Embracing n-ary Relations in Network Science

Jeffrey H. Johnson(✉)

Topdrim Project, The Open University, Milton Keynes MK7 6AA, UK
jeff.johnson@open.ac.uk
http://www.topdrim.eu

Abstract. Most network scientists restrict their attention to relations between pairs of things, even though most complex systems have structures and dynamics determined by n-ary relation where n is greater than two. Various examples are given to illustrate this. The basic mathematical structures allowing more than two vertices have existed for more than half a century, including hypergraphs and simplicial complexes. To these can be added hypernetworks which, like multiplex networks, allow many relations to be defined on the vertices. Furthermore, hypersimplices provide an essential formalism for representing multilevel part-whole and taxonomic structures for integrating the dynamics of systems between levels. Graphs, hypergraphs, networks, simplicial complex, multiplex network and hypernetworks form a coherent whole from which, for any particular application, the scientist can select the most suitable.

Keywords: n-ary relation · Graph · Hypergraph · Network · Simplicial complex · Multiplex network · Hypernetwork

1 Introduction

Given the success of graph and network theory since computers became available to scientists in the nineteen sixties, it is remarkable that the majority of the research done in network science has remained focussed on edges representing binary relations between two vertices. If all relations were binary relations this would be understandable. However, the structure and dynamics of many systems depend on relations between many things.

For example, the participants in a dinner party do not just interact in pairs. Nor do the member of a team or a committee. The members of a choir are not singing pairwise with the others. A great part of the dynamics of social and biological systems involves interactions between many individuals and many things. Surely a science of multidimensional universe should not be constrained to representing it solely through one dimensional objects.

Supported by the European *Dynamics of Multi-Level Complex Systems* (DyM-CS) FP7 FET programme, http://cordis.europa.eu/fp7/ict/fet-proactive/dymcs_en.html.

A. Wierzbicki et al. (Eds.): NetSci-X 2016, LNCS 9564, pp. 147–160, 2016.
DOI: 10.1007/978-3-319-28361-6_12

This is not to criticise networks in any way. As will become clear, they are part of a wider story that extends to hypergraphs, simplicial complexes and hypernetworks. It begins with graphs.

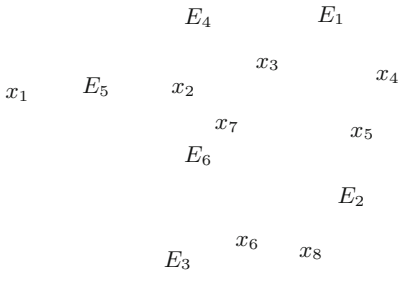
In the literature the terminology for graph theory varies considerably. Here, let a *graph*, G, be defined to be a set, V with elements called *vertices* and a set, E, of pairs of vertices called edges. Write $G = (V, E)$. Let a and b be vertices and let $\{a, b\}$ be an edge Graphs are usually drawn with dots such as a and b representing vertices and, for example, a line joint a and b to represent the edge $\{a, b\}$. Usually the edges in graphs represent *binary relations* between the vertices. To go beyond binary relation something else is required.

2 Hypergraphs

Hypergraphs represent an early attempt to allow graph edges to have more than two vertices [7]. Berge writes 'The idea of looking at a family of sets from this standpoint took shape around 1960. In regarding each set as a "generalised edge" and in calling the family itself a "hypergraph", the initial idea was to try to extend certain classical results of Graph Theory. ... Next it was noticed that this generalisation often led to simplification; moreover, one single statement ... could unify several theorems on graphs" [8]. In his 1969 paper [7] he gives the following definition. 'Let $X = \{x_1, x_2, ..., x_n\}$ be a finite set. A *hypergraph* on X is a family $H = (E_1, E_2, ..., E_m)$ of subsets of X such that

(1) $E_i \neq \emptyset \qquad (i = 1, 2, ..., m)$
(2) $\bigcup_{i=1}^{m} = X.$

The elements $x_1, x_2, ..., x_m$ are called *vertices* and the sets $E_1, E_2, ..., E_m$ are the edges of the hypergraph.' Berge gives the example shown in Fig. 1 where the relationship between the vertices and edges is given as an incidence matrix.



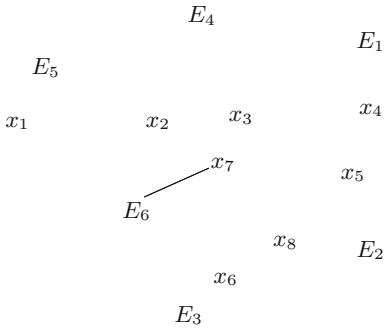
(a)

$$V = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{array} \begin{array}{c} \begin{array}{cccccc} E_1 & E_2 & E_3 & E_4 & E_5 & E_6 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

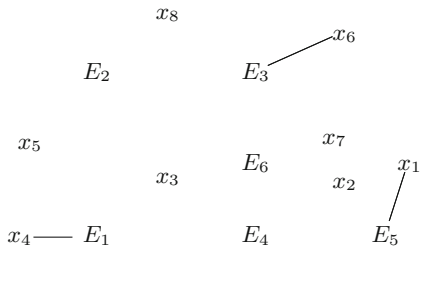
(b)

Fig. 1. The Berge hypergraph

Berge's method of drawing hypergraphs is a hybrid between graph-theoretic links and loops, and hypergraph-theoretic sets. Figure 2(a) shows the Berge hypergraph drawn entirely as sets. Here the edges corresponding to pairs of



(a) $H_E(V, R)$



(a) $H_V(E, R^{-1})$

Fig. 2. The dual Berge hypergraphs

vertices are shown as sets, namely $\{x_1, x_2\}$ and $\{x_5, x_8\}$, and the loop from x_7 to itself is draw as a singleton set $\{x_7\}$ which is the edge E_6.

