Math. Appl. **2** (2013), 153–167
DOI: 10.13164/ma.2013.12

MATHEMATICAL EVOLUTION IN DISCRETE NETWORKS

JOHN L. PFALTZ

Abstract. This paper provides a mathematical explanation for the phenomenon of "triadic closure" so often seen in social networks. It appears to be a natural consequence when network change is constrained to be continuous. The concept of chordless cycles in the network's "irreducible spine" is used in the analysis of the network's dynamic behavior.

A surprising result is that as networks undergo random, but continuous, perturbations they tend to become more structured and less chaotic.

1. Introduction

In this paper we explore the behavior of networks as they change under a sequence of simple, "continuous" transformations. We will start with a random network and apply a sequence of transformations which randomly adds edges (or connections) to it and randomly deletes edges. Periodically we observe the resulting network. The structure, or topology, of these show considerable similarity to that appearing in social networks. The rather surprising conclusion will be that "random change, provided it is continuous, leads to a less chaotic and more regular structure". This provides a purely mathematical basis for the presence of "triadic closure" in social networks.

In Section 2, we formally establish criteria for edge, or link, addition and deletion to be "continuous". Then in Section 2.2, we apply these criteria to a concrete application, the evolution of a random network. Section 3 employs two different network properties to compare the original network with its "evolved" form.

Both "continuity" and network "structure" will be expressed in terms of closed sets.

1.1. Closure and closed sets

The traditional approach to understanding the behavior of smooth, or continuous, phenomena in a Euclidean world is based on real numbers and an open-set topology [18, 21, 26]. But, we are interested in *discrete* systems, which by definition do not support the kinds of limit operations found in open-set topologies [4]. The integers are one example of a discrete set. Consequently, the traditional delta/epsilon interpretation of continuity based on open sets fails in a discrete world. To compensate, this paper makes use of the concept of "closed sets" to analyze dynamic discrete systems, such as social networks.

MSC (2010): primary 90B10, 90B15.
Keywords: discrete network, system evolution, closure operator.

Let $\mathcal{S}$ be any set system. An operator $\mathcal{S} \xrightarrow{\varphi} \mathcal{S}$ is said to be a *closure operator* if for all sets $X, Y \subseteq \mathcal{S}$, φ is:

(a) *expansive*, or $Y \subseteq Y.\varphi$;

(b) *monotone*, or $X \subseteq Y$ implies $X.\varphi \subseteq Y.\varphi$;

(c) *idempotent*, or $Y.\varphi.\varphi = Y.\varphi$.[1]

A set $Y \subset \mathcal{S}$ is said to be *closed* if $Y.\varphi = Y$.

A different, but equivalent definition, capitalizes on the fact that the intersection of closed sets must be closed. We let $\mathcal{C} \subseteq \mathcal{S}$ be a collection of subsets with the property that $X, Y \in \mathcal{C}$ implies that $X \cap Y \in \mathcal{C}$. These then are the closed sets of $\mathcal{S}$ and the operator φ is defined to be $X.\varphi = \bigcap_{X \subseteq Y} Y \in \mathcal{C}$.

Still, a third formulation of the closure concept is based on "path independence" which is encountered in the economic literature [24, 31, 41, 42], particularly those employing choice functions. An operator α is *path independent* if

$$(X \cup Y).\alpha = (X.\alpha \cup Y.\alpha).\alpha$$

An expansive (C1), monotone (C2) operator α is a closure operator, φ, if and only of it is path independent [32].

This abundance of definitions suggests the variety of applications in which closure can be employed.

A closure operator can satisfy other axioms depending on the mathematical discipline. A *topological closure* is grounded and closed under finite union, that is

(C0) $\emptyset.\varphi = \emptyset$,

(C4) $(Y \cup Z).\varphi = Y.\varphi \cup Z.\varphi$. The closure operator of linear systems, often called the *spanning operator*, satisfies the Steinitz-MacLane *exchange axiom*

(C5) if $x, z \notin Y.\varphi$, $x \neq z$, and $z \in (Y \cup \{x\}).\varphi$ then $x \in (Y \cup \{z\}).\varphi$.

Such closure systems are called *matroids*; they are generalized matrices [23, 46]. Still other closure operators may satisfy an *anti-exchange axiom*

(C6) if $x, z \notin Y.\varphi$, $x \neq z$, and $z \in (Y \cup \{x\}).\varphi$ then $x \notin (Y \cup \{z\}).\varphi$.

These closure operators, which include the geometric convex hull operator, are said to be *antimatroid closure* operators and the corresponding closure systems are called *antimatroids* or *convex geometries*, [3, 11, 13]. We also have "greedoids" which model greedy algorithms [22]. In the following sections we will employ a closure operator that seems to be more appropriate for our study of networks; it satisfies none of these more restrictive axioms.

1.2. Network closure

Graphs, or networks, constitute one class of discrete systems; we treat these terms as synonyms. There is a ground set N of nodes, or elements, or individuals, and a relationship between the nodes which we normally model as a set E of edges. There are a number of different closure operators that can be associated with discrete network systems. If a network, $\mathcal{N}$, is directed, one can define $\{y\}.\varphi_1$ as the set of all z reachable by a directed path from y, and then extend it so that $Y.\varphi_1 = \bigcup_{y \in Y} \{y\}.\varphi_1$. If the directed system is a partial order this closure can be expressed as $Y.\varphi_1 = \{z | \exists y \in Y, y \leq z\}$. Alternatively, one could define

[1] Because closure operators, and later network transformations, are set valued, we use suffix notation to distinguish them from numerically valued functions.

$Y.\varphi_2 = \{x | \exists y \in Y, x \leq y\}$, or $Y.\varphi_3 = \{x | \exists y_1, y_2 \in Y, y_1 \leq x \leq y_2\}$. All three closures are explored in [37]. Note that in this case, $\varphi_3 = \varphi_1 \cap \varphi_2$.

When the network is undirected, it would be natural to define a closure using the concept of "convexity", that is to define $Y.\varphi_4$ to be the set of all nodes on shortest paths between nodes in Y. Unfortunately, Farber and Jamison show this to be mathematically unsound in [14].

