707.000
Web Science and Web Technology
„Affiliation Networks“

How can we identify and analyze subgroups in affiliation networks?

Markus Strohmaier

Univ. Ass. / Assistant Professor
Knowledge Management Institute
Graz University of Technology, Austria

e-mail: markus.strohmaier@tugraz.at
web: http://www.kmi.tugraz.at/staff/markus

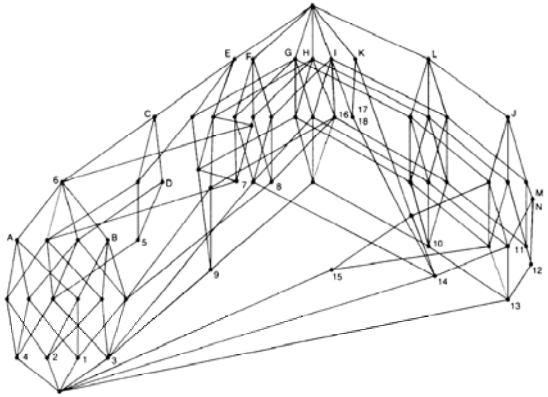


FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

[Freeman White 1993]

Centrality and Prestige in Undirected Social Graphs [Wasserman Faust 1994]

Examples and Simulation:

file:///M:/mydocs/courses/SS2010/707.000%20Web%20Science%20and%20Web%20Technology/measure/measure.html

Overview

Today‘s Agenda:

Analysis of Afilliation Networks

- A (very brief) repetition of Affiliation Networks
- Properties of Affiliation Networks
  - Properties of One-Mode Networks derived from Affiliation Networks
- Galois Lattices for Affiliation Networks

How can we identify groups and subgroups in a social graph?

Subgroups in Co-Affiliation Networks

Borgatti 1997

- We talked about cliques, clans and clubs in 1-mode
- The obvious next step would be to try to identify these subgroups in co-affiliation networks.
  - For example, we can search for cliques, n-cliques, n-clans, n-clubs.
- Unfortunately, these methods are not well suited for analysing a bipartite graph.
  - In fact, bipartite graphs contain no cliques
  - In contrast, bipartite graphs contain too many 2-cliques and 2-clans.
  - One of the problems is that, in the bipartite graph, all nodes of the same type are necessarily two links distant.

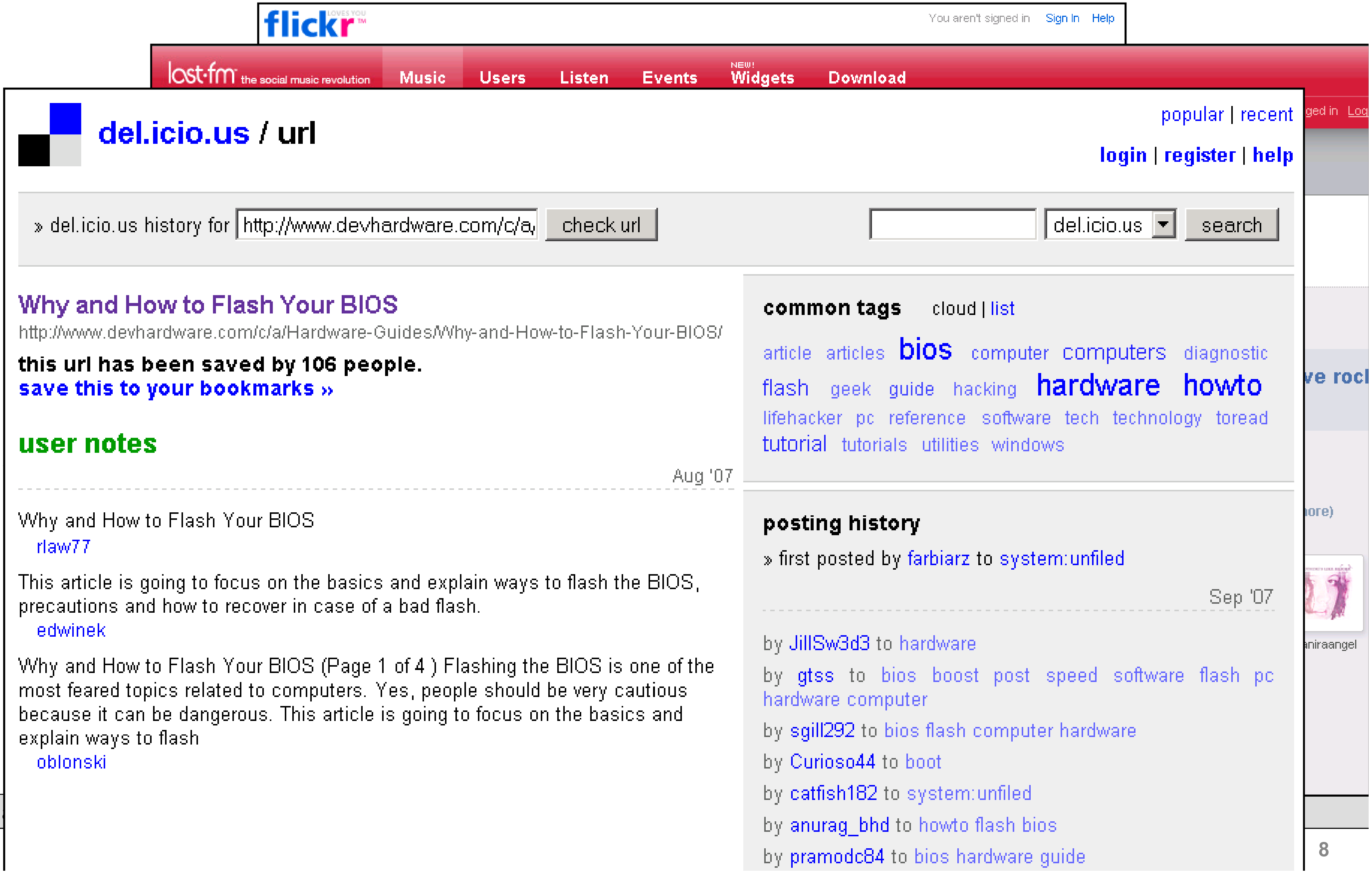
➔ we need to consider special types of subgraphs which are more appropriate for two-mode data.

Subgroups in Co-Affiliation Networks

Borgatti 1997

- **Clearly, we can define extensions of n-cliques, n-clubs and n-clans to n-bicliques, n-biclubs and n-biclans.**
- **But, the extensions would in many senses be unnatural since n would need to be odd.**

Reminder: Social Networks Examples



del.icio.us / url

popular | recent

login | register | help

» del.icio.us history for http://www.devhardware.com/c/a/ check url

del.icio.us search

Why and How to Flash Your BIOS

http://www.devhardware.com/c/a/Hardware-Guides/Why-and-How-to-Flash-Your-BIOS/

this url has been saved by 106 people.

save this to your bookmarks »

user notes

Aug '07

Why and How to Flash Your BIOS
rlaw77

This article is going to focus on the basics and explain ways to flash the BIOS, precautions and how to recover in case of a bad flash.
edwinek

Why and How to Flash Your BIOS (Page 1 of 4) Flashing the BIOS is one of the most feared topics related to computers. Yes, people should be very cautious because it can be dangerous. This article is going to focus on the basics and explain ways to flash
oblonski

common tags cloud | list

article articles bios computer computers diagnostic flash geek guide hacking hardware howto lifehacker pc reference software tech technology toread tutorial tutorials utilities windows

posting history

» first posted by farbiarz to system:unfiled

Sep '07

by JillSw3d3 to hardware

by gtss to bios boost post speed software flash pc hardware computer

by sgill292 to bios flash computer hardware

by Curioso44 to boot

by catfish182 to system:unfiled

by anurag_bhd to howto flash bios

by pramodc84 to bios hardware guide

Transforming Two Mode Networks into One Mode Networks

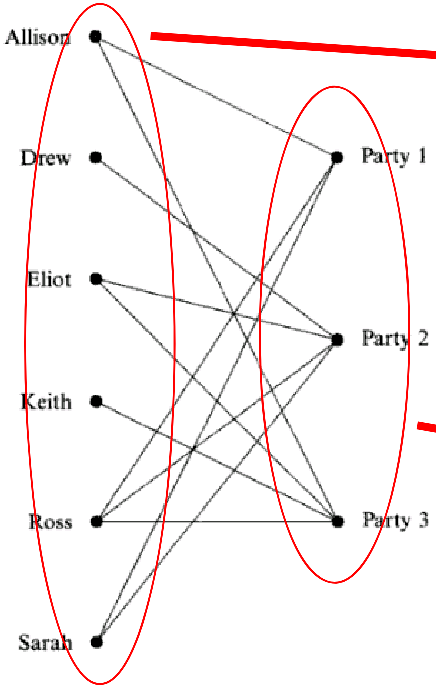
[Wasserman Faust 1994]

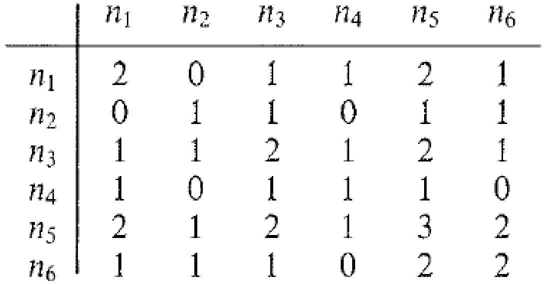
$$\boxed{M_P = M_{PC} * M_{PC}'}$$

C…Children

P…Party

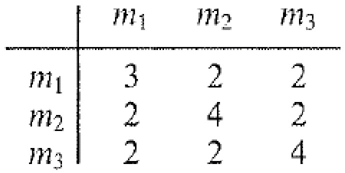
- Two one mode (or co-affiliation) networks (folded from the children/party affiliation network)





	n_1	n_2	n_3	n_4	n_5	n_6
n_1	2	0	1	1	2	1
n_2	0	1	1	0	1	1
n_3	1	1	2	1	2	1
n_4	1	0	1	1	1	0
n_5	2	1	2	1	3	2
n_6	1	1	1	0	2	2

Fig. 8.5. Actor co-membership matrix for the six children



	m_1	m_2	m_3
m_1	3	2	2
m_2	2	4	2
m_3	2	2	4

Fig. 8.6. Event overlap matrix for the three parties

[Images taken from Wasserman Faust 1994]

Properties of Affiliation Networks

[Wasserman Faust 1994]

- Properties of Actors and Events
  - Rates of Participation
  - Size of Events
- Properties of One-Mode Networks that are derived from Affiliation Networks
  - Density
  - Reachability
  - Connectedness
  - Diameter
  - Cohesive Subsets of Actors or Events
  - (Reachability for Pairs of Actors)

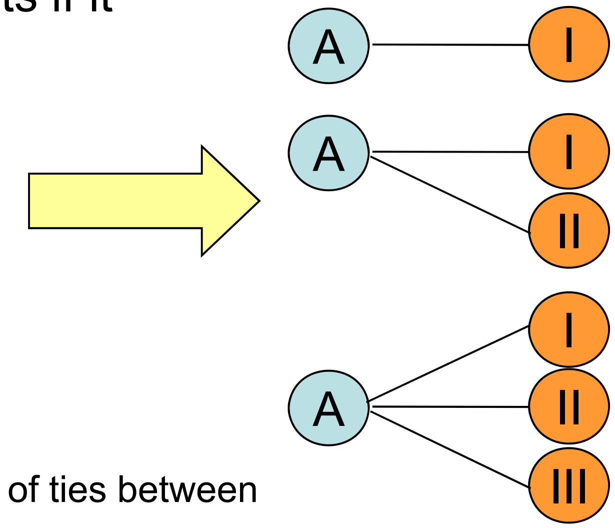
+ visual representations for affiliation networks

Density

[Wasserman Faust 1994]

Reminder: Density in regular networks is the ratio of edges to vertices

- The **density** of a one-mode network derived from an affiliation network is a **function of the pairwise ties between actors or between events**.
- The number of overlap ties between events is, in part, a function of the number of events to which actors belong
- An actor only creates a tie between a pair of events if it belongs to both events.
- An actor who belongs to
  - only one event creates no overlap ties between events
  - exactly two events creates a single tie
  - three events creates three ties
  - …
- In General
  - An actor who belongs to n events creates n*(n-1)/2 ties
- Thus,
  - the rates of membership for actors influence the number of ties between events, and
  - the sizes of the events influnce the number of ties between actors.



