707.000
Web Science and Web Technology
„Affiliation Networks“

How can we visualize and analyze 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

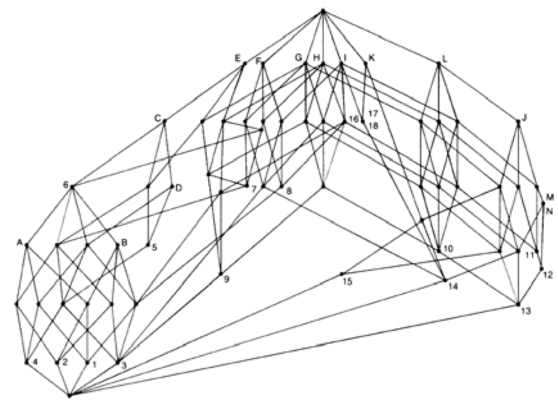


FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

[Freeman White 1993]

Home Assignments

Results HA1-2:

http://kmi.tugraz.at/staff/markus/courses/papers/results/WSWT08_HA1-2.pdf

Available this week.

Overview

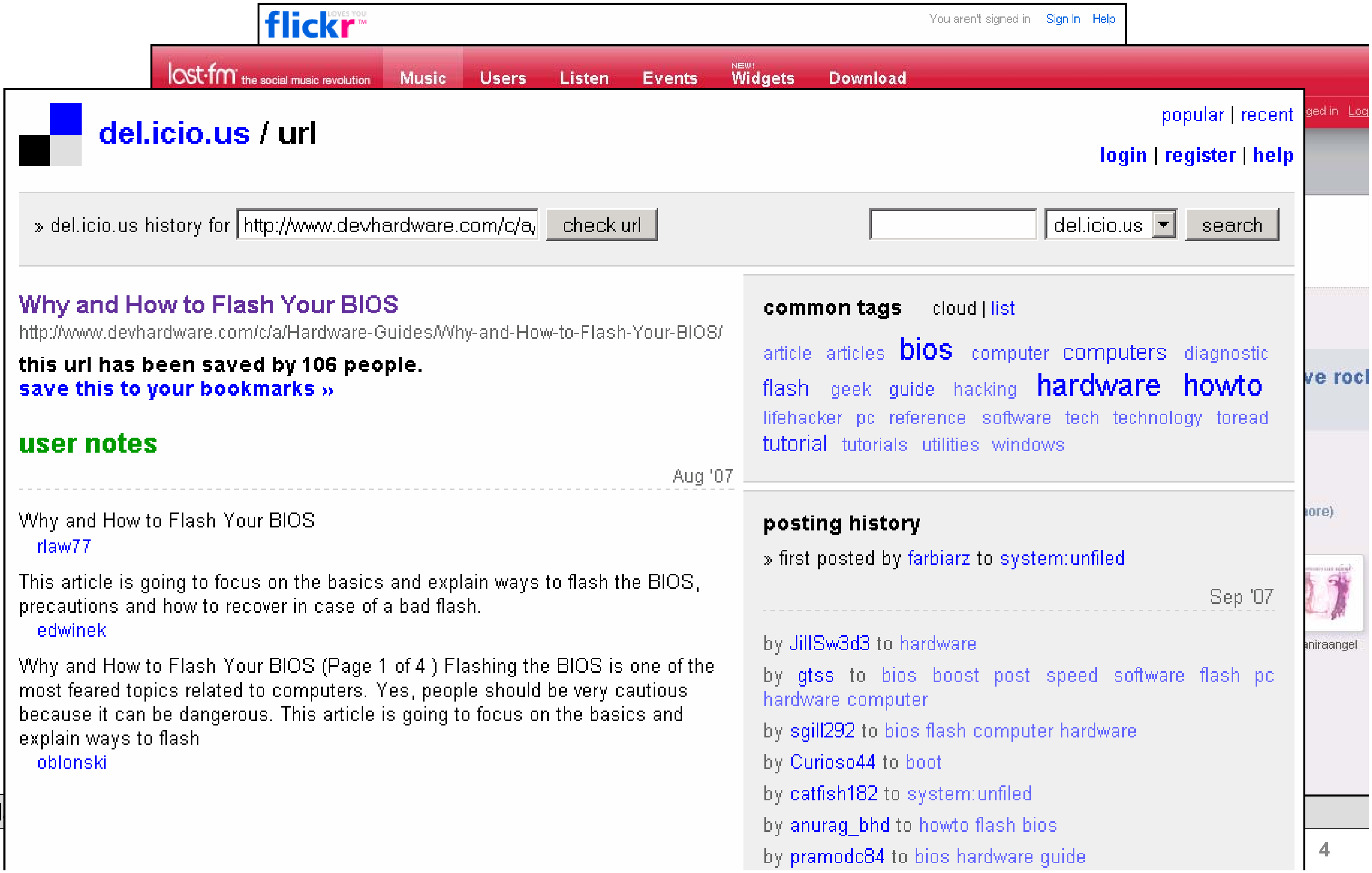
New course content!

Today‘s Agenda:

Analysis of Afilliation Networks

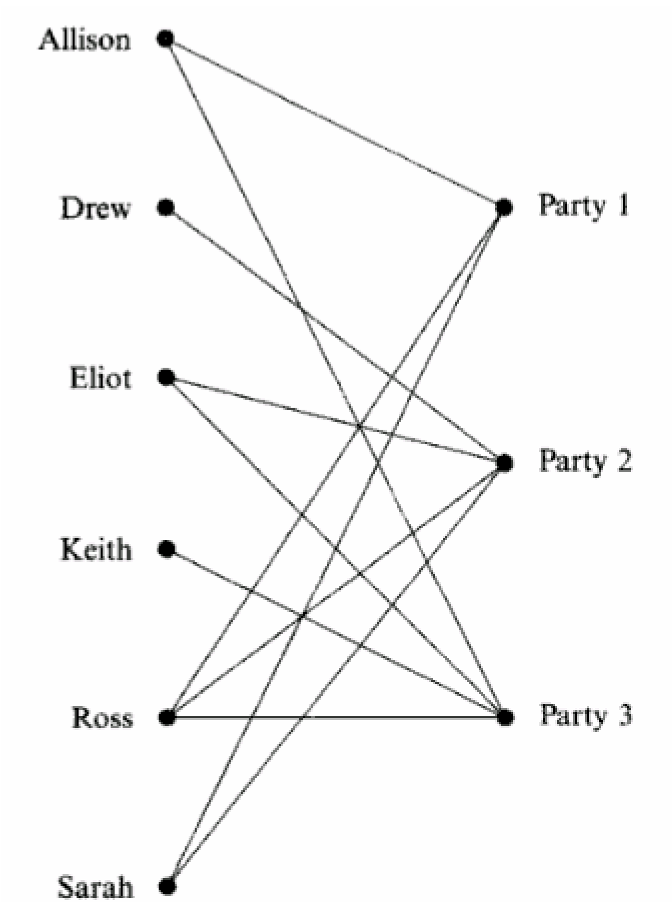
- A (very brief) recapitulation of Affiliation Networks
- Properties of Affiliation Networks
  - Properties of Actors and Events
  - Properties of One-Mode Networks derived from Affiliation Networks
- Galois Lattices for Affiliation Networks

Reminder: Social Networks Examples



Representing Affiliation Networks As Two Mode Sociomatrices

[Wasserman Faust 1994]



General form:
$$\begin{pmatrix} 0 & A \\ A' & 0 \end{pmatrix}$$

Affiliation matrix

	Allison	Drew	Eliot	Keith	Ross	Sarah	Party 1	Party	arty 3
Allison	-	0	0	0	0	0	1	0	1
Drew	0	-	0	0	0	0	0	1	0
Eliot	0	0	-	0	0	0	0	1	1
Keith	0	0	0	-	0	0	0	0	1
Ross	0	0	0	0	-	0	1	1	1
Sarah	0	0	0	0	0	-	1	1	0
Party 1	1	0	0	0	1	1	-	0	0
Party 2	0	1	1	0	1	1	0	-	0
Party 3	1	0	1	1	1	0	0	0	-

actors

events

Fig. 8.3. Sociomatrix for the bipartite graph of six children and three parties

Transforming Two Mode Networks into One Mode Networks

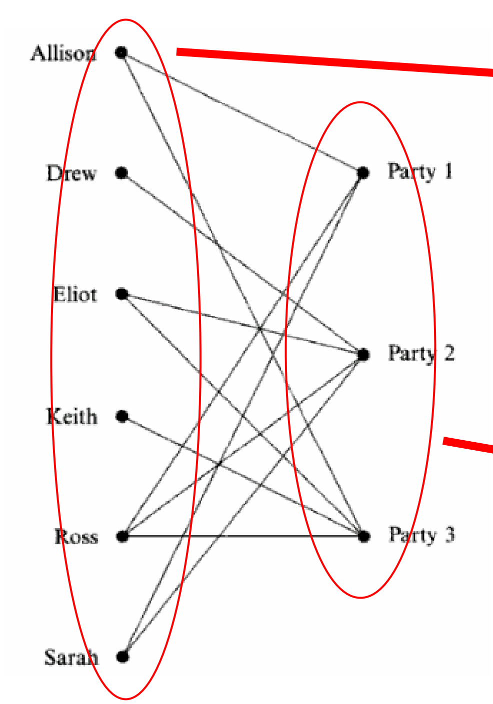
[Wasserman Faust 1994]

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

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

C…Children

P…Party



	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]

Subgroups in Co-Affiliation Networks

In the last lecture, we discussed

- cliques
- n-cliques
- n-clans
- n-clubs
- bi-cliques

as instruments to analyze one-mode networks, but …

Subgroups in Co-Affiliation Networks

Borgatti 1997

- 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.**

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

Rates of Participation

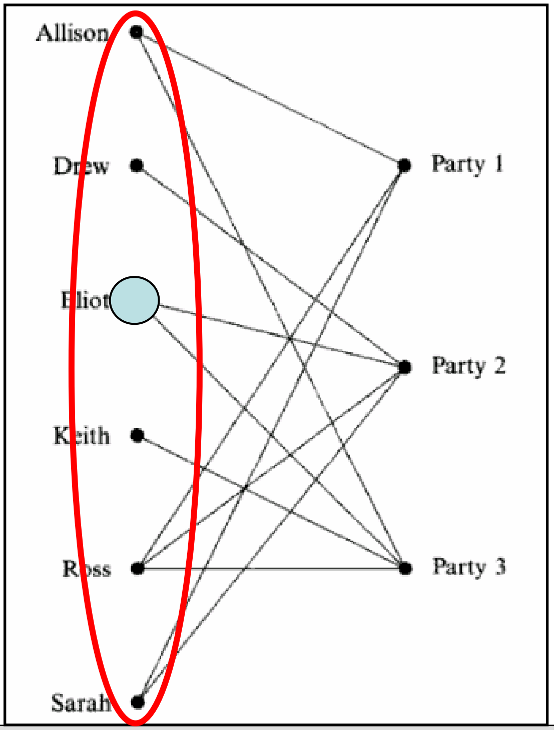
[Wasserman Faust 1994]

Example:

	Party 1	Party 2	Party 3
Allison	1	0	1
Drew	0	1	0
Eliot	0	1	1
Keith	0	0	1
Ross	1	1	1
Sarah	1	1	0

- The number of events with which each actor is affiliated.
- These quantities are either given by
  - the row totals of affiliation matrix A or
  - the entries on the main diagonal of the one-mode socio-matrice X^N
- Thus, the number of events with which actor i is affiliated is equal to the degree of the node representing the actor in the bipartite graph.
- Also interesting: **Average rate of participation**

Examples: What does the rate of participation relate to in the Netflix / Amazon bipartite graph of customer/movies or customer/products?



