Chapter 8

Networks Beyond Pairwise interactions

Summary

- High-order relations
- Simplicial Complexes
- Hypergraphs
- Revising Basilar Concepts
- Outlook

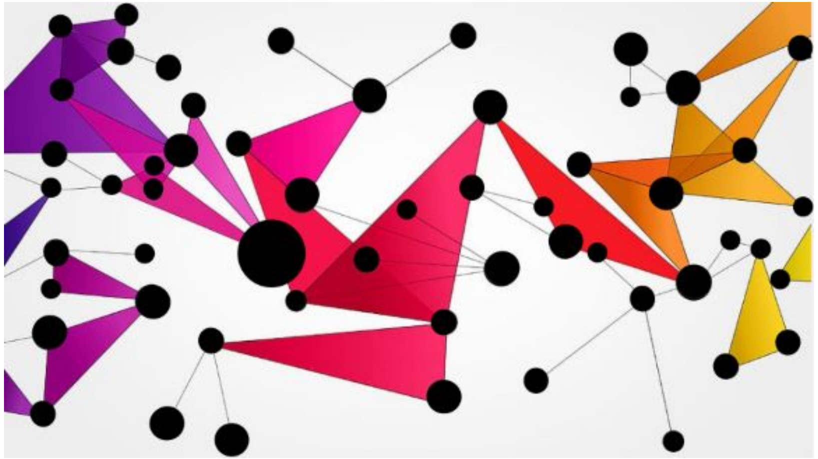
Reading

- See suggested papers

High-order

As all models, graphs are just proxies we can use to reason on real phenomena.

As such, sometimes, they are not expressive enough, thus making impossible to describe specific patterns.



Graphs describes “pairwise” interactions among nodes…

- What about many-to-many interactions?
- Are “cliques” enough?

Examples of

High-order interactions

Type: Online Debate
Nodes: Individuals
Links: Group Discussions

Type: Scientific Collaborations
Nodes: Researchers
Links: Co-Authorships

Type: Actor connectivity
Nodes: Actors
Links: Cast jointly

Type: Mobility
Nodes: Individuals
Links: Co-Location

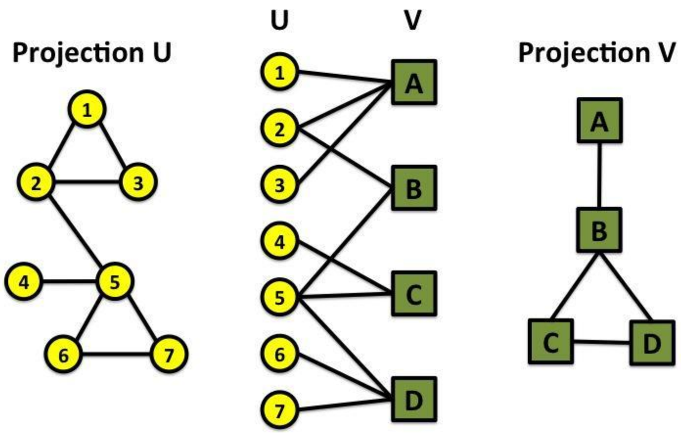
Recap:

Bipartite Graphs

Bipartite graph (or bigraph)

a graph whose nodes can be divided into two disjoint sets *U* and *V* such that every link connects a node in *U* to one in *V*.

Examples
Hollywood actor network
Collaboration networks
Disease network (diseasome)

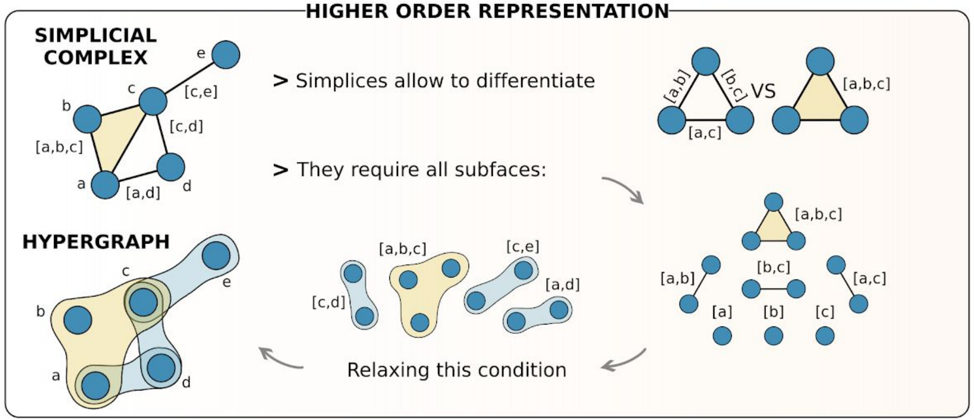


Projection
Two nodes of the **same class** are connected by a (weighted) edge if they share at least a **common neighbor**

High-order modeling

Two major frameworks:

- Simplicial Complexes
- Hypergraphs



High-order Interaction:

a set I = $[p_0, p_1, \ldots, p_{k-1}]$ containing an arbitrary number k of basic elements

