a way of functionally deleting

elements from a discrete domain  $N$ . Similarly  $\emptyset.\tau = X'$  is a well-defined function introducing new elements to the codomain  $N'$ .

A transformation,  $\tau$ , is said to be *monotone* if for all non-empty  $X, Y$ ,<sup>5</sup>  $X \subseteq Y$  implies  $X.\tau \subseteq Y.\tau$ . All operators and transformations considered in this paper will be monotone. Without monotonicity, very little can be proved. Observe that an operator corresponding to a binary relation is always monotone.

We will say that a transformation  $\tau : (N, \eta) \rightarrow (N', \eta')$  is *continuous* with respect to closure operators  $\varphi, \varphi'$  on  $\mathcal{N}$  and  $\mathcal{N}'$ , respectively, if for all subsets  $Y \subseteq N$ ,

$$Y.\varphi.\tau \subseteq Y.\tau.\varphi'. \quad (2.1)$$

This definition can also be found in [21, 26, 31]. Continuous transformations preserve the closure structure of the network.

The first five propositions are true for all closure operators,  $\varphi$ . In Section 2.1 we specifically address  $\varphi_\eta$ . Here,  $(N, \varphi)$  denotes any set  $N$  of elements with any closure operator  $\varphi$ . The following proposition is proved in [31].

**Proposition 2.1.** *Given networks  $\mathcal{N} = (N, \varphi), \mathcal{N}' = (N', \varphi'), \mathcal{N}'' = (N'', \varphi'')$ . Let  $\mathcal{N} \xrightarrow{\sigma} \mathcal{N}'$ ,  $\mathcal{N}' \xrightarrow{\tau} \mathcal{N}''$  be monotone transformations. If both  $\sigma$  and  $\tau$  are continuous, then so is  $\mathcal{N} \xrightarrow{\sigma.\tau} \mathcal{N}''$ .*

Continuous transformations of discrete spaces exhibit many of the properties of continuous real functions with which we are more familiar. For example, let  $f$  be a function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ; if  $f$  is continuous and  $X'$  is open/closed in  $\mathbb{R}$ , then  $f^{-1}(X')$  is open/closed in  $\mathbb{R}$ .

**Proposition 2.2.** *Let  $(N, \varphi) \xrightarrow{\tau} (N', \varphi')$  be monotone and continuous with  $Y \subseteq N$ . If  $Y' = Y.\tau$  is closed then  $Y.\varphi.\tau = Y'$ .*

*Proof.* Let  $Y.\tau$  be closed in  $N'$ . Because  $\tau$  is continuous  $Y.\varphi.\tau \subseteq Y.\tau.\varphi' = Y.\tau$ , since  $Y.\tau$  is closed. By monotonicity,  $Y.\tau \subseteq Y.\varphi.\tau$ , so  $Y.\varphi.\tau = Y.\tau$ .  $\square$

One can also consider *closure preserving* transformations which map closed sets in  $N$  onto closed sets in  $N'$ .<sup>6</sup> It is apparent that the composition of closure preserving transformations is closure preserving. The following two propositions can be found in [31].

**Proposition 2.3.** *A transformation  $(N, \varphi) \xrightarrow{\tau} (N', \varphi')$  is closure preserving if and only if for all  $Y \subseteq N$ ,  $Y.\tau.\varphi' \subseteq Y.\varphi.\tau$ .*

Consequently,

**Proposition 2.4.** *A monotone transformation  $(N, \varphi) \xrightarrow{\tau} (N', \varphi')$  is closure preserving and continuous if and only if, for all  $Y \subseteq N$ ,  $Y.\varphi.\tau = Y.\tau.\varphi'$ .*

<sup>5</sup>This condition is important because although we may want  $\emptyset.\tau = X'$  to be non-empty, we generally do not want to have  $X' \subseteq Y' = Y.\tau$  just because for all  $Y$ ,  $\emptyset \subseteq Y$ .

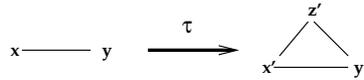
<sup>6</sup>The topological term ‘‘closed’’ is traditional for structure preserving maps, whether expressed in terms of open sets or closed sets. But, it is most unfortunate in this context, where the multiple meanings can lead to confusion.

A common way of defining a network transformation  $(N, \eta) \xrightarrow{\tau} (N', \eta')$  is to first define  $\{y\}.\tau$  for all singleton sets in  $N$  and then extend this to all  $Y \subseteq N$  by  $Y.\tau = \bigcup_{y \in Y} \{y\}.\tau$ . That is, we can consider extensible transformations, just as we have extensible operators. We call an extensible transformation,  $\tau$ , *complete* if  $N.\tau = N'$ . Any extensible transformation is, by construction, monotone.

**Proposition 2.5.** *If an extensible transformation,  $(N, \eta) \xrightarrow{\tau} (N', \eta')$  is complete, then for all  $y' \in Y' = Y.\tau$ , there exists  $y \in Y$  such that  $y' \in \{y\}.\tau$ .*

*Proof.* Let  $y' \in Y'$ . By construction  $Y' = \bigcup_{y \in Y} \{y\}.\tau$ , hence,  $y' \in \{y\}.\tau$  for some  $y \in Y$ .  $\square$

Note that this is quite different from asserting a true inverse existence, that for all  $y' \in Y'$ , there exists some  $y \in Y$  such that  $\{y\}.\tau = \{y'\}$ . To get some sense of the import of this “weak inverse existence” proposition, consider the simple transformation  $\tau$  of Figure 2.



**Figure 2.** A transformation  $\tau$  with multiple definitions.

If we define  $\tau$  on  $N$  by  $\{x\}.\tau = \{x'\}$  and  $\{y\}.\tau = \{y'\}$ , then by extension  $\{xy\}.\tau = \{x'y'\}$  and  $\{z'\}$  has no pre-image; so  $N.\tau \neq N'$ . However, if we let  $\{x\}.\tau = \{x'z'\}$ ,  $\{y\}.\tau = \{y'z'\}$ , then  $\{xy\}.\tau = \{x'y'z'\}$ . Now  $N.\tau = N'$ , so  $\tau$  is a complete transformation. Proposition 2.5 is clearly satisfied. A complete transformation  $\tau$  need not be surjective. In the example above,  $\{x', y'\} \in 2^{N'}$  has no preimage in  $2^N$ .

## 2.1. Neighborhood transformations

The preceding assertions are true for all transformations,  $\tau$ . Now we focus specifically on neighborhood transformations. It is the neighborhood,  $Y.\eta$ , which is central. In Section 1, extensibility was necessary to ensure that  $\varphi_\eta$  was a closure operator; it is equally important here.

**Proposition 2.6.** *If  $\eta$  is extensible, the following three statements are equivalent.*

- (a)  $X \subseteq Y.\varphi_\eta$ ,
- (b)  $X.\eta \subseteq Y.\eta$ ,
- (c)  $X.\eta - Y \subseteq Y.\eta - Y$ .