Figure 2(a) shows the hypergraph with the columns of the incidence matrix as the edges. The *dual hypergraph* has sets of edges corresponding to the vertices as shown in Fig. 2(c). Looking along the rows, each vertex is related to a set of edges, for example x_7 is related to the set of edges $\{E_3, E_4, E_6\}$ This is a 'dual' edge in the *dual hypergraph*, as shown in Fig. 2(b).

The Galois Lattice. Figure 3 shows a set of arches, $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7\}$ with each arch made from a subset of the blocks $B = \{b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9, b_{10}, b_{11}, b_{12}\}$. Let arch a_i be R-related to block b_j if it contains that block. This bipartite relation can be represented by the incidence matrix shown in Fig. 4. The entry in the i^{th} row and the j^{th} column of the matrix is one if a_i is related to b_j, and it zero otherwise. Let $E(a_i)$ be the set of blocks related to arch a_i. Then:

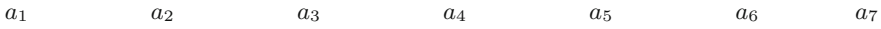


Fig. 3. Arches related to the blocks used to construct them

$E(a_1) = \{b_1, b_3, b_4\}$ $E(a_2) = \{b_2, b_3, b_4\}$ $E(a_3) = \{b_3, b_4, b_5\}$
$E(a_4) = \{b_4, b_5, b_6, b_7\}$ $E(a_5) = \{b_7, b_8, b_9, b_{10}\}$ $E(a_6) = \{b_7, b_8, b_9, b_{11}\}$
$E(a_7) = \{b_7, b_8, b_9, b_{12}\}$.

Apart from these 'first order' edges it is interesting to generate 'higher order' edges from all their intersections:

$E(a_1) \cap E(a_2) \cap E(a_3) = \{b_3, b_4\}$ $E(a_1) \cap E(a_2) \cap E(a_3) \cap E(a_4) = \{b_4\}$
$E(a_3) \cap E(a_4) = \{b_4, b_5\}$ $E(a_4) \cap E(a_5) \cap E(a_6) \cap E(a_7) = \{b_7\}$
$E(a_5) \cap E(a_6) \cap E(a_7) = \{b_7, b_8, b_9\}$

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}
a_1	1	0	1	1	0	0	0	0	0	0	0	0
a_2	0	1	1	1	0	0	0	0	0	0	0	0
a_3	0	0	1	1	1	0	0	0	0	0	0	0
a_4	0	0	0	1	1	1	1	0	0	0	0	0
a_5	0	0	0	0	0	0	1	1	1	1	0	0
a_6	0	0	0	0	0	0	1	1	1	0	1	0
a_7	0	0	0	0	0	0	1	1	1	0	0	1

Fig. 4. Maximal rectangles in the arch-block structure

Let the set of first order and higher order edges be called the *augmented hypergraph* for the relation in Fig. 5. The edges of the augmented dual hypergraph can be found in a similar way:

$$E(b_1) = \{a_1\} \quad E(b_2) = \{a_2\} \quad E(b_3) = \{a_1, a_2, a_3\}$$
$$E(b_4) = \{a_1, a_2, a_3, a_4\} \quad E(b_5) = \{a_3, a_4\} \quad E(b_6) = \{a_4\}$$
$$E(b_7) = \{a_4, a_5, a_6, a_7\} \quad E(b_8) = \{a_5, a_6, a_7\} \quad E(b_9) = \{a_5, a_6, a_7\}$$
$$E(b_{10}) = \{a_5\} \quad E(b_{11}) = \{a_6\} \quad E(b_{12}) = \{a_7\}$$

$$E(b_1) \cap E(b_3) \cap E(b_4) = \{a_1\} \qquad E(b_2) \cap E(b_3) \cap E(b_4) = \{a_2\}$$
$$E(b_3) \cap E(b_4) = \{a_1, a_2, a_3\} \qquad E(b_3) \cap E(b_4) \cap E(b_5) = \{a_3\}$$
$$E(b_4) \cap E(b_5) = \{a_3, a_4\} \qquad E(b_4) \cap E(b_5) \cap E(b_6) \cap E(b_7) = \{a_4\}$$
$$E(b_7) \cap E(b_8) \cap E(b_9) = \{a_5, a_6, a_7\} \qquad E(b_7) \cap E(b_8) \cap E(b_9) \cap E(b_{10}) = \{a_5\}$$
$$E(b_7) \cap E(b_8) \cap E(b_9) \cap E(b_{11}) = \{a_6\} \qquad E(b_7) \cap E(b_8) \cap E(b_9) \cap E(b_{12}) = \{a_7\}$$

Bringing together the sets in the augmented hypergraphs shows that they can be put is one-to-one correspondence. This is known as the *Galois connection* and the *Galois pairs* can be listed as:

$$\{b_1, b_3, b_4\} \leftrightarrow \{a_1\} \qquad \{b_7, b_8, b_9, b_{10}\} \leftrightarrow \{a_5\} \qquad \{b_3, b_4\} \leftrightarrow \{a_1, a_2, a_3\}$$
$$\{b_2, b_3, b_4\} \leftrightarrow \{a_2\} \qquad \{b_7, b_8, b_9, b_{11}\} \leftrightarrow \{a_6\} \qquad \{b_7, b_8, b_9\} \leftrightarrow \{a_5, a_6, a_7\}$$
$$\{b_3, b_4, b_5\} \leftrightarrow \{a_3\} \qquad \{b_7, b_8, b_9, b_{12}\} \leftrightarrow \{a_7\} \qquad \{b_4\} \leftrightarrow \{a_1, a_2, a_3, a_4\}$$
$$\{b_4, b_5, b_6, b_7\} \leftrightarrow \{a_4\} \qquad \{b_4, b_5\} \leftrightarrow \{a_3, a_4\} \qquad \{b_7\} \leftrightarrow \{a_4, a_5, a_6, a_7\}$$