Reachability, Connectedness, Diameter

[Wasserman Faust 1994]

Reminder: Two nodes in a graph are adjacent if there is a line between them

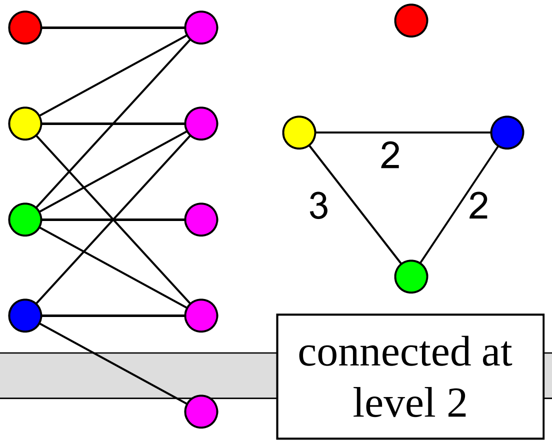
- In an affiliation network,
  - no pair of actors is adjacent
  - No pair of events is adjacent
- no paths of length 1 between actors, but potentially paths of some longer length
- **Reachability** corresponds to path lengths between nodes
- An affiliation network is **connected** when all pairs of nodes (both actors and events) are reachable
- The **diameter** of an affiliation network is the length of the longest shortest path between any pair of nodes.

Reachability for Pairs of Actors

[Wasserman Faust 1994]

- In a valued graph, we can **define connectedness at level c** as the subsets of actors all of whom are connected at some minimum level c
- Two nodes are **c-connected** (or reachable at level c) if there is a path between them in which all lines have a value of no less than c.

- Basis for the k-neighbourhood graph G_k / KNC Plot



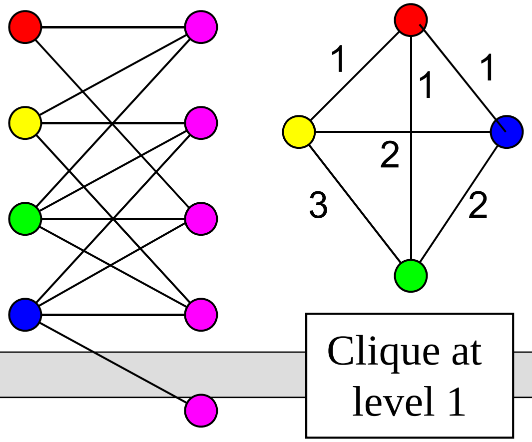
connected at level 2

Cohesive Subsets of Actors or Events

[Wasserman Faust 1994]

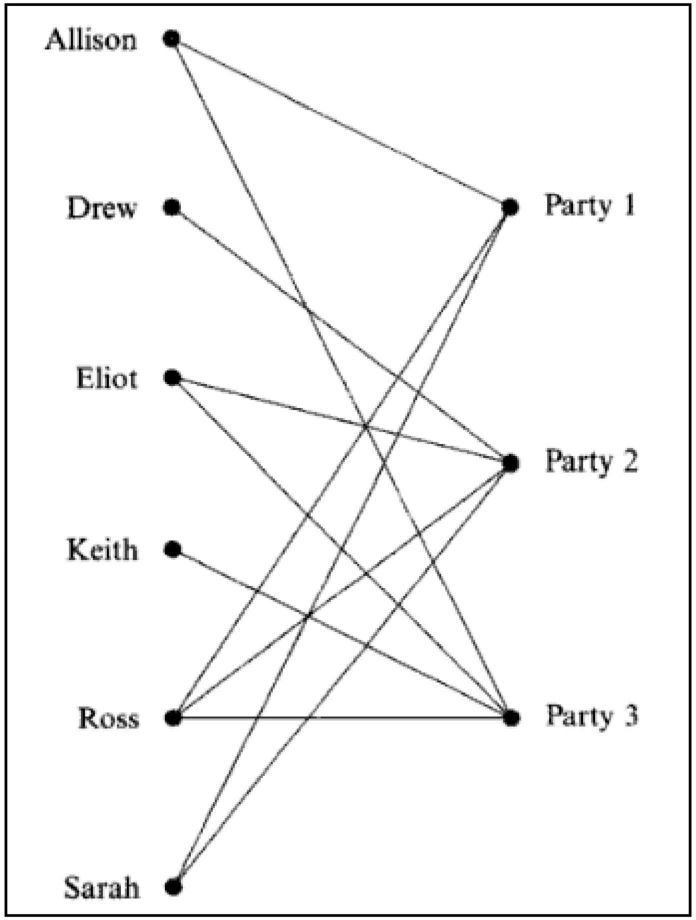
Reminder: a clique is a maximal complete subgraph of three or more nodes

- In a valued graph, we can define a **clique at level c** as a maximal complete subgraph of three or more nodes, all of which are adjacent at level c
- That is all pairs of nodes have lines between them with values that are greather than or equal to c. *By increasing c, we can locate more cohesive subgroups.*
- A **clique at level c** is a subgraph in which all pairs of actors share memberships in no fewer than c events
- Basis for the k-neighbourhood graph G_k / KNC Plot



Affiliation Networks

[Wasserman Faust 1994]

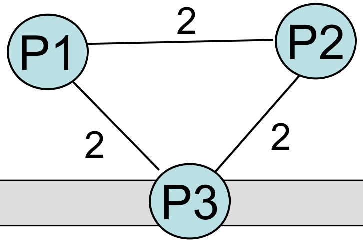


„**Two mode**“ networks

- Links only between the two modes
- Folding
- K-neighbourhood graph
- KNC Plot

But:

By folding, information is lost



Redundancy

[latapy 2006, p 22]

Neighbours (v)

Another node (not v)

That shares links with at least two neighbours

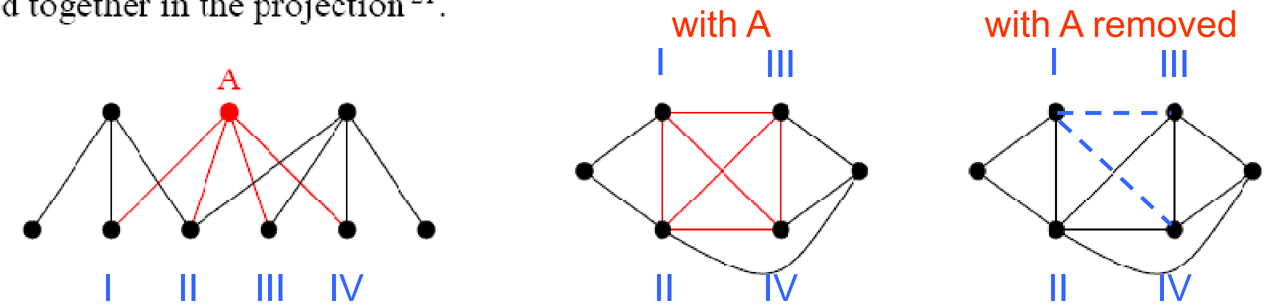
Number of links between neighbours after deleting (v)

Number of possible links between neighbours

redundancy coefficient of v:

$$\mathrm{rc}(v) = \frac{|\{\{u,w\} \subseteq N(v),\ \exists v' \neq v,\ (v',u) \in E \text{ and } (v',w) \in E\}|}{\frac{|N(v)|(|N(v)|-1)}{2}}.$$

In other words, the redundancy coefficient of v is the fraction of pairs of neighbours of v linked to another node than v. In the projection, these nodes would be linked together even if v were not there, see Figure 9; this is why we call this the redundancy. If it is equal to 1 then the projection would be exactly the same without v; if it is 0 it means that none of its neighbours would be linked together in the projection[21].

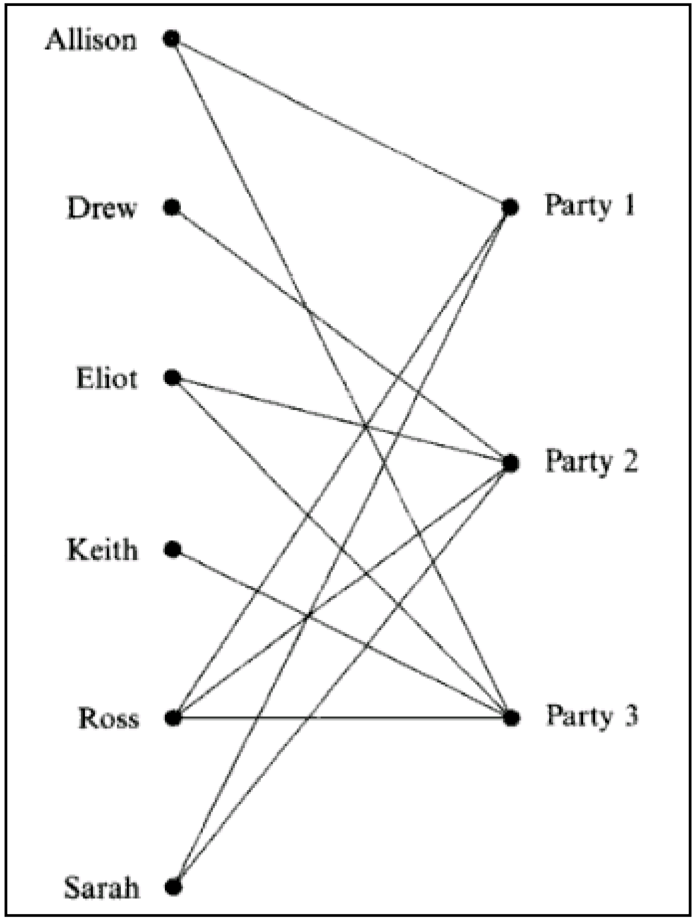


Rc(A) = 4/6

Figure 9: Example of redundancy computation. From left to right: a bipartite graph, its ⊥-projection, and the ⊥-projection obtained if the node A is first removed. Only two links disappear, leading to $\mathrm{rc}(A) = \frac{4}{6} = 0.666\cdots$.

Affiliation Networks

[Wasserman Faust 1994]



So:

How can we simultaneously analyze

Actors AND Events

in Affiliation Networks?

Can we show both,

- the relationships among the entities within each mode, and also
- how the two modes are associated with each other?

Galois Lattices

[Freeman White 1993]

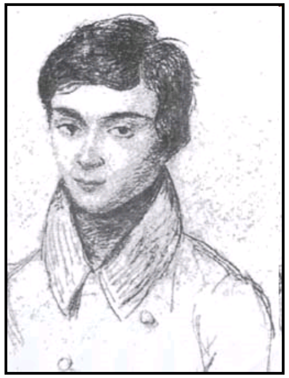
A satisfactory representation should facilitate the visualization of three kinds of patterning:

1. the actor-event structure,
2. the actor-actor structure, and
3. the event-event structure

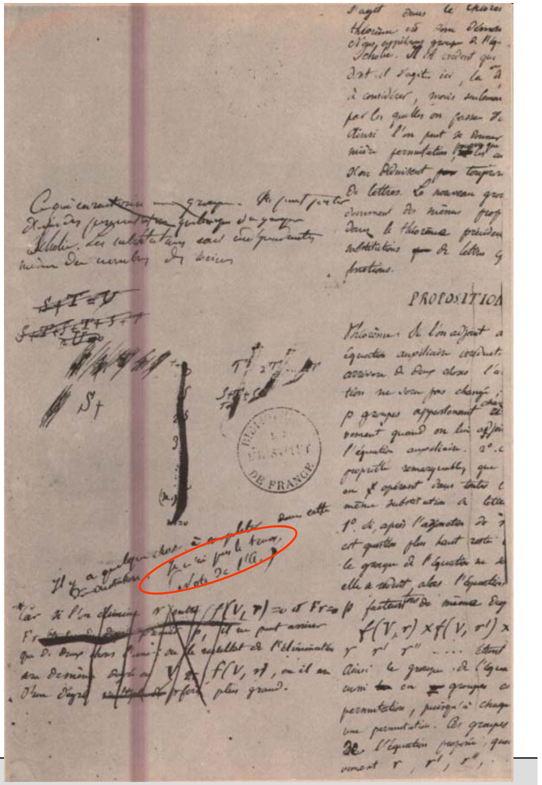
at the **same time**.