*Proof.* By extension, the first equivalence is simply a restatement of the definition of neighborhood closure. Proof of the second equivalence is straightforward.  $\square$

Thus, besides  $Y$  itself, the haloes of all elements  $x$  comprising  $Y.\varphi_\eta$  must be contained in the halo of  $Y$ .

By (2.1), continuity is defined in terms of subsets  $Y$ . The following proposition, similar to one in [25], shows that we need only consider singleton subsets.

**Proposition 2.7.** *Let  $(N, \eta) \xrightarrow{\tau} (N', \eta')$  be extensible. If  $\tau$  is not continuous, there exists  $Y \subseteq N$ , and  $z \in Y.\eta$  such that either*

$$(1) \{z'\} \notin Y.\tau.\eta'$$

or

$$(2) \{z\}.\eta \subseteq Y.\eta \text{ and } \{z'\}.\eta' \not\subseteq Y.\tau.\eta'$$

where  $\{z'\} = \{z\}.\tau$ .

*Proof.* Since  $\tau$  is not continuous, there exists  $Y$  such that  $Y.\varphi_\eta.\tau \not\subseteq Y.\tau.\varphi_{\eta'}$ . Thus,  $\exists z' \in Y.\varphi_\eta.\tau, z' \notin Y.\tau.\varphi_{\eta'}$ . By, Prop. 2.5,  $\exists z \in Y.\varphi_\eta$  such that  $z' \in \{z\}.\tau$ .  $z \notin Y$  else  $z' \in Y.\tau$ . Consequently,  $z \in Y.\eta$  and  $\{z\}.\eta \subseteq Y.\eta$ . Now, since  $z' \notin Y.\tau.\varphi_{\eta'}$  we know that either (1)  $z' \notin Y.\tau.\eta'$  or (2)  $\{z'\}.\eta' \not\subseteq Y.\tau.\eta'$ .  $\square$

This proposition establishes that if  $\tau$  is discontinuous anywhere, then it will be discontinuous at, or near, a single element. One need not consider all subsets of  $2^N$ . Just as is the case with classical function theory, discontinuity, and thus continuity, is a local phenomena. Moreover, Proposition 2.7 provides conditions (1) and (2) which are necessary, but not sufficient to demonstrate discontinuity. If, for a point  $z \in N$ , neither condition (1) nor (2) holds, we say that  $\tau$  is *continuous at  $y$* . If either condition holds, other criteria must be used, such as Proposition 2.8 below.

We have said that a transformation  $N \xrightarrow{\tau} N'$  is monotone if for all non-empty  $X, Y, X \subseteq Y$  implies  $X.\tau \subseteq Y.\tau$ . Let  $(N, \eta) \xrightarrow{\tau} (N', \eta')$  be a transformation between two networks. The transformation  $\tau$  is said to be *neighborhood monotone* if  $X.\eta \subseteq Y.\eta$  implies  $X.\tau.\eta' \subseteq Y.\tau.\eta'$ . A transformation that is monotone need not be neighborhood monotone, and conversely.

**Proposition 2.8.** *Let  $(N, \eta) \xrightarrow{\tau} (N', \eta')$  be extensible, then  $\tau$  is continuous if and only if  $\tau$  is neighborhood monotone.*

*Proof.* Let  $\tau$  be continuous and let  $X.\eta \subseteq Y.\eta$ . By Proposition 1.3,  $X \subseteq X.\varphi_\eta \subseteq Y.\varphi_\eta$ . Thus,  $X.\tau \subseteq Y.\varphi_\eta.\tau \subseteq Y.\tau.\varphi_{\eta'}$  by continuity. So  $X.\tau.\eta' \subseteq Y.\tau.\eta'$ .

Conversely, let  $\tau$  be neighborhood extensible. Since  $x.\eta \subseteq Y.\eta \Rightarrow x.\tau.\eta' \subseteq Y.\tau.\eta'$  we have

$$\begin{aligned} u \in Y.\varphi_\eta.\tau &= \{x | \{x\}.\eta \subseteq Y.\eta\}.\tau \\ &\subseteq \{x | \{x\}.\tau.\eta' \subseteq Y.\tau.\eta'\}.\tau, \end{aligned}$$

thus there exists  $x$  such that  $u \in \{x\}.\tau$  and  $\{x\}.\tau.\eta' \subseteq Y.\tau.\eta'$  thus  $\{u\}.\eta' \subseteq \{x\}.\tau.\eta' \subseteq Y.\tau.\eta'$ . Therefore,  $u \in Y.\tau.\varphi_{\eta'}$ .  $\square$

### 3. THE OMEGA TRANSFORMATION, $\omega$

Because the structure of large networks can be so difficult to comprehend, it is natural to seek techniques for reducing their size, while still preserving certain essential properties [1, 14], often by selective sampling [19]. Our approach is somewhat different. We view “structure” through the lens of neighborhood closure, which we then use to find the unique irreducible sub-network  $\mathcal{T} \subseteq \mathcal{N}$ .

A graph, or network, is said to be *irreducible* if every singleton subset  $\{y\}$  is closed. A node  $z$  is *subsumed* by a node  $y$  if  $z \in \{y\}.\varphi_\eta$ , or equivalently, because we are concerned with neighborhood closure, if  $\{z\}.\eta \subseteq \{y\}.\eta$ . Since in this case,

$z$  contributes very little to our understanding of the closure structure of  $\mathcal{N}$ , its removal will result in little loss of information. If a node  $z$  is subsumed by  $y$  ( $z \in \{y\}.\varphi_\eta$ ), we say that  $z$  *belongs* to  $y$ , and denote the set of all nodes belonging to  $y$  by  $\{y\}.\beta$ . Readily,  $y \in \{y\}.\beta$ .

The  $\omega$  transformation, which has been previously described in [24, 25, 29], maps any symmetric network  $\mathcal{N} = (N, \eta)$  onto an irreducible subgraph. It is iteratively defined as  $\omega = \omega_1.\omega_2.\dots.\omega_k$  where,  $\omega_i : \mathcal{N}_i \rightarrow \mathcal{N}_{i+1}$ . Here  $\mathcal{N}_{i+1} = (N_{i+1}, \eta_{i+1})$  is a subgraph of  $(N_i, \eta_i)$  consisting of chosen nodes which are not mutually subsumed, but any other node is subsumed by some of them. The binary relation  $\eta_{i+1}$  is the restriction of  $\eta_i$  on this subset. Thus, the transformation  $\omega_i$  is determined as the mapping  $2_i^N \rightarrow 2_{i+1}^N$  given by the full preimage of the embedding  $N_{i+1} \rightarrow N_i$ , i.e.  $Y.\omega_i = \{x \in Y \mid x \in N_{i+1}\} = Y \cap N_{i+1}$ . The composition terminates when  $\omega_i = \omega_{i+1}$ . Since this process is dependent on the choice of elements, single steps are not uniquely determined. However, as will be further proved, the final transformation  $\omega$  is independent of this choice.