In a Galois pair $A' \leftrightarrow B'$ every a in A' is R-related to every b in B'. Therefore the rows and columns of the matrix can be rearranged so that all the a_i in A' are contiguous and all the b_j in B' are contiguous, with the corresponding rectangle of entries in the matrix all ones. For example, let $A' = \{a_1, a_2, a_3\}$ and $B' = \{b_3, b_4\}$. Then as shown in Fig. 4 the corresponding rectangle is filled with ones because each of a_1, a_2 and a_3 is related to b_3 and b_4.

The rectangle corresponding to $A' = \{a_1, a_2, a_3\} \leftrightarrow B' = \{b_3, b_4\}$ is *maximal*. Two other maximal rectangles are shown in Fig. 4 corresponding to the Galois

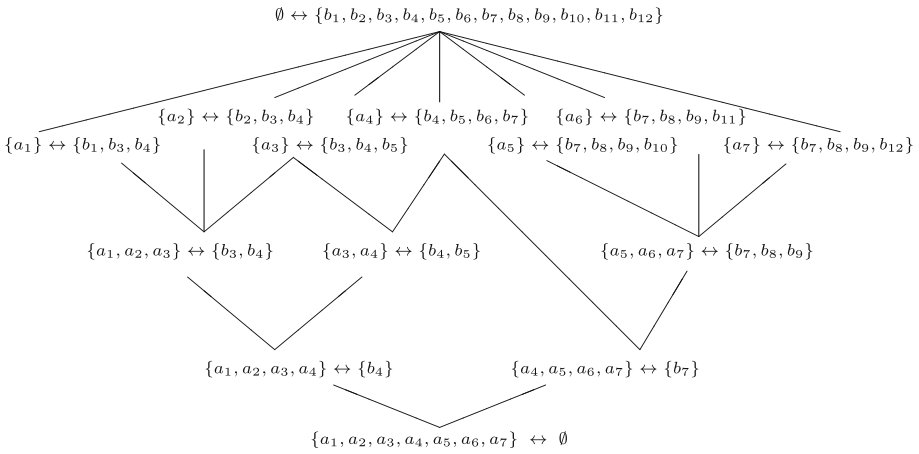


Fig. 5. The Galois Lattice for the arch-block relation of Fig. 4

pairs $\{a_3, a_4\} \leftrightarrow \{b_4, b_5\}$ and $\{a_5, a_6, a_7\} \leftrightarrow \{b_7, b_8, b_9\}$. The maximal rectangles $A' \leftrightarrow B'$ where A' has just one element or B' has just one element are not shown.

The Galois pairs can be arranged as a *Galois lattice* [13] with upwards set inclusion on the left and downward set inclusion on the right (Fig. 5).

Galois pairs are particularly interesting, since they are sites of relatively high connectivity. However for relations between large sets there can be a combinatorial explosion of Galois pairs making computation difficult. Nonetheless Galois pairs play an important role in hypernetwork theory [17].

Hypergraphs are an excellent first step towards mathematical structure able to represent n-ary relations. However they are essentially set-theoretic and have no orientation. Simplicial complexes provide this.

3 Simplicial Complexes

In the nineteen fifties C.H. Dowker published the paper *The homology groups of relations* [11] which showed that relations between n things could be represented by multidimensional polyhedra with n vertices, such as those shown in Fig. 6. This idea lay dormant for a quarter of a century until in the nineteen sixties R.H. Atkin introduced the revolutionary idea that social relations could be represented by polyhedra. For example, a business deal between three people can be represented by a triangle, written as $\langle a, b, c\rangle$, the relation of four people playing music together can be represented by a tetrahedron, $\langle a, b, c, d\rangle$, and the relationship between five people working together as a team can be represented by a 5-hedron, $\langle a, b, c, d, e\rangle$. This idea is entirely compatible with network theory since, for example, a relationship between two people having a conversation can be represented by a polyhedron with two vertices, namely a line or an edge, $\langle a, b\rangle$. These ideas first appeared in the article *A mathematical approach towards a social science*, published in the *Essex Review* in 1968 [1].

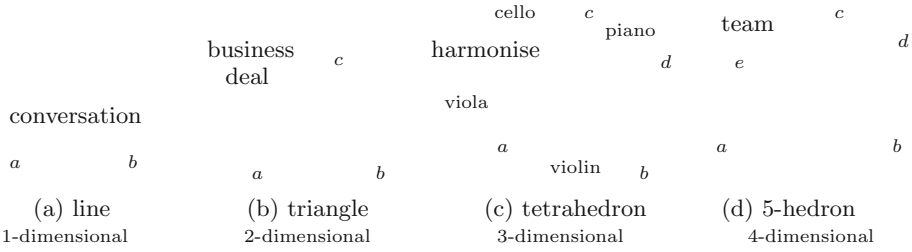


Fig. 6. Simplices can represent relations between two or more things

Polyhedra are the geometric realisation of more abstract objects called *simplices*. Let V be a set of vertices. An abstract *p-simplex* is determined by a set of $p+1$ vertices, written as $\langle v_0, v_1, ..., v_p \rangle$. Simplices are often represented by the symbol σ.