Simplicial Complexes

Terminology

Simplicial complexes are collections of simplices σ

A collection of n simplices

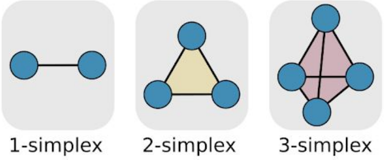
$$K = \{\sigma_0, \sigma_1 \ldots \sigma_n\}$$

is a valid simplicial complex if, for every k-simplex

$$\sigma = [p_0, p_1, \ldots, p_k] \in K,$$

all its subfaces of any dimensions belong to K too

For example, if the triangle [a, b, c] ∈ K, then we also require [a],[b],[c],[a, b],[a, c],[b, c] to belong to K.



		Symbol
Components	Simplices	σ
System	Simplicial Complex	K

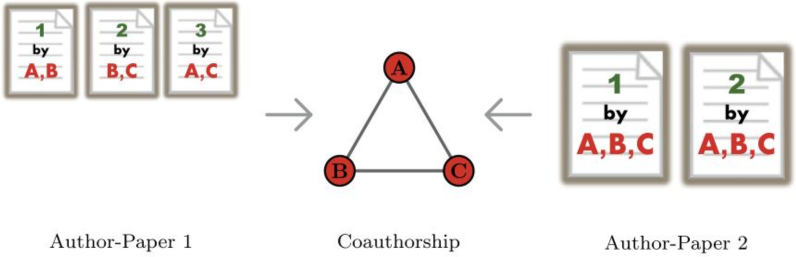
What is a Simplicion?

Simplifying: Cliques (almost correct)

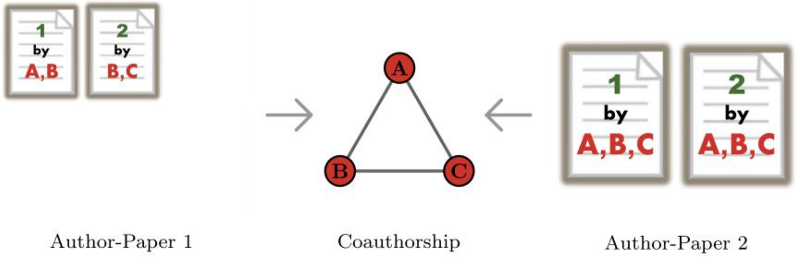
Cliques make impossible to distinguish the two cases in which a sub-face is present or not (clique complex)

Exact intuition:
Cliques that guarantee all induced sub-cliques exist.

Example- Collaborations in scientific papers



[A,B,C] **is** a simplicion

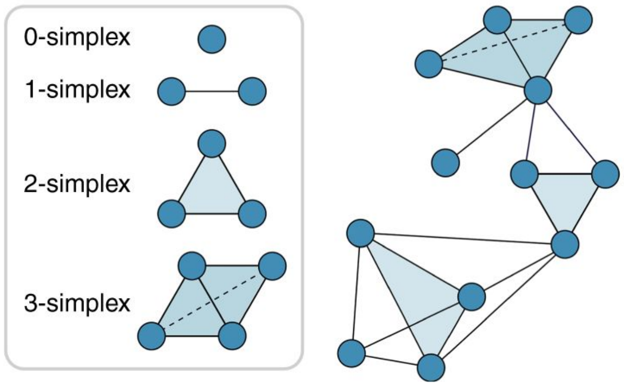


[A,B,C] **is not** a simplicion

Strong requirements...

Although simplicial complexes overcome some of the problems encountered by other lower dimensional representations, they are still quite limited by the requirement on the existence of all subfaces.

- Subfaces information often unavailable
- Time may play a role...



Hypergraphs

(let's relax a little bit...)

Terminology

Nodes:
individual system entities

Hyperedge:
relation involving a nonempty set of nodes

k-uniform hypergraph:
hypergraph having all hyperedges of size k

Conversely from Simplicial Complexes,
Hypergraphs can include the k-hyperlinks (e.g., [a, b, c])
without any requirement on the existence of (k-1)-hyperlinks
(e.g., [a, b],[a, c] and [b, c])

		Symbol
Components	nodes, vertices	N
Interactions	hyperedges, hyperlinks	L
System	hypergraph	H=(N,L)

Adjacency matrix

Paper #	Authors
1	Andrews, Davis
2	Andrews, Carter, Davis
3	Davis
4	Andrews, Bailey
5	Bailey, Carter

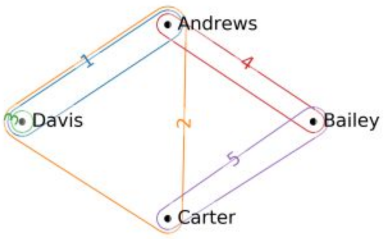
	Andrews	Bailey	Carter	Davis
Andrews		Y	Y	Y
Bailey	Y		Y	
Carter	Y	Y		Y
Davis	Y		Y	

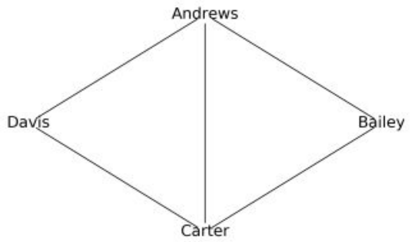
Same of “standard” graphs

It encodes pairs of adjacent nodes within the hypergraph.