Évariste Galois

1811-1832



- A republican (fighting the french king)
- Not allowed to enter Ecole Polytechnique *twice*
- No recognition of his work as a mathematician
  (not considered for the Academy of Sciences Grand Prize in Mathematics for his ideas on solutions for quintic equations)
- Then he focused on politics / sentenced to prison for marching against the king
- Romance with a mysterious woman who was engaged
- Died in a pistol duel with her fiance at the age of 20
- A letter to his friends
  „I beg my patriots, my friends, not to reproach me for dying otherwise than for my country. I died the victim of an infamous coquette and her two dupes. It is in a miserable piece of slander that I end my life. Oh! Why die for something so little, so contemptible? I call on heavon to witness that only under compulsion and force have I yielded to a provocation which I have tried to avert by every means". [Fermat's Last Theorem]
- Spent the night before the duel writing down his mathematical achievements *(„I have no time!" see image to the right)*



Galois Lattices (or Galois Connections)

[Wasserman Faust 1994, Freeman White 1993]

- A long history in Mathematics
- Introduced by Birkhoff in 1940 (cf. Birkhoff „Lattice Theory“ 1967)
- Affiliation networks focus on **subsets** and the **duality of the relationship** between actors and events
- The idea of **subsets** refers both to subsets of actors contained in events and **subsets** of events that actors attend.
- The idea of **duality** refers to the complementary perspectives of relations
  - *between actors as participants in events*, and
  - *between events as collections of actors*.
- **Galois lattices** incorporate both ideas.

- Galois lattices are based on the kind of triple *(A,E,I)* defined by two mode social network data. *A* and *E* are finite nonempty sets and I (or lambda „λ“) is a binary relation in *A* x *E.*

A Lattice

[Wasserman Faust 1994]

- Galois lattices are special kind of lattices

Consider a set of elements N = {n_1, n_2, …n_g} and a binary relation „less than or equal“ (≤) that is reflexive, antisymmetric and transitive.

Formally

- Reflexive: $n_i \leq n_i$
- Antisymmetric: $n_i \leq n_j$ and $n_j \leq n_i$ iff $n_j = n_i$
- Transitive: $n_i \leq n_j$ and $n_j \leq n_k$ implies $n_i \leq n_k$

Such a system defines a **partial order** on the set N.

(cf.Partially Ordered Sets, POS, poset)

A Lattice

[Wasserman Faust 1994]

For any pair of elements, n_i, $n_{j,}$ we define their

- **lower bound** as that element n_k such that $n_k \leq n_i$ and $n_k \leq n_j$
- **upper bound** as that element n_k such that $n_i \leq n_k$ and $n_j \leq n_k$

With that,

- A lower bound is called a **greatest lower bound** n_k (or meet/infimum) of n_i, n_j if $n_l \leq n_k$ for all lower bounds n_l of n_i, n_j

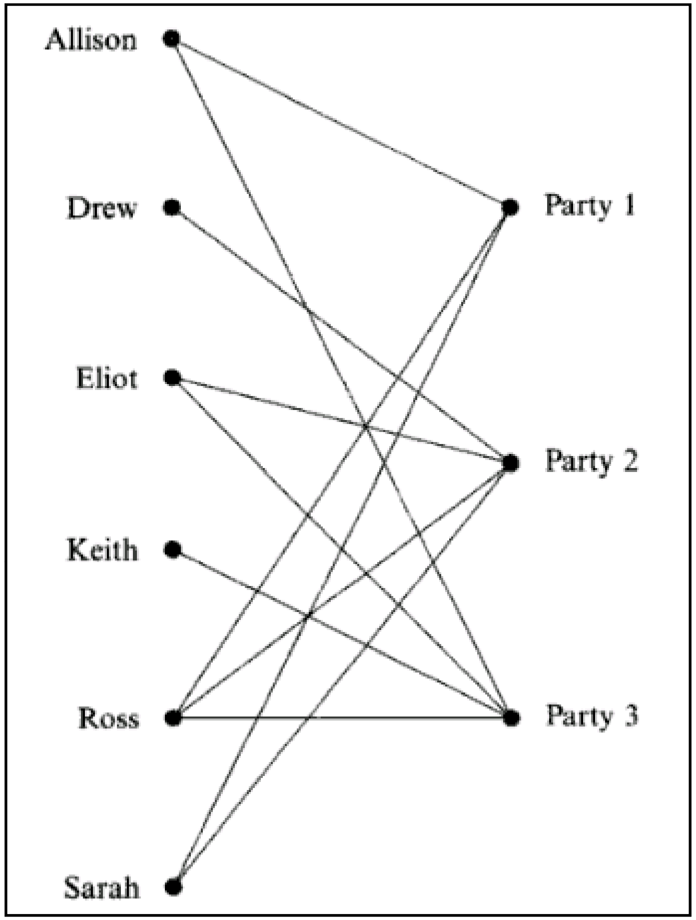
- An upper bound is called a **least upper bound** n_k (or join/supremum) of n_i, n_j if $n_k \leq n_l$ for all upper bounds n_l of n_i, n_j

A **lattice** consists of a set of elements N, a binary relation ≤ that is reflexive, antisymmetric and transitive and each pair of elements n_i, $n_{j,}$ has both a least upper bound and a greatest lower bound.

A *lattice* is thus a partially ordered set in which each pair of elements has both a meet and a join.

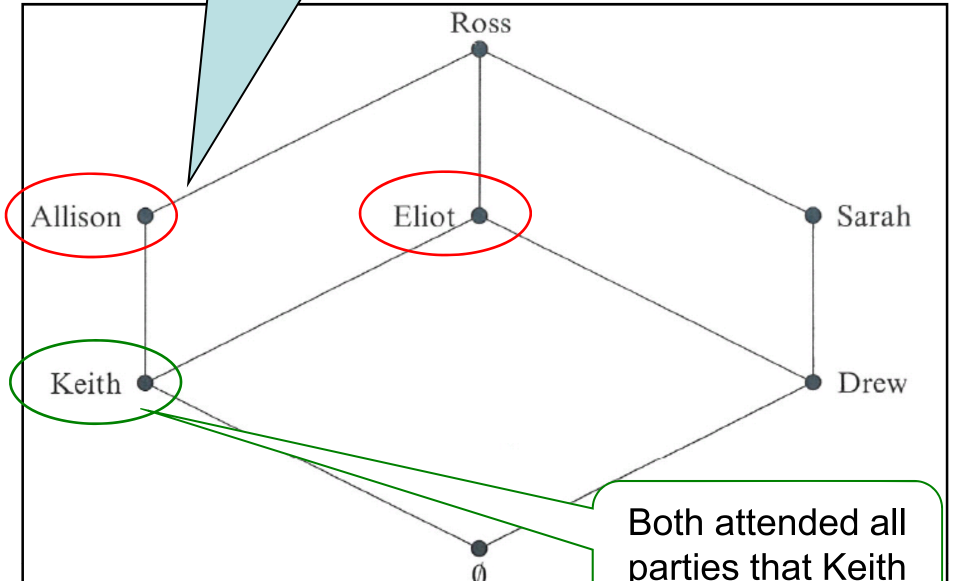
“⊆“ as our relation

[Wasserman Fau 94]



Each point represents a **subset of parties**

What is the greatest lower bound (meet) of Allison and Eliot?
And what does it mean?



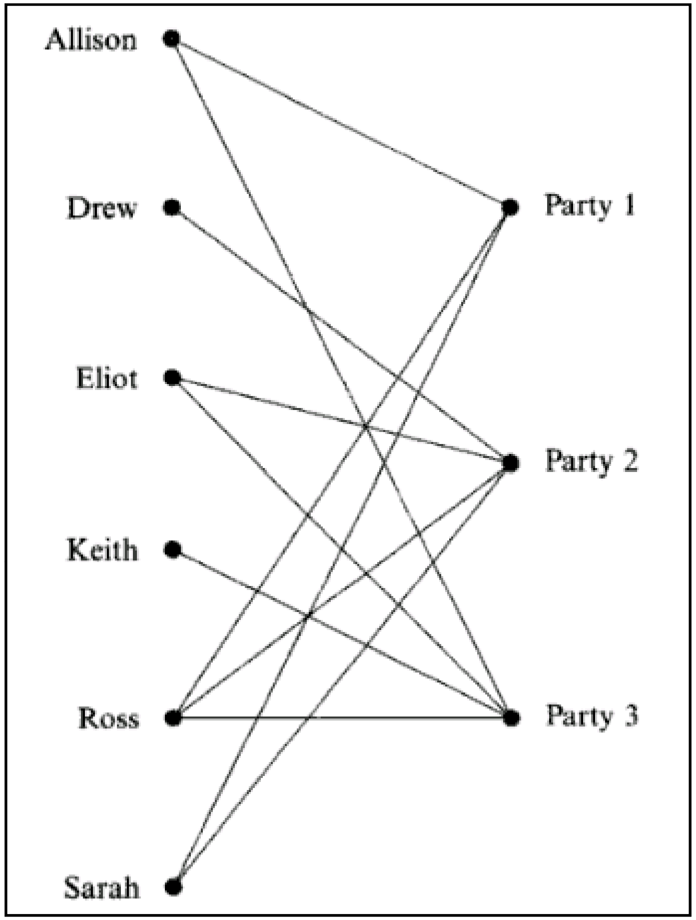
Both attended all parties that Keith attended (P3).

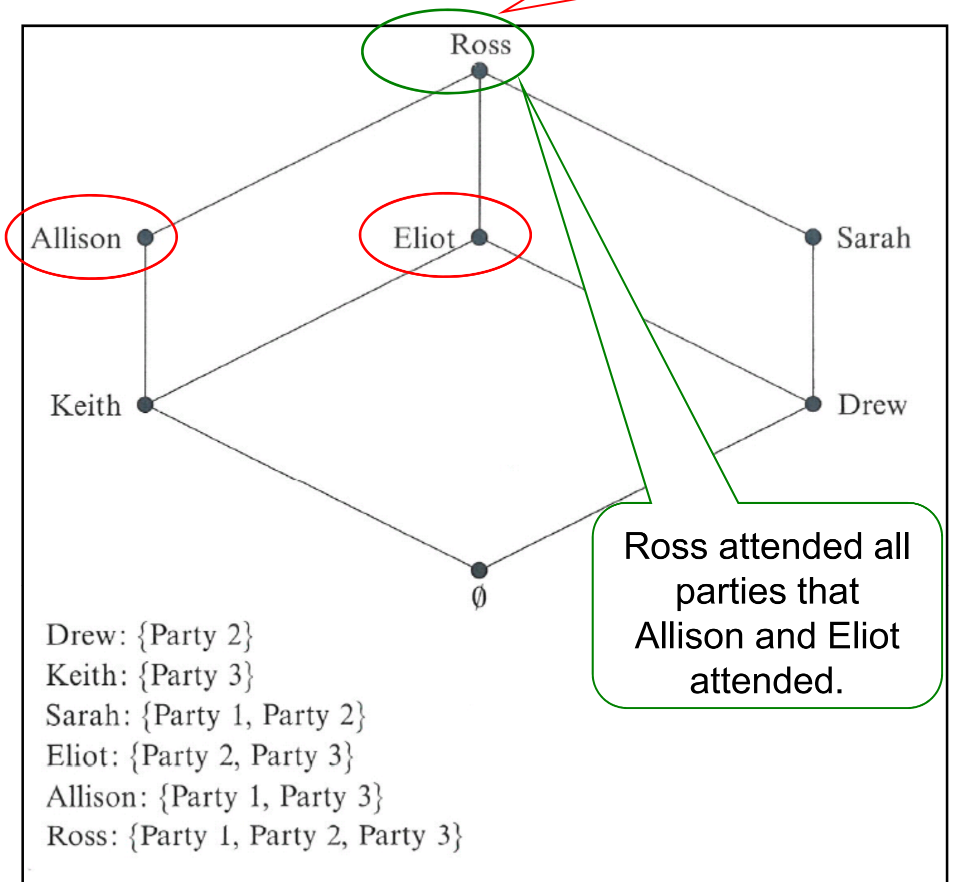
Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

Fig. 8.10. Relationships among children as subsets of birthday parties

Lattice

[Wasserman Faust 1994]





Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

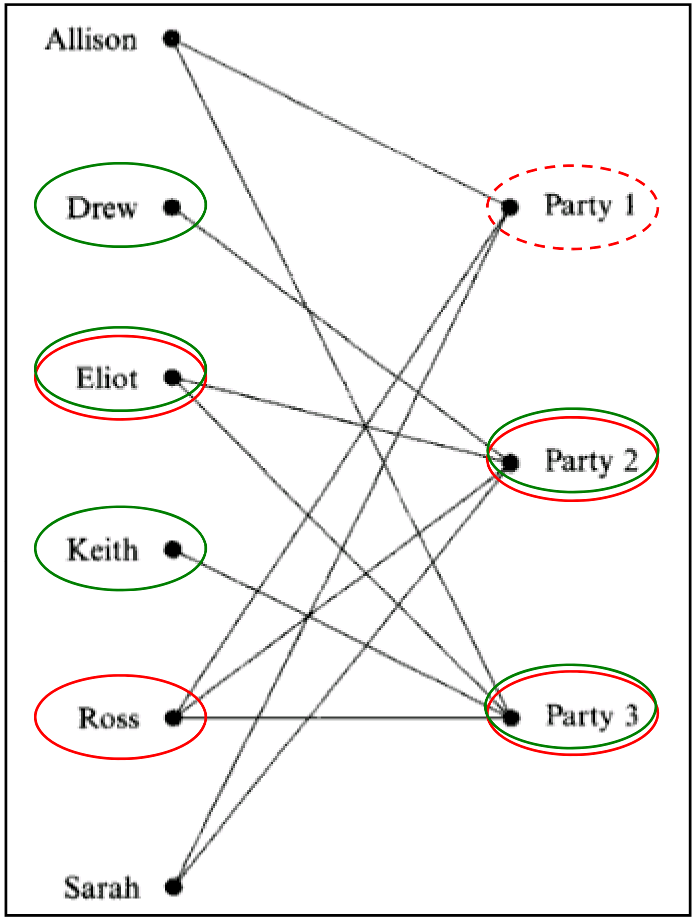
Fig. 8.10. Relationships among children as subsets of birthday parties

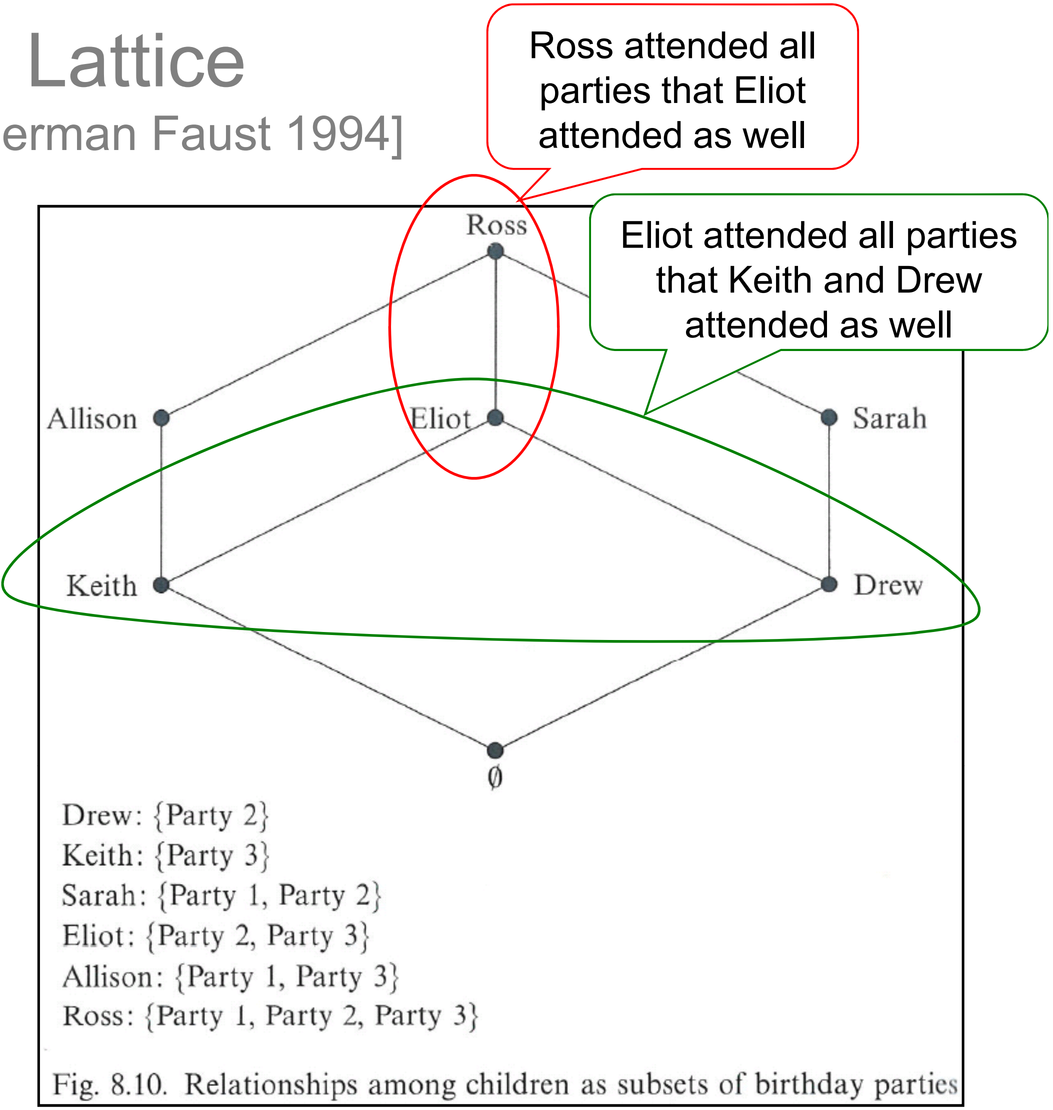
What is the least upper bound (join) of Allison and Eliot?
And what does it mean?

Ross attended all parties that Allison and Eliot attended.

Lattice

[Wasserman Faust 1994]





Ross attended all parties that Eliot attended as well

Eliot attended all parties that Keith and Drew attended as well

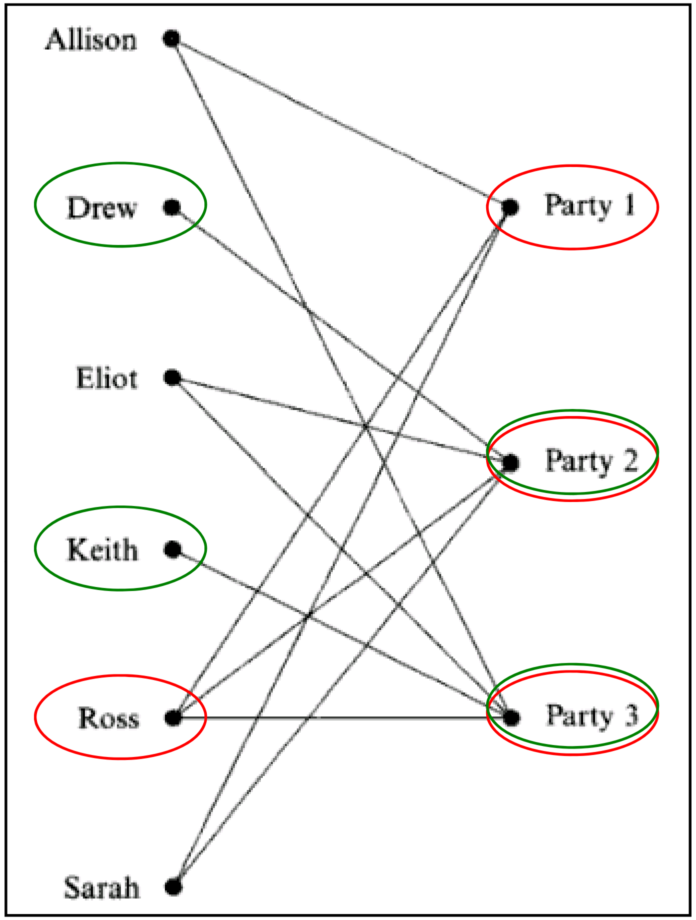
Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

Fig. 8.10. Relationships among children as subsets of birthday parties

Lattice

[Wasserman Faust 1994]

Who attended the most parties?



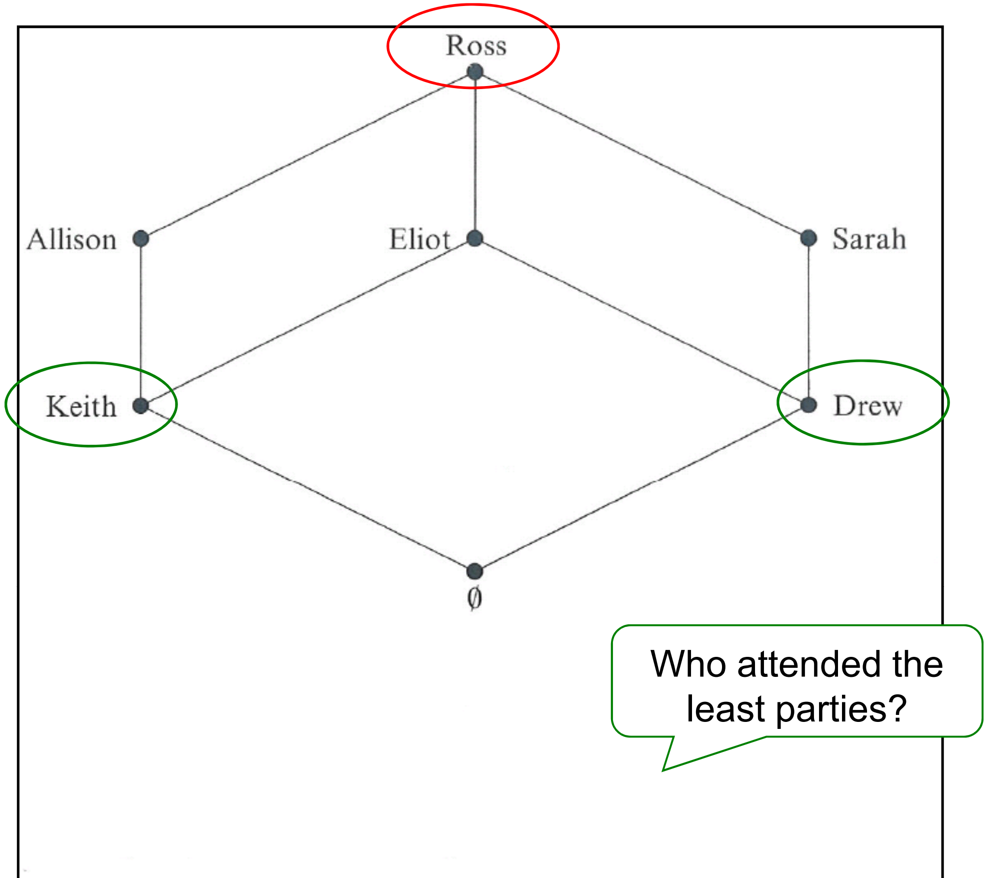


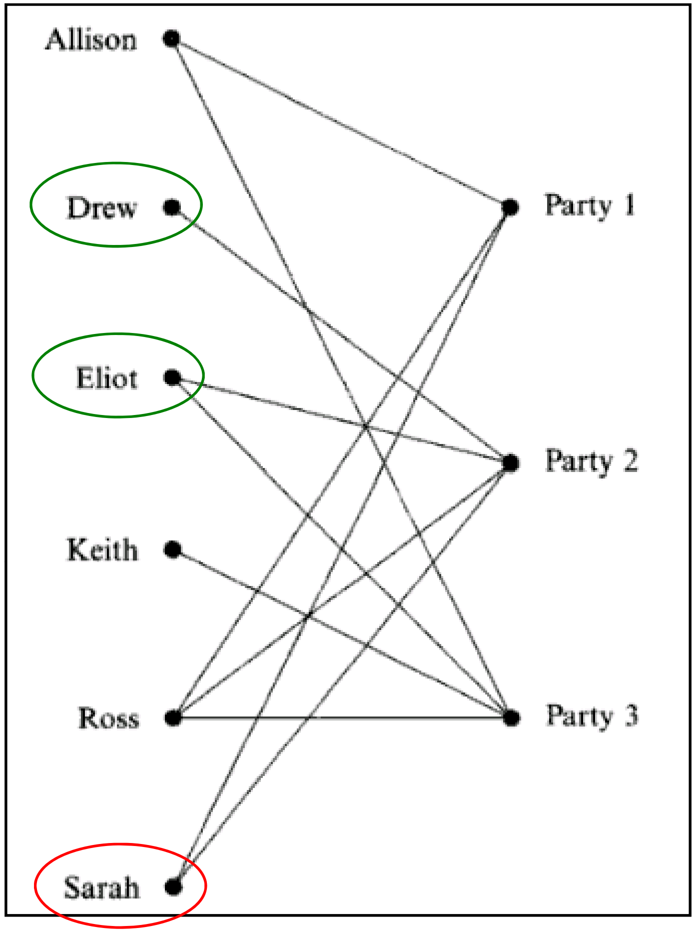
Fig. 8.10. Relationships among children as subsets of birthday parties

Who attended the least parties?