A simplex σ is a *face* of a simplex σ', $\sigma \lesssim \sigma'$, if every vertex of σ is also a vertex of σ'. For example the 2-dimensional simplex $\langle x_1, x_2, x_3 \rangle$ is a triangular face of the tetrahedron representing the 3-dimensional simplex $\langle x_1, x_2, x_3, x_4 \rangle$. A set of simplices with all its faces is called a *simplicial complex*.

Algebraic Topology. In algebraic topology simplices provide an algebraic way of calculating the topological invariants of spaces. The ideas will be briefly and informally sketched here. Figure 7 shows a complex made up of three triangles with all their faces (lines and vertices). This complex has the topological feature of a *hole* surrounded by the triangles.

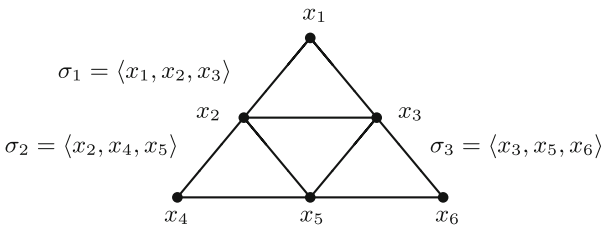


Fig. 7. A hole in a simplicial complex.

A q-dimension chain is an expression of the form $\Sigma_{i \in I}\, n\, \sigma_i$ where n is a number. The *boundary* operator, ∂, maps a simplex to its boundary according to the rule $\partial\langle x_0, ..., x_p \rangle = \Sigma_{i=0}^{p}(-1)^i \langle x_0, ..., \hat{x_i}, ..., x_p \rangle$, where $\hat{x_i}$ means omit the i^{th} entry along, counting from zero. For example, $\partial\langle x_1, x_2, x_3 \rangle = \langle x_2, x_3 \rangle - \langle x_1, x_3 \rangle + \langle x_1, x_2 \rangle$. This chain of 1-simplices is called a *cycle*.

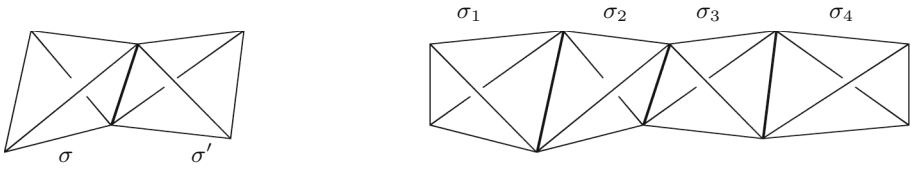
In algebraic topology switching a pair of vertices changes the sign (and orientation) of a simplex, so $-\langle x_1, x_3 \rangle = \langle x_3, x_1 \rangle$. Thus the cycle can be written as $\langle x_2, x_3 \rangle + \langle x_3, x_1 \rangle + \langle x_1, x_2 \rangle$. In this case it is a *bounding cycle* because it is a closed loop of 1-simplices that goes round the shaded 2-dimensional triangle. It starts at $\langle x_2 \rangle$ and goes to $\langle x_3 \rangle$ along the oriented edge $\langle x_2, x_3 \rangle$, goes to x_1 along

the oriented edge $\langle x_3, x_1\rangle$ and back to close the loop at x_2 along the oriented edge $\langle x_3, x_2\rangle$.

The boundary operator is *nilpotent*, *i.e.* when applied twice it gives zero. For example, $\partial^2\langle x_1, x_2, x_3\rangle = \partial\langle x_2, x_3\rangle - \partial\langle x_1, x_3\rangle + \partial\langle x_1, x_2\rangle = \langle x_3\rangle - \langle x_2\rangle - \langle x_3\rangle + \langle x_1\rangle + \langle x_2\rangle - \langle x_1\rangle = 0$.

Any chain c with $\partial c = 0$ is defined to be a *cycle*. Apart from bounding cycles as seen above, there can be *non-bounding cycles*. For example consider $c = \langle x_2, x_5\rangle + \langle x_5, x_3\rangle + \langle x_3, x_2\rangle$. Then $\partial c = \langle x_5\rangle - \langle x_2\rangle + \langle x_3\rangle - \langle x_5\rangle + \langle x_2\rangle - \langle x_3\rangle = 0$ and c is a cycle. However there is no 2-dimensional chain c' with $\partial c' = c$ so c is a non-bounding cycle. In general, non-bounding cycles correspond to holes, in this case exactly the hole bounded by c.

Atkin's Q-analysis. In the early seventies Atkin and coworkers investigated the topological properties of relations in the context of town planning. Atkin suggested a new kind of connectivity based on the shared faces of social polyhedra [3–5].



(a) σ and σ' are 1-near (b) σ_1 and σ_4 are 1-connected

Fig. 8. q-connected polyhedra

Two simplices are *q-near* if they share a q-dimensional face. Two simplices are *q-connected* if there is a chain of pairwise q-near simplices between them. The tetrahedra σ and σ' are 1-near in Fig. 8(a) because they share an edge, or 1-dimensional face. In Fig. 8(b) the tetrahedra σ_1 and σ_4 are 1-connected, since σ_1 is 1-near σ_2, σ_2 is 1-near σ_3, and σ_3 is 1-near σ_4. A *Q-analysis determines* classes of *q-connected components*, sets of simplices that are all q-connected. An early application of Q-analysis studied land uses in Colchester [6].