The "neighborhood closure" which we use in this paper is applicable to either directed or undirected networks, however all our examples will be of undirected networks, that is $(x, y) \in E$, implies $(y, x) \in E$. Let $Y.\eta$ denote the *neighborhood* of Y, that is all nodes $z \notin Y$ such that $(y, z) \in E$ is an edge with $y \in Y$. It is convenient to let $Y.\rho$ denote $Y \cup Y.\eta$ or the *region* dominated by Y.[2] By the *closure* of Y, we mean the set, $Y.\varphi$, of all z such that $\{z\}.\rho \subseteq Y.\rho$. It is not hard to show this is really a closure operator, [38]. An equivalent, but more useful formulation is

$$Y.\varphi = Y \cup \{z | z \in Y.\eta, \{z\}.\eta \subseteq Y.\rho\} \tag{1.1}$$

All operators, such as η and φ and the transformations in the next section, are defined over sets, so an expression such as $z.\varphi$ is technically incorrect; it should be $\{z\}.\varphi$, as in (1.1). But, we will often use the abbreviated form when dealing with singleton sets.

In Section 2, we use a 35 node network for a running illustration. It is large enough to be non-trivial, but small enough to be comprehensible.

2. The evolution process

Evolution is a process which changes the structure of a discrete system $\mathcal{S}$. We model such processes by mathematical "transformations". Let $\mathcal{S} = (S, E)$ and $\mathcal{S}' = (S', E')$ be undirected networks. A *transformation* $\mathcal{S} \xrightarrow{f} \mathcal{S}'$ is a function $f : \mathcal{P}_2(S) \to \mathcal{P}_2(S')$ between the sets of ground sets' subsets of the cardinality less or equal to 2. To every subset Y of $\mathcal{S}$, $|Y| \leq 2$, it assigns a unique set $Y.f \subseteq \mathcal{S}'$, $|Y.f| \leq 2$. Each such function can be extended by unions of images to a mapping $\mathcal{P}(S) \to \mathcal{P}(S')$. We presume that the structure of $\mathcal{S}'$ is in some fashion different from that of $\mathcal{S}$; the functions plays a role of modification of the set of nodes (in form of the singletons) and the edges simultaneously.

Consider a transformation $\mathcal{S} \xrightarrow{f} \mathcal{S}'$. Then f is called

- *an addition of the node* y if $y \notin S$, $S' = S \cup \{y\}$, $E' = E$ and $K.f = K$ for $\emptyset \neq K$ and $\emptyset.f = \{y\}$,
- *an addition of the edge* $\{x, y\}$ if $S' = S$, $\{x, y\} \notin E$, $E' = E \cup \{x, y\}$ and $K.f = K$ for $\emptyset \neq K$ and $\emptyset.f = \{x, y\}$,
- *a deletion of the node* y if $y \in S$, $S' = S \setminus \{y\}$, $E' = E \cap S \times S$ and $K.f = K$ for $K \not\ni y$ and $K.f = \emptyset$ for $K \ni y$,
- *a deletion of the edge* $\{x, y\}$ if $S' = S$, $\{x, y\} \in E$, $E' = E \setminus \{x, y\}$ and $K.f = K$ for $K \neq \{y\}$ and $\{x, y\}.f = \emptyset$.

[2]Some graph theory texts, e.g. [1, 2, 20], use $N(Y)$, or "open" neighborhood, to denote $Y.\eta$, and $N[Y]$, or "closed" neighborhood, to denote $Y.\rho$. But, they don't define "neighborhoods" as operators as we do here.

In the case of transformation $\mathcal{S} \xrightarrow{f} \mathcal{S}'$ we adopt the notation of the set operators from the Section 1.2 exactly for the network $\mathcal{S}$ and with $'$-notation for the network $\mathcal{S}'$ (e.g., φ').

2.1. Continuous transformations

A transformation, f, is said to be *monotone* if $X \subseteq Y$ implies $X.f \subseteq Y.f$. All transformations will be assumed to be monotone.

A transformation f is said to be *continuous* if for all subsets $Y \subseteq \mathcal{S}$,

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

If this inclusion is valid for $Y = \{x\}$, $x \in S$, we say f is *continuous at* x. Although this definition of "continuity" has been well established, c.f. [10, 35, 47], it might merit some further justification. A continuous transformation f in a Euclidean world can be described by "for any set Y, if O' is an open set such that $Y.f \subseteq O'$, then there exists an open set O such that $Y \subseteq O$, and $O.f \subseteq O'$ ". In a discrete world, the notion of an "open" set is largely meaningless; but closure is not. In a discrete world we say f is continuous if "for any set Y, if C' is a closed set such that $Y.f \subseteq C'$ then there exists a closed set C such that $Y \subseteq C$ and $C.f \subseteq C'$". The expression $Y.\varphi.f \subseteq Y.f.\varphi'$ encapsulates this.

Proposition 2.1. *The composition $(f \cdot g)$ of continuous monotone transformations f and g is continuous.*

Proof can be found in [36, 38], where it is also shown that g must be monotone to ensure continuity.

Monotone, continuous transformations are mathematically quite well defined. They constitute a concrete category *MCont* with direct products [36]. They share some of the characteristics of continuous functions in the real Euclidean world; for example, the inverse image of a closed set in $\mathcal{S}'$ is closed in $\mathcal{S}$, or more precisely if $Y.f$ is closed in $\mathcal{S}'$, then $Y.\varphi.f = Y.f$ [38].

Suppose a transformation f is discontinuous, that is, there exists a set Y such that $Y.\varphi.f \not\subseteq Y.f.\varphi'$. Trying to test every such subset Y would be an exponential task. However, it was shown in [38] for extended (i.e. union-preserving) transformations,

Proposition 2.2. *If there exists Y such that $Y.\varphi.f \not\subseteq Y.f.\varphi'$ then there exists a singleton set $\{y\} \subseteq Y.\rho$ such that $\{y\}.\varphi.f \not\subseteq \{y\}.f.\varphi'$.*

This makes testing for continuity viable when the transformation is local. The notion of "continuity" is independent of the system $\mathcal{S}$ or the closure operator φ. However from now on, we will assume that $\mathcal{S}$ is an discrete network $\mathcal{N}$ and that φ is the neighborhood closure defined in (1.1). We will also assume that discrete, continuous functions are "evolutionary", that they are the "smooth", "low energy" transformations of the discrete world.