Issue:

it does not allow to efficiently discriminate the graph from the hypergraph representation





Incidence matrix

Let's give a name to each hyperedge!

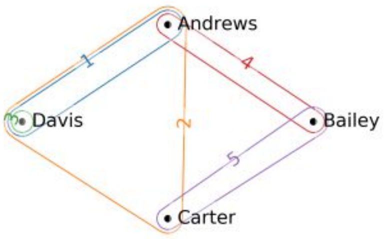
Nodes = {a, b, c, d}
Hyperedges = {1, 2, 3, 4, 5}

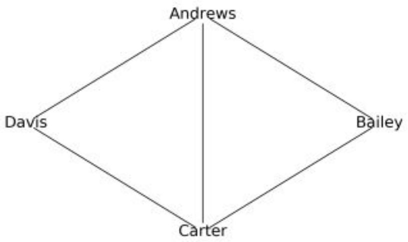
Incidence matrix associate each node to the set of hyperedges it belongs to.

Columns describe the hyperedge components.

Paper #	Authors
1	Andrews, Davis
2	Andrews, Carter, Davis
3	Davis
4	Andrews, Bailey
5	Bailey, Carter

	1	2	3	4	5
a	1	1	0	1	0
b	0	0	0	1	1
c	0	1	0	0	1
d	1	1	1	0	0





Relevant Transformations

Dual Hypergraph (H*)

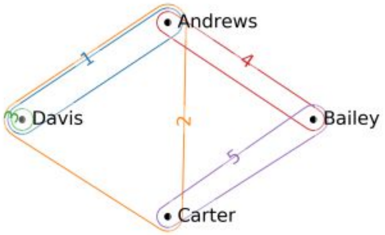
- H* nodes are H hyperedges;
- H* nodes are connected if they share in H a same subset of elements

Projections

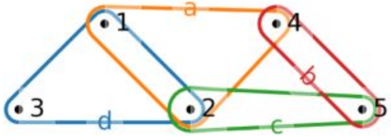
- Projections of the bi-partite graph encoding the hypergraph

Induced Hypergraphs

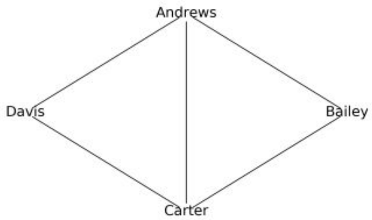
- Nodes (H_n):
  subset of H composed by all those hyperedges that contains $N_c \subseteq N$
- Hyperedges (H_h):
  subset of H composed by all those hyperedges $L_h \subseteq L$



H

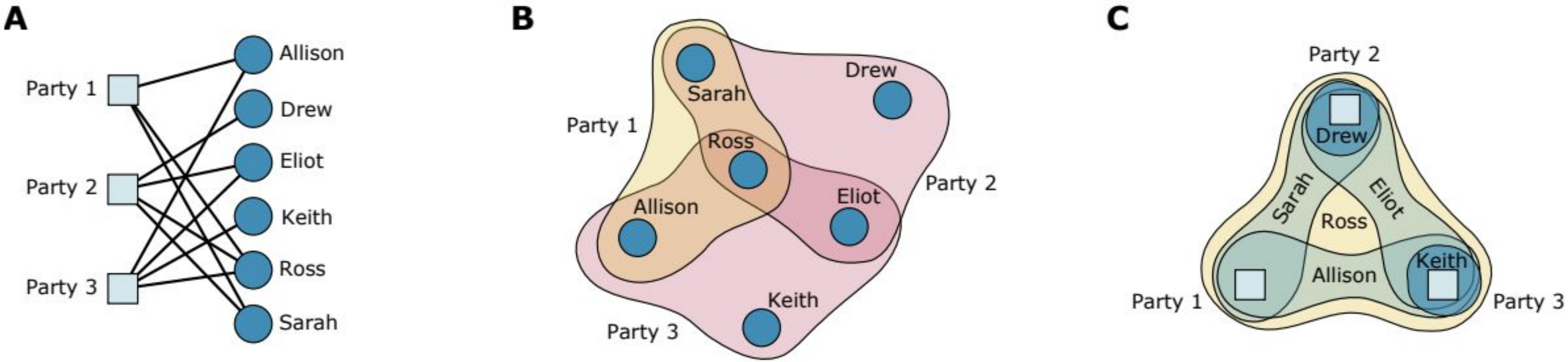


H*



H
left-projection

Affiliation network of six children and three parties



Example: Understanding Projections, H and H*

s-Analysis Framework

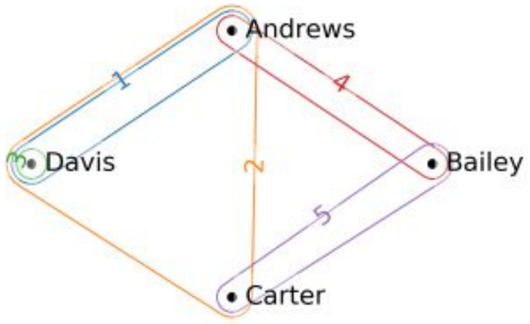
Goal:

Extend results from “classic” graph theory to hypergraphs.