Size of Events

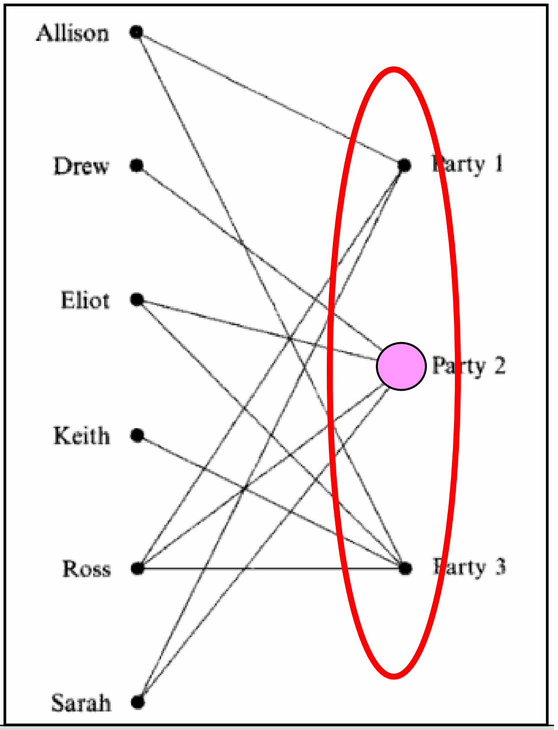
[Wasserman Faust 1994]

- The number of actors participating in each event.
- The size of each event is given by either
  - the column totals of the affiliation matrix A or
  - the entries on the main diagonal of the one-mode sociomatrix X^M.
- Thus, the size of each event is equal to the degree of the node representing the event in the bipartite graph.
- Also interesting: **Average size of events**
  - Sometimes useful to study average size of clubs or organizations
- Size of events might be constrained:
  - E.g. board of company directors are made up of a fixed number of people

Examples: What does the rate of participation relate to in the Netflix / Amazon bipartite graph of customer/movies or customer/products?

Example:

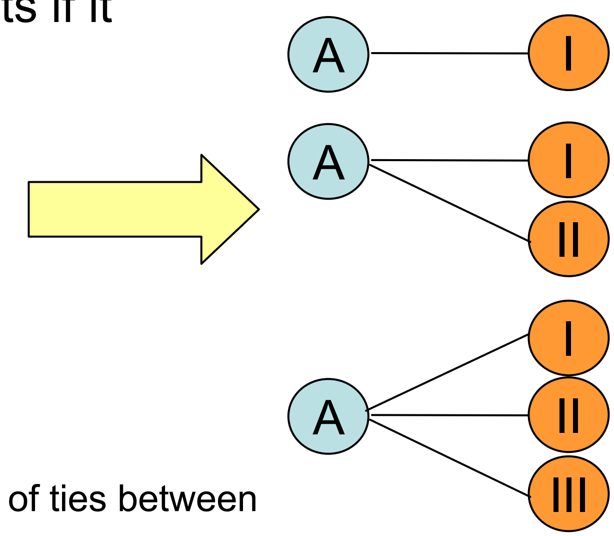
	Party 1	Party 2	Party 3
Allison	1	0	1
Drew	0	1	0
Eliot	0	1	1
Keith	0	0	1
Ross	1	1	1
Sarah	1	1	0



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

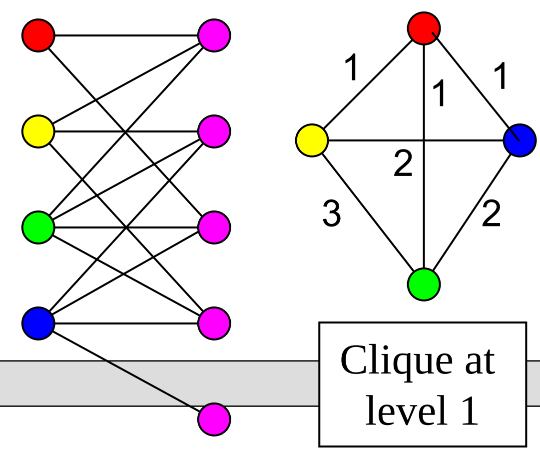
- In an affiliation network,
  - no pair of actors is adjacent
  - No pair of events is adjacent
- no paths of length 1 between actors, but potentially paths of some longer length
- **Reachability** corresponds to path lengths between nodes
- An affiliation network is **connected** when all pairs of 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.

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

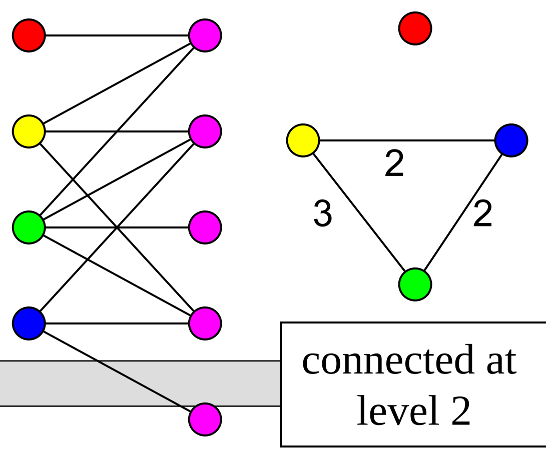


Clique at level 1

Reachability for Pairs of Actors
[Wasserman Faust 1994]

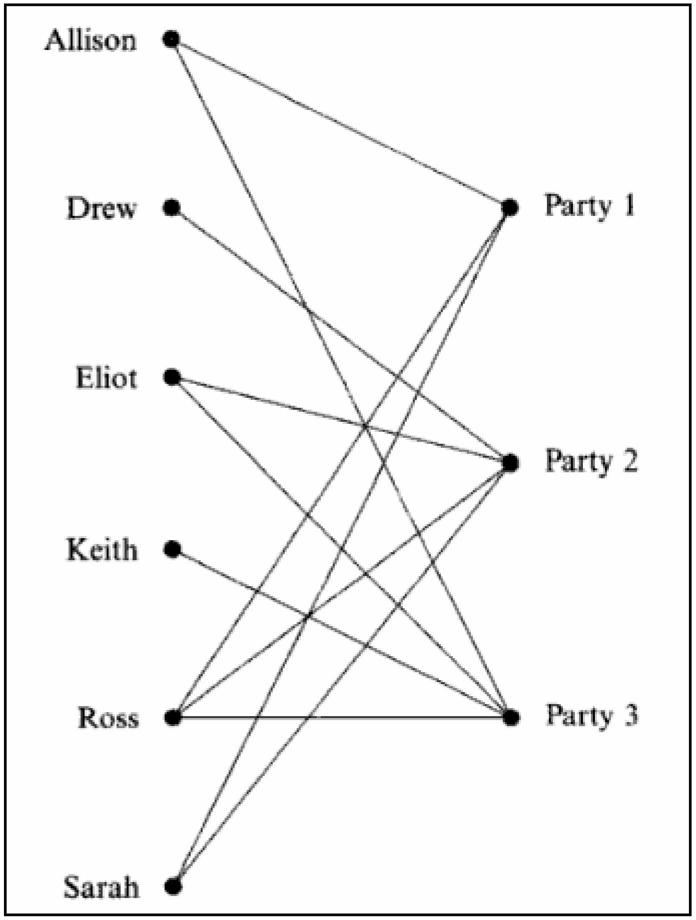
- In a valued graph, we can **define connectedness at level c** as the subsets of actors all of whom are connected at some minimum level c
- Two nodes are **c-connected** (or reachable at level c) if there is a path between them in which all lines have a value of no less than c.

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



Affiliation Networks

[Wasserman Faust 1994]

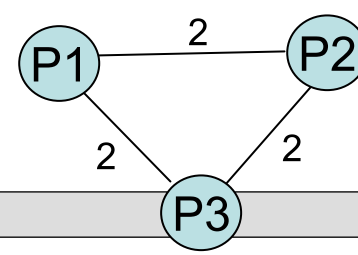


„**Two mode**“ networks

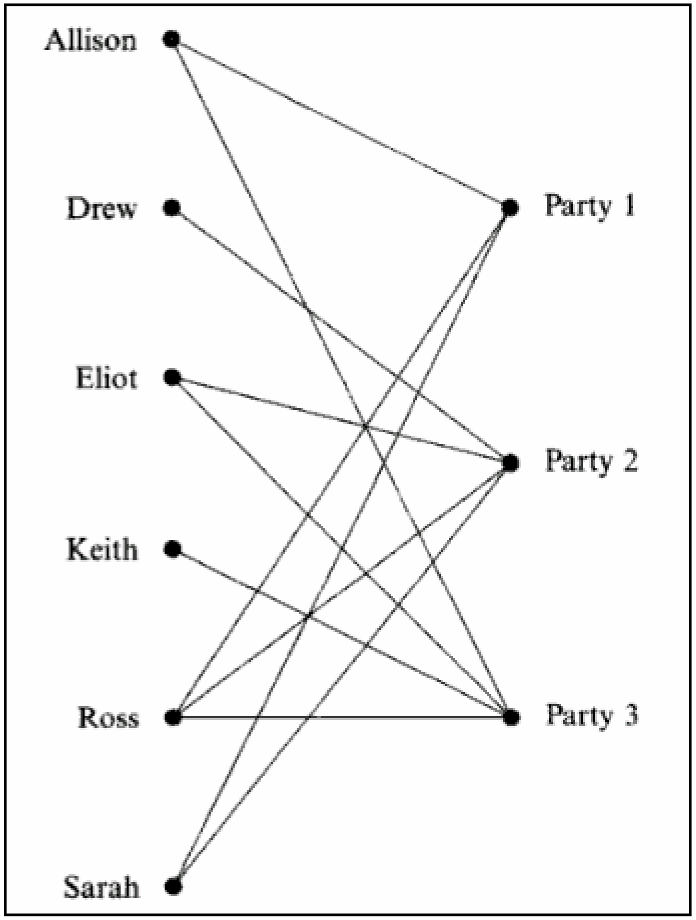
- Links only between the two modes
- Folding
- K-neighbourhood graph
- KNC Plot

But:

By folding, information is lost



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**.