Proposition 2.3. *Let $\mathcal{S} = (S, E)$ be an undirected network. Let $x, y, z \in S$ and $\{x, z\} \notin E$. Then the transformation f which adds the edge $\{x, z\}$ is continuous at y if and only if ($y \in x.\eta$ and $x \in y.\varphi$) implies $z \in y.\eta$.*

Proof. Assume that $y \in x.\eta, x \in y.\varphi$ but $z \notin y.\eta$. Since $x \in y.\varphi$, $x.\eta \subseteq y.\rho$. But, because $z \notin y.\eta$, $x.\eta' \not\subseteq y.\rho'$ and $y.\varphi.f \not\subseteq y.f.\varphi'$, thus f is discontinuous at y. Conversely, assume f is discontinuous at $y \in x.\eta$; that is, $\exists w \in y.\varphi$ such that $w \notin y.\varphi'$, because $w.\eta' \not\subseteq y.\rho'$. Readily w is x or z. Since $z \notin x.\eta$ ($\{x,z\} \notin E$), $z \notin y.\varphi$. Hence $w \neq z$ and $w = x$. After adding the edge $\{x,z\}$, $x.\eta' \not\subseteq y.\rho'$ only if $z \notin y.\eta$, that is $z \notin y.\eta$. □

Proposition 2.4. *Let $\mathcal{S} = (S,E)$ be an undirected network and $\{x,z\} \in E$. Then the transformation f which deletes the edge $\{x,z\}$ is discontinuous at x if and only if either*

(a) $z \in x.\varphi$

(b) *there exists $y \in x.\varphi$, $y \neq x$, with $z \in y.\eta$.*

Proof. Suppose (a), $z \in x.\varphi$. Since $\{x,z\}$ is being deleted $z \notin x.\eta'$. Consequently, $\{x\}.\varphi.f \not\subseteq \{x\}.f.\varphi'$.
Suppose (b) and negation of (a). Hence $\exists y \in x.\varphi$, $y \neq x$, with $z \in y.\eta$ and since $z \notin x.\varphi$ we have $y \neq z$. Hence $y.\eta = y.\eta'$. Then $z \notin x.\eta'$ and $z \in y.\eta'$ implies that $y.\eta' \not\subseteq x.\eta'$, hence $y \notin x.\varphi' = x.f.\varphi'$. Now, $\{x\}.\varphi.f \not\subseteq \{x\}.f.\varphi'$, and f is discontinuous at x.
Conversely, suppose f is not continuous at x, i.e., $x.\varphi \not\subseteq x.\varphi'$. Hence there exists $y \in x.\varphi \setminus x.\varphi'$. Hence $y \neq x$ and either $y = z$ or $y \neq z$. In the former case we get $z \in x.\varphi$ (the condition (a)) and in the latter we have $y \in x.\varphi$ such that $z \in y.\eta$ (the condition (b)) since the edge $\{x,z\}$ makes the only difference of the networks and hence of their closures. □

Formal proofs of propositions 2.3 and 2.4 can also be found in [38]; however, understanding their content is more easily achieved by example. Suppose in Figure 1(a) that f introduces a new edge/relationship denoted by the dashed line. Here, $y_1 \in x.\eta$ and $x \in y_1.\varphi$, but $z \notin y_1.\eta$, so the condition for con-

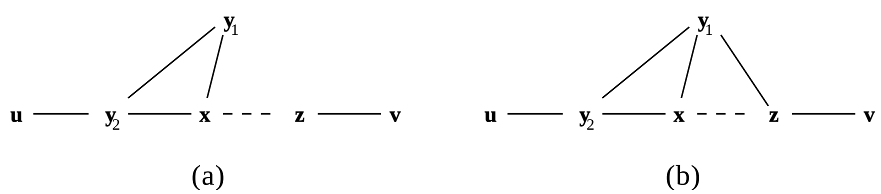

Figure 1. Two discontinuous transformations.

tinuity in Proposition 2.3 is not satisfied. We see that $y_1.\varphi = \{x, y_1\}$ since $x.\eta = \{y_1, y_2\} \subseteq \{x, y_1, y_2\} = y.\rho$. But, if f adds $\{x,z\}$ then $y_1.f.\varphi' = \{y_1\}$ since $x.f.\eta' = \{y_1, y_2, z\} \not\subseteq \{x, y_1, y_2\} = y_1.f.\rho'$. So $y_1.\varphi.f \not\subseteq y_1.f.\varphi'$ and f is not continuous.

Now suppose in Figure 1(b) that f deletes the edge $\{x,z\}$. By Proposition 2.4(b), f should be discontinuous. We verify by observing that $x.\varphi = \{x, y_1\}$ while $x.f.\varphi' = \{x\}$ because $y_1.\eta = \{x, y_2, z\} \not\subseteq \{x, y_1, y_2\} = x.\rho$. So $x.\varphi.f \not\subseteq x.f.\varphi'$.

2.2. An evolutionary example

By a *random network*, we mean one in which the probability of any pair of nodes, x and z, being connected by an edge $\{x, z\}$ is uniform. This is the Gilbert's classic definition [16]. Other definitions can be found in [5,6]. We consider such a random network to be the archetype of a *chaotic* network. While random networks are "chaotic", there can be chaotic networks that were not randomly generated.