Lattice

[Wasserman Faust 1994]

Did Sarah attend all parties that Eliot and Drew attended?



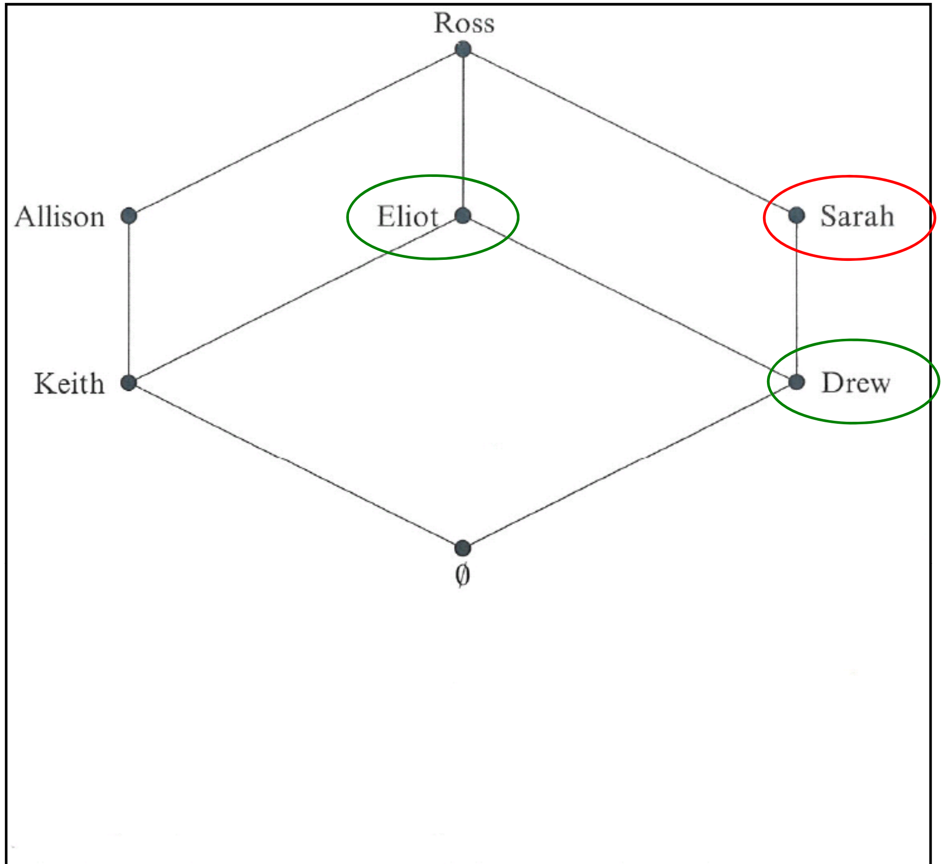
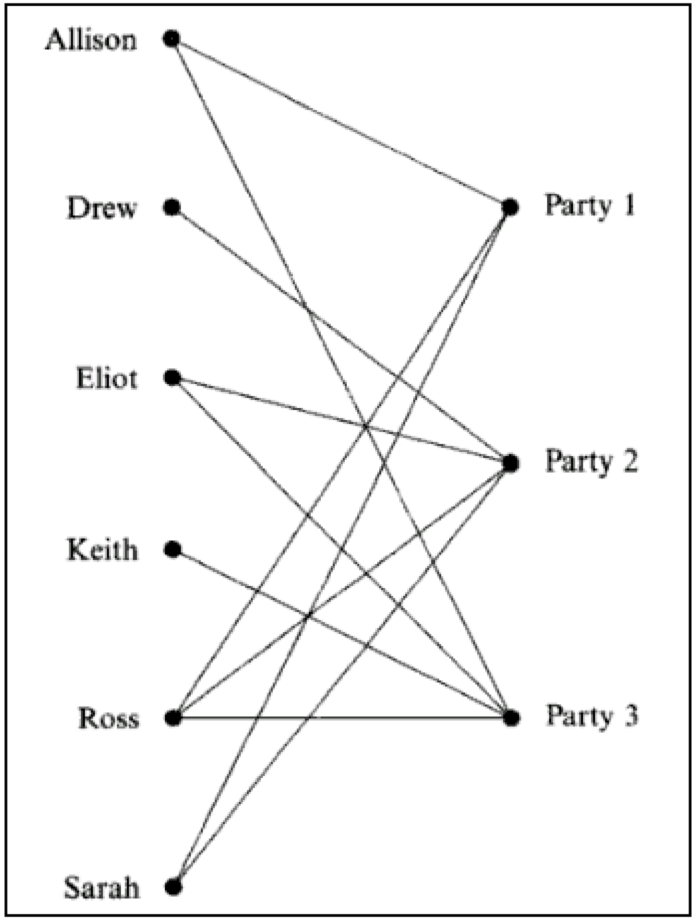


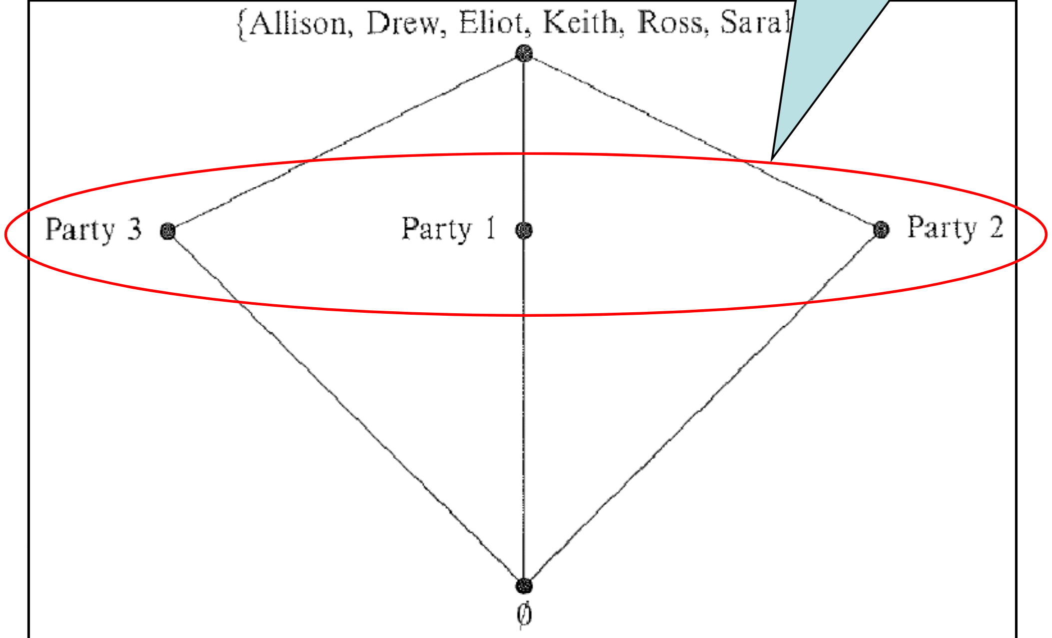
Fig. 8.10. Relationships among children as subsets of birthday parties

Lattice

[Wasserman Faust 1994]

Each point represents a **subset of children**





Party 1: {Allison, Ross, Sarah}
Party 2: {Drew, Eliot, Ross, Sarah}
Party 3: {Allison, Eliot, Keith, Ross}

Fig. 8.9. Relationships among birthday parties as subsets of children

A Galois Lattice

[Wasserman Faust 1994]

A **Galois lattice** (also called a *Galois connection*)

focuses on the **relation between two sets**.

- A relation λ is defined on pairs from the Cartesian product N x M.
- λ is thus defined on pairs, a relation $n_i \in N \lambda m_j \in M$

We let the sets N and M be the set of actors and the set of events, and let λ be the relation of affiliation.

- Thus, $n_i \lambda m_j$ if actor i is affiliated with event j.

We also have λ^{-1} where $m_j \lambda^{-1} n_i$ if event j contains actor i.

A Galois Lattice

[Wasserman Faust 1994]

Just as we have considered an *individual* actor and the subset of event with which it is affiliated, we can also consider a *subset of actors* and the *subset of event* with which all of these actors are affiliated.

We can define two mappings:

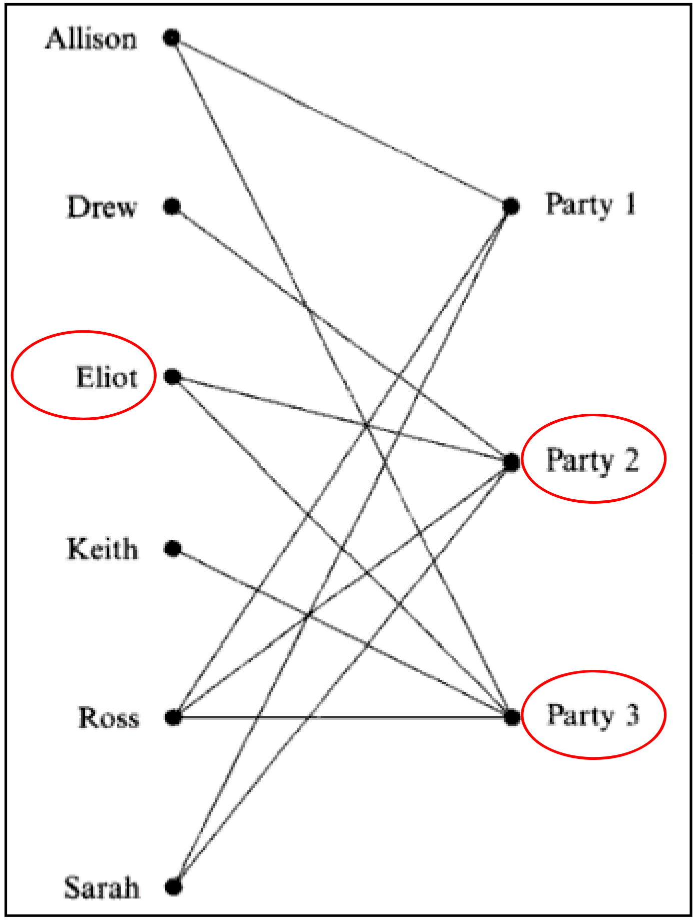
- $\uparrow$: N_s -> M_s from a subset of actors $N_s \subseteq N$ to a subset of events $M_S \subseteq M$ such that $n_i \lambda m_j$ for all $n_i \in N_i$ and all $m_j \in M_j$.
  - In terms of an affiliation network, the ↑ **mapping** goes from a subset of actors to that subset of events **with which all of the actors in the subset are affiliated**.
  - For example, if there is no event with which all actors in subset N_s are affiliated, then $\uparrow(N_S) = 0$
- ↓ mapping can be defined analogously.

↑ mapping

Galois Lattice

[Wasserman Faust 1994]

Which parties did Eliot attend?