Galois Lattices

[Wasserman Faust 1994, Freeman White 1993]

- A long history in Mathematics
- Introduced by Birkhoff in 1940
- 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 $N = \{n_1, n_2, \ldots n_g\}$ 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.

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) 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) 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.

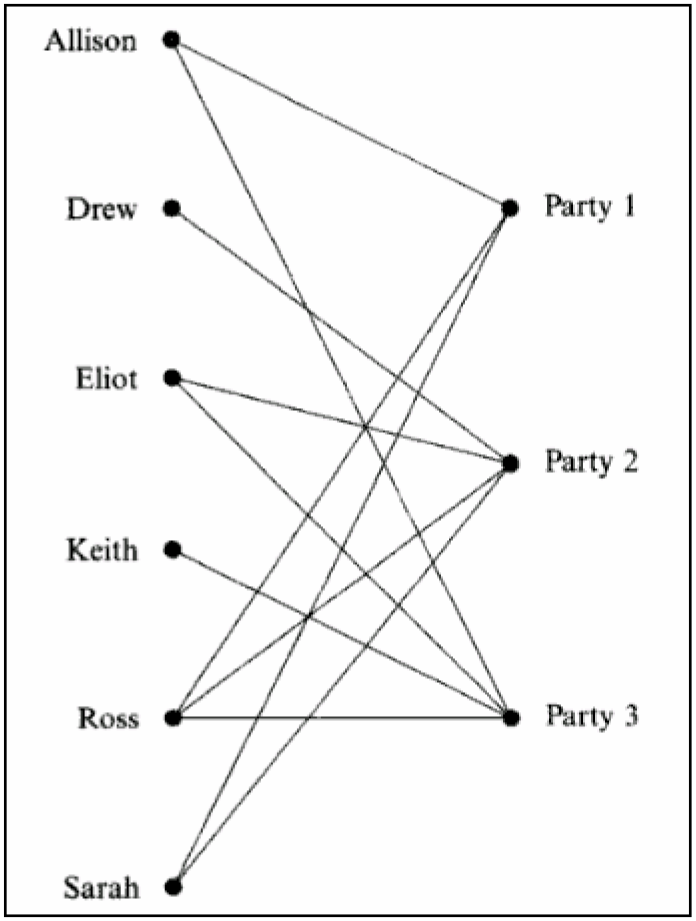
A *lattice* is thus a partially ordered set in which each pair of elements has both a meet and a join.

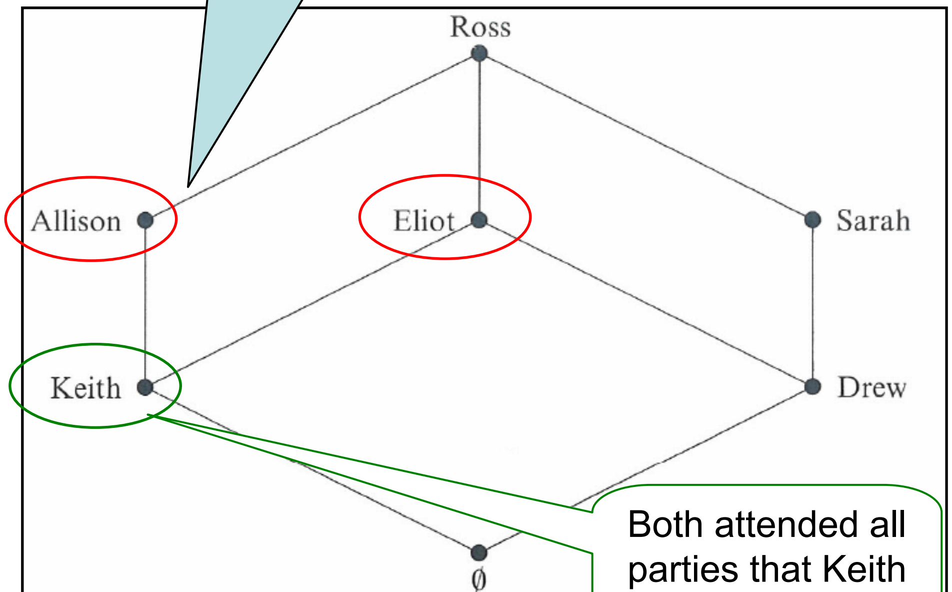
“⊆“ as our relation

[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.

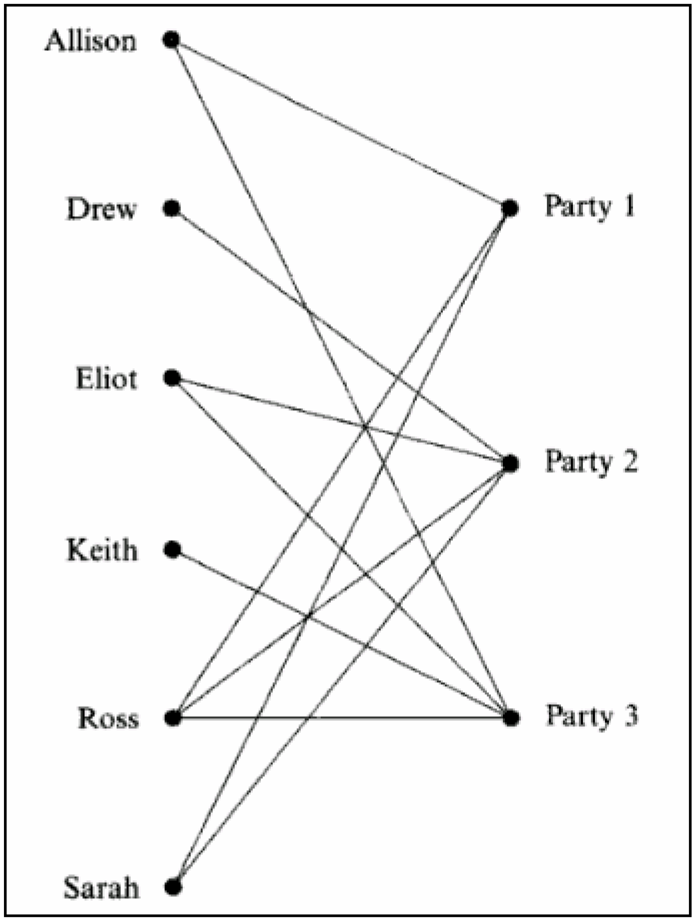
Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

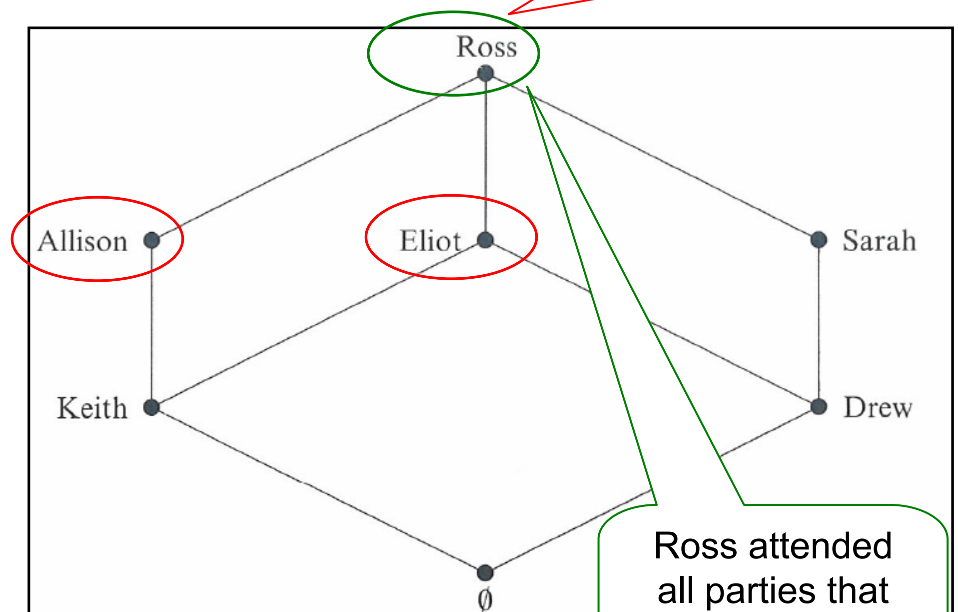
Fig. 8.10. Relationships among children as subsets of birthday parties

Lattice

[Wasserman Faust 1994]

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





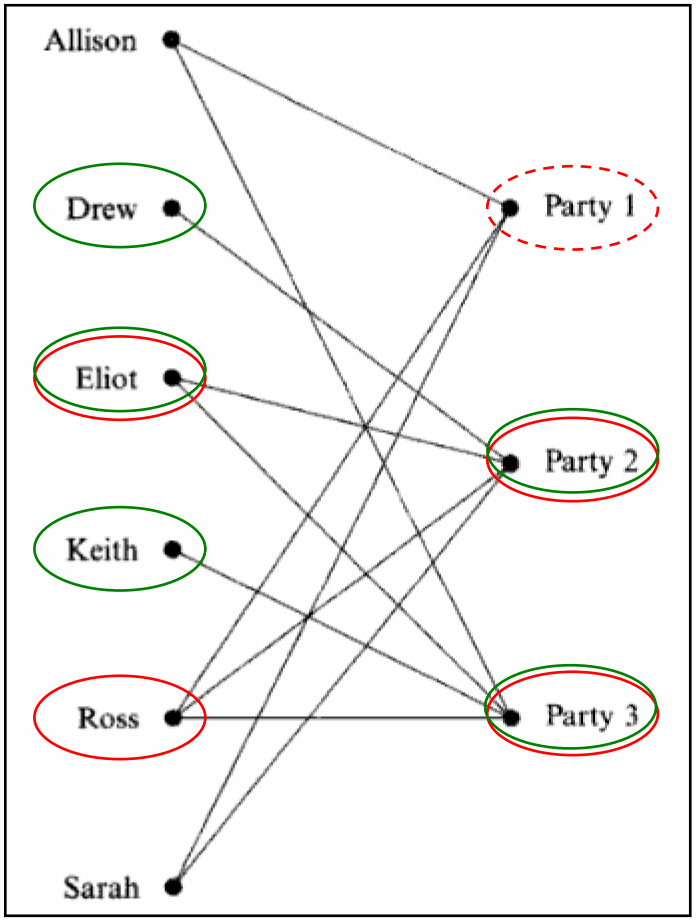
Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