To simulate evolutionary transformation we first generate a random network, $\mathcal{N}$, on n nodes with e edges. Figure 2 is a small such network with 56 randomly generated edges. We will use this network as a running example in the rest of this section. Given an initial random network, we then invoke a procedure E which

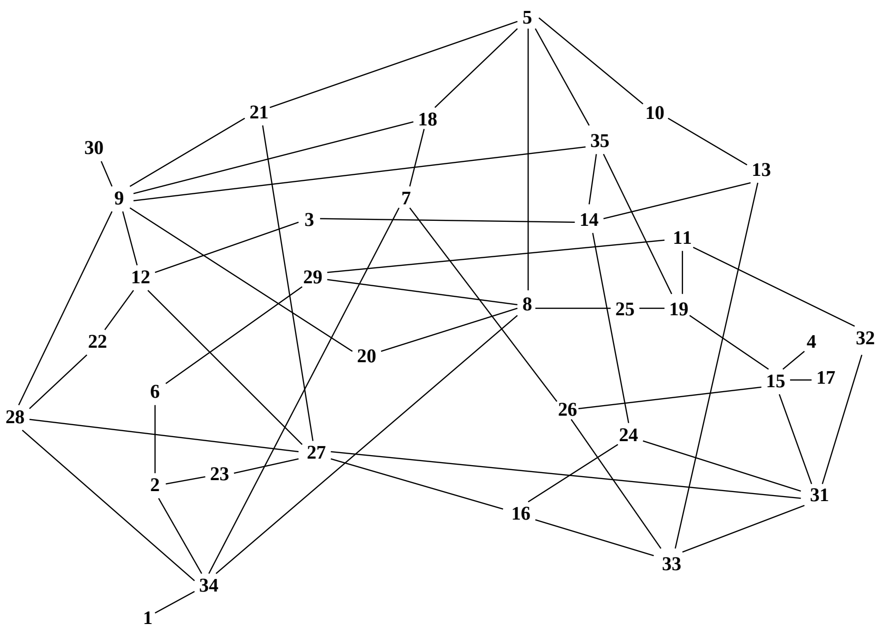

Figure 2. A randomly generated initial 35 point network $\mathcal{N}_0$.

deletes a randomly chosen edge $\{u, v\}$ and inserts a new edge $\{x, z\}$ between two randomly chosen points — subject only to the constraint that the addition and the deletion be continuous. If addition/deletion of the random edge would be discontinuous, as defined in propositions 2.3 and 2.4, another random edge is selected. Thus, both the number of points and number of edges always remains constant. This process, E, defines a transformation, $\mathcal{N} \xrightarrow{E} \mathcal{N}'$, which, by construction and Proposition 2.1, is continuous. By E^i, we mean E applied i times, that is $\mathcal{N}.E^i = \mathcal{N}.E.E.\ldots.E$.

In Figure 3 (a) and (b), we show our running example, $\mathcal{N}_0$ after iterating E 5 times and then 15 times.

To get a sense of this process, the five edges that have been deleted in the previous 5 iterations are dashed, and the five edges that have been added are

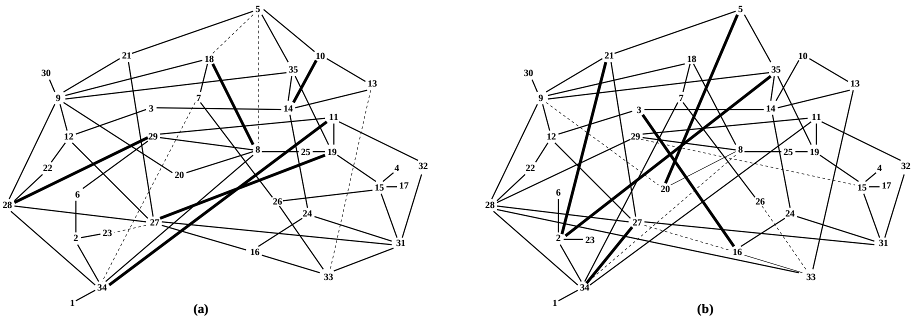

Figure 3. (a) $\mathcal{N}_0.E^5$; (b) $\mathcal{N}_0.E^{15}$; in both figures, recent additions are bold, recent deletions are dashed.

bolder. Node positions have been kept unchanged, even though the resulting diagrams have become less visually pleasing.

In Figure 4(a), we display the

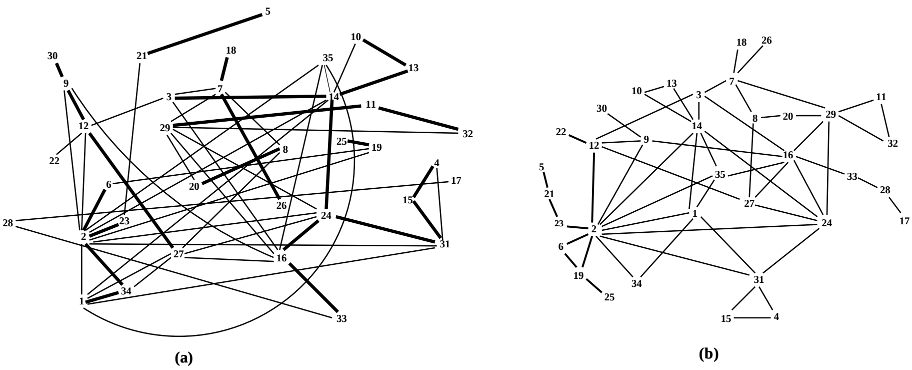

Figure 4. (a) $\mathcal{N}_0.E^{50}$; (b) $\mathcal{N}_0.E^{50}$ redrawn for greater clarity.

final version after 50 iterations; now the bolder edges denote those that were present in Figure 2, the original random network. Of the original 56 edges, only 25 remain.[3] It suggests the total amount of change. The *degree*, $d(y)$, of a node y is the number of incident edges. Since all versions of this graph have 56 edges, the expected $d(y) = (56/35) \times 2 = 3.2$. The degree of node 2, $d(2)$, has increased from 3 to 11; that of node 5 has dropped from 5 to just 1.

Figure 4(b) is a simpler rendition of the same graph. Only the nodes have been re-arranged for greater clarity. It can be difficult to verify that Figure 4(b) is identical to Figure 4(a); yet it is the clearer representation. In the next section we consider other methods for assessing network change.

[3]Some edges have been added and deleted multiple times. For example, $(13, 33)$ has been deleted in E^5 and re-instated in E^{15}, then deleted again in E^{50}!

3. Two network properties

Transformations, especially those which leave the number of nodes and connections unchanged, must in some way alter the internal structure of the network. In this section we consider two network properties that can reveal this change. They are the number of embedded triangles and the network's irreducible spine.