- Multiple strategies (i.e., frameworks) to approach that goal
- Complete “conservative” extensions are not always possible

Idea:

- Leverage the incidence matrix;
- Measuring functions parametric on the number (“s”) of nodes shared by pairs of hyperedges.



Example:

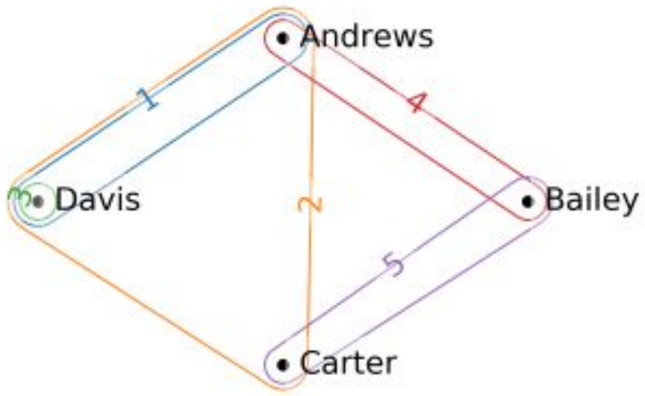
- 1 and 2 are 2-incident (they share {Davis, Andrews})
- 5 and 4 are 1-incident (they share {Bailey})

s-Neighbors

Neighbors that share at least “s” hyperedges with the considered node

In the example:

- Γ(Davis, 1) = {Andrews, Carter}
- Γ(Davis, 2) = {Andrews}
- Γ(Andrews, 1) = {Davis, Carter, Bailey}
- Γ(Andrews, 2) = {Davis}



Line graph L(H)

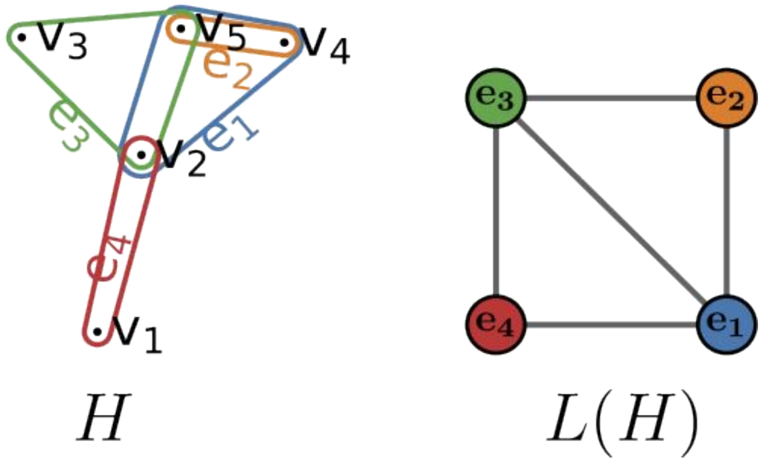
Right projection of H
(i.e., projection on the hyperedges)

- Projection nodes are the hyperedges;
- Two nodes are connected if they are incident

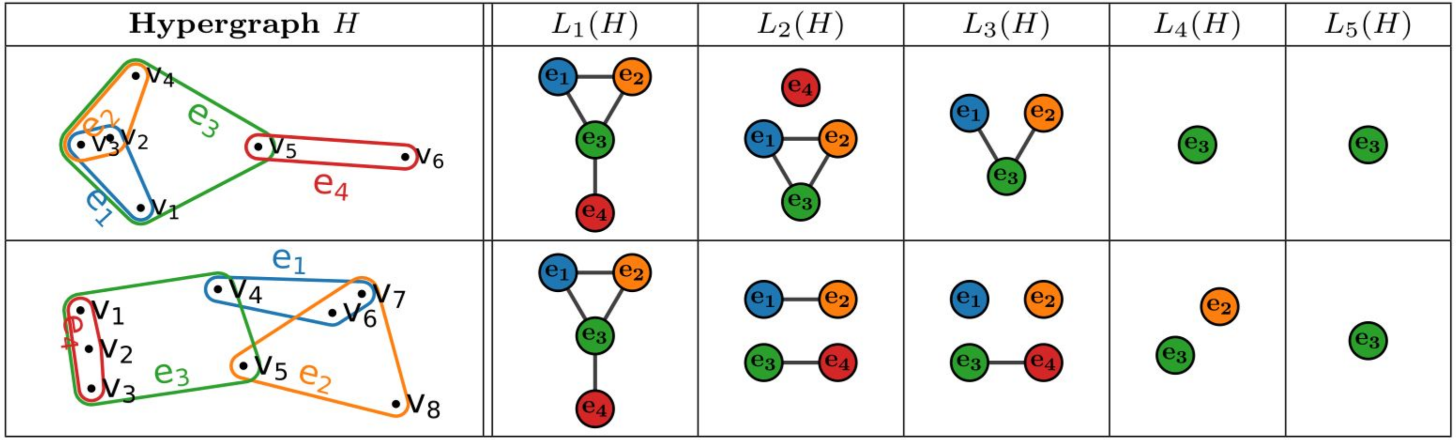
or, equivalently,
2-section of the dual hypergraph H*
(i.e., the graph reconstructed by the adjacency matrix of H*)