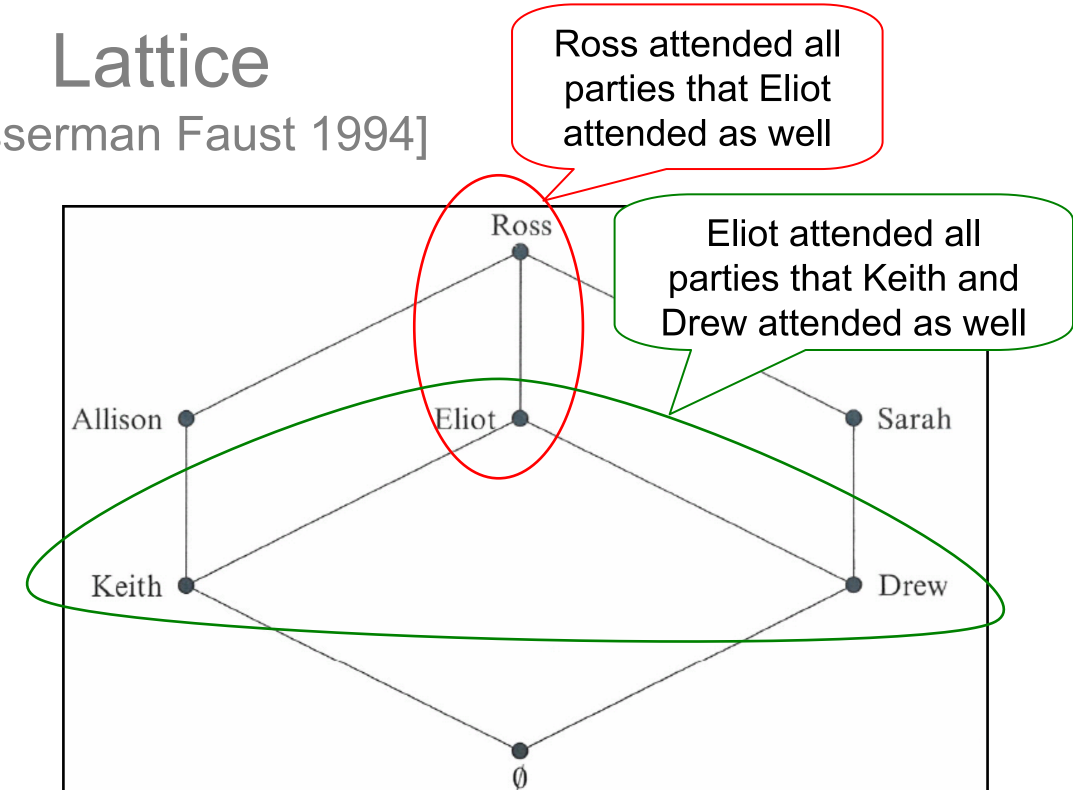
Fig. 8.10. Relationships among children as subsets of birthday parties

Ross attended all parties that Allison and Eliot attended.

Lattice

[Wasserman Faust 1994]





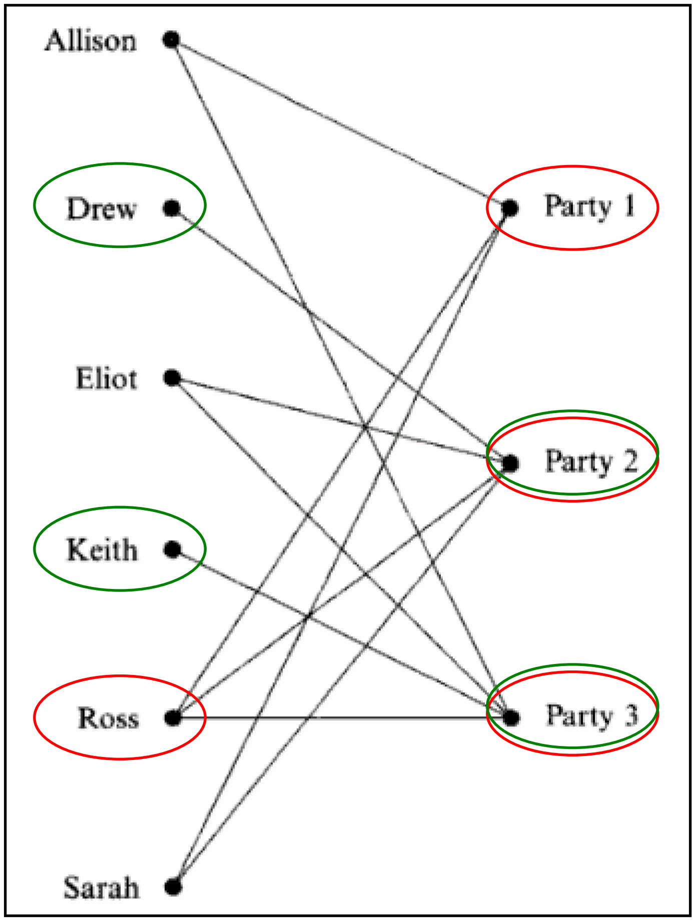
Drew: {Party 2}
Keith: {Party 3}
Sarah: {Party 1, Party 2}
Eliot: {Party 2, Party 3}
Allison: {Party 1, Party 3}
Ross: {Party 1, Party 2, Party 3}

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

Lattice

[Wasserman Faust 1994]

Who attended the most parties?



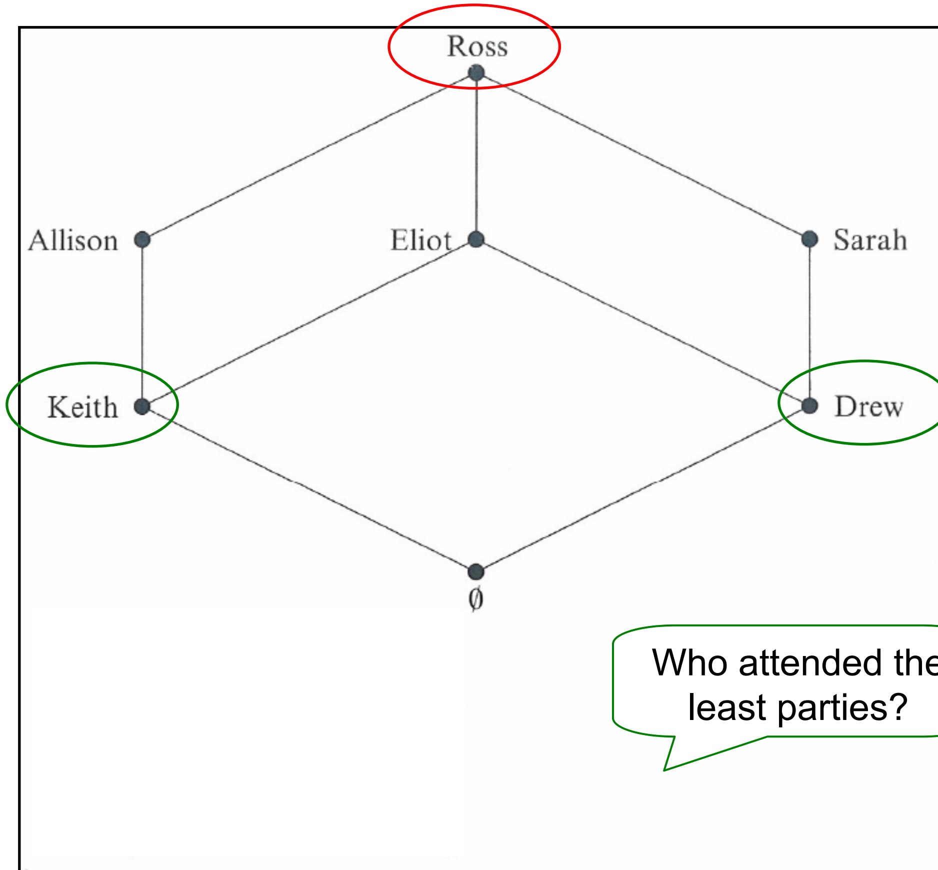


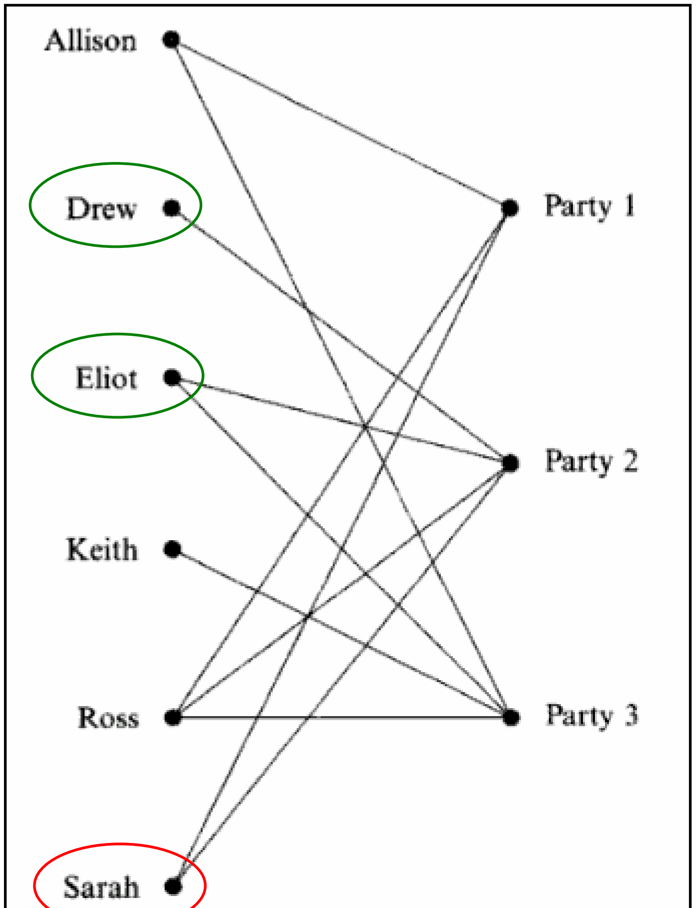
Fig. 8.10. Relationships among children as subsets of birthday parties

Who attended the least parties?

Lattice

[Wasserman Faust 1994]

Did Sarah attend all parties that Eliot and Drew attended?