This  $\omega$  transformation is easily programmed as a *while* loop shown in Figure 3. Operationally, it is easiest to search the halo  $\{y\}.\eta - \{y\}$  of each node  $y$ , and test whether  $\{z\}.\eta \subseteq \{y\}.\eta$  as shown in the code fragment. It requires initial order on elements which determine the order of subsumption and, therefore, the choice of elements. This iterative reduction process we denote by  $\omega$ . Since this

```

while there exist subsumable nodes
{
  for_each y in N
  {
    get {y}.nbhd
    for_each z in {y}.nbhd - {y}
    {
      if ({z}.nbhd contained_in {y}.nbhd
        {          // z is subsumed by y
          remove z from network
          union {z}.beta with {y}.beta
        }
      }
    }
  }
}

```

**Figure 3.** Pseudo code for the inner loop of  $\omega$ .

process must loop over each node in the network, and since one can construct pathological examples such as Figure 4 in which only one node is subsumed with each iteration, the theoretical order of complexity is  $O(\omega) = n^2$ , if nodes are encountered in subscript order. So worst case behavior is  $O(n^2)$ . However, in practice it is much better. On a variety of networks with  $n \geq 1,000$  nodes, the maximum number of iterations required has been 6, with each succeeding iteration over many fewer nodes. We do not know its expected order of complexity, but it appears to be fast and efficient.

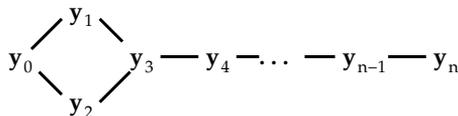


Figure 4. Reduction,  $\omega$ , has worst case  $O(n^2)$  behavior.

Readily, the iterative loop in  $\omega$  terminates when every singleton node  $\{y\}$  is closed. Therefore,  $\mathcal{N}.\omega$  is irreducible. We call  $\mathcal{N}.\omega$  the *irreducible trace* of  $\mathcal{N}$  and denote it by  $\mathcal{T}$ .<sup>7</sup>

The behavior of  $\omega$  may be more easily understood by an example. Figure 5(a) is an 18 node network that was presented in [33] to test matrix-oriented eigenvector algorithms. The dashed lines of Figure 5(b) delimit  $\{y\}.\beta$ , that is, sets of nodes that have the same closure properties. For example, node 9 is subsumed by 3,

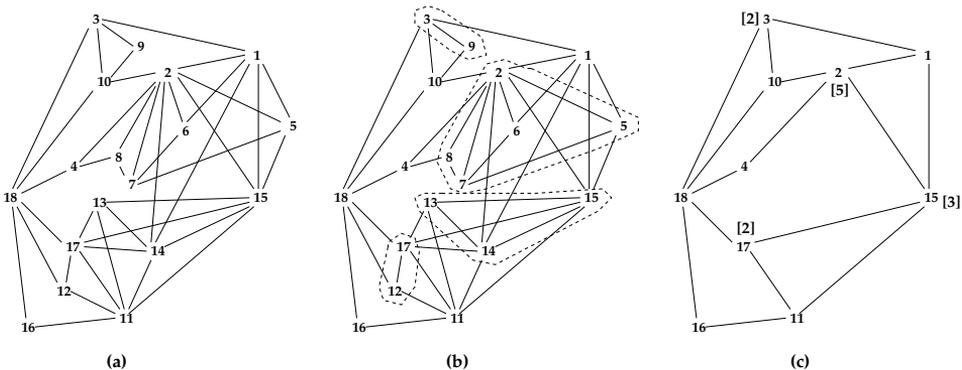


Figure 5. Stages in the execution of  $\omega$ .

because  $\{9\}.\eta = \{3, 9, 10\} \subseteq \{1, 3, 9, 10, 18\} = \{3\}.\eta$ . So  $9 \in \{3\}.\beta$ . Similarly, 5 is subsumed by 2 because  $\{5\}.\eta = \{1, 2, 5, 7, 15\} \subseteq \{1, 2, 4, 5, 6, 7, 8, 15\} = \{2\}.\eta$ , whence  $5 \in \{2\}.\beta$ .

Figure 5(c) illustrates the trace,  $\mathcal{T}$ , of Figure 5(a). The figures  $[n]$  in brackets denote those cardinalities with  $|\{y\}.\beta| > 1$ .

The order in which nodes, or more accurately the singleton subsets, of  $\mathcal{N}$  are encountered in the outer loop of Figure 3 arbitrarily determines which points are subsumed and subsequently deleted. For example, node 9 could also have been subsumed by 10, instead of 3. Nevertheless, we show below that the reduced trace,  $\mathcal{T} = \mathcal{N}.\omega$ , will be unique, up to isomorphism. Two networks are *isomorphic*, denoted  $\mathcal{N} = (N, \eta) \cong \mathcal{N}' = (N', \eta')$ , if there exists a 1 – 1 function  $i$  where  $i : N \rightarrow N'$  and for all  $x, y \in N$ ,  $x \in \{y\}.\eta$  if and only if  $i(x) \in i(\{y\}.\eta')$ .

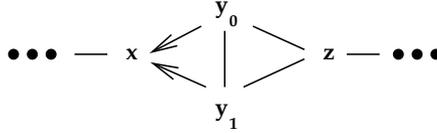
**Proposition 3.1.** *Let  $\mathcal{T} = \mathcal{N}.\omega$  and  $\mathcal{T}' = \mathcal{N}'.\omega'$  be two irreducible subsets of a finite network  $\mathcal{N}$ , then  $\mathcal{T} \cong \mathcal{T}'$ .*

<sup>7</sup> In [25],  $\mathcal{N}.\omega$  was called the “irreducible spine” of  $\mathcal{N}$ .

*Proof.* Suppose  $\mathcal{T} \neq \mathcal{T}'$ . Let  $y_0 \in \mathcal{N}$ ,  $y_0 \notin \mathcal{T}'$ , so  $\exists y_1 \in \mathcal{T}$  where  $y_0 \in \{y_1\} \cdot \varphi_\eta$ , or equivalently,  $\{y_0\} \cdot \eta \subseteq \{y_1\} \cdot \eta$ .  $y_1 \notin \mathcal{T}$  because  $y_0 \in \{y_1\} \cdot \varphi_\eta$  would contradict the irreducible nature of  $\mathcal{T}$ . Since  $y_1 \notin \mathcal{T}$ , there must exist  $y_2 \in \mathcal{T}$  such that  $y_1 \in \{y_2\} \cdot \varphi_\eta$ . If  $y_2 = y_0$  we are done, and  $i(y_0) = y_2$  is part of the desired isomorphism,  $i$ .