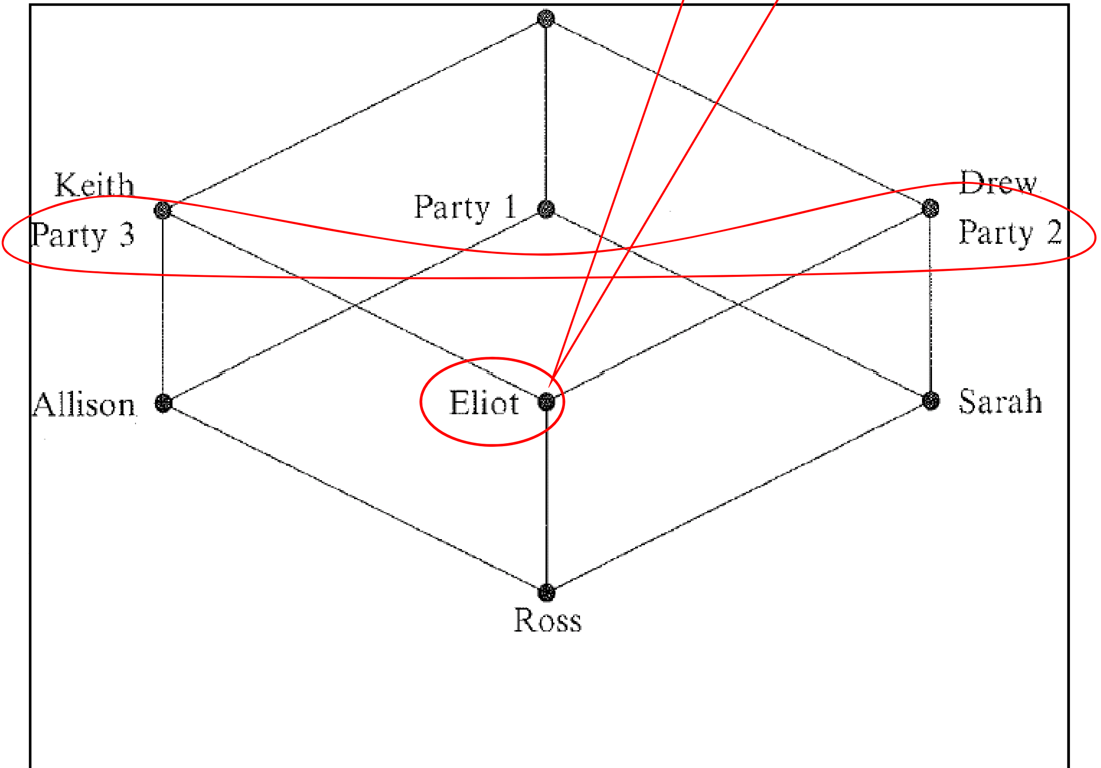
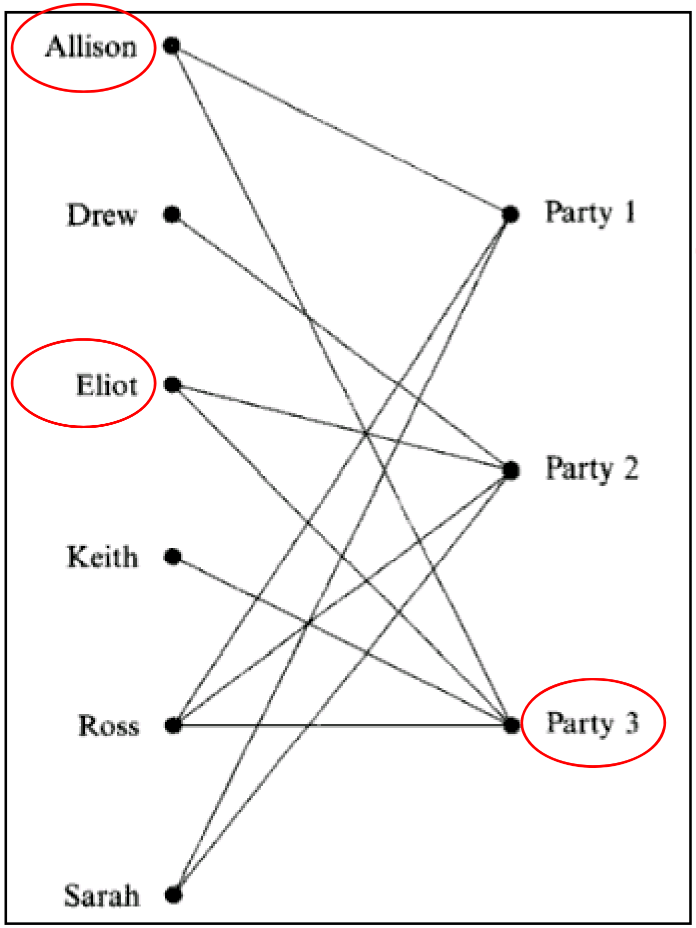


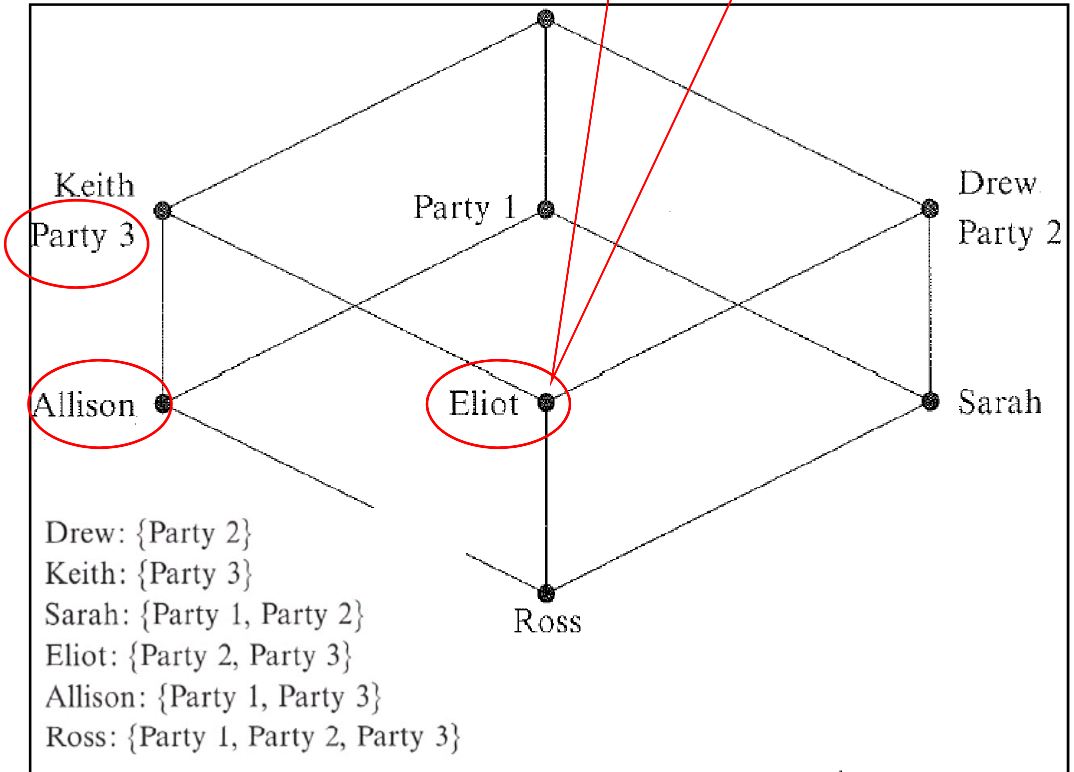
Fig. 8.11. Galois lattice of children and birthday parties

↑ mapping

Galois Lattice

[Wasserman Faust 1994]





Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

Fig. 8.11. Galois lattice of children and birthday parties

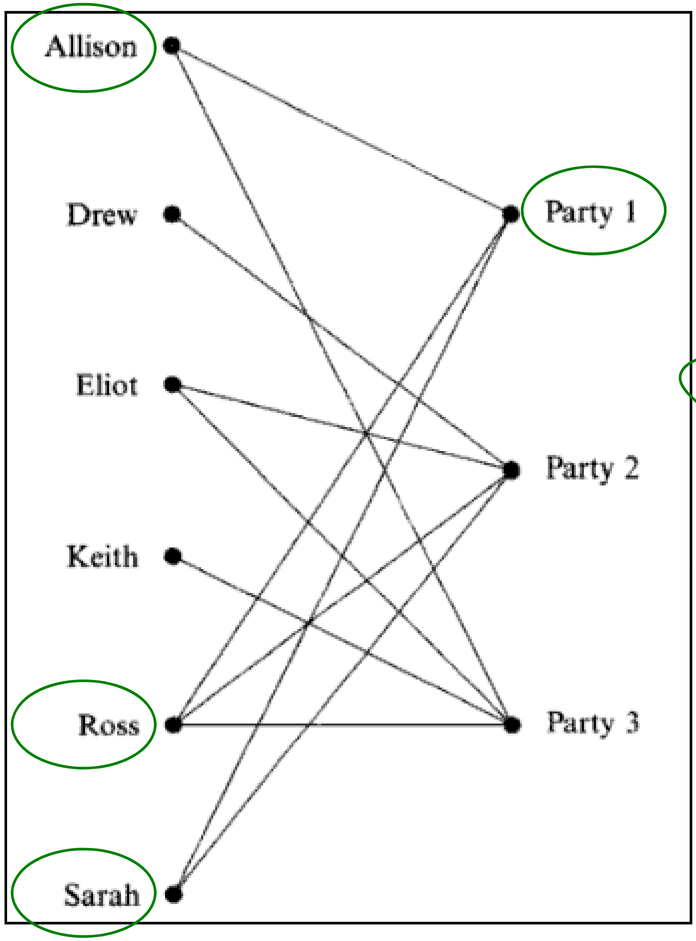
Which parties did both Eliot and Allison attend?

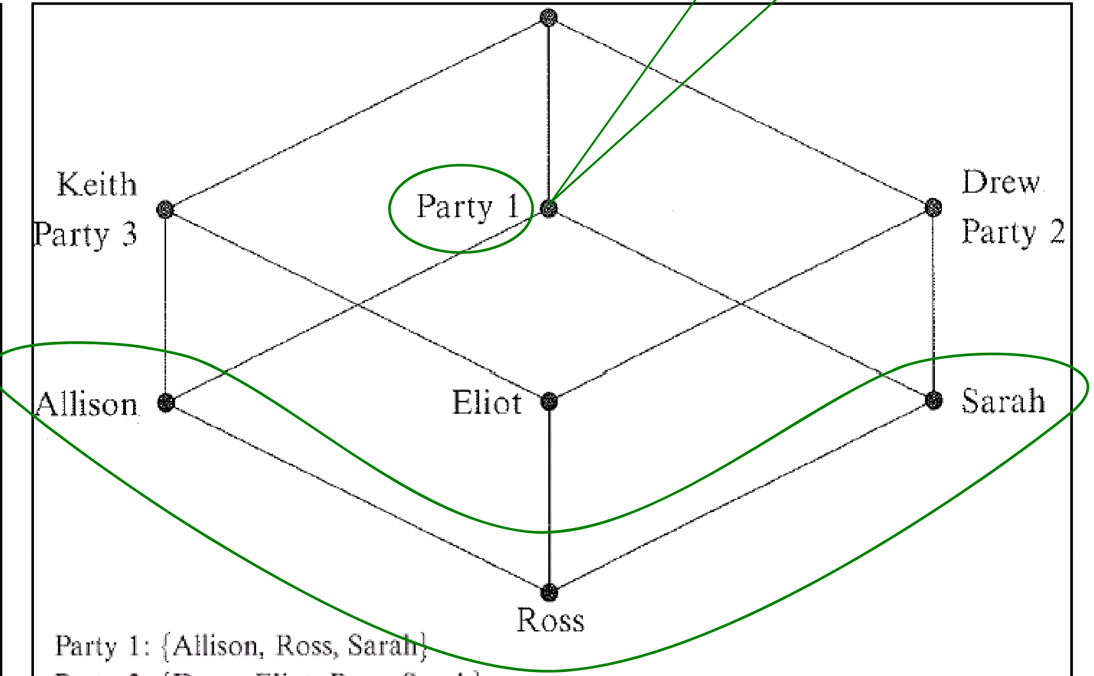
↓ mapping

Galois Lattice

[Wasserman Faust 1994]

Who attended party 1?