3.1. Triangle counts

A characteristic feature of random networks is the number of embedded triangles. Counting embedded triangles in networks has a rich literature, e.g. [43, 45]. As can be seen from Table 1, the expected number of embedded

e/n	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6
n									
35	1.25	2.00	3.10	5.85	7.70	10.65	13.85	17.60	22.55
100	0.90	1.45	3.30	4.25	7.05	9.40	12.15	16.45	21.35
500	1.36	2.45	3.50	5.73	8.16	11.55	14.50	19.25	24.40
1000	1.20	2.50	3.25	4.20	6.75	9.20	13.55	16.75	20.60

Table 1. Expected number of embedded triangles, $\exp(\tau(\mathcal{N}))$, in random networks.

triangles, $\tau(\mathcal{N})$, is essentially a function of the ratio e/n.[4] For relatively sparse networks, with $1 \leq e/n \leq 2$, we have $1 \leq \exp(\tau(\mathcal{N})) \leq 10$, apparently independent of n. For these sparse networks, $2 \leq \exp(d(y)) \leq 4$. When $e/n \geq 2$, $\exp(d(y)) \geq 4$, so every node is expected to be connected to at least 4 others, and $\exp(\tau(\mathcal{N}))$ begins to climb rather sharply. This is a familiar property of random graphs; below a certain threshold the probability of each additional edge introducing a specific property is low, after the threshhold it becomes high [6].

We observe that there are no triangles in the random graph of Figure 2, and that, after 5 iterations of the evolutionary process, one appears in Figure 3(a). It is $\{10, 13, 14\}$. After 15 iterations, we have 2 triangles in Figure 3(b); they are $\{10, 13, 14\}$ and $\{27, 28, 34\}$. There are 15 triangles in Figure 4(b), which shows the configuration after 50 iterations.

This is not an isolated phenomena. As shown in Table 2, as a network evolves continuously, it appears that the number of triangles strongly tends to increase, even though e/n remains constant. As Table 2 also shows, triangles can decrease during random evolution. But this seldom occurs since the edge $\{x, z\}$ in a triangle $\{x, y, z\}$ can be continuously deleted only if $\{y, z\} \subset \{x\}.\eta$, $\{x, z\} \subset \{y\}.\eta$, and $\{x, y\} \subset \{z\}.\eta$, which is relatively uncommon, because all three containments must be strict.

Each line of Table 2 denotes the behavior of a single randomly generated network. Consequently, we see the considerable variance natural in such random processes. Averaging over multiple executions, as in Table 1 would mask this variance. The key observation is that any process, continuous or not, which increases

[4]Entries in this table were created using a small sample size, $s = 20$. There was considerable variance. It is not definitive; it is only meant to be illustrative.

the number of edges, or connections, would be expected to increase the number of triangles. However, here the number of edges has been held constant, so it must be the constraint of continuity that induces the triangle increase.

E^i	$i =$	0	10	20	30	40	50
n	e/n						
35	1.2	1	3	4	7	9	18
35	1.4	4	5	8	15	17	17
35	1.6	6	3	10	16	17	18
35	1.8	4	9	14	12	15	17
35	2.0	11	13	18	19	25	26
E^i	$i =$	0	20	40	60	80	100
n	e/n						
100	1.2	1	3	9	14	16	18
100	1.4	3	5	6	9	11	13
100	1.6	4	5	13	15	14	20
100	1.8	8	12	18	20	21	24
100	2.0	6	11	14	13	19	21
E^i	$i =$	0	200	400	600	800	1000
n	e/n						
1000	1.2	3	3	6	9	18	22
1000	1.4	0	3	4	8	12	11
1000	1.6	5	5	7	9	12	19
1000	1.8	6	13	11	12	14	21
1000	2.0	14	17	19	18	20	19

Table 2. Observed numbers of embedded triangles, $\tau(\mathcal{N})$, $n = 35, 100, 1000$, after E^i iterations.

3.2. Irreducible spines

A reduction process R was introduced in [38, 39] as a way to eliminate local well-structured portions of a network, especially triangles, while retaining its global properties of connectivity. It is a transformation $\mathcal{N} \xrightarrow{R} \mathcal{N}'$ composed of iterative steps in which, for some chosen node y, all elements from $y.\varphi \setminus \{y\}$ are deleted. After such a process we end up with an "irreducible" sub-network $\mathcal{I} = \mathcal{N}.R \subseteq \mathcal{N}$.[5] A network is *irreducible* if every node y is closed, that is $\{y\}.\varphi = \{y\}$. In Figure 4(b) node 29 is not closed because $\{29\}.\varphi = \{11, 29, 32\}$. Node 29 is said to subsume 11 and 32; both of which will be deleted. In [39], it is shown that for all $\mathcal{N}$, its *irreducible spine*, $\mathcal{I}$, obtained by the reduction R is unique, upto isomorphism.

In Figure 5, the nodes x and z subsume each other. If $y_1, y_2 \in \mathcal{I}$ then either x or z will be in $\mathcal{I}$, but not both. Consequently, if $\mathcal{I}_1$ and $\mathcal{I}_2$ are both irreducible

[5]In [28], Lin, Soulignac and Szwarcfiter, speak of a "*dismantling* of a graph G as a graph H obtained by removing one dominated vertex of G, until no more dominated vertices remain". This is equivalent to R.

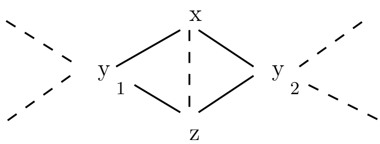

Figure 5. Mutual subsumption of two nodes.

spines of $\mathcal{N}$, then $x \in \mathcal{I}_1, x \notin \mathcal{I}_2$ implies $z \in \mathcal{I}_2.\eta$. Where not identical, they are "close", as are x and z in Figure 5. Thus $\mathcal{I}$ can be regarded as a true surrogate of $\mathcal{N}$.

An important property of the irreducible spine is that if $< x, \ldots, y, \ldots, z >$ is a path from x to z through y in $\mathcal{N}$ then there exists a path $< x', \ldots, y', \ldots, z' >$ in $\mathcal{I}$, and conversely. Because of this, it can be shown in [40] that, subject to reasonable constraints, the centers of the network, whether defined with respect to shortest path distance, or to betweenness, c.f. [7, 8, 15] will be located in $\mathcal{I}$.