In fact, we must have  $y_2 = y_0$ , because supposing  $y_0 \in \{y_1\} \cdot \varphi_\eta$  implies  $\{y_0\} \cdot \eta \subseteq \{y_1\} \cdot \eta$ ;  $y_1 \in \{y_2\} \cdot \varphi_\eta$  implies  $\{y_1\} \cdot \eta \subseteq \{y_2\} \cdot \eta$ ; so  $\{y_0\} \cdot \eta \subseteq \{y_2\} \cdot \eta$  and thus  $y_0 \in \{y_2\} \cdot \varphi_\eta$ . But this contradicts the assumption that  $y_0, y_2 \in \mathcal{T}$  which is irreducible.  $\square$

It is instructive to see just how these isomorphic copies arise. Consider Figure 6. Here,  $\{y_0\} \cdot \varphi_\eta = \{y_1\} \cdot \varphi_\eta$  independently of what other nodes might be connected to



**Figure 6.**  $\{y_0\} \cdot \varphi_\eta = \{y_1\} \cdot \varphi_\eta$ . Either can subsume the other.

$x$  and  $z$ . Thus, either could subsume the other depending on which is encountered first in the outer loop of Figure 3.

We characterize the nature of the irreducible trace,  $\mathcal{T} = \mathcal{N} \cdot \omega$ , of any finite network,  $\mathcal{N}$  in the following proposition.

**Proposition 3.2.** *Let  $y \in \mathcal{T} = \mathcal{N} \cdot \omega$  with  $N$  finite. If  $z \in \{y\} \cdot \eta$ , then there exists a path  $\langle y, z, \dots \rangle$  terminating in a chordless cycle of length  $\geq 4$ .*

*Proof.* Let  $z \in \{y\} \cdot \eta$ . Let  $y_0 = y, y_1 = z$ . Since  $\mathcal{T}$  is irreducible,  $y_1 \notin \{y_0\} \cdot \varphi_\eta$ , so  $\exists y_2 \in \{y_1\} \cdot \eta$ ,  $y_2 \notin \{y_0\} \cdot \eta$ . Again, since  $\mathcal{T}$  is irreducible,  $\exists y_3 \in \{y_2\} \cdot \eta$ ,  $y_3 \notin \{y_1\} \cdot \eta$ . else  $\{y_2\} \cdot \eta \subseteq \{y_1\} \cdot \eta$  contradicting irreducibility. But, possibly,  $y_3 \in \{y_0\} \cdot \eta$  in which case  $\langle y_0, y_1, y_2, y_3, y_0 \rangle$  is the desired chordless cycle.

If  $y_3 \notin \{y_0\} \cdot \eta$ , irreducibility requires that  $\exists y_4 \in \{y_3\} \cdot \eta$ ,  $y_4 \notin \{y_2\} \cdot \eta$ . Again possibly  $y_4 \in \{y_1\} \cdot \eta$  or  $y_4 \in \{y_0\} \cdot \eta$ , yielding a chordless cycle of length 4, or of length 5.

Otherwise  $\exists y_5 \in \{y_4\} \cdot \eta$ ,  $y_5 \notin \{y_3\} \cdot \eta$ , and so on. Since  $N$  is finite, this construction must terminate with the desired chordless cycle.  $\square$

We would observe that if  $\mathcal{T}$  is irreducible and  $\{y\} \cdot \eta = \{y\}$ , then  $y$  is an isolated element of  $\mathcal{T}$ .

Figure 5(c) which illustrates the irreducible trace  $\mathcal{T}$  of Figure 5(a), is symmetric. It is easier to visualize closure in symmetric networks. The irreducible trace  $\mathcal{T}$  of Figure 1 is shown in Figure 7. Again we denote  $|\{y\} \cdot \beta|$  in square brackets. Observe that, after  $j$  has been subsumed by  $h$  in the first iteration (and  $g$  subsumed by  $d$ ),  $h$  could in turn be subsumed by either  $d$ ,  $f$  or  $i$ .

Proposition 3.2 assumes that  $\eta$  might be directed (asymmetric); the following proposition assumes symmetry.

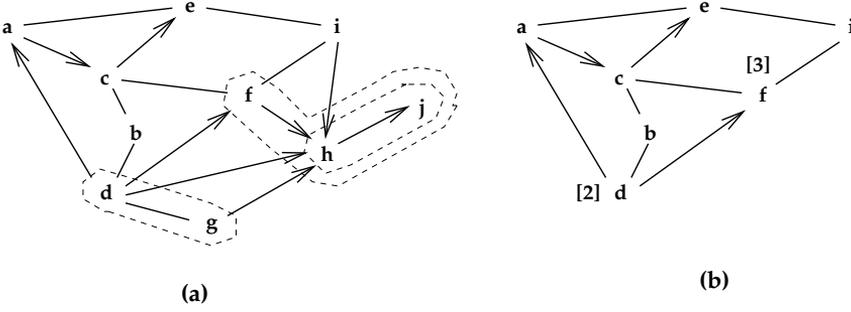


Figure 7. Closed sets and the irreducible trace,  $\mathcal{T}$ , of Figure 1.

**Proposition 3.3.** *If  $\eta$  is a symmetric (undirected) relation on  $\mathcal{N}$  and  $\mathcal{T} = \mathcal{N}.\omega$ , then for all  $y \in \mathcal{N}$ , either*

- (a)  *$y$  is an element in a chordless cycle of length  $\geq 4$ , or*
- (b)  *$y$  is an element in a path between two chordless cycles of length  $\geq 4$ .*

*Proof.* A proof similar to Proposition 3.2 can be found in [25].  $\square$

The  $\omega$  transformation has a number of valuable properties. In particular,

**Proposition 3.4.** *The  $\omega$  transformation is continuous.*

*Proof.* First, consider any singleton node  $\{z\}$  that is subsumed by another node  $\{y\}$ . So  $\{z\}.\eta \subseteq \{y\}.\eta$  and  $z \in \{y\}.\varphi_\eta$ . Call this step  $\omega_z$ . This  $\omega_z$  can be defined for nodes by  $\{z\}.\omega_z = \emptyset$  and  $\{x\}.\omega_z = \{x'\}$  for  $x \neq z$ , and for relation  $\eta$  by  $\{x\}.\omega_z.\eta' = \{x\}.\eta - \{z\}$  and  $\omega_z$  is the identity map for all nodes except those  $x$  that interact with  $z$ . So  $\{x\}.\eta.\omega_z \subseteq \{x'\}.\eta'$  with equality if  $z \notin \{x\}.\eta$ .

Now consider any set  $Y$ ,

$$\begin{aligned}
 Y.\varphi_\eta.\omega_z &= \bigcup_{x \in Y.\eta} \{x\} \{ \{x\}.\eta \subseteq Y.\eta \}.\omega_z && \text{because } \omega_z \text{ is extensible} \\
 &\subseteq \bigcup_{x \in Y.\eta} \{x\} \{ \{x\}.\eta.\omega_z \subseteq Y.\eta.\omega_z \} && \text{because } \omega_z \text{ is monotone} \\
 &\subseteq \bigcup_{x \in Y.\eta'} \{ \{x\}.\eta' \subseteq Y.\eta' \} && \text{because } \{x\}.\eta.\omega_z \subseteq \{x\}.\eta' \\
 &= Y.\omega_z.\varphi_{\eta'}
 \end{aligned}$$

So  $\omega_z$  is continuous. By Proposition 2.1,  $\omega$  is continuous.  $\square$

### 3.1. Centers and centrality

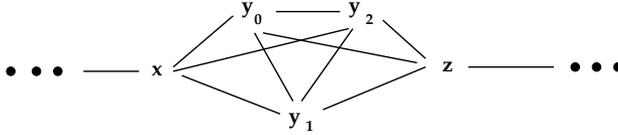
Concepts of “centrality” seem to be essentially meaningless unless the relationships are symmetric. Consequently, in this section, and for the rest of the paper we will assume that  $\eta$  is a symmetric relation.