Backcloth and Traffic. The vertices and edges of networks often have numbers associated with them. For example in a social network the vertices may be associated with the amount of money a person has and the edges may be associated with how much money passes between pairs of people. In electrical networks the vertices have voltage associated with them and the edges have current. Although the network's voltages and currents may change, the network itself does not. Similarly in a road network the daily traffic flows may vary but usually the network infrastructure does not. The same holds for simplicial complexes when there are patterns of numbers across the vertices and the simplices. The numbers may change when the underlying simplicial complex does not.

Atkin suggested that the relatively unchanging network or simplicial complex structure be called a *backcloth* and that the numbers be called the *traffic* of activity on the backcloth. As an example, the airline network acts as a backcloth

to the traffic of airline passengers. The term backcloth comes from the scenery painted on large canvas sheets used in theatres as a static backdrop behind the actors.

Atkin first used simplicial complexes to characterise a wide variety of phenomena in physics by his *Cocycle Law* that the space-time backcloth supporting many physical phenomena has no holes. His conceptual leap "from cohomology in physics to q-connectivity in social science" was published in 1972 [2].

Flows and q-transmission as Multidimensional Percolation. Networks are excellent for representing and calculating the dynamics of flows, from electricity to oil to cars and sentiments. Simplicial complexes are multidimensional networks and they too can carry equally diverse traffic flows. Generally the q-connectivity of the underlying backcloth constraints the dynamics of the flows. This has been called *q-transmission* and has been described as a multidimensional analogue analogue to percolation in networks [17].

Example: Road Accidents. A study of road accidents illustrates the combinatorial nature of simplices [17]. Drivers who had been involved in accidents were interviewed to find out the possible causes. The telephone interviews were unstructured with the interviewer eliciting the causes from the interviewees, *e.g.* interviewees would often would volunteer that they were going too fast for the conditions. Some typical examples of the 57 reported accident simplices are:

⟨mechanical failure, need to stop, lack anticipation, stress; R_1⟩
⟨carelessness, unexpected manoeuvre; R_8⟩
⟨change in road layout, poor signposting, bad visibility; R_{16}⟩
⟨speed, lack of concentration; R_{23}⟩
⟨inexperienced driver, car in wrong position; R_{31}⟩
⟨poor visibility, lack of caution, road wet; R_{23}⟩
⟨not paying attention, to near/too fast, brakes poor, unexpected manoeuvre; R_{51}⟩
⟨narrow road, speed R_{53}⟩

These combinations of causes were expressed in everyday language. The data was analysed according to the classes:

D1–Stress	D2–carelessness	D3–Poor anticipation
D4–Too close	D5–Looking wrong way	D6–Alcohol
D7–Health/Tiredness	D8–Young male ego	D9–Inexperience
D10–Unfamiliarity with vehicle	D11–Cyclist blind	D12–In a hurry
D13–Unfamiliar with road	D14–Speed	D15–Mistaken priority
V1–Mechanical failure	R1–Difficult configuration	R2–Poor visibility
R3–Poor signposting	R4–Difficult surface	R6–Heavy traffic
A1–Unexpected event	A2–Slow vehicle in front	

Like hypergraphs, simplicial complex also have Galois pairs:

⟨D2–Carelessness, R1–Difficult configuration⟩	⟨2, 5, 9, 12, 35, 40, 42, 51, 57⟩
⟨D1–Stress, R1–Difficult configuration⟩	⟨1, 2, 20, 26, 34, 51, 52⟩
⟨D2–Carelessness, R2–Poor visibility⟩	⟨2, 3, 4, 35, 38, 40⟩
⟨D14–Speed, R1–Difficult configuration⟩	⟨10, 12, 22, 39, 43, 53⟩

⟨D1–Stress, R2–Poor visibility⟩	⟨2, 3, 11, 13, 26⟩
⟨R1–Difficult configuration, R2–Poor visibility⟩	⟨2, 26, 35, 40, 43⟩
⟨R2–Poor visibility, R4–Difficult road surface⟩	⟨11, 13, 26, 36, 38⟩
⟨R2–Poor visibility, A1–Unexpected event⟩	⟨11, 13, 16, 36, 54⟩
⟨R2–Poor visibility, R3–Poor signposting⟩	⟨2, 16, 26, 56⟩
⟨D1–Stress, D13–Unfamiliar with road⟩	⟨2, 3, 25, 52⟩
⟨D2–Carelessness, A1–Unexpected event⟩	⟨1, 9, 10, 41⟩
⟨R2–Poor visibility, R4–Difficult road surface⟩	⟨11, 13, 26⟩
⟨R2–Poor visibility, R4–Difficult road surface, A1–Unexpected event⟩	⟨11, 13, 36⟩
⟨D2–Carelessnes, R1–Difficult configuration, R2–Poor visibility⟩	⟨2, 35, 40⟩

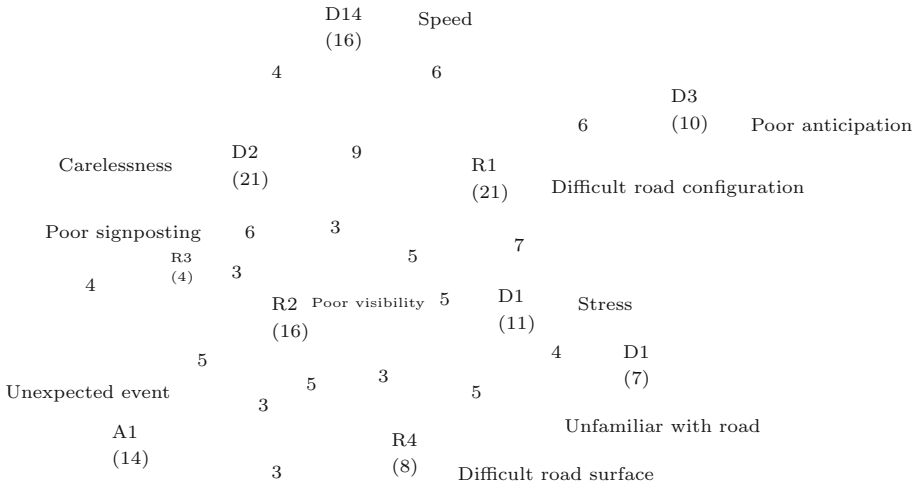


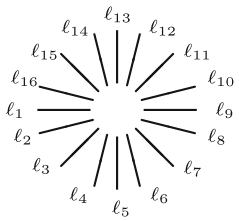
Fig. 9. Frequencies of occurrences of accident factors