So the irreducible spine of a network can convey considerable information about its global structure. But, not if $\mathcal{N}$ is chaotic. The nodes connected by dotted edges have been removed, the remaining solid lines of Figure 6(a) illustrate the irreducible spine of our running example as initially generated, see Figure 2. Figure

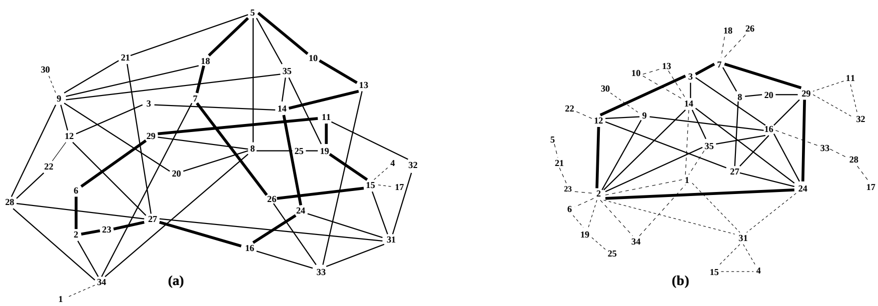

Figure 6. Irreducible spines of the initial random network of Figure 2, and evolved network of Figure 4(b). Dashed lines denote deleted sub-networks.

4(b) shows the network resulting from 50 iterations of the evolutionary process, and Figure 6(b) shows its irreducible spine, $\mathcal{I}$. The bolder portions of each figure denote a cycle of *maximal* length. In the former case the irreducible spine consists of almost the entire network.

One way of interpreting the irreducible spine of a network is by means of its cycle structure. A cycle C is *chordal* if for every path sequence $< \ldots, w, x, y, z, \ldots >$ in C, either $\{w, y\}$ or $\{x, z\} \in E$. Chordal graphs support simplicial ordering and greedy algorithms [1, 27, 29]. Our interest, however, is in "chordless" cycles. All nodes in an irreducible spine lie on a chordless cycle C_i of length $k \geq 4$, or on a unique path between two such cycles. The bold cycle of length 17 in Figure 6(a) is a chordless cycle of maximal length. There are 3 such cycles. The bold cycle of length 6 in Figure 6(b) is also maximal. Again, there happen to be three 6-cycles.

The increase in triangles reveals change in the "local structure" as a consequence of evolution. If we use the irreducible spine to represent global connectivity, then the change created by evolution is shown by comparing the resulting chordless cycle distributions. Figure 7 compares two histograms depicting the cycle structures of

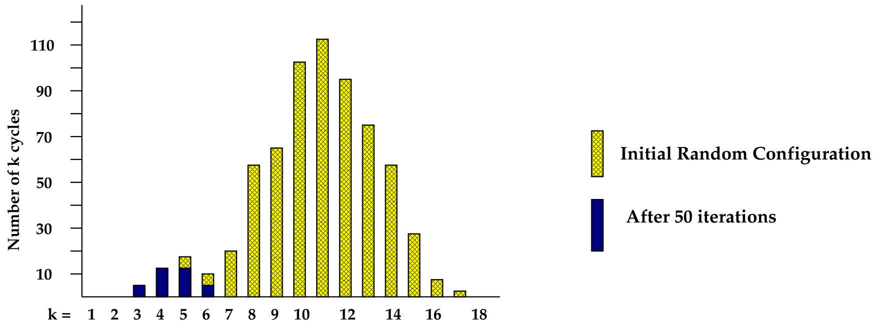

Figure 7. Distribution of chordless cycles in the irreducible spines of Figure 6(a) and Figure 6(b).

Figure 6 (a) and (b). We know of no way to quantify this change; but it is striking nevertheless.

For Figure 8, we first generated a random 100 node network with 140 edges. As expected, it had 3 embedded triangles; after 100 iterations the evolved network

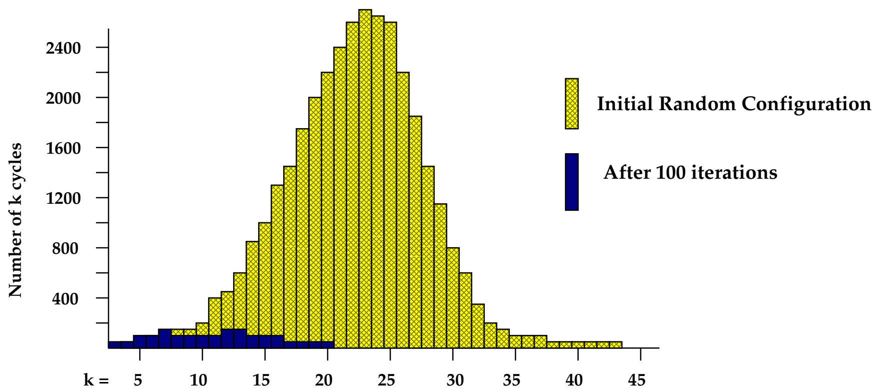

Figure 8. Distribution of chordless cycles in a 100 node network with 140 edges as originally generated and after executing E^{100}.

had 16 embedded triangles, while its irreducible spine had 6. Figure 8 then illustrates its cycle structure histograms. The distinctive, strongly unimodal initial distribution appears to be a characteristic feature of "chaotic" networks.

While it is not clear how to compare the cycle structures of the irreducible spine of a network, they do possess a curious mathematical property. Consider the nodes in a cycle C_i as a set. Since the cycles are chordless, $C_i \not\subseteq C_j$, for all $i \neq j$. Consequently, the collection $\mathcal{C} = \{C_1, \ldots, C_n\}$ of chordless cycles comprising an irreducible spine constitute an antichain. Anti-chains have been related to chains in a partially ordered set by Dilworths's theorem [5, 12].