Let  $\sigma(x, z)$  denote a *shortest path* (w.r.t.  $\eta$ ) between  $x$  and  $z$ , that is  $\langle x, \dots, z \rangle$  is of minimal length.  $\sigma(x, z)$  need not be unique, thus by the notation  $\sigma(x, y) = \langle x, \dots, z \rangle$  we mean the latter is some shortest path from  $x$  to  $z$ . By the *distance between*  $x$  and  $z$ ,  $d(x, z)$ , we mean the length of any shortest path  $\sigma(x, z)$ .

**Proposition 3.5.** *Let  $\sigma(x, z)$  denote a shortest path between  $x$  and  $z$ . If  $x$  (or  $z$ ) subsumes  $y$ ,  $y \notin \{x, z\}$ , then  $y \notin \sigma(x, z)$ .*

*Proof.* Suppose that  $\sigma(x, z) = \langle x, y, z_0, \dots, z_n = z \rangle$ . Since  $x$  subsumes  $y$ ,  $y \in \{x\}.\varphi_\eta$ , or  $\{y\}.\eta \subseteq \{x\}.\eta$  implying  $z_0 \in \{x\}.\eta$  yielding a shorter path  $\langle x, z_0, \dots, z \rangle$ .  $\square$

In other words,  $y$  can be removed from  $\mathcal{N}$  with the certainty that if there was a path from some node  $x$  to  $z$  through  $y$ , it was not a shortest path. Such subsumed nodes can be iteratively removed from  $\mathcal{N}$  without changing connectivity, or distances, between retained nodes. In Figure 8,  $d(x, z) = 2$ . Any two of  $\{y_0, y_1, y_2\}$  may be subsumed by the remaining node yielding 3 possible isomorphic



**Figure 8.** Shortest path  $\sigma(x, z)$  where  $x, z \in \mathcal{T}$ .

traces,  $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2$ , while retaining a shortest path. Compare this with Figure 6.

A central quest in the analysis of social networks is the identification of its “important” nodes. In social networks, “importance” may be defined with respect to the path structure [13].

Those nodes  $C_D = \{y \in \mathcal{N}\}$  for which  $\delta(y) = \sum_{s \neq y} d(s, y)$  is *minimal* have traditionally been called the *center* of  $\mathcal{N}$  [15], they are “closest” to all other nodes. It is well known that this subset of nodes must be edge connected. One may assume that these nodes, in the “center”, of a network are “important” nodes.

Alternatively, one may consider those nodes which “connect” many other nodes, or clusters of nodes, to be the “important” ones. Let  $\sigma_{st}(y)$  denote the number of shortest paths  $\sigma(s, t)$  containing  $y$ ; then those nodes  $y$  for which  $\sigma_{st}(y)$  is *maximal* are those nodes that are involved in the most connections. Let  $C_B = \{y \in \mathcal{N}\}$ , for which  $\sigma_{st}(y)$  is maximal. This is sometimes called *betweenness centrality* [4, 13]. (Note: traditionally, centrality measures are normalized to range between 0 and 1, but we will not need this for this paper.)

In the following sequence, we consider how reduction by  $\omega$  affects the paths in  $\mathcal{N}$ , and especially the shortest paths on which notions of centrality are based.

**Lemma 3.6.** *Let  $y \in \mathcal{T}$  with  $z \in \{y\}.\beta$  and let  $\sigma(x, z)$  be a shortest path (in  $\mathcal{N}$ ) where  $x \notin \{y\}.\beta$ . There exists a path  $\sigma(x, y)$  such that  $|\sigma(x, y)| \leq |\sigma(x, z)|$ .*

*Proof.* We may assume  $\sigma(x, z) = \langle x, \dots, u, y_i, \dots, z \rangle$ , where  $y_i \in \{y\}.\beta$ . Since  $\eta$  is symmetric,  $u \in \{y_i\}.\eta$ , and we have  $\sigma(x, y) = \langle x, \dots, u, y \rangle$ ,  $|\sigma(x, y)| \leq |\sigma(x, z)|$  with equality if and only if  $\sigma(x, z) = \langle x, \dots, u, z \rangle$  and  $z \in \{y\}.\eta$ .  $\square$

**Lemma 3.7.** *Let  $\sigma(x, z)$  be a path where  $x \in \{w\}.\beta$ ,  $z \in \{y\}.\beta$ ,  $w \neq y$ . There exists  $u \in \sigma(x, z)$ ,  $u \in \mathcal{T}$ .*

*Proof.* Suppose that  $\sigma(x, z) = \langle x = x_0, x_1, \dots, x_k, z_0, z_1, \dots, z_m = z \rangle$ , where  $x_i \in \{w\}.\beta$ ,  $0 \leq i \leq k$  and  $z_j \in \{y\}.\beta$ ,  $0 \leq j \leq m$ .  $z_0 \in \{x_k\}.\eta$  implies  $\{z_0\}.\eta \subseteq \{x_k\}.\eta$ , and repeating this observation indicates  $z_j \in \{w\}.\beta$ ,  $0 \leq j \leq m$ . Similarly  $x_i \in \{y\}.\beta$ ,  $0 \leq i \leq k$ , contradicting the assumption that  $w \neq y$ .

So we have  $\sigma(x, z) = \langle x = x_0, \dots, x_k, u, z_0, \dots, z_m = z \rangle$ . If  $u \notin \mathcal{T}$ ,  $u \in \{v\}.\beta$  for some  $v \neq w, y \in \mathcal{T}$  and the same argument is repeated.  $\square$

**Proposition 3.8.** *Let  $\sigma(x, z)$  be a shortest path in  $\mathcal{N}$ , and let  $\sigma(x', z') \in \mathcal{T}$ , where  $x' = \{x\}.\omega$  and  $z' = \{z\}.\omega$ . Then,  $|\sigma(x', z')| \leq |\sigma(x, z)|$ .*

*Proof.* If  $x, z \in \mathcal{T}$ , the result is trivial.

If either  $x$  or  $z \notin \mathcal{T}$ , then the result follows from Lemma 3.6.