Line graph, L(H)

- Projection nodes are the hyperedges;
- Two nodes are connected if they are “s”-incident



s-Line graph



s-Line graph, $L_s(H)$

- Projection: nodes are the hyperedges;
- Two nodes are connected if they are “s”-incident

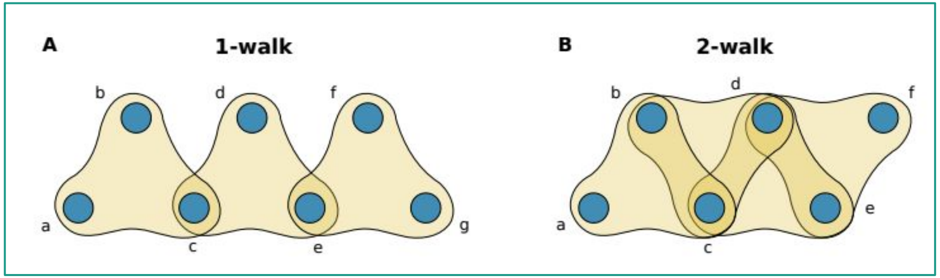
s-walks

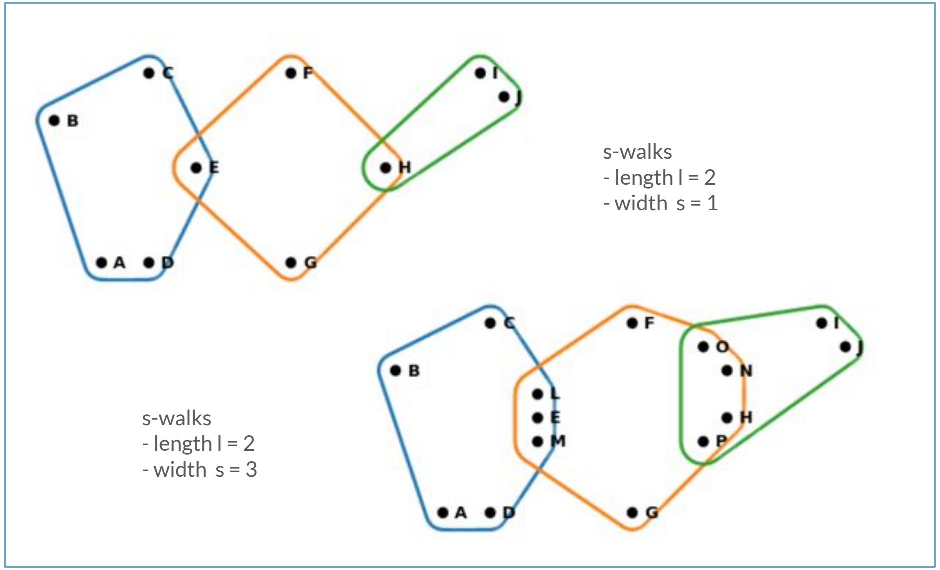
s-walk

a sequence of edges $e_0, e_1, \ldots, e_l$ such that

$$s \leq inc(e_i, e_{i+1}) \text{ for all } 0 \leq i \leq l-1$$

Walks in hypergraphs are characterized not only by length l, indicating the distance of interaction, but also by "width" s, indicating a strength of interaction.





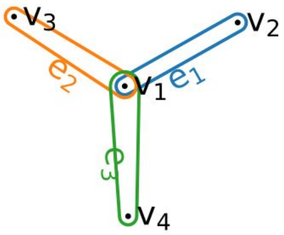
s-walks and L(H)

Issue:

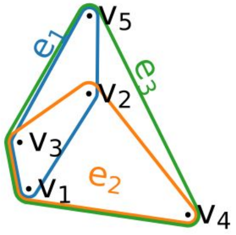
- Different Hypergraphs can have a same line graph
- s-walks on L(H) may describe different structures

An s-walk is called

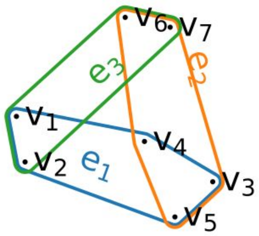
- s-trace: if all hyperedges are distinct by labels
- s-meander: if is an s-trace and all intersection are pairwise distinct
- s-path: if is an s-meander in which no intersection is included in another



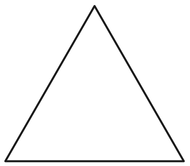
1-trace



3-meander



2-path



L(H)

Distances in Hypergraphs

Fixed s > 0, we define

s-distance

$d_s(e,f)$ between two edges e, f ∈ E as the length of the shortest s-walk between them, or infinite if there is none

s-diameter

the maximum s-distance between any two edges

s-component

a set of edges all connected pairwise by an s-walk

How to compute them?