Party 1: {Allison, Ross, Sarah}
Party 2: {Drew, Eliot, Ross, Sarah}
Party 3: {Allison, Eliot, Keith, Ross}

Fig. 8.11. Galois lattice of children and birthday parties

Ga

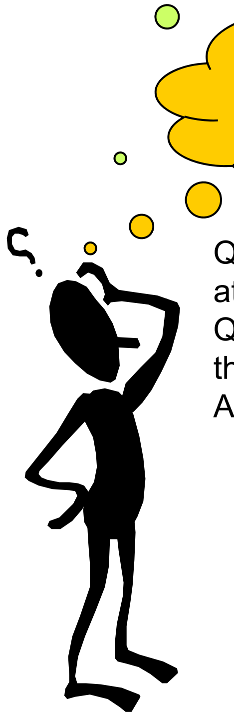
[Wasserman Faust 19

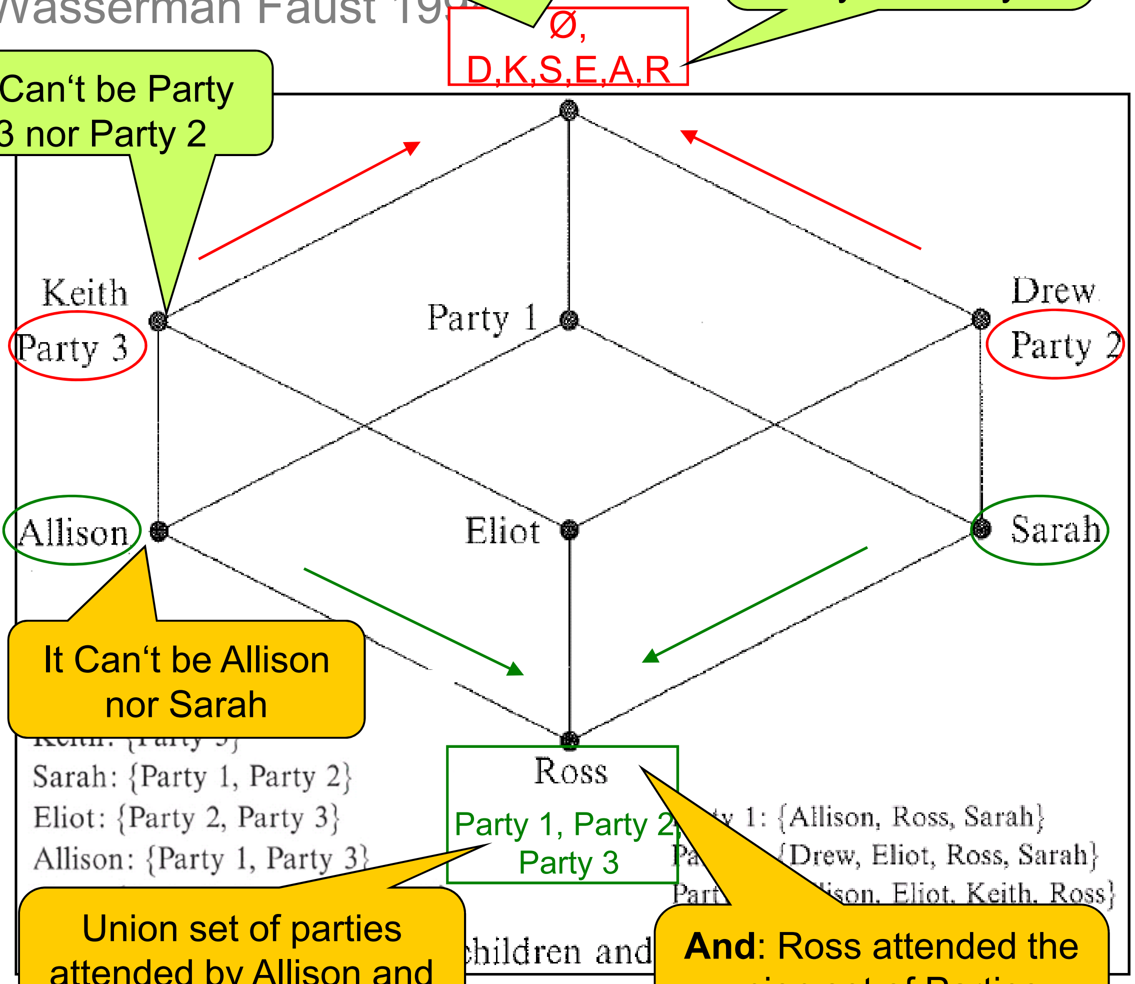
What is the least upper bound of Party 2 and Party 3?

Q1: What is the intersection set of children attending Party 2 & 3?
Q2: Is there a party that was attended by the intersection set of children attending Party 2 and 3?

What is the greatest lower bound of Allison and Sarah?

Q1: What is the union set of parties attended by Allison and Sarah?
Q2: Is there a child that attended the union set of Parties attended by Allison and Sarah?





But: there is no party which was attended by the intersection set of children attending Party 2 and Party 3

Intersection set of children attending Party 2 & Party 3

It Can‘t be Party 3 nor Party 2

It Can‘t be Allison nor Sarah

Union set of parties attended by Allison and Sarah

And: Ross attended the union set of Parties attended by A & S

In terms of an a ping goes f at
subset of events with which all of the acto d.

Galois Lattice

[Wasserman Faust 1994]

Consider
- N_s = {Allison, Sarah}

Since
- Allison {Party 1, Party 3}
- Sarah {Party 1, Party 2}

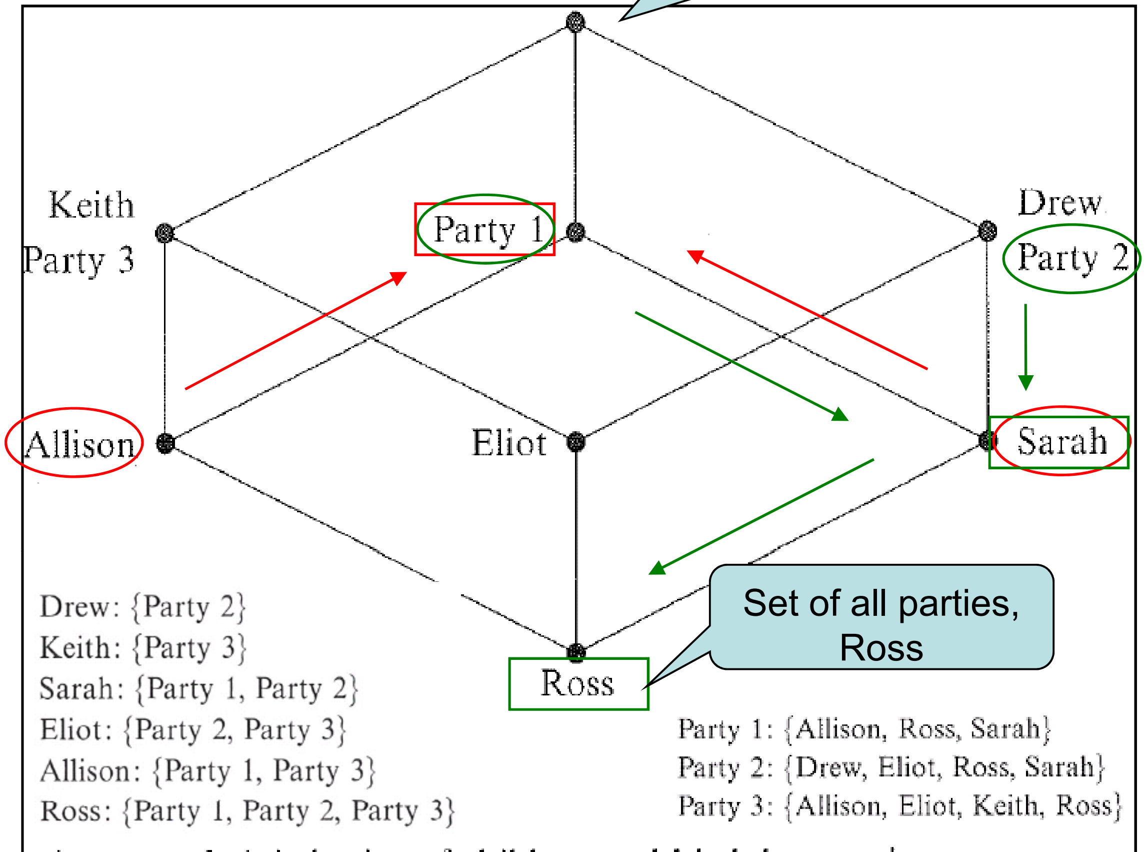
It follows that
- M_s = {Party 1} and
- $\uparrow(N_s) = M_s$ = {Party 1}

Consider
- M_s = {Party 1, Party 2}

Since
- Party 1 {Allison, Ross, Sarah}
- Party 2 {Drew, Eliot, Ross, Sarah}

It follows that
- N_s = {Ross, Sarah} and
- $\downarrow(M_s) = N_s$ = {Ross, Sarah}

Empty set of parties, set of all actors



Set of all parties, Ross

Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

Party 1: {Allison, Ross, Sarah}
Party 2: {Drew, Eliot, Ross, Sarah}
Party 3: {Allison, Eliot, Keith, Ross}

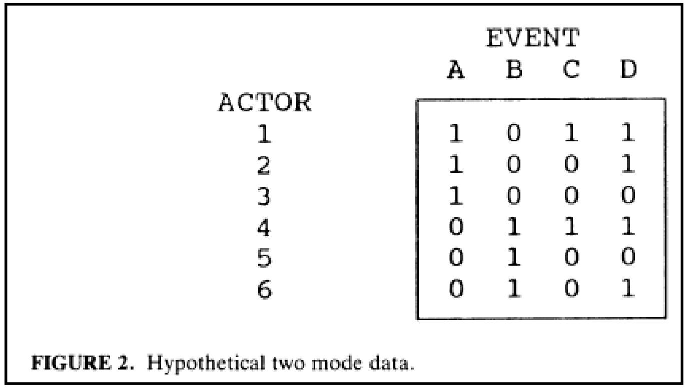
Fig. 8.11. Galois lattice of children and birthday parties

In terms of an affiliation network, the ↑ mapping goes from a subset of actors to that subset of events with which all of the actors in the subset are affiliated.

Galois Lattice

[Freeman White 1993]

- **Reduced** and **Full** Labeling



	EVENT			
ACTOR	A	B	C	D
1	1	0	1	1
2	1	0	0	1
3	1	0	0	0
4	0	1	1	1
5	0	1	0	0
6	0	1	0	1

FIGURE 2. Hypothetical two mode data.

Full labeling

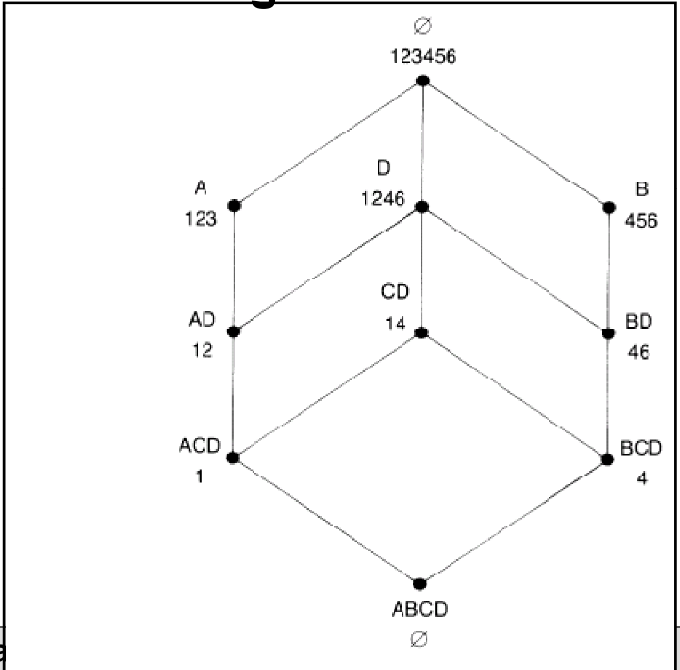


FIGURE 3. Lattice of the data of Figure 2—full labeling.

equivalent

Reduced labeling

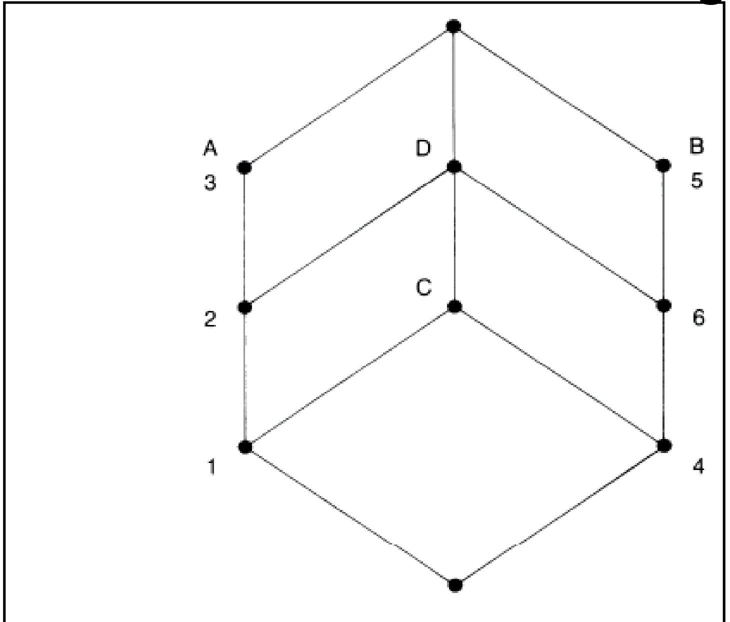
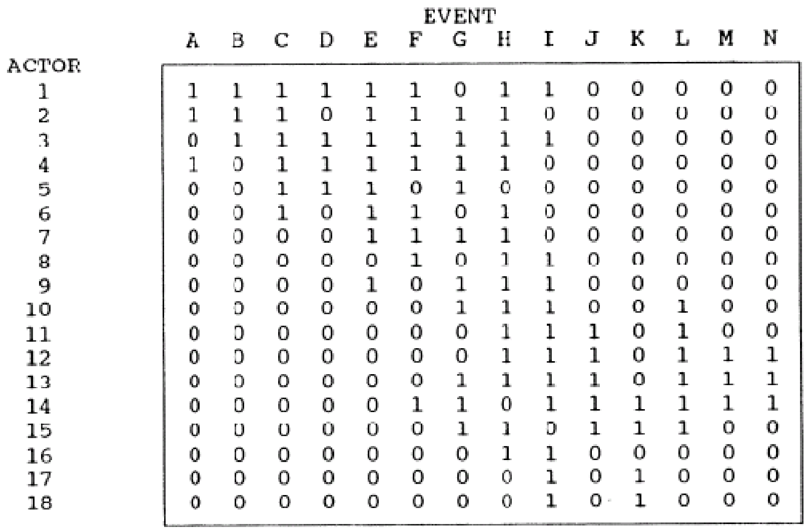


FIGURE 4. Lattice of the data of Figure 2—reduced labeling.