If both  $x$  and  $z \notin \mathcal{T}$ , then by Lemma 3.7  $\exists u \in \mathcal{T}$ ,  $u \in \sigma(x, z)$  and by Lemma 3.6,  $|\sigma(x', z')| = |\sigma(x', u)| + |\sigma(u, z')| \leq |\sigma(x, u)| + |\sigma(u, z)| = |\sigma(x, z)|$ .  $\square$

Thus, paths of shortest length in  $\mathcal{N}$  are preserved as paths that have no longer length in  $\mathcal{T}$ ; even though they may no longer be “shortest”. One would expect, therefore, that nodes with minimal distance and maximal betweenness measures will almost always be found in the irreducible trace  $\mathcal{T}$ . This is non-trivial because it need not always be true. One problem is that we may have several isomorphic traces,  $\mathcal{T}_1, \dots, \mathcal{T}_k$ , so we can only assert that  $C_D \cap \mathcal{T}_j$  and  $C_B \cap \mathcal{T}_j$  are non-empty for some  $1 \leq j \leq k$ . Second, there exist pathological cases where the centers are disjoint from  $\mathcal{T}$ . The network of Figure 4 is an example. If  $n = 8$ , then  $C_D = C_B = y_4$  because  $18 = \delta(y_4) < \delta(y_3) = \delta(y_5) = 19$ , and  $24 = \sigma_{st}(y_4) > \sigma_{st}(y_3) = \sigma_{st}(y_5) = 23$ . But,  $y_4 \notin \mathcal{T}$ . Still, for most networks,  $\mathcal{N}$ , we have:

**Proposition 3.9.** *The centers  $C_D$  and  $C_B$  of a network  $\mathcal{N}$  are retained in  $\mathcal{T} = \mathcal{N}.\omega$  provided that  $\mathcal{N}$  is reasonably large and balanced.*

*Proof.* An unwieldy, sufficient condition for  $\mathcal{N}$  to be balanced can be found in [25]; but this condition is not at all necessary, and not worth repeating.  $\square$

The irreducible trace,  $\mathcal{T} = \mathcal{N}.\omega$ , is a rather good abstraction of a symmetric network,  $\mathcal{N}$ .

#### 4. THE EXPANSION TRANSFORMATION, $\varepsilon$

A *phrase-structured grammar*,  $\mathcal{G}$ , is a system of transformations of the form  $A \rightarrow \gamma$ , in which a non-terminal symbol  $A$  is rewritten as a substring  $\gamma$ . These have been widely studied as a mechanism for generating linear strings, especially computer code [7, 17, 37]. There have also been efforts to employ grammars as a mechanism for generating/describing multi-dimensional structures [30, 32, 34, 35]. These latter efforts have been far less successful because of the difficulty of embedding the re-written structure  $\gamma$  within the structure surrounding  $A$ . Readily, we expect  $\gamma.\eta = A.\eta$ , but this still leaves the actual embedding of  $\gamma$  under-defined.

Another approach, called a *neighborhood expansion grammar* was explored in [23]. Here, a subset  $Y$  of a growing structure is identified to be the neighborhood of a new element  $p'$ . That is  $\{p'\}.\eta' = Y'$  in the rewritten structure. It was felt that this kind of grammar better explained how we naturally express ourselves by adding thoughts incrementally rather than rewriting “placeholder” thoughts, *i.e.*

non-terminal thoughts. Using this mechanism, one can functionally define various network structures, such as the class of all chordal graphs [23].

We will use an expansion grammar to define  $\mathcal{T}.\omega^{-1}$  for any irreducible trace  $\mathcal{T}$ . We expand  $\mathcal{T}$  by one node  $p$  at a time, where we choose an arbitrary element  $y \in N$ , then choose an arbitrary subset  $Y \subseteq \{y\}.\eta$  and let  $\{p\}.\eta = Y \cup \{p\}$ . We denote this non-deterministic *expansion transformation* by  $\varepsilon$ .

More formally, let  $\mathcal{N}_i = (N_i, \eta_i)$  be any network and let  $y_i \in N_i$ .  $\mathcal{N}_{i+1} = (N_{i+1}, \eta_{i+1}) = \mathcal{N}_i.\varepsilon$  where  $N_{i+1} = N_i \cup \{p_{i+1}\}$  and  $\eta_{i+1}|_{N_i} = \eta_i$ ,  $\{p_{i+1}\}.\eta \subseteq \{y_i\}.\eta$ .

**Proposition 4.1.** *Let  $\mathcal{T}$  be irreducible and let  $\mathcal{T} = \mathcal{N}_0, \mathcal{N}_1, \dots, \mathcal{N}_i, \dots$  be any sequence of networks such that  $N_{i+1} = N_i.\varepsilon$ . For all  $i, k$ ,  $\mathcal{N}_i.\omega = \mathcal{N}_k.\omega = \mathcal{T}$ .*

*Proof.* By Proposition 3.1, the order in which nodes are subsumed during reduction by  $\omega$  is irrelevant. Consequently, we may assume that, since  $\{p\}_{i+1}.\eta \subseteq \{y_i\}.\eta$ ,  $p_{i+1}$  will be subsumed by  $y_i$  in the first step of the reduction by  $\omega$ , yielding  $\mathcal{N}_i$ . The result follows by finite induction.  $\square$

This  $\varepsilon$  expansion transformation has been implemented using the following pseudocode in Figure 9. This implementation code is a bit more constrained than that

```

for all y in N
{
  while (|{y}.beta| > 1)
  {
    create new node p;
    Y = choose_random_in ({y}.nbhd);
    {p}.nbhd = Y;
    k = random_int(1, |{y}.beta|-1);
    |{y}.beta| = |{y}.beta| - k;
    |{p}.beta| = k;
    add p to N;
  }
}

```

**Figure 9.** Pseudocode implementing  $\varepsilon$  expansion.

described above where  $y$  can be any element in  $N$ . Since the reduction code of Figure 3 records  $\{y\}.\beta$ , this code constrains  $y$  to be one in which  $|\{y\}.\beta| > 1$ . It then decrements this “ $\beta$ -count” and randomly distributes part of the remainder (if any) to the new node  $p$ .

Figure 10 shows three randomly generated networks,  $\mathcal{N}_a, \mathcal{N}_b, \mathcal{N}_c$  produced by the code operating on the irreducible trace,  $\mathcal{T}$ , of Figure 5(c). In each case, 8 new nodes  $\{a, b, \dots, g, h\}$  were added. Observe that only nodes 2, 3, 15, and 17 serve as  $y_i$  in these expansions. In each network  $|N_a| = |N_b| = |N_c| = 18$ , the size of Figure 5(a); but the numbers of edges are 42, 48, and 55 respectively.

The reduction operator,  $\omega$ , is continuous (Proposition 3.4), but the expansion operator need not be.