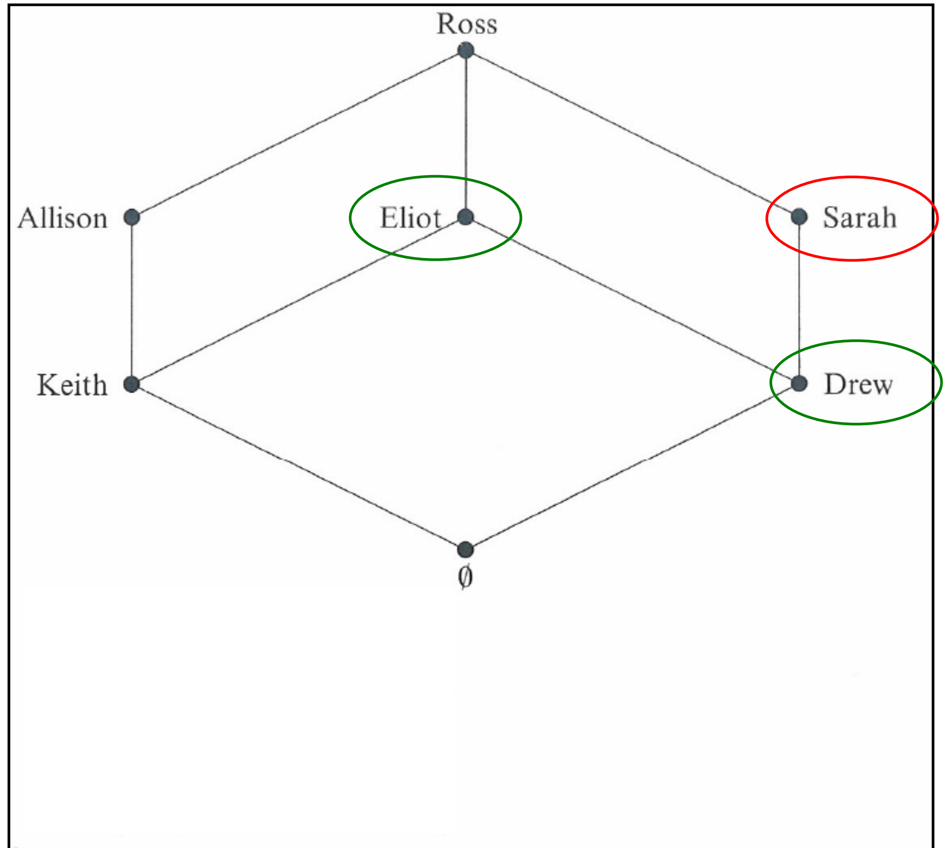
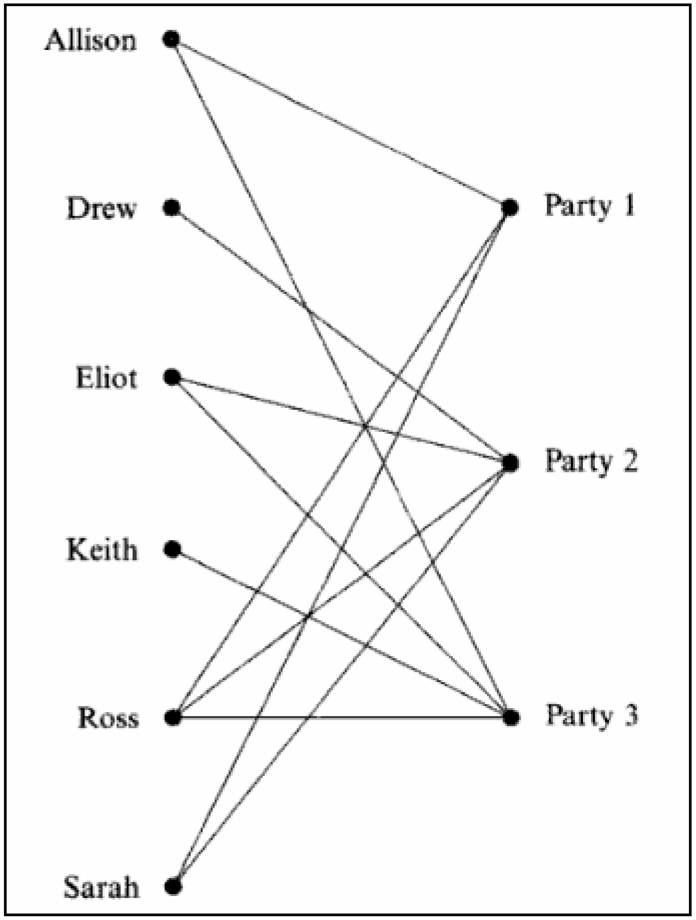


Fig. 8.10. Relationships among children as subsets of birthday parties

Lattice

[Wasserman Faust 1994]

Each point represents a **subset of children**



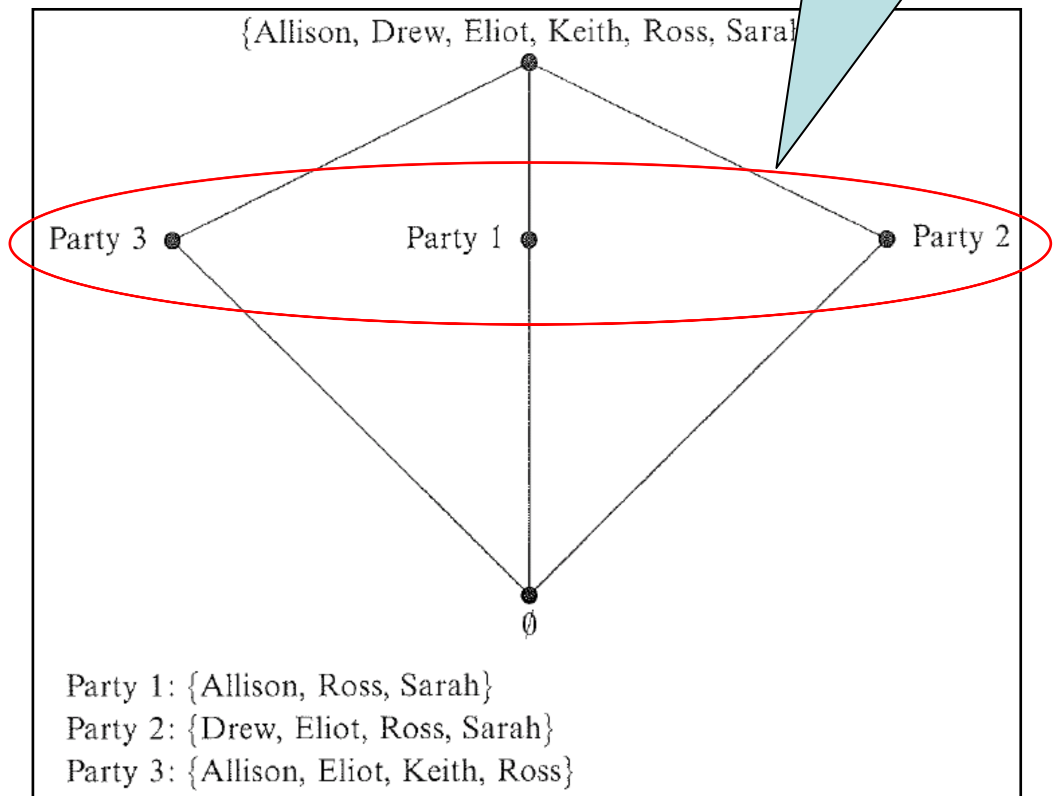


Fig. 8.9. Relationships among birthday parties as subsets of children

A Galois Lattice

[Wasserman Faust 1994]

A **Galois lattice** 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:

- ↑: 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 $n_i \in N$ and all $m_j \in M$.
  - 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 ↑(N_s) = 0
- ↓ mapping can be defined analogously.

Q: Why do we need the top-most and bottom most nodes?

A: each pair of elements ni, nj, must have both a least upper bound and a greatest lower bound

...attice

[Wasserman Faust 19...4]

Consider

- N_s = {Keith, Drew}

Since

- Keith {Party 3}
- Drew {Party 2}

It follows that

- M_s = {Ø} and
- $\uparrow(N_s) = M_s$ = {Ø}

Consider

- M_s = {Party 1, Party 2, Party 3}

Since

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

It follows that

- N_s = {Ross} and
- $\downarrow(M_s) = N_s$ = {Ross}

In terms of attendence Party 3 is a subset of attendees of Ø

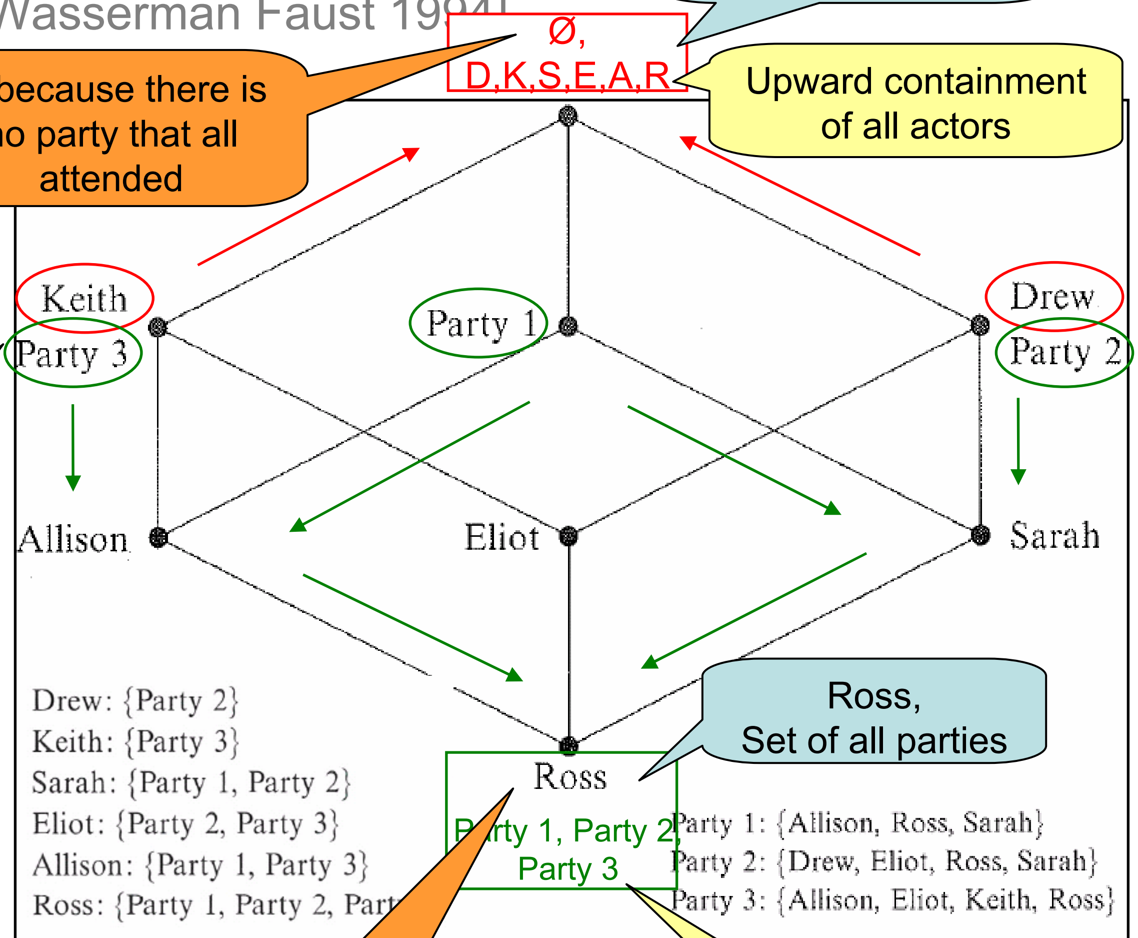


Fig. 8.11. Galois lat... of children and bi... day parties

In terms of an affiliation network ... mapping goes from a ... set of actors to that su... ll of the actors in ... ed.