- Given H, build L(H,s);
- Apply standard graph analysis measures.

s-Clustering Coefficient

All local indicators are computed on hyperedges, not nodes

s-triangle: a closed s-path of length 3
s-wedge: an s-path of length 2

s-Clustering: extension of classical definition(s)

NB: Remember that L(G) allows only the identification of s-walks

Local Clustering Coefficient
of hyperedge f (e_1,f,e_2)

$$s\text{-LCC}(f) = \begin{cases} \dfrac{\text{number of } s\text{-triangles containing } f}{\text{number of } s\text{-wedges centered at } f} & \text{if } f \text{ is the center of an } s\text{-wedge,} \\ 0 & \text{otherwise.} \end{cases}$$

Global Clustering Coefficient of H

$$s\text{-GCC}(H) = \frac{3 \cdot \text{total number of } s\text{-triangles}}{\text{total number of } s\text{-wedges}}.$$

s-Centralities?

As for the Clustering coefficient,

all centralities (degree, katz, harmonic...) can be redefined on hypergraphs.

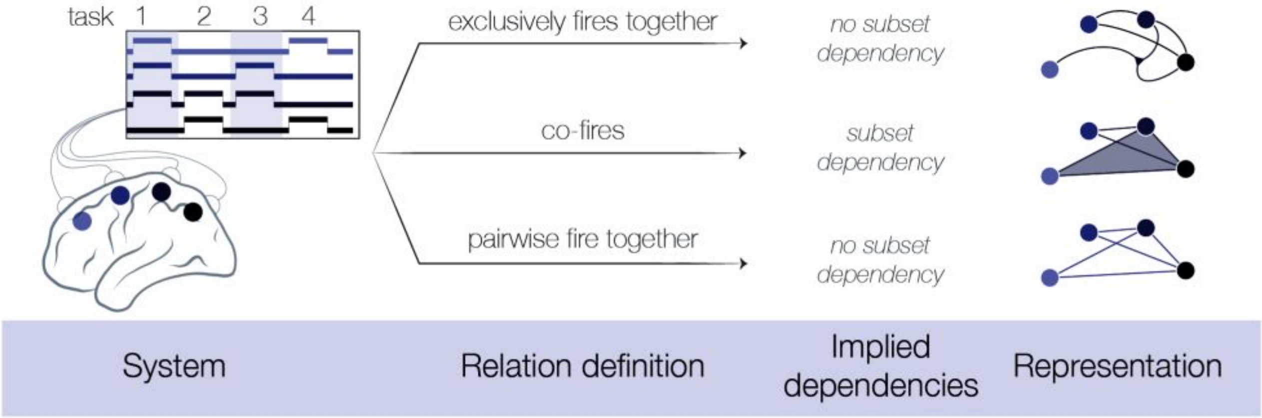
s-centralities are usually computed for hyperedges (starting from L(H))

Node s-centralities can be defined, using the same approach, on L(H*) - i.e., on the node projection of the bipartite graph.