**Proposition 4.2.** *An expansion transformation  $(N, \varphi_\eta) \xrightarrow{\varepsilon} (N', \varphi_{\eta'})$  with*

$$\{p\}.\eta' = Y \cup \{p\},$$

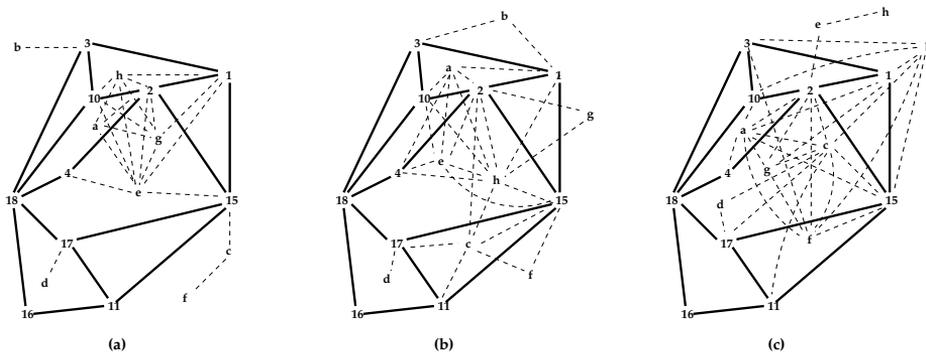


Figure 10. Three expansions of the trace  $\mathcal{T}$  of Figure 5(c).

for given  $p \notin N$ ,  $N' = N\{p\}$ ,  $Y \subseteq N$ , is discontinuous if and only if there exists  $x \notin Y$  such that  $\{x\}.\varphi_\eta \cap Y \neq \emptyset$ .

*Proof.* Suppose  $\varepsilon$  is discontinuous. Then,  $\exists X$  such that  $X.\varphi_\eta.\varepsilon \not\subseteq X.\varepsilon.\varphi_\eta$ . Because  $\varepsilon$  is an extensible transformation,  $\exists x \in X$  such that  $\{x\}.\varphi_\eta.\varepsilon \not\subseteq \{x\}.\varepsilon.\varphi_\eta$ .  $x \notin Y$  because if so  $\{x\}.\eta' = \{x\}.\eta.\varepsilon \cup \{p\}$  and so  $\{x\}.\varphi_\eta \subseteq \{x\}.\varphi_\eta$ .

Readily, if the second condition holds, then since  $\{x\}.\varphi_\eta \cap Y \neq \emptyset$ ,  $\exists y \in Y$  such that  $y \in \{x\}.\eta$ , but  $p \notin \{x\}.\eta$ . Hence,  $y \notin \{x\}.\varphi_\eta$ . So  $\varepsilon$  is discontinuous.  $\square$

The following Figure 11 illustrates such a discontinuity. Here the expansion

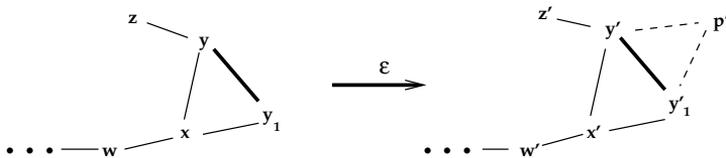


Figure 11. A discontinuous  $\varepsilon$  transformation.

neighborhood is  $Y = \{y, y_1\}$ ; but  $y_1 \in \{x\}.\varphi_\eta$ , so by Proposition 4.2,  $\varepsilon$  should be discontinuous. In fact,  $\{x\}.\varphi_\eta.\varepsilon = \{x', y'\} \not\subseteq \{x'\} = \{x\}.\varepsilon.\varphi_\eta$ , violating (2.1). This example also illustrates Proposition 2.8, because  $\varepsilon$  is not neighborhood monotone.  $\{y_1\}.\eta = \{x, y, y_1\} \subseteq \{w, x, y, y_1\} = \{x\}.\eta$ , but  $\{y_1\}.\eta.\varepsilon = \{x', y', y_1, p'\} \not\subseteq \{w', x', y', y_1\} = \{x\}.\eta.\varepsilon$ .

Two networks  $\mathcal{N}_1, \mathcal{N}_2$  are said to be *structurally similar* if  $\mathcal{N}_1.\omega = \mathcal{N}_2.\omega$ , that is, they have the same irreducible trace,  $\mathcal{T}$ . Readily, structural similarity is an equivalence relation. If, in addition,  $|\mathcal{N}_1| = |\mathcal{N}_2|$ , we say that they are strongly similar. The three networks,  $\mathcal{N}_a, \mathcal{N}_b, \mathcal{N}_c$  of Figure 10 are strongly similar.

By Proposition 4.1, for all networks,  $\mathcal{N}$ , we have  $\mathcal{N}.\omega.\varepsilon.\omega = \mathcal{N}.\omega$ , or more simply  $\omega.\varepsilon.\omega = \omega$ . Thus, over the subspace of irreducible networks,  $\mathcal{T}$ , we have  $\mathcal{T}.\varepsilon.\omega = \mathcal{T}$ , so  $\varepsilon$  functions as a left inverse of  $\omega$ .

## 5. DISCUSSION AND POSSIBLE APPLICATION

It is not unusual to write computer codes that analyze, or otherwise manipulate, collections of related elements such as those of graphs, or networks. But, to the author's knowledge, the code of Figure 3 is the first to implement a well-defined, network-valued function,  $\omega$ , over the space of all networks. Real-valued functions of a real variable can also be defined by computable expressions. In some cases, particularly if the expression involves an infinite sum, such as the Riemann zeta function

$$\zeta(x) = \sum_{k=1}^{\infty} k^{-x},$$

it may be difficult to determine whether the computation converges to a unique value. Establishing that the computation of Figure 3 yields a unique network (up to isomorphism) and is, therefore, a true function for all argument networks is fundamental; as is establishing that it is continuous. Moreover, we have shown that it is, by means of an expansion grammar, invertible over the space of symmetric networks.

This  $\omega$  transformation has several attractive properties. Even though it is defined and implemented locally (the computer process need never access elements more than two steps away), it nevertheless preserves global properties such as path connectedness, and global centers with respect to distance and betweenness.

We have tried to be quite general in framing our definitions, such as the neighborhood operator,  $\eta$ , which defines the relationships between the elements of the network. However, to obtain many results we have explicitly acknowledged the need for all operators to be monotone and extensible. Moreover, the centrality properties of  $\omega$  assume symmetric (undirected) neighborhoods, even though  $\omega$  is well-defined for both directed and undirected networks.

Networks are widely used to model the relationships found in many situations. Closure concepts are thought to be an important mathematical tool for dealing with these relational networks. The results of this paper would seem to support this contention. For example, closure is thought to be relevant to cognitive processing [9, 27, 28]. It is accepted that human cognition is realized by means of neural circuitry, or neural networks, of some form. It has been proposed that such cognitive events are preserved in long-term memory by chordless cycles such as those that characterize the irreducible trace. It may well be that  $\omega$  models what psychologists call *consolidation*, a period when thoughts and experiences are processed for long-term storage [18].

Psychologists also recognize that memories are not a mere recapitulation of some stored structure, but rather an active neural process that embellishes whatever was initially stored [2, 8]. The  $\varepsilon$  process feels very much like that. Neither it, nor a memory of an event, may precisely duplicate the event (or neural network representing it), but they will be very similar.

Finally, there is a growing thought that long-term memories may need to be repeatedly recalled and rewritten. The observation that  $\omega \cdot \varepsilon \cdot \omega = \omega$  seems quite congruent to this idea.

There is, as yet, no firm evidence that human memory can actually be modeled by these abstract mathematical transformations; but the parallels are striking.