We have been illustrating this "evolution" transformation with a small 35 node network, because it is hard to visualize in any detail the larger thousand node networks to which we have applied our procedure. To partially compensate for this, we present in Figure 9 the irreducible spine of a 200 node network with 320 edges that had initially been generated by our random process, then transformed by E^{1000}. The nodes removed by the reduction process, R, have been denoted

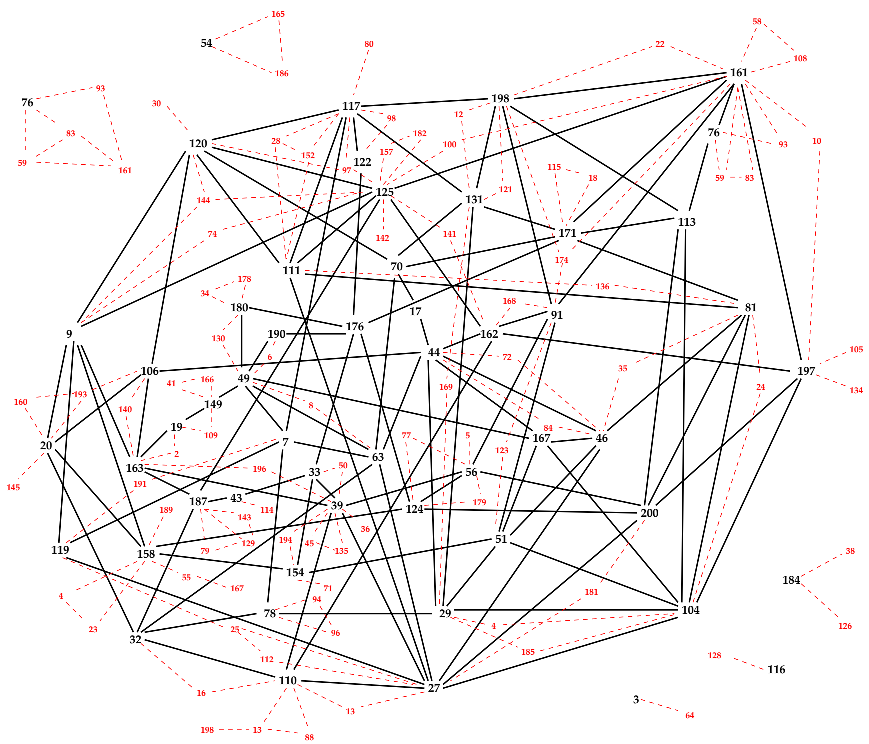

Figure 9. The irreducible spine of a random 200 node network after E^{1000}.

by a smaller font and connected by dashed lines. Neither the original random network, nor the evolved network, $\mathcal{N}.E^{1000}$, were connected; but only one main component had chordless k-cycles, $k \geq 4$. The other five components reduced to single points. As initially generated this network had 3 triangles, slightly fewer than the 4 to 5 suggested by Table 1. After E^{1000}, it had 99 triangles.

4. Discussion

The preceding sections describe a small experiment which, first, defines a non-trivial continuous network transformation, and second, seeks to establish some measures of its effect. We now reflect on its results.

The network structures we see around us are dynamic. There is no way of knowing whether the "evolutionary" process, E, which alternately adds and deletes connections, or edges, actually corresponds to any real such process. However, the resulting structures have a similarity to some social network structures. We have seen that when creating a relationship, or edge $\{x, z\}$, the existence of $y \in x.\eta$ with $y \in z.\eta$ is sufficient to ensure continuity, that is x and z are already related through y. This is commonly known as *triadic closure*. Triads were studied by Granovetter in [17], although he did not use the term "closure". It is not truly a closure operator (it is not idempotent); however, it appears to be a frequently occurring process in dynamic social systems [19,25,30,34]. The concept of continuity appears to provide a mathematical basis for this common observation. It should be contrasted with the approach of Newman, in [33], where a statistical generating function is used to model this kind of phenomena.

There is some literature on social network evolution [9, 19, 25]. It seems evident that in real life, much of an evolving social network may be shaped by external considerations, e.g. "proximity" or "social pressure". Nevertheless, the closure conditions of continuity may have real social significance.

Random neuronal activity is highlighted in the survey article "Perchance to Prune" [44]. The authors claim that such random neuronal firing while we sleep helps to better organize memory by weakening some connections and strengthening others. It appears to be similar to the "evolutionary" processes we have been exploring here.

The term "random" appears throughout this paper; we have random networks, random connections, additions and deletions. But, although we use a reasonably good linear congruence random number generator with various seeds to make these decisions, we make no pretense of ever achieving a uniform distribution. Our method of selecting edges to delete by first selecting a node, x, and then another node, z, within its neighborhood, is manifestly non-uniform over the set of all connections. Perhaps a better, if less euphonious, word would have been "unpredictable". At no point in the paper do we claim to make true probabilistic assertions.

Nevertheless, we believe that we have established that closure and continuity are interesting network concepts that provide a different approach to discrete network analysis.

Acknowledgement. The author would like to thank an anonymous referee for a number of insightful comments which were valuable additions to this paper.

References

[1] G. Agnarsson and R. Greenlaw, *Graph Theory: Modeling, Applications and Algorithms*, Prentice Hall, Upper Saddle River, NJ, 2007.

[2] M. Behzad and G. Chartrand, *Introduction to the Theory of Graphs*, Allyn and Bacon, Boston, 1971.

[3] M. Berger, *Convexity*, Am. Math. Mon. **97** (1990), 650–678. DOI: 10.2307/2324573

[4] N. Biggs, *Discrete Mathematics*, Clarendon Press, Oxford, 1989.

[5] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998.

[6] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics **73**, Cambridge Univ., 1998.

[7] S. Borgatti and M. Everett, *A graph-theoretic perspective on centrality*, Soc. Networks **28** (2006), 466–484. DOI: 10.1016/j.socnet.2005.11.005

[8] U. Brandes, *A faster algorithm for betweeness centrality*, J. Math. Sociol. **25** (2001), 163–177.

[9] U. Brandes, J. Lerner and T. Snijders, *Networks evolving step by step: Statistical analysis of dyadic event data*, in: ASONAM 2009, Athens, Greece, 2009, 200–205. DOI: 10.1109/ASONAM.2009.28

[10] G. Castellini, *Categorical Closure Operators*, Birkhäuser, Boston, 2003. DOI: 10.1007/978-1-4612-1370-3_8

[11] W. Coppel, *Foundations of Convex Geometry*, Australian Mathematical Society Lecture Series **12**, Cambridge Univ. Press, 1998.