Examples and Case Study

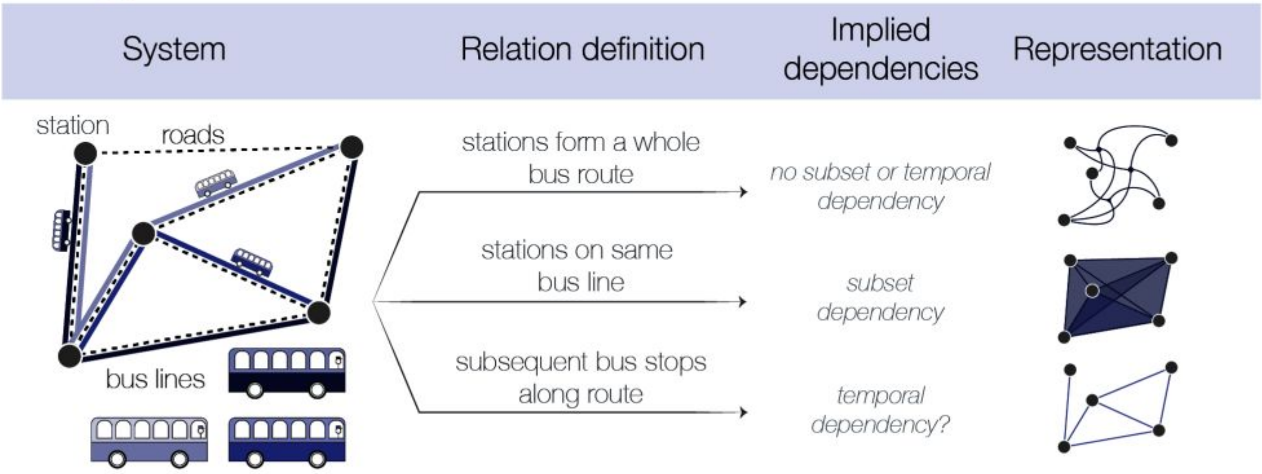
Hypernetwork Science

Different constraints, different modeling



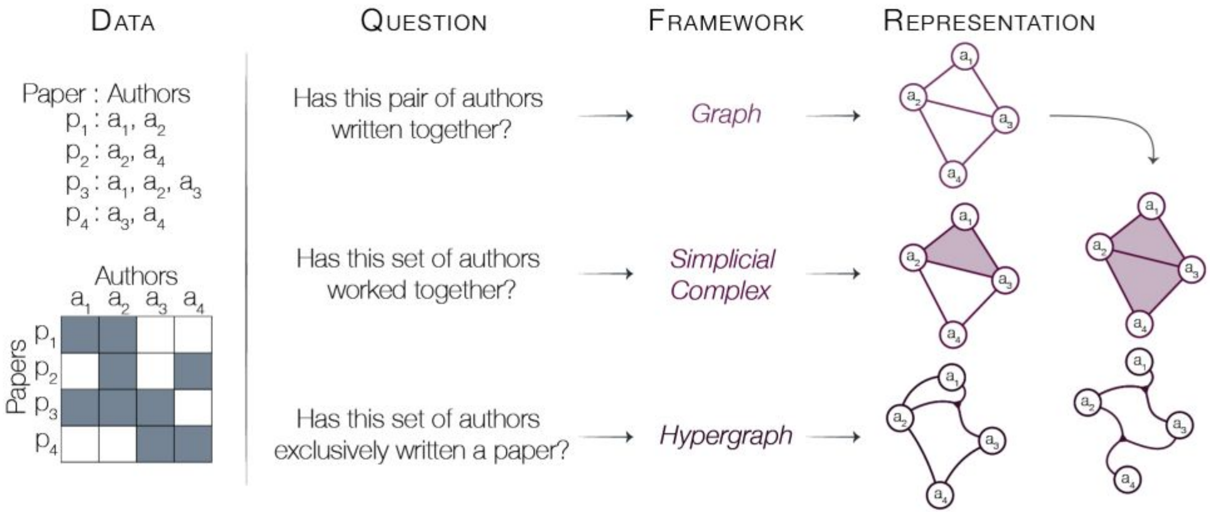
We might record the on/off activity of four brain regions in each of four tasks.
Depending on the definition of the relation chosen, we may or may not record a dependency in a representation