Figure 9 gives a graphical summary of the Galois pairs and the numbers accidents associated with the simplices. The interviewees were asked to rate the importance of the factors on a five-point low-high scale. For example, σ(Accident-2) = ⟨D1–Stress(5), D2–Careless(3), D13–Unfamiliar road(5), D15–Mistaken priority(5), R1–Difficult config(5), R2–Poor visibility(3), R3–Poor signposting(5)⟩, and σ(Accident–2) = ⟨D1–Stress(5), D2–Careless(4), D6–Alcohol(1), D7–Tired(5), D13–Unfamiliar road(3), D15–Speed(3) R2–Poor visibility(2)⟩. Let $\mu(v_i)$ be the weighting given to accident factor v_i, $\mu(v_i)$. A value on the whole simplex, the *fuzzy conjunction*, can be defined as $\mu\sigma = \min\{\mu(v_i) \,|\, v_i \lesssim \sigma\}$. Then for a fuzzy value of 3, σ(Accident-2) and σ(Accident-3) share the face ⟨D1-Stress, D2-Careless, D13-Unfamiliar road⟩, and they are 3-fuzzy 2-near.

4 Hypernetworks

Figure 10(a) shows the lines $\ell_1, ..., \ell_{16}$ arranged in a circle by the relation R_1. The resulting structure $\langle \ell_1, ..., \ell_{16}; R_1 \rangle$ has the emergent property that most people see a white disk at the centre of the lines, the so-called *sun illusion*.

Figure 10(b) shows the same set of lines assembled under a different relation, R_2. Now there is no disk but a rectangle shape emerges. This example illustrates that the same ordered set of elements can be the subject of more than one relation, and that the simplex notation $\langle \ell_1, ..., \ell_{16} \rangle$ cannot discriminate these very different cases.



(a) The sun illusion
$\sigma_1 = \langle \ell_1, ..., \ell_{16}; R_1 \rangle$

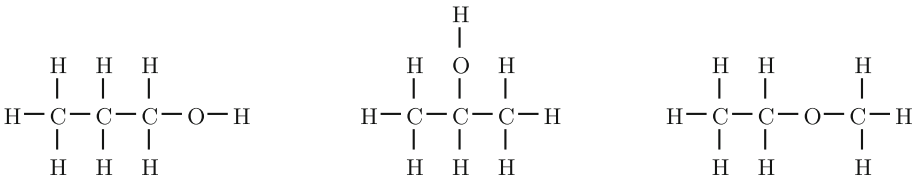
(b) the rectangle illusion
$\sigma_2 = \langle \ell_1, ..., \ell_{16}; R_2 \rangle$

Fig. 10. The lines $\ell_1, ..., \ell_{16}$ organised by two different relations, R_1 and R_2

In order to do this another symbol is necessary to represent the relation. We write $R_1 : \langle \ell_1, ..., \ell_{16} \rangle \rightarrow \langle \ell_1, ..., \ell_{16}; R_1 \rangle$ and $R_2 : \langle \ell_1, ..., \ell_{16} \rangle \rightarrow \langle \ell_1, ..., \ell_{16}; R_2 \rangle$. Let σ_1 represent the sun configuration and σ_2 represent the rectangle configuration. Then σ_1 and σ_2 are examples of *relational simplices*, or *hypersimplices*. Now the notation enables σ_1 to be discriminated from σ_2, since $\sigma_1 \neq \sigma_2$.

In general a *hypernetwork* is defined to be any collection of hypersimplices. This definition is deliberately undemanding, so that almost anything can be a hypersimplex, and any collection of hypersimplices can be a hypernetwork.

Example: Chemical Molecules. Chemical molecules illustrate the idea of hypersimplices. For example, propanol assembles three carbon atoms with eight hydrogen atoms and one oxygen atom, written as C_3H_8O or C_3H_7OH. Figure 11 shows the atoms of propanol arranged in a variety of ways. The first two show the isomers *n*-propyl alcohol and isopropyl alcohol. The oxygen atom is attached to an end carbon in the first isomer and to the centre carbon in the second, but the C-O-H hydroxyl group substructure is common to both. The rightmost isomer of C_3H_8O, methoxyethane, has the oxygen atom connected to two carbon atoms and there is no C-O-H substructure. This makes it an ether, methyl-ethyl-ether, rather than an alcohol. Thus the relational simplices of the isomers have