Galois Lattice - Example

[Freeman White 1993]



	EVENT													
ACTOR	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	1	1	1	1	1	1	0	1	1	0	0	0	0	0
2	1	1	1	0	1	1	1	1	0	0	0	0	0	0
3	0	1	1	1	1	1	1	1	1	0	0	0	0	0
4	1	0	1	1	1	1	1	1	0	0	0	0	0	0
5	0	0	1	1	1	0	1	0	0	0	0	0	0	0
6	0	0	1	0	1	1	0	1	0	0	0	0	0	0
7	0	0	0	0	1	1	1	1	0	0	0	0	0	0
8	0	0	0	0	0	1	0	1	1	0	0	0	0	0
9	0	0	0	0	1	0	1	1	1	0	0	0	0	0
10	0	0	0	0	0	0	1	1	1	0	0	1	0	0
11	0	0	0	0	0	0	0	1	1	1	0	1	0	0
12	0	0	0	0	0	0	0	1	1	1	0	1	1	1
13	0	0	0	0	0	0	1	1	1	1	0	1	1	1
14	0	0	0	0	0	1	1	0	1	1	1	1	1	1
15	0	0	0	0	0	0	1	1	0	1	1	1	0	0
16	0	0	0	0	0	0	0	1	1	0	0	0	0	0
17	0	0	0	0	0	0	0	0	1	0	1	0	0	0
18	0	0	0	0	0	0	0	0	1	0	1	0	0	0

FIGURE 5. Davis, Gardner, and Gardner's two mode data.

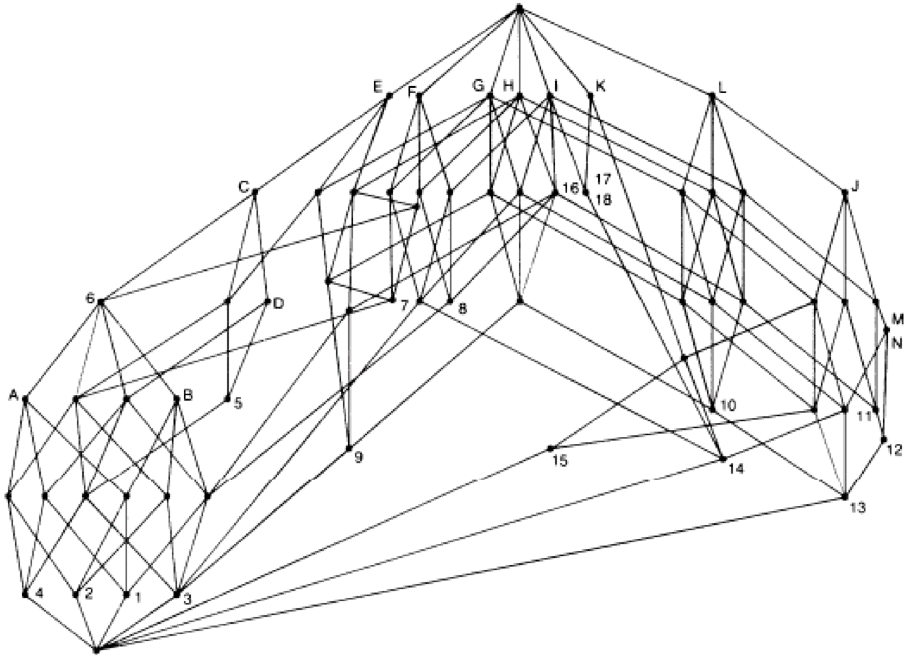


FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

Galois Lattice

[Freeman White 1993]

What can we do with Galois Lattices?

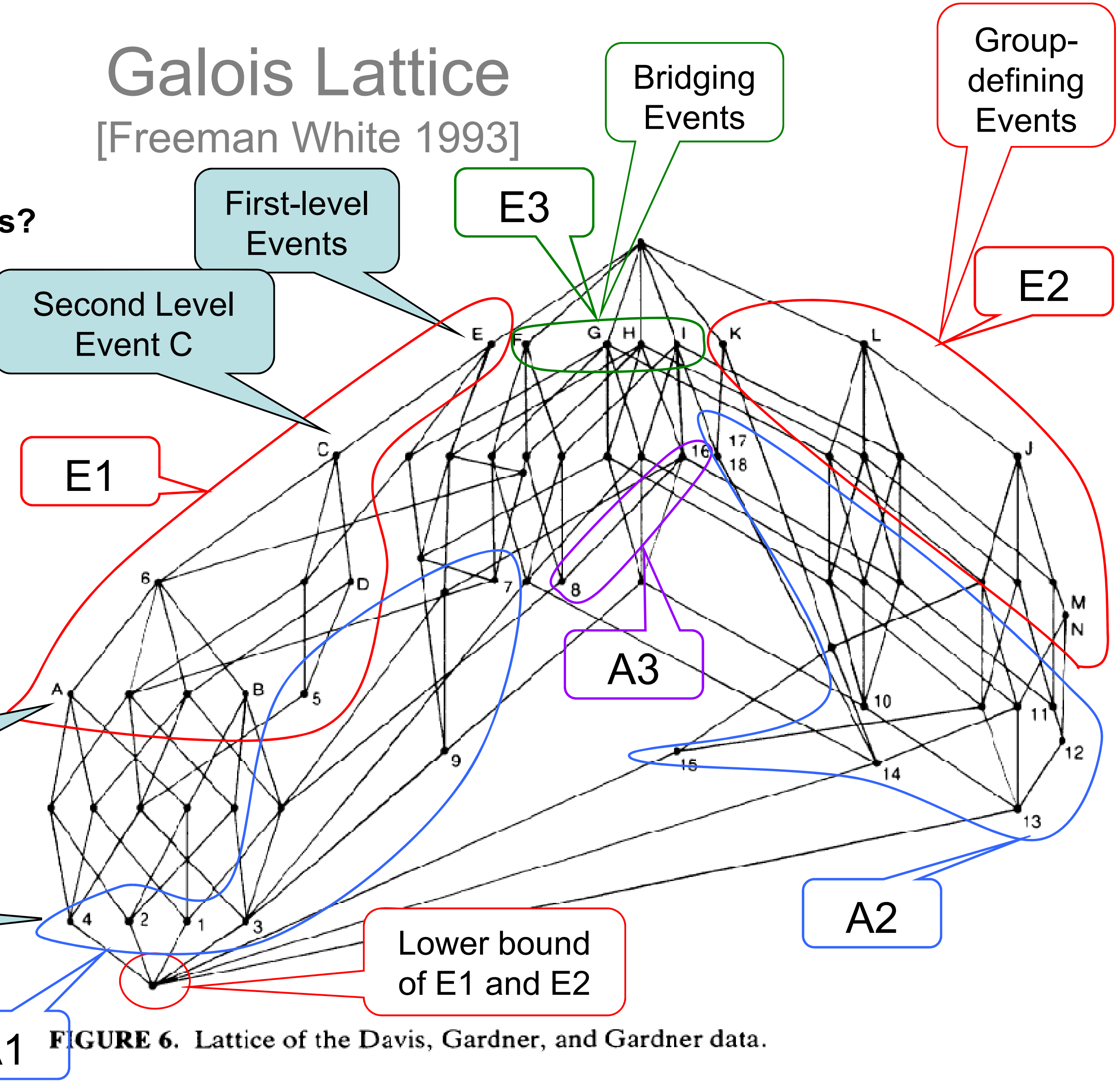
1. We can see the pattern of participation of actors in events.
   - Each actor (or set of actors) participated in those events labeled at or above her labeled point in the line diagram and each event (or set of events) included all the actors labeled at or below its point.
   - Thus the relation λ (I) is displayed, and the original data are **completely recoverable from the diagram**.
2. We can see the downward containment structures of events.
   - The **uppermost set** of seven labeled events (E, F, G, H, I, K, and L) are the events that involved the largest sets of actors.
   - Other events are contained in the lower intersections (meets) of these events. Event C is a second level event: It is contained in event E, and events A, B, and D are. in turn, third level events; they are contained in C (and therefore in E).
   - Similarly, event J is **second level**, contained in L, and M and N are **third level**, contained in J.
3. We can see the upward containment structures of actors.
   - The lowest labeled actors (1, 2, 3, 4, 13, 14, and 15) are **primary**. They are the actors who were active in the largest sets of events.
4. We can distinguish classes of events.
   - Two sets of events E1 = {A, B, C, D, E} and E2 = {J, K, L, M, N} share no common actor. This is shown by the fact that their lower bound falls at the bottommost point, the point that contains no common actors. Therefore. E1 and E2 are **group-defining events**. In contrast. The four events E3 = {F,G, H, I) each share at least one actor with events in E1, and at least one actor with events in E2; they might be called **bridging events**.
5. We can see the segregation of actors by the event classes.
   - The nonoverlapping event sets E1 and E2 segregate all but two of the actors into two sets A1 = {1, 2, 3, 4, 5, 6, 7, 9) and A2 = (10, 11, 12, 13, 14, 15, 17, 18). Actors from these different subsets never interact in the non-overlapping events.

Galois Lattice

[Freeman White 1993]

What can we do with Galois Lattices?

1. We can see the pattern of participation of actors in events.
2. We can see the downward containment structures of events.
3. We can see the upward containment structures of actors.
4. We can distinguish classes of events.
5. We can see the segregation of actors by the event classes.



Bridging Events

Group-defining Events

First-level Events

E3

E2

Second Level Event C

E1

Third-level Events A,B, D

A3

Primary Actors

Lower bound of E1 and E2

A2

A1

FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

Advantages of Galois Lattices

[Wasserman Faust 1994]

- Focus on subsets
  - Especially appropriate for representing affiliation networks
- Complementary relationships between actors and events displayed at the same time
- Patterns in the relationships between actors and events may be more apparent in the Galois lattice

➔Galois lattice serves much the same function as a graph or sociogram (which serves as a representation of a one-mode network)

Shortcomings of Galois Lattices

[Wasserman Faust 1994]

- Visual display of Galois Lattices can become quite complex
- No unique „best“ visual representation for a given Galois lattice
- Although the vertical dimension represents the degrees of subset inclusion relationships, the horizontal dimension is arbitrary.
- Properties and further analyses of Galois lattices (unlike networks) are not well developed

➔ Galois lattices are primarily an exploratory representation of an affiliation network, from which one might be able to see patterns in the data.

ConExp

Download: http://sourceforge.net/projects/conexp

Project Website: http://conexp.sourceforge.net/

Documentation:
http://conexp.sourceforge.net/users/documentation/

- Import your network data using .csv format
  (see „Opening existing documents“ in the documentation)
- Experiment with „Drawing Options“ and layouting the lattice manually
- Show „multi-labels“ for objects and attributes
- Interpret the lattice
- Save it in .cex file

Any questions?

See you in two weeks!