Ross, because there is a child that attended all parties

Downward containment of all events

Ga…

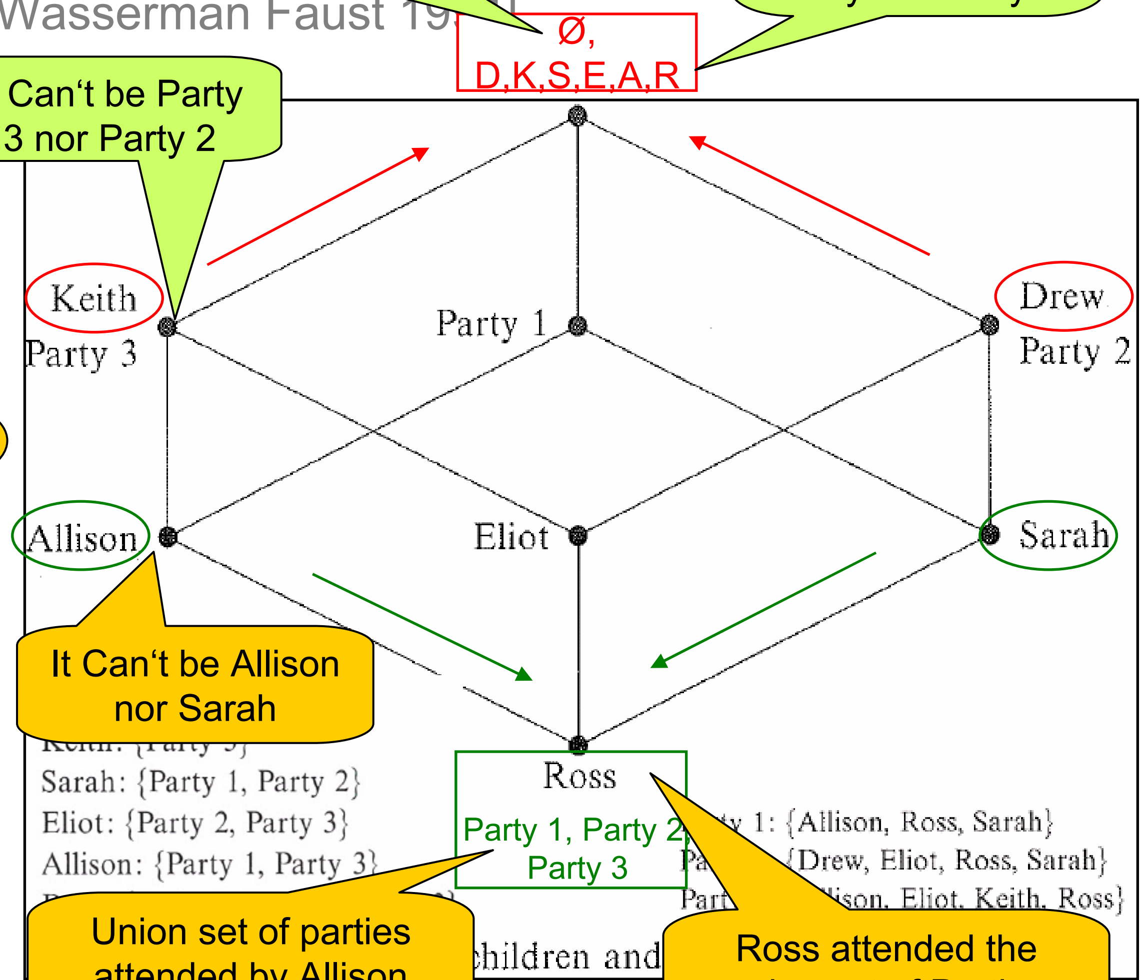
[Wasserman Faust 19…4]

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

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



In terms of an a… ping goes f… at subset of events with which all of the actor… ed.

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}

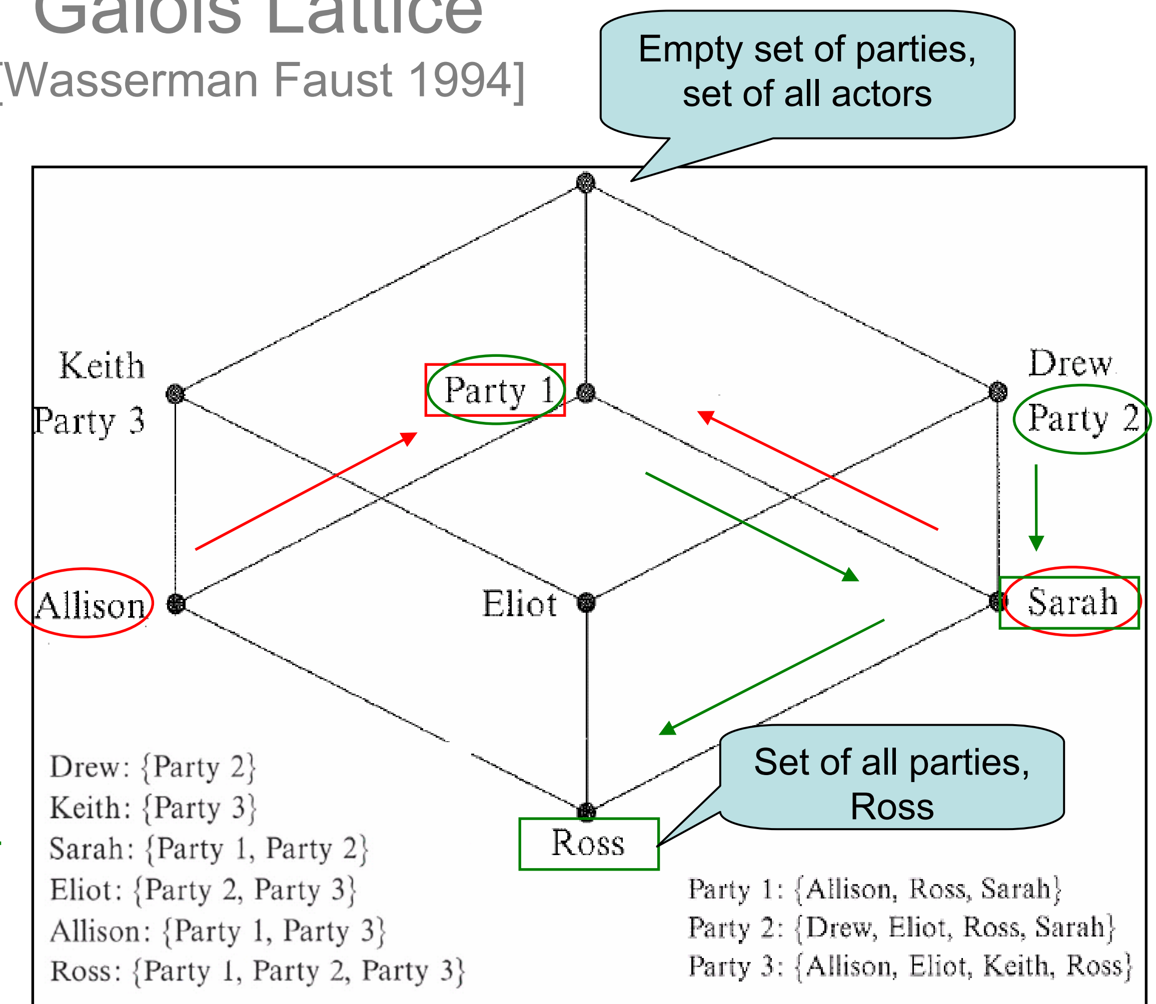


Fig. 8.11. Galois lattice of children and birthday parties

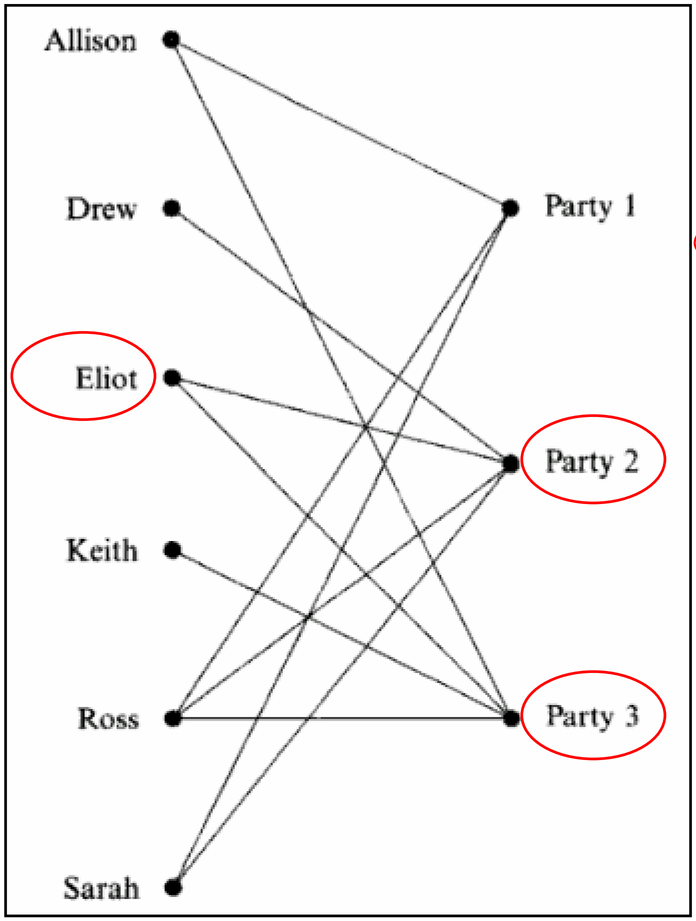
In terms of an affiliation network, the ↑ mapping goes from a subset of actors to that subset of events with which all of the actors in the subset are affiliated.

↑ mapping

Galois Lattice

[Wasserman Faust 1994]

Which parties did Eliot attend?