(a) *n*-propyl alcohol

(b) isopropyl alcohol

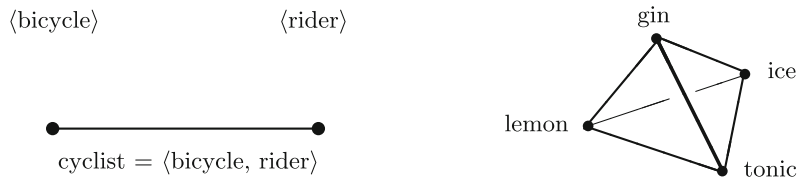
(c) methyl-ethyl-ether

Fig. 11. Chemical isomers as relational simplices

the same vertices, but the assembly relations are different. n-propyl alcohol and isopropyl alcohol share the hydroxyl group substructure C-O-H and are similar, but methyl-ethyl-ether does not and has different properties. Thus

$$\begin{aligned}
&\langle \text{C, C, C, H, H, H, H, H, H, H, H, O}\,;\, R_{n-\text{propylalcohol}}\rangle \quad \neq \\
&] \\
&\langle \text{C, C, C, H, H, H, H, H, H, H, H, O}\,;\, R_{\text{isopropylalcohol}}\rangle \quad \neq \\
&\langle \text{C, C, C, H, H, H, H, H, H, H, H, O}\,;\, R_{\text{methyl}-\text{ethyl}-\text{ether}}\rangle
\end{aligned}$$

The Vertex Removal Test for n-ary Relations. The essential feature of a polyhedron is that it ceases to exist if any of the vertices are removed. For example, consider a cyclist represented as the combination $\langle$rider, bicycle; $R_{\text{riding}}\rangle$. Remove either the man or the bicycle and what is left ceases to be a cyclist. Removing a vertex is like sticking a pin in a balloon, causing the structure to collapse and whatever is left is not the whole simplex. Remove any vertex from $\langle$gin, tonic, ice, lemon; $R_{\text{mixed}}\rangle$ and it ceases to be the perfect gin and tonic. Generalising edges to polyhedra allows a distinction to be made between the *parts* of things represented by vertices, and *wholes* represented by hypersimplices. Using this test it is easy to find many examples of n-ary relations, *e.g.* a path with n edges in a network forms a hypersimplex - remove an edge and the path ceases to exist; four bridge players form a hypersimplex - remove one and the game collapses; and a car and its wheels are 5-ary related - without any of them it won't work.



(a) Remove a vertex and the cyclist simplex ceases to exist

(b) Remove a vertex and the perfect gin and tonic ceases to exist

Fig. 12. Remove a vertex and the simplex ceases to exist.

5 Hypernetworks and Multilevel Structure

Hypersimplices enable the definition of multilevel part-whole structures, *e.g.* the four blocks assembled by the 4-ary relation R to form an arch in Fig. 13. Here the whole has the emergent property of a gap not possessed by any of its parts. If the parts exist in the system at an arbitrary *Level N* then the whole exists at a higher level, here shown as *Level N+1*. Thus assembly relations provide an immutable upwards arrow for the definition of multilevel structure.

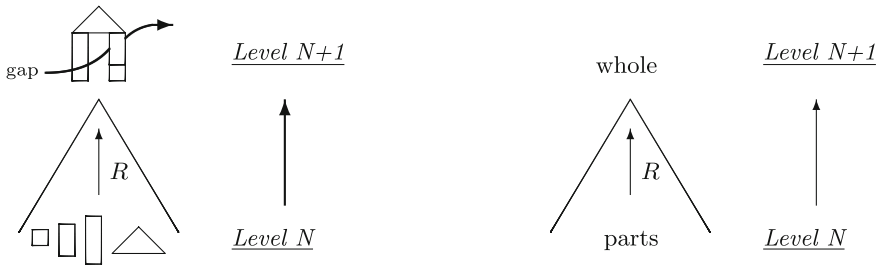


Fig. 13. The fundamental part-whole diagram of multilevel aggregation

Part-whole aggregations are interleaved with taxonomic aggregations, as shown in Fig. 14. The aggregation between *Level N* and *Level N+1* combines graphical parts to form faces. The aggregation between *Level N+1* and *Level N+2* establishes classes of faces in a taxonomy. Such aggregations depend on the purpose of the taxonomy. For example, there is no class of 'frowny' faces because, for the purpose here, it is not required. Note that part-whole aggregation require *all* the parts. In contrast taxonomic aggregations require just one example to aggregate. For example, the round smiley face is sufficient for there to be a smiley face, irrespective of whether or not there is a square smiley face.

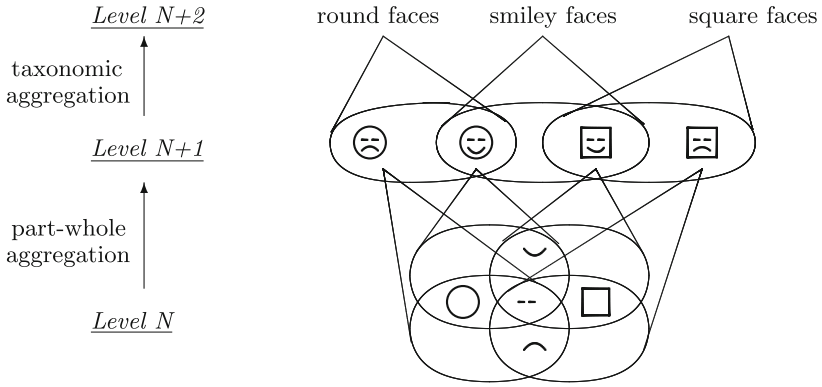


Fig. 14. Part-whole and taxonomic aggregation

6 Embracing *n*-ary Relations in Network Science

Despite the mathematics literature on multi-vertex relational structure dating back at least to the 1950s, and despite the efforts of visionaries such as Berge and Atkin in the 1960s, today many scientists still shy away from relations between more than two things. It is all the more remarkable because graph theorists have known about this mathematics but not adopted it, *e.g.* in his classic book on graph theory, Harary [14] quotes Veblen's 1922 book [19] as a source for his definition of simplicial complex but, frustratingly, notes in passing that a graph is a

one-dimensional simplicial complex, even though Veblen explicitly considers two-dimensional simplicies in the second chapter of his book. In contrast, computer science recognises the importance of n-ary relations, *e.g.* Codd [9] uses them in his seminal paper on relational data structures, and the WC3 consortium defines their use in the semantic web [15].