[12] R. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. Math. **51** (1950), 161–166. DOI: 10.2307/1969503

[13] P. Edelman and R. Jamison, *The theory of convex geometries*, Geometriae Dedicata **19** (1985), 247–270. DOI: 10.1007/BF00149365

[14] M. Farber and R. Jamison, *Convexity in graphs and hypergraphs*, Siam. J. Algebra Discr. **7** (1986), 433–444. DOI: 10.1137/0607049

[15] L. Freeman, *Centrality in social networks, conceptual clarification*, Soc. Networks **1** (1978/79), 215–239. DOI: 10.1016/0378-8733(78)90021-7

[16] E. Gilbert, *Random graphs*, Annals of Math. Statistics **30** (1959), 1141–1144. DOI: 10.1214/aoms/1177706098

[17] M. Granovetter, *The strength of weak ties*, Am. J. Sociol. **78** (1973), 1360–1380. DOI: 10.1086/225469

[18] D. Hall and G. Spencer, *Elementary Topology*, Wiley, New York, 1955.

[19] N. Hanaki, A. Peterhansl, P. Dodds and D. Watts, *Cooperation in evolving social networks*, Manage. Sci. **53** (2007), 1036–1050. DOI: 10.1287/mnsc.1060.0625

[20] F. Harary, *Graph Theory*, Addison-Wesley, 1969.

[21] J. Kelley, *General Topology*, Van Nostrand, 1955.

[22] B. Korte and L. Lovász, *Greedoids, as structural framework for the greedy algorithm*, in: W. R. Pulleyblank (ed.), Progress in Combinatorial Optimization., Proc. of the Silver Jubilee Conf. on Combinatorics, Waterloo, Academic Press, 1984, 221–243.

[23] B. Korte and L. Lovász, *Intersections of matroids and antimatroids*, Discrete Math. **73** (1988/89), 143–157. DOI: 10.1016/0012-365X(88)90142-2

[24] G. Koshevoy, *Choice functions and abstract convex geometries*, in: Jean-Paul Doignon (ed.), Ordinal and Symoblic Data Analysis (OSDA 2000), Brussels, Belgium, 2000.

[25] G. Kossinets and D. Watts, *Empirical analysis of an evolving social network*, Science **311** (2006), 88–90. DOI: 10.1126/science.1116869

[26] S. Krantz, *A Guide to Topology*, Math. Assoc. of America, Washington, DC, 2009. DOI: 10.5948/UPO9780883859179

[27] R. Laskar and R. Jamison-Waldner, *Elimination orderings of chordal graphs*, Proc. Indian Stat. Inst., 1982, 192–200, in: Combinatorics and Applications, in honor of Prof. S.S. Shrikhande.

[28] M. Lin, F. Soulignac and J. Szwarcfiter, *Arboricity, h-index, and dynamic algorithms*, arXiv:1005.2211v1, 2010, 1–19.

[29] T. McKee, *How chordal graphs work*, Bull. Inst. Combin. Appl. **9** (1993), 27–39.

[30] G. Mollenhorst, B. Völker and H. Flap, *Shared contexts and triadic closure in core discussion networks*, Soc. Networks **34** (2012), 292–302. DOI: 10.1016/j.socnet.2011.09.001

[31] B. Monjardet, *Closure operators and choice operators: A survey*, in: Fifth Int. Conf. on Concept Lattices and their Applications, Montpellier, Lecture notes, France, 2007.

[32] B. Monjardet, *The duality between the antiexchange closure operators and the path independent choice operators on a finite set*, Math. Soc. Sci. **41** (2001), 131–150. DOI: 10.1016/S0165-4896(00)00061-5

[33] M. Newman, *Random graphs with clustering*, Phys. Rev. Lett. **103** (2009), 058701. DOI: 10.1103/PhysRevLett.103.058701

[34] T. Opsahl, *Triadic closure in two-mode networks: Redefining the global and local clustering coefficients*, arXiv:1006.0887v3 **27** (2011), 1–20.

[35] O. Ore, *Mappings of closure relations*, Ann. Math. **47** (1946), 56–72. DOI: 10.2307/1969035

[36] J. Pfaltz and J. Šlapal, *Transformations of discrete closure systems*, Acta Math. Hung. **138** (2013), 386–405. DOI: 10.1007/s10474-012-0262-z

[37] J. Pfaltz, *Closure lattices*, Discrete Math. **154** (1996), 217–236. DOI: 10.1016/0012-365X(96)00043-X

[38] J. Pfaltz, *Mathematical continuity in dynamic social networks*, in: A. Datta, S. Shulman, B. Zheng, S. Lin, A. Sun and Ee-P. Lim (eds.), Third Int. Conf. on Social Informatics, SocInfo2011, Lect. Notes Comput. Sci. **6984**, 2011, 36–50. DOI: 10.1007/978-3-642-24704-0_9

[39] J. Pfaltz, *Finding the mule in the network*, in: R. Alhajj and B. Werner (eds.), Int. Conf. on Advances in Social Network Analysis and Mining, ASONAM 2012, Istanbul, Turkey, 2012, 667–672. DOI: 10.1109/ASONAM.2012.111

[40] J. 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, Lect. Notes Comput. Sci. **6984**, Nanjing, PRC, 2013, 104–117.

[41] C. Plott, *Path independence, rationality and social choice*, Econometrica **41** (1973), 1075–1091. DOI: 10.2307/1914037

[42] A. Sen, *Social choice theory: A re-examination*, Econometrica **45** (1977), 53–88. DOI: 10.2307/1913287

[43] S. Suri and S. Vassilvitskii, *Counting triangles and the curse of the last reducer*. in: WWW 2011, Hyderabad, India, 2011. DOI: 10.1145/1963405.1963491

[44] G. Tononi and C. Cirelli, *Perchance to prune*, Sci. Am. **309** (2013), 34–39. DOI: 10.1038/scientificamerican0813-34

[45] C. Tsourakakis, U. Kang, G. Miller and C. Faloutsos, *Doulion: Counting triangles in massive graphs with a coin*, in: *KDD'09*, Paris, France, June 2009. DOI: 10.1145/1557019.1557111

[46] W. Tutte, *Introduction to the Theory of Matroids*, Amer. Elsevier, 1971.

[47] J. Šlapal, *A Galois correspondence for digital topology*, in: K. Denecke, M. Erné and S. L. Wismath (eds.), Galois Connections and Applications, Kluwer Academic, Dordrecht, 2004, 413–424.

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