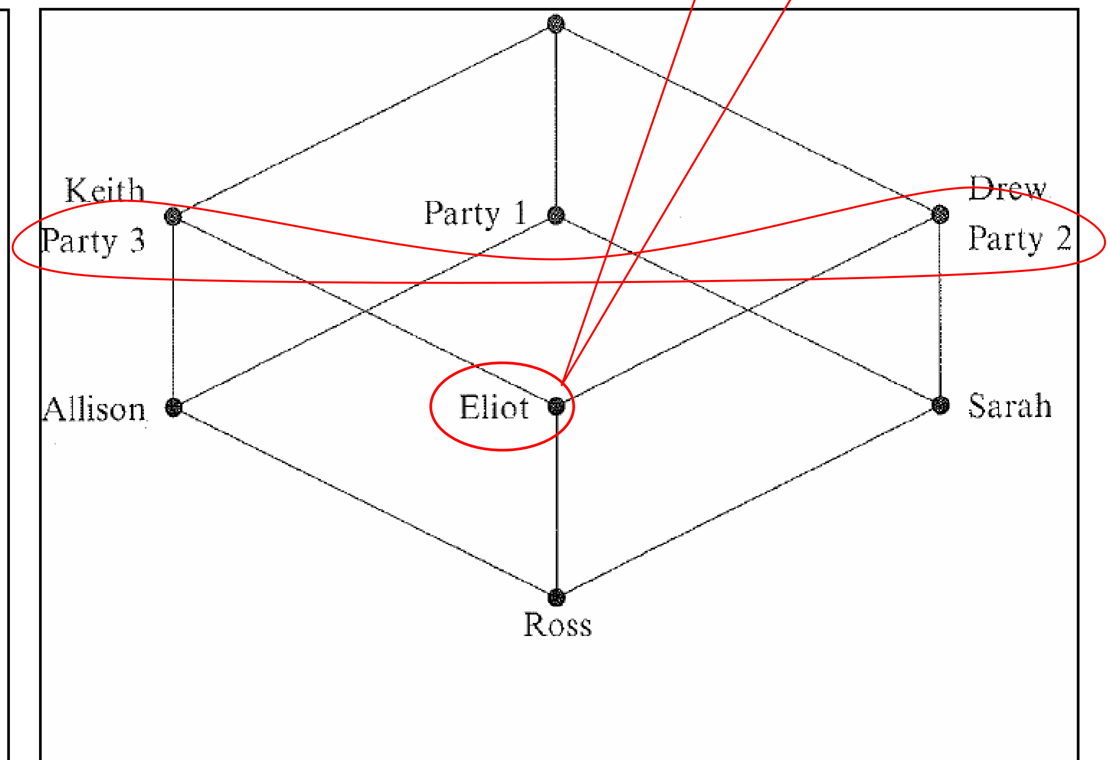


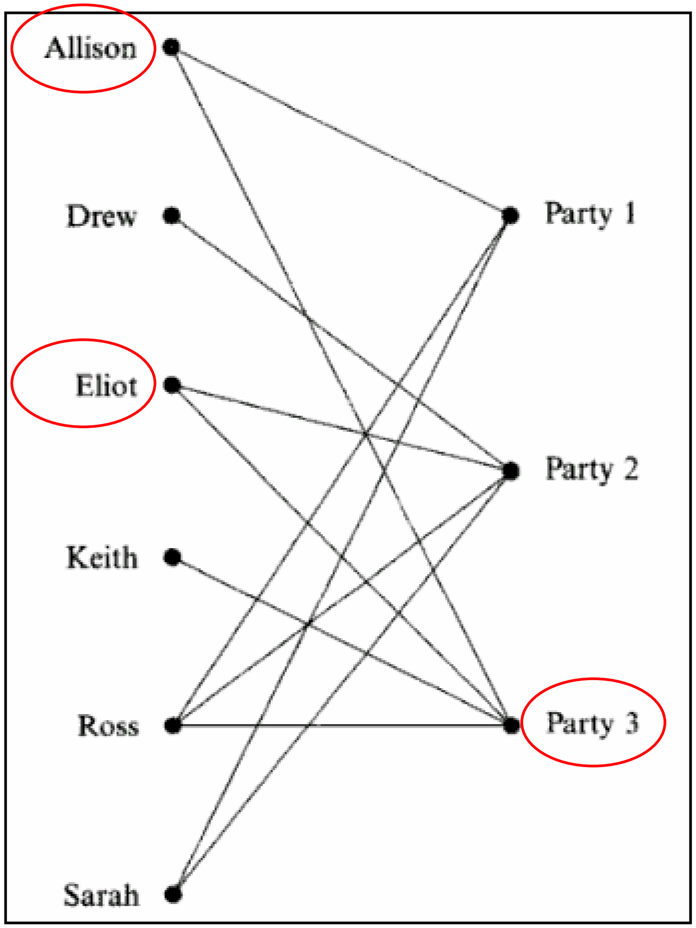
Fig. 8.11. Galois lattice of children and birthday parties

↑ mapping

Galois Lattice

[Wasserman Faust 1994]

Which parties did both Eliot and Allison attend?



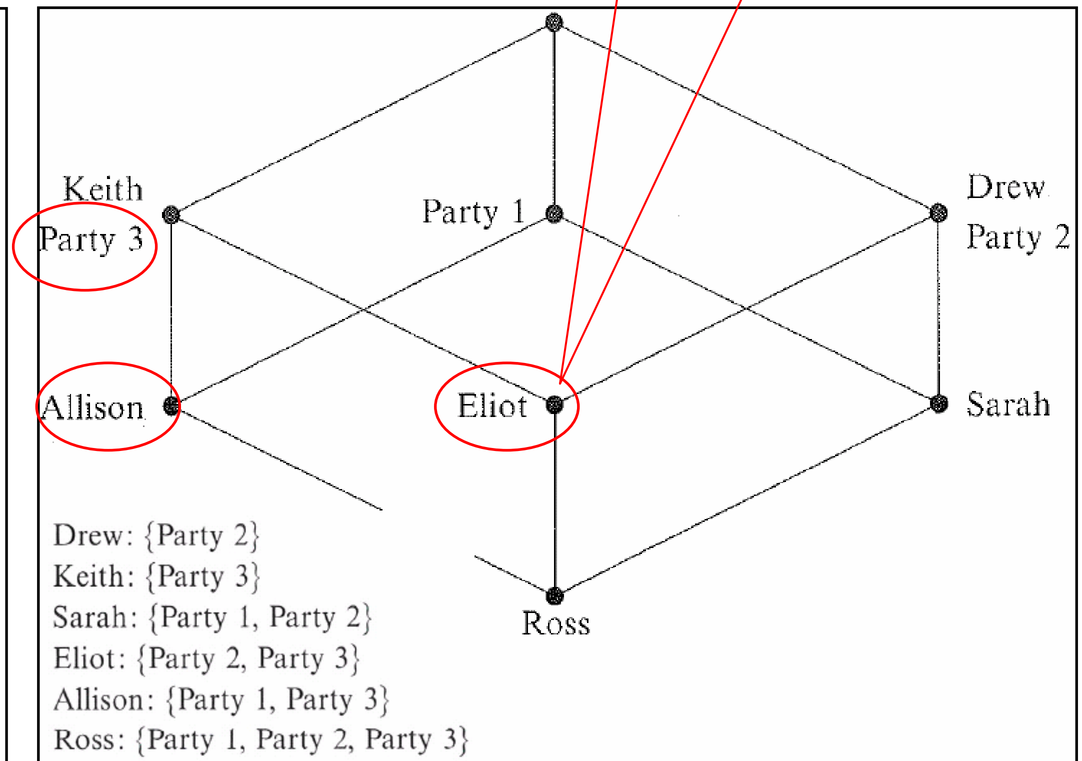


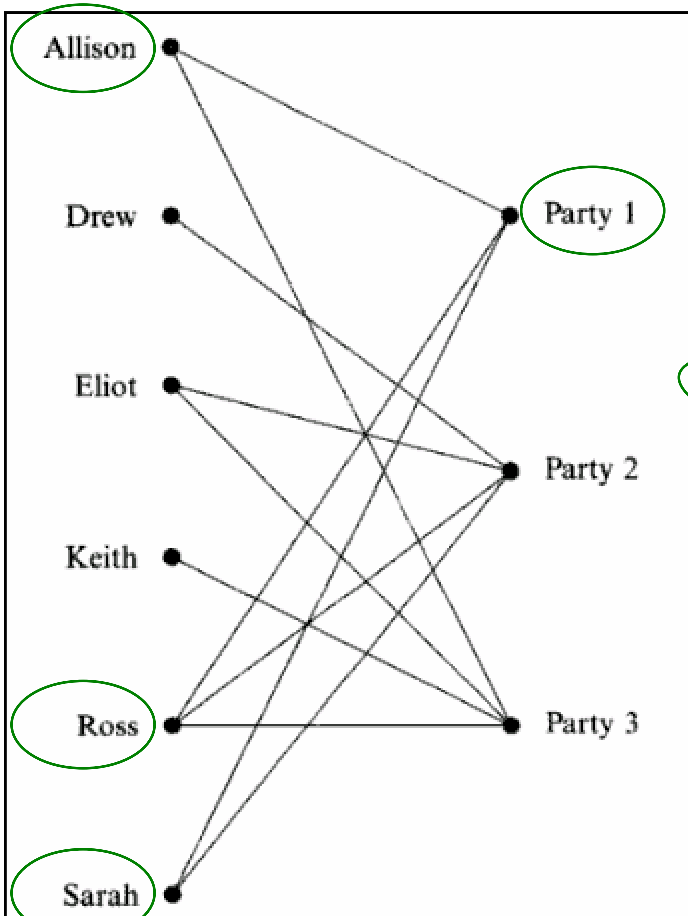
Fig. 8.11. Galois lattice of children and birthday parties

↓ mapping

Galois Lattice

[Wasserman Faust 1994]

Who attended party 1?