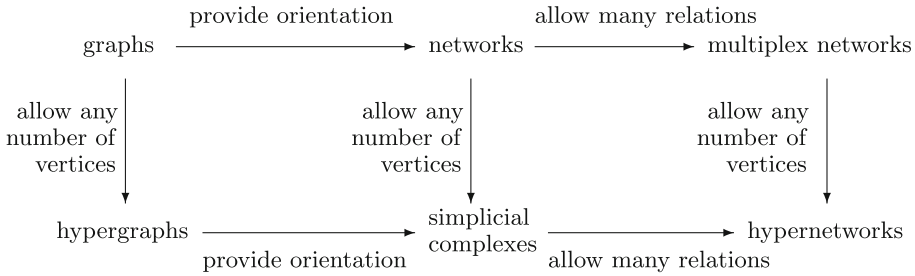


Fig. 15. The natural family of network structures embraces n-ary relations

It is unfortunate that network scientists should neglect n-ary relations since they are part of a natural family of network structures (Fig. 15). Assuming appropriate definitions, providing orientation makes a non-oriented graph into a network, and allowing pairs of vertices to support many relations makes multiplex networks. Vertically, allowing edges to have many vertices generalises graphs to hypergraphs, allowing oriented edges to have many vertices generalises networks to simplicial complexes, and allowing oriented edges supporting many relations to have many vertices generalises multiplex networks to hypernetworks. Horizontally, orienting the edges of hypergraphs creates simplicial families and complexes, and allowing a simplex to support many relations creates hypernetworks. Thus the diagram in Fig. 15 commutes and these structures form a natural family by adding structure from top left to bottom right.

Hopefully this paper will stimulate more interest in n-ary relations in network science:

- many systems involve n-ary relations – ignoring this misrepresents them
- n-ary relations are essential for representing part-whole structures and related dynamics in multilevel systems
- there is a rich and coherent mathematical theory for n-ary relations - with many remaining challenges and opportunities for the network community.

References

1. Atkin, R.H., Bray, R., Cook, I.: 'A mathematical approach towards a social science', The Essex Review, University of Essex, Autumn 1968, no. 2, 3–5 (1968)
2. Atkin, R.H.: From cohomology in physics to Q-connectivity in social science. Int. J. Man-Mach. Stud. **4**(2), 139–167 (1972)
3. Atkin, R.H.: Mathematical Structure in Human Affairs. Heinemann Educational Books, London (1974)

4. Atkin, R.H.: Combinatorial Connectivities in Social Systems. Birkhäuser (Basel), Basel (1974)
5. Atkin, R.H.: Multidimensional Man. Penguin Books, Harmondsworth (1981)
6. Atkin, R.H., Johnson, J.H., Mancini, V.: An analysis of urban structure using concepts of algebraic topology. Urban Stud. **8**, 221–242 (1971)
7. Berge, C.: Sur certains hypergraphes généralisant les graphes bipartites. In: Erdös, P., Rhényi, A., Sós, V.T. (eds.) Combinatorial Theory and its Applications I, (Proceedings of the Colloquium on Combinatorial Theory and its Applications, 1969), pp. 119–133, North-Holland (1970)
8. Berge, C.: Hypergraphs: Combinatorics of Finite Sets. Elseiver, Amsterdam (1989)
9. Codd, E.F.: A relational model of data for large shared data banks. Commun. ACM **16**(6), 377–387 (1970)
10. De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivela, M., Moreno, Y., Porter, M.A., Gómez, S., Arenas, A.: Mathematical formulation of multilayer networks. Phys. Rev. X **3**, 041022 (2013). http://journals.aps.org/prx/pdf/10.1103/PhysRevX.3.041022
11. Dowker, C.H.: The homology groups of relations. Ann. Math. **56**(1), 84–95 (1952)
12. Freeman, L.C., White, D.R., Romney, A.K.: Research Methods in Social Network Analysis. Transaction Publishers, New Brunswick (1991)
13. Freeman, L.C., White, D.R.: Using Galois lattices to represent network data. Sociol. Methodol. **23** (1993). ISBN 1-55786-464-0, ISSN 0081-1750, http://eclectic.ss.uci.edu/~drwhite/pw/Galois.pdf
14. Harary, F.: Graph Theory, (third printing 1972). Addison-Wesley, Reading (1969)
15. Hayes, P., Welty, C.,: Defining N-ary relations on the semantic web. W3C Working Group Note (12 April 2006). http://www.w3.org/TR/swbp-n-aryRelations
16. Johnson, J.H.: Hypernetworks for reconstructing the dynamics of multilevel systems. In: European Conference on Complex Systems 2006 (25–29 September 2006), Oxford. http://oro.open.ac.uk/4628/1/ECCS06-Johnson-R.pdf
17. Johnson, J.H.: Hypernetworks in the Science of Complex Systems. Imperial College Press, London (2014)
18. Lee, K.-M., Mina, B., Gohb, K.-I.: Towards real-world complexity: an introduction to multiplex networks. Eur. Phys. J. B **88**, 48 (2015). doi:10.1140/epjb/e2015-50742-1
19. Veblen, O.: Analysis Situs. American Mathematical Society Colloquium Lectures, Volume 5, The Cambridge Colloquium, 1916. Part II. Reprinted facsimile, Leopold Classic Library. http://www.leopoldclassiclibrary.com