**Acknowledgment.** The author would like to thank an unknown referee for several helpful modifications to this paper.

## REFERENCES

- [1] C. C. Aggarwal and H. Wang, *On dimensionality reduction of massive graphs for indexing and retrieval*, in: S. Abiteboul, K. Böhem, C. Koch and K.-L. Tan (eds.), 27th International Conference on Data Engineering (ICDE), IEEE, Hanover, Germany, 2011, 1091–1102.
- [2] F. C. Bartlett, *Remembering: A Study in Experimental and Social Psychology*, Cambridge University Press, Cambridge, UK, 1932.
- [3] M. Behzad and G. Chartrand, *Introduction to the Theory of Graphs*, Allyn and Bacon, Boston, 1971.
- [4] U. Brandes, *A faster algorithm for betweenness centrality*, The Journal of Mathematical Sociology **25** (2001), 163–177.
- [5] N. Caspard and B. Monjardet, *The lattices of closure systems, closure operators and implicational systems on a finite set: A survey*, Discrete Applied Mathematics **127** (2003), 241–269.
- [6] G. Castellini, *Categorical Closure Operators*, Birkhäuser, Boston, 2003.
- [7] N. Chomsky, *Formal properties of language*, in: R. D. Luce, R. R. Bush and E. Galanter (eds.), Handbook of Mathematical Psychology, Wiley, New York, 1963.
- [8] M. A. Conway, *Remembering: A process and a state*, in: H. L. Roediger III, Y. Dudai and S. M. Fitzpatrick (eds.), Science of Memory: Concepts, Oxford University Press, New York, 2007, 237–241.
- [9] J.-P. Doignon and J.-C. Falmagne, *Knowledge Spaces*, Springer, Berlin, 1999.
- [10] P. H. Edelman and R. E. Jamison, *The theory of convex geometries*, Geometriae Dedicata **19** (1985), 247–270.
- [11] K. Engle, *Sperner theory*, in: M. Hazewinkle (ed.), Encyclopedia of Mathematics, Springer, Berlin, 2001.
- [12] M. Farber and R. E. Jamison, *Convexity in graphs and hypergraphs*, SIAM Journal on Algebraic and Discrete Methods **7** (1986), 433–444.
- [13] L. C. Freeman, *Centrality in social networks, conceptual clarification*, Social Networks **1** (1978/79), 215–239.
- [14] A. C. Gilbert and K. Levchenko, *Compressing network graphs*, in: Proceedings LinkKDD’04, Seattle, WA, 2004.
- [15] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.
- [16] R. E. Jamison-Waldner, *A perspective on abstract convexity: Classifying alignments by varieties*, in: D. C. Kay and M. Breen (eds.), Convexity and Related Combinatorial Geometry, Marcel Dekker, Inc., New York, 1982.
- [17] D. E. Knuth, *Semantics of context free languages*, Mathematical Systems Theory **2** (1968), 127–145.
- [18] J. E. LeDoux, *Consolidation: Challenging the traditional view*, in: H. L. Roediger III, Y. Dudai and S. M. Fitzpatrick (eds.), Science of Memory: Concepts, Oxford University Press, New York, 2007, 171–175.
- [19] J. Leskovec and C. Faloutsos, *Sampling from large graphs*, in: Proceedings of the 12th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining (KDD’06), Philadelphia, PA, 2006, 631–636.
- [20] N. M. Martin and S. Pollard, *Closure Spaces and Logic*, Kluwer Academic Publishers, Boston, 1996.
- [21] O. Ore, *Mappings of closure relations*, Annals of Mathematics **47** (1946), 56–72.
- [22] J. L. Pfaltz, *Closure lattices*, Discrete Mathematics **154** (1996), 217–236.

- [23] J. L. Pfaltz, *Neighborhood expansion grammars*, in: G. Goos, J. Hartmanis and J. van Leeuwen (eds.), International Workshop on Theory and Application of Graph Transformation (TAGT'98), Paderborn, Germany, 1998, Lecture Notes in Computer Science **1764**, Springer, 1998, 30–44.
- [24] J. L. Pfaltz, *Finding the mule in the network*, in: R. Alhajj and B. Werner (eds.), International Conference on Advances in Social Network Analysis and Mining (ASONAM 2012), Istanbul, Turkey, 2012, IEEE, 2012, 667–672.
- [25] J. L. Pfaltz, *The irreducible spine(s) of discrete networks*, in: X. Li, Y. Manolopoulos, D. Srivastava and G. Huang (eds.), Web Information Systems Engineering (WISE 2013), Nanjing, PRC, 2013, Lecture Notes in Computer Science **8181**, Part 2, Springer, 2013, 104–117.
- [26] J. L. Pfaltz, *Mathematical continuity in dynamic social networks*, Social Network Analysis and Mining **3** (2013), 863–872.
- [27] J. L. Pfaltz, *The Role of continuous processes in cognitive development*, Mathematics for Applications **4** (2015), 129–152.
- [28] J. L. Pfaltz, *Using closed sets to model cognitive behavior*, in: T. Ray, R. Sarker and X. Li (eds.), *Proc. Australian Conference on Artificial Life and Computational Intelligence (ACALCI 2016)*, Canberra, ACT, 2016, Lecture Notes in Computer Science **9592**, Springer, 2016, 13–26.
- [29] J. L. Pfaltz, *Computational processes that appear to model human memory*, in: D. Figueiredo, C. Martín-Vide, D. Pratas and M. A. Vega-Rodríguez (eds.), Algorithms for Computational Biology, 4th International Conference, AlCoB 2017, Aveiro, Portugal, 2017, Proceedings, Lecture Notes in Bioinformatics **10252**, Springer, 85–99.
- [30] J. L. Pfaltz and A. Rosenfeld, *Web grammars*, in: Proceedings of the First International Joint Conference on Artificial Intelligence, Washington, DC, 1969, 609–619.
- [31] John Pfaltz and Josef Šlapal, *Transformations of discrete closure systems*, Acta Mathematica Hungarica **138** (2013), 386–405.
- [32] T. W. Pratt, *Pair grammars, graph languages, string to graph translations*, Journal of Computer and System Sciences **5** (1971), 560–595.
- [33] W. Richards and A. Seary, *Eigen analysis of networks*, Journal of Social Structure **1** (2000), 1–16.
- [34] A. Rosenfeld, *Isotonic grammars, parallel grammars, and picture grammars*, Machine Intelligence **6** (1971), 281–294.
- [35] H. J. Schneider and H. Ehrig, *Grammars on partial graphs*, Acta Informatica **6** (1976), 297–316.
- [36] Josef Šlapal, *Complete-lattice morphisms compatible with closure operators*, Thai Journal of Mathematics **8** (2010), 255–262.
- [37] B. Stroustrup, *The C++ Programming Language*, 2nd. ed., Addison–Wesley, Reading, MA, 1991.

John L. Pfaltz, Department of Computer Science, University of Virginia, 85 Engineer's Way, Charlottesville, VA 22904, USA  
*e-mail*: jlp@cs.virginia.edu