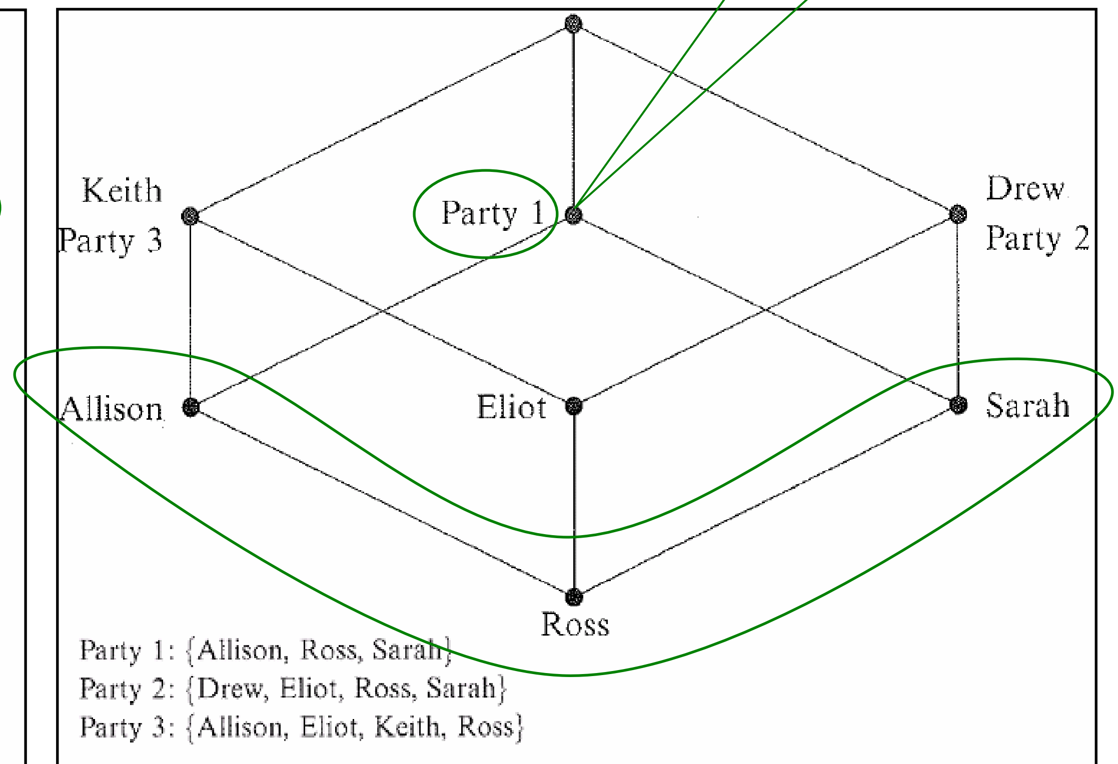


Fig. 8.11. Galois lattice of children and birthday parties

Galois Lattice

[Freeman White 1993]

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

ACTOR	EVENT 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

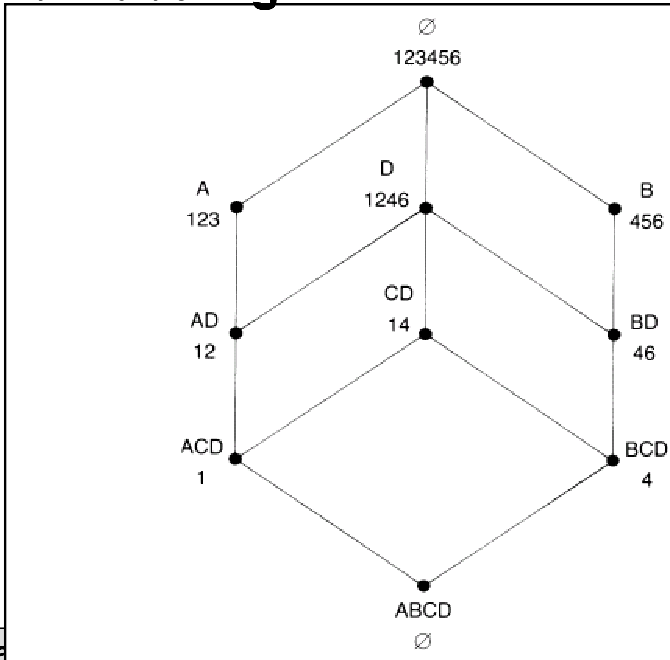
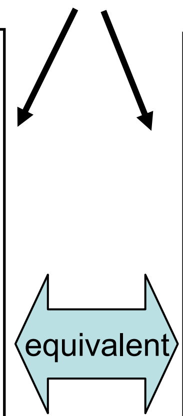


FIGURE 3. Lattice of the data of Figure 2—full labeling.



Reduced labeling

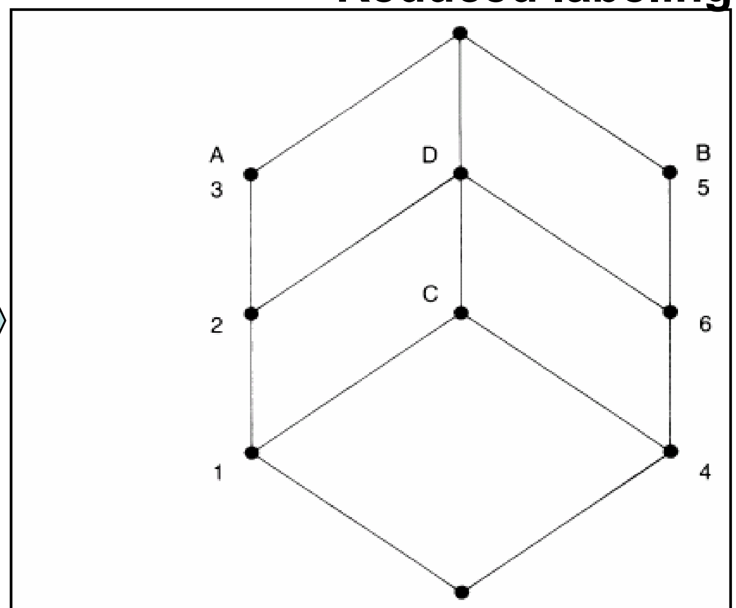


FIGURE 4. Lattice of the data of Figure 2—reduced labeling.

Galois Lattice - Example

[Freeman White 1993]

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.

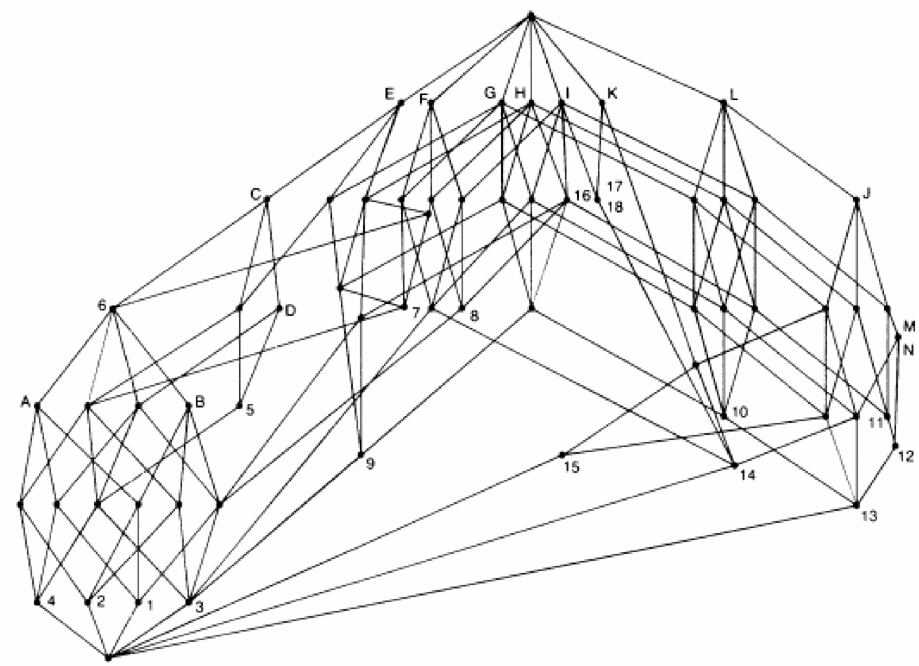


FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

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.

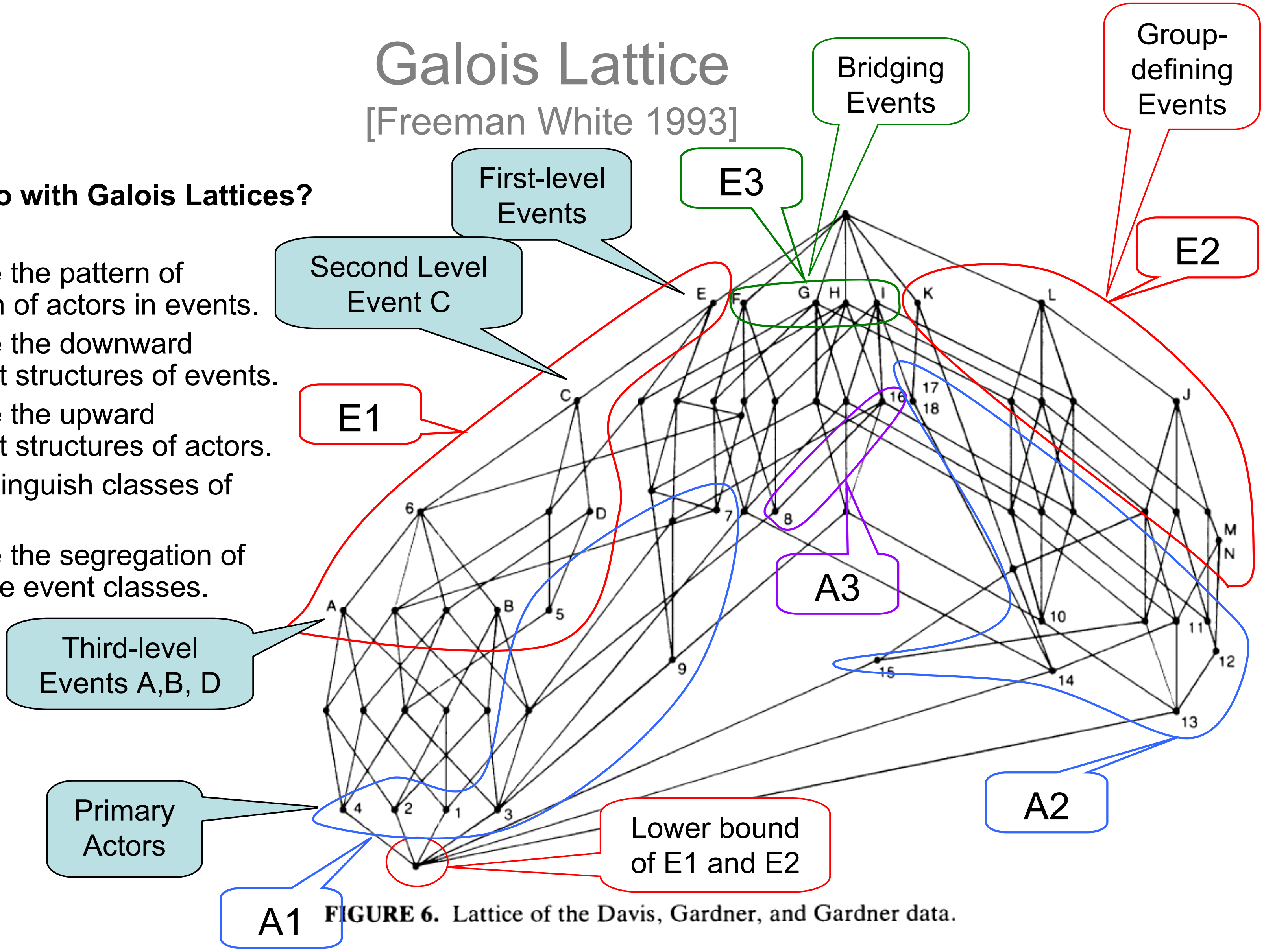


FIGURE 6. Lattice of the Davis, Gardner, and Gardner data.

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.

Any questions?

See you next week!