Different constraints, different modeling

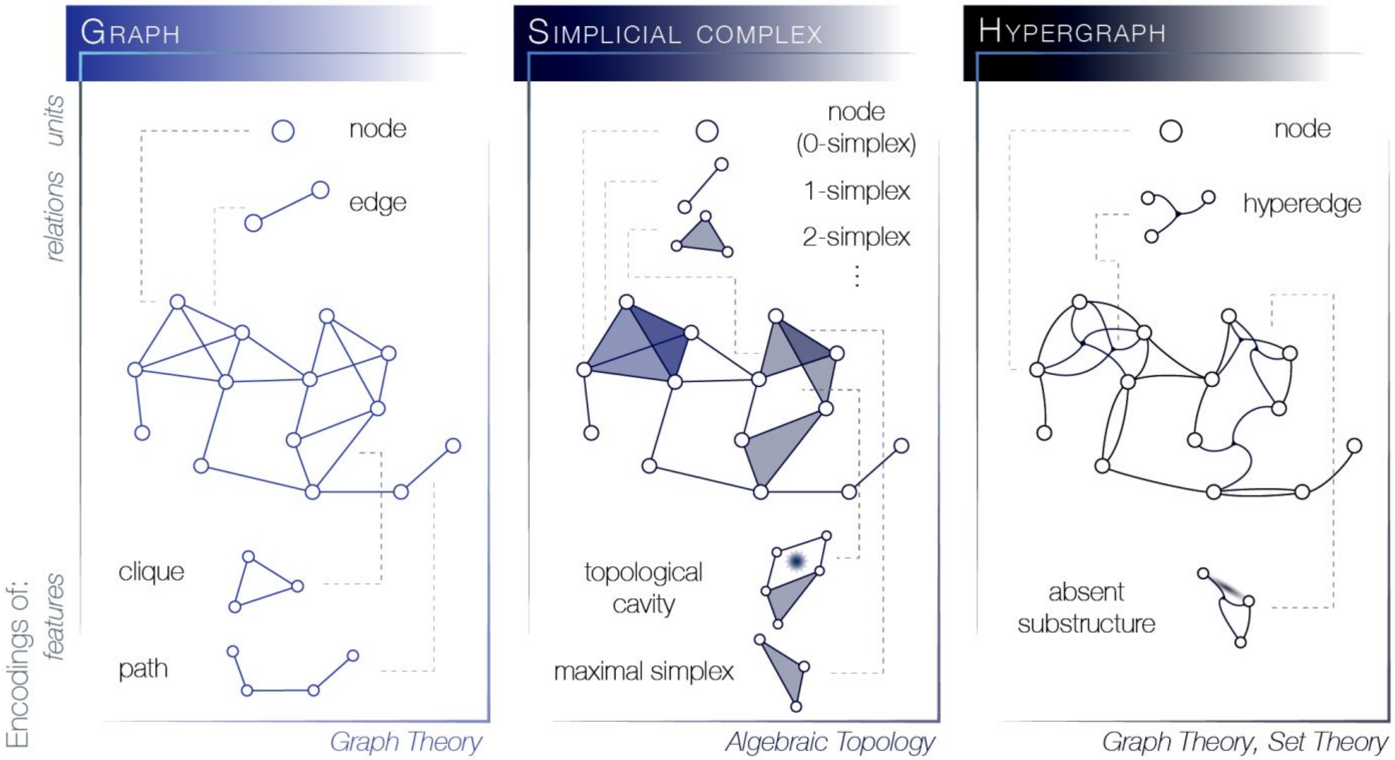


Given five bus stations placed along a set of roads we observe three bus lines that connect the stations.
Depending on the definition of the relation chosen, we might include a subset or temporal dependency which we would want to capture in our representation of the system

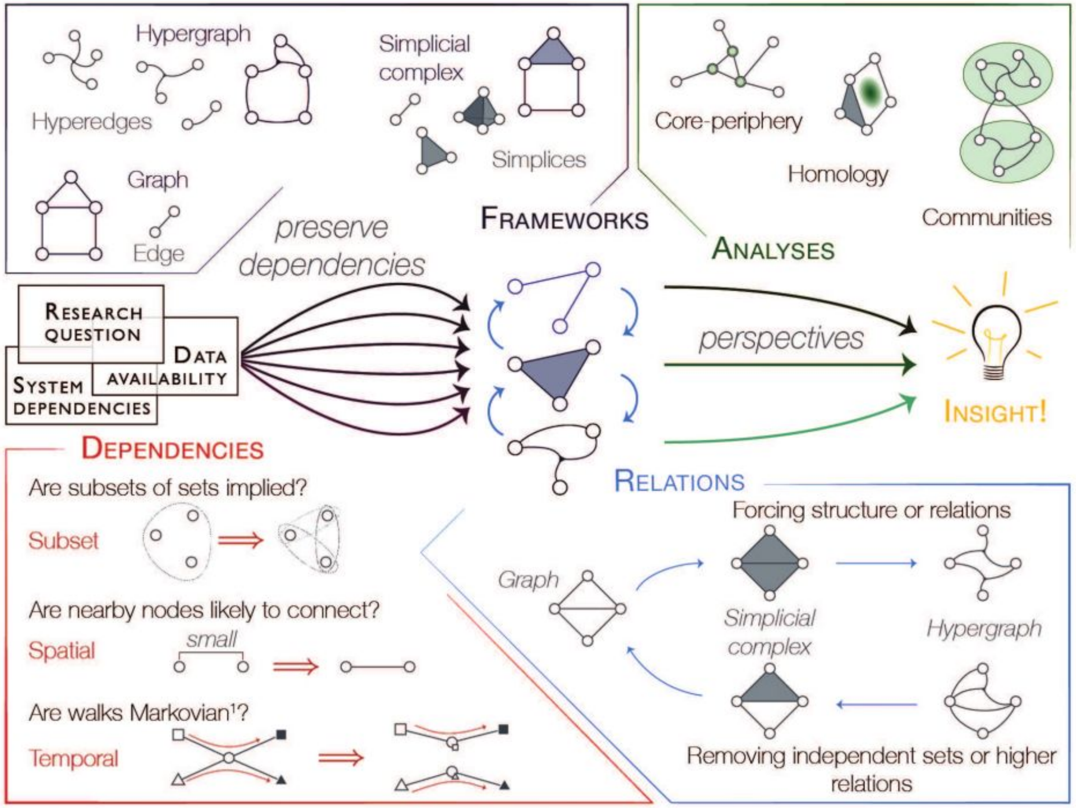
Different questions, different modeling



Summarizing...



Different complex systems require different modeling frameworks…



… explicitating peculiar entities' dependencies and enabling specific analysis!

Chapter 8

Conclusion

Take Away Messages

1. Pairwise interaction are, sometimes, not enough
2. Different models, allow answering different questions
3. s-analysis is an extension of classical graph analysis

Suggested Readings

- Battiston, Federico, et al. "Networks beyond pairwise interactions: structure and dynamics."
- Joslyn, Cliff A., et al. "Hypernetwork science: from multidimensional networks to computational topology."
- Torres, Leo, et al. "The why, how, and when of representations for complex systems."
- Aksoy, Sinan G., et al. "Hypernetwork science via high-order hypergraph walks."

What's Next

Chapter 9:
Community Discovery