Finding Cohesive Subgroups in Social Networks

Illya V. Hicks
Computational and Applied Mathematics
Rice University

RUSMP

Fall Networking Conference

September 19, 2009

Outline

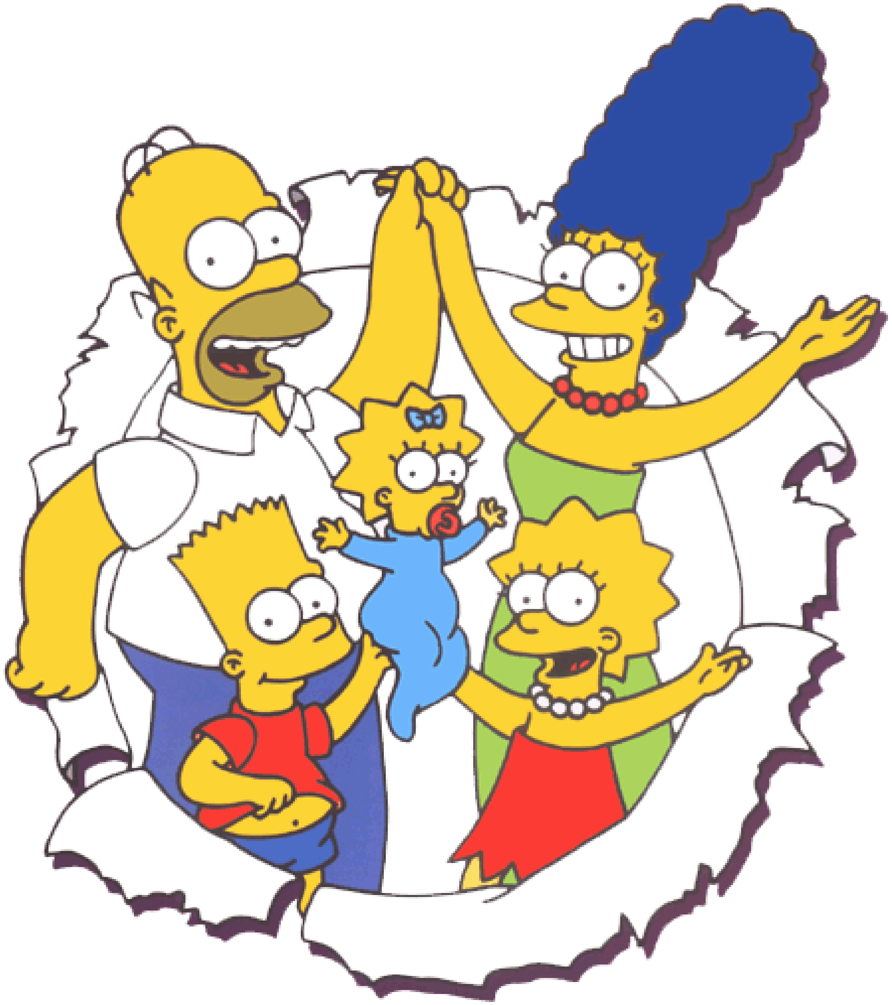
I. Basic Definitions

II. Social Networks

III. *k*-plexes and co-*k*-plexes

IV. Conclusions & Future Work

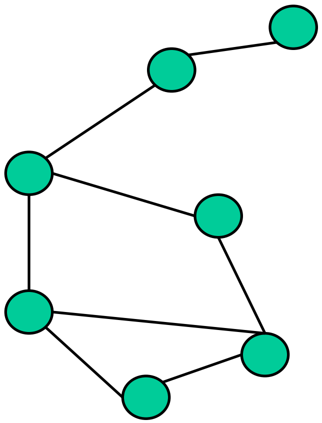
Our Helpers For Today



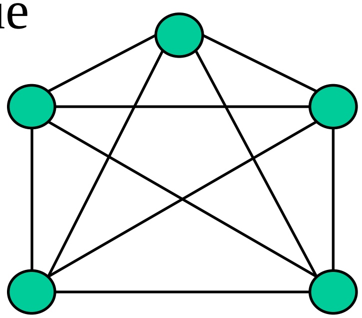
Graphs

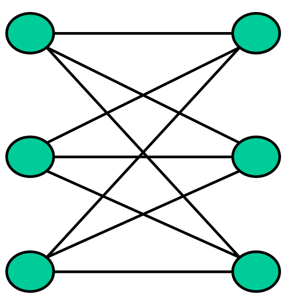
Graph G=(V, E)

- Vertex set V is finite
- Edges E = {uv : u,v $\in$ V}
- Undirected (for this talk)



clique





Network Applications

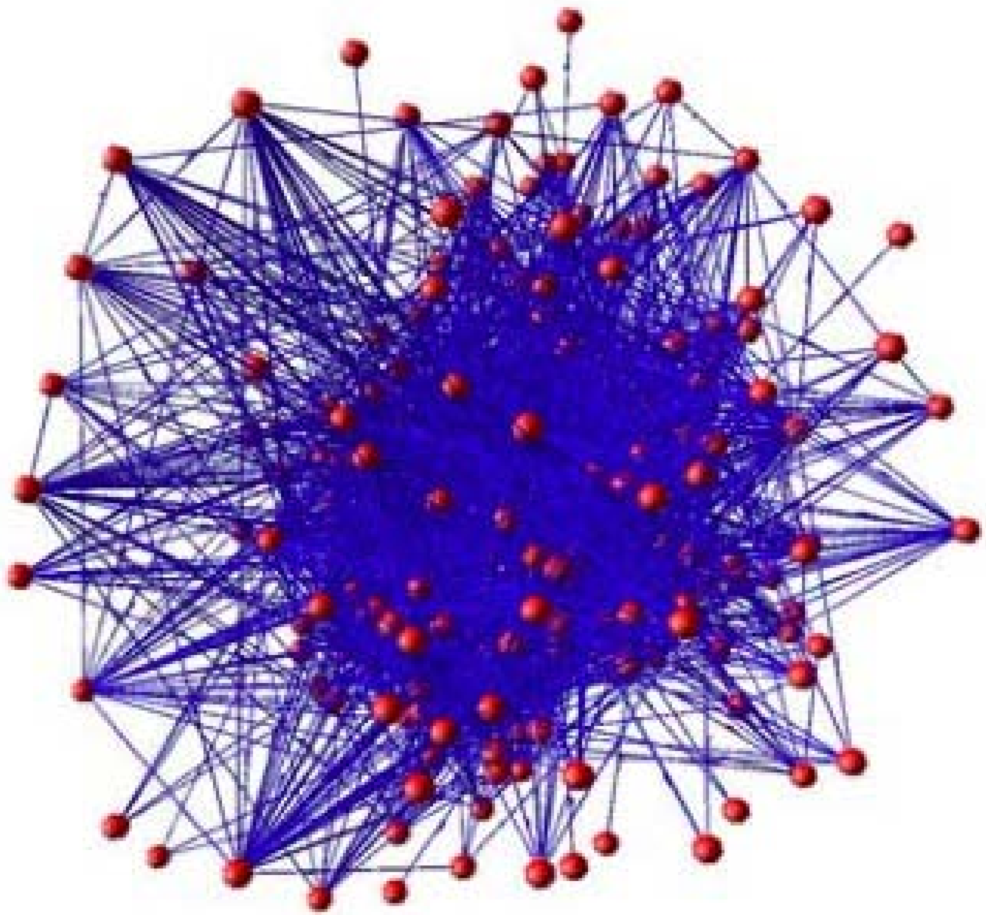
- vertices represent actors: people, places, companies
- edges represent ties or relationships
- Applications (cohesive subgroups)
  - Criminal network analysis
  - Data mining
  - Wireless Networks
  - Genes Therapy
  - Biological Neural Networks

Gene Co-expression Networks

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

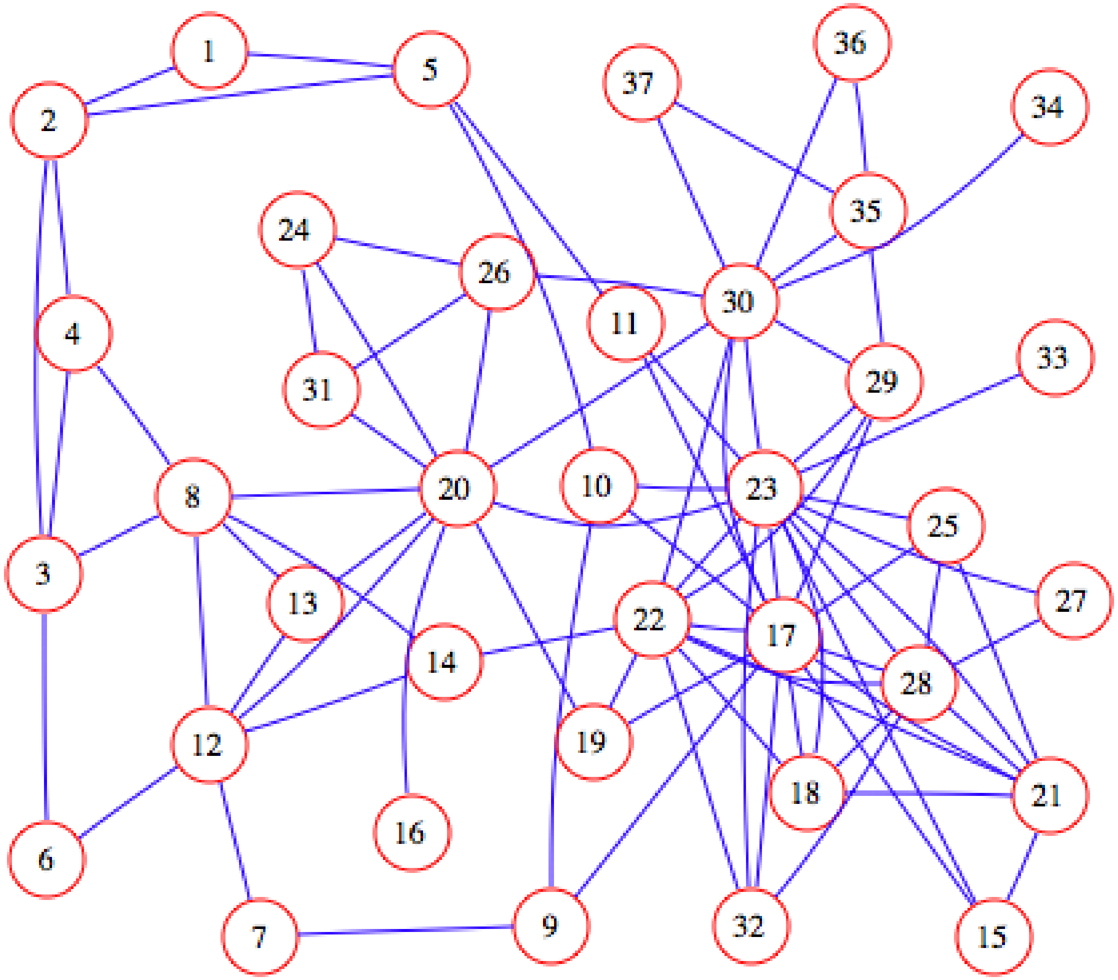
vertices represent genes
edges represent high correlation between genes
(Carlson et al. 2006)

Biological Neural Networks



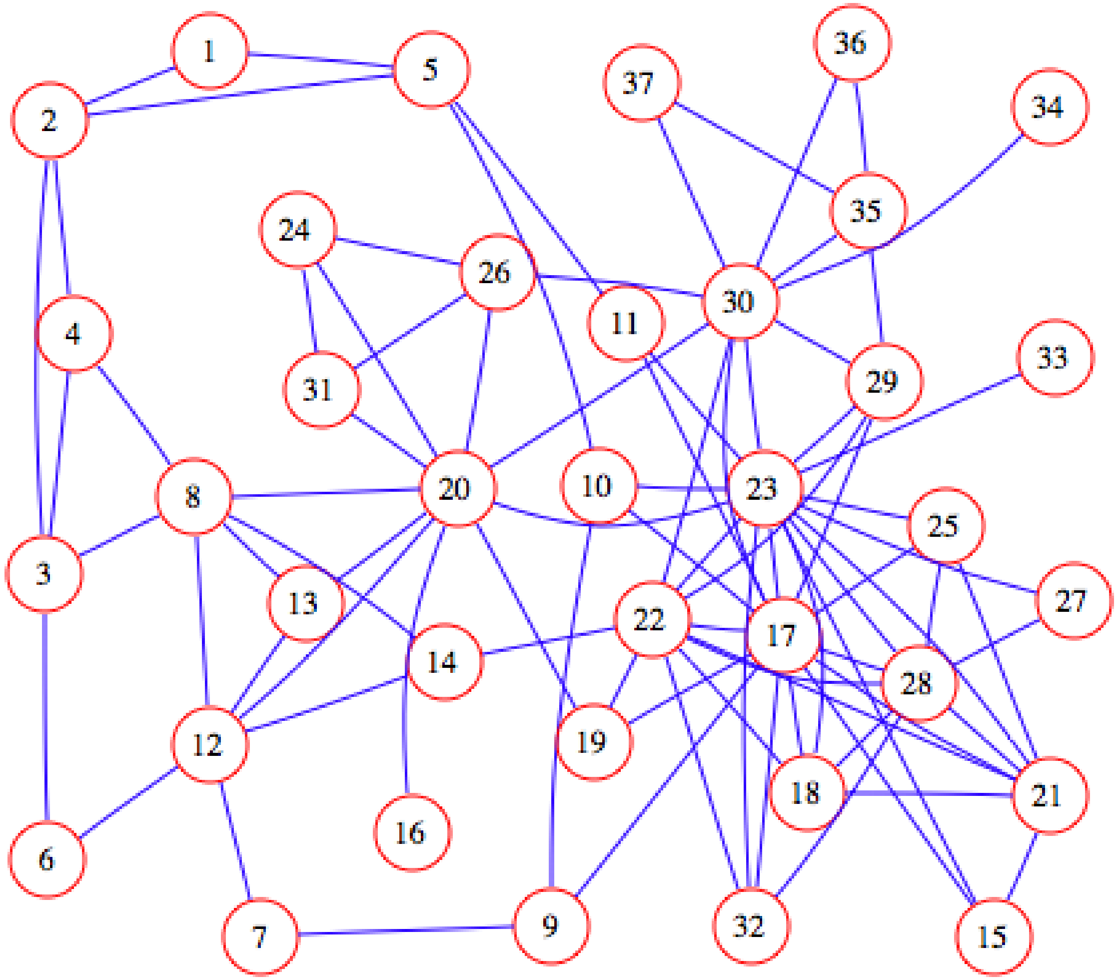
vertices represent neurons
(Berry and Temman 2005)

Social Network Pop Quiz



9-11Terrorist Network

1) Alshehri
2) Sugami
3) Al-Marabh
4) Hijazi
5) W. Alshehri
6) A. Alghamdi
7) M. Alshehri
8) S. Alghamdi
9) Ahmed
10) Al-Hisawi
11) Al-Omari
12) H. Alghamdi
13) Alnami
14) Al-Haznawi
15) Darkazanli
16) Abdi
17) Al-Shehhi
18) Essabar
19) S. Alhazmi



20) N. Alhazmi
21) Bahaji
22) Jarrah
23) Atta
24) Shaikh
25) El Motassadeq
26) Al-Mihdhar
27) Moussaoui
28) Al-Shibh
29) Raissi
30) Hanjour
31) Awadallah
32) Budiman
33) Al-ani
34) Moqed
35) Abdullah
36) Al Salmi
37) Alhazmi

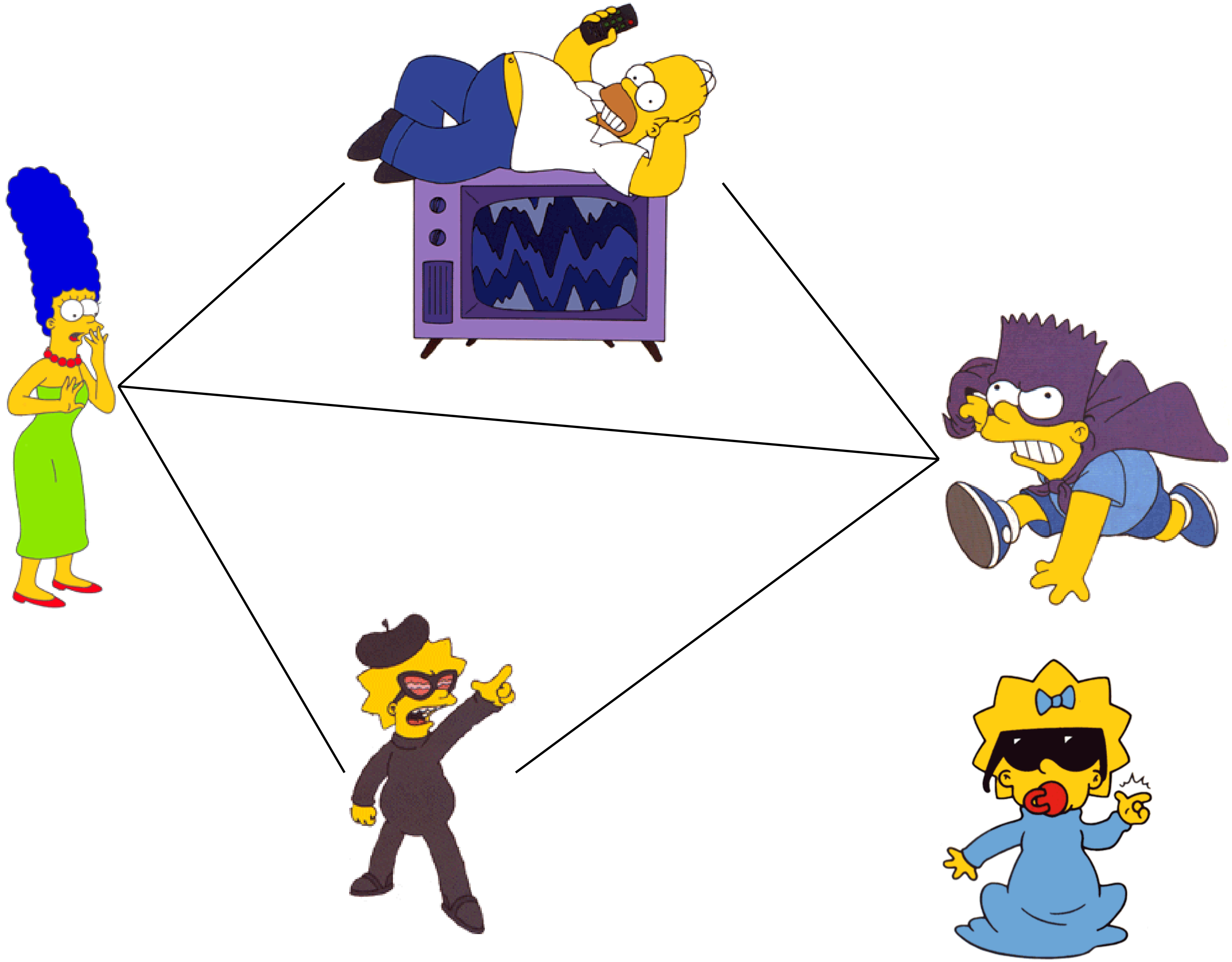
Another Example: The Simpsons

- Homer: Marge? Since I am not talking to Lisa, would you please ask her to pass me the syrup?
- Marge: Dear, please pass your father the syrup, Lisa.
- Lisa: Bart, tell Dad I will only pass the syrup if it won't be used on any meat products.
- Bart: You dunkin' your sausage in that syrup homeboy?
- Homer: Marge, tell Bart I just want to drink a nice glass of syrup like I do every morning.
- Marge: Tell him yourself, you're ignoring Lisa, not Bart.

Another Example: The Simpsons

- Homer: Bart, thank your mother for pointing that out.
- Marge: Homer, your not not-talking to me and secondly, I heard what you said.
- Home: Lisa, tell your mother to get off my case.
- Bart: Uh, Dad, Lisa’s the one you’re not talking to.
- Homer: Bart, go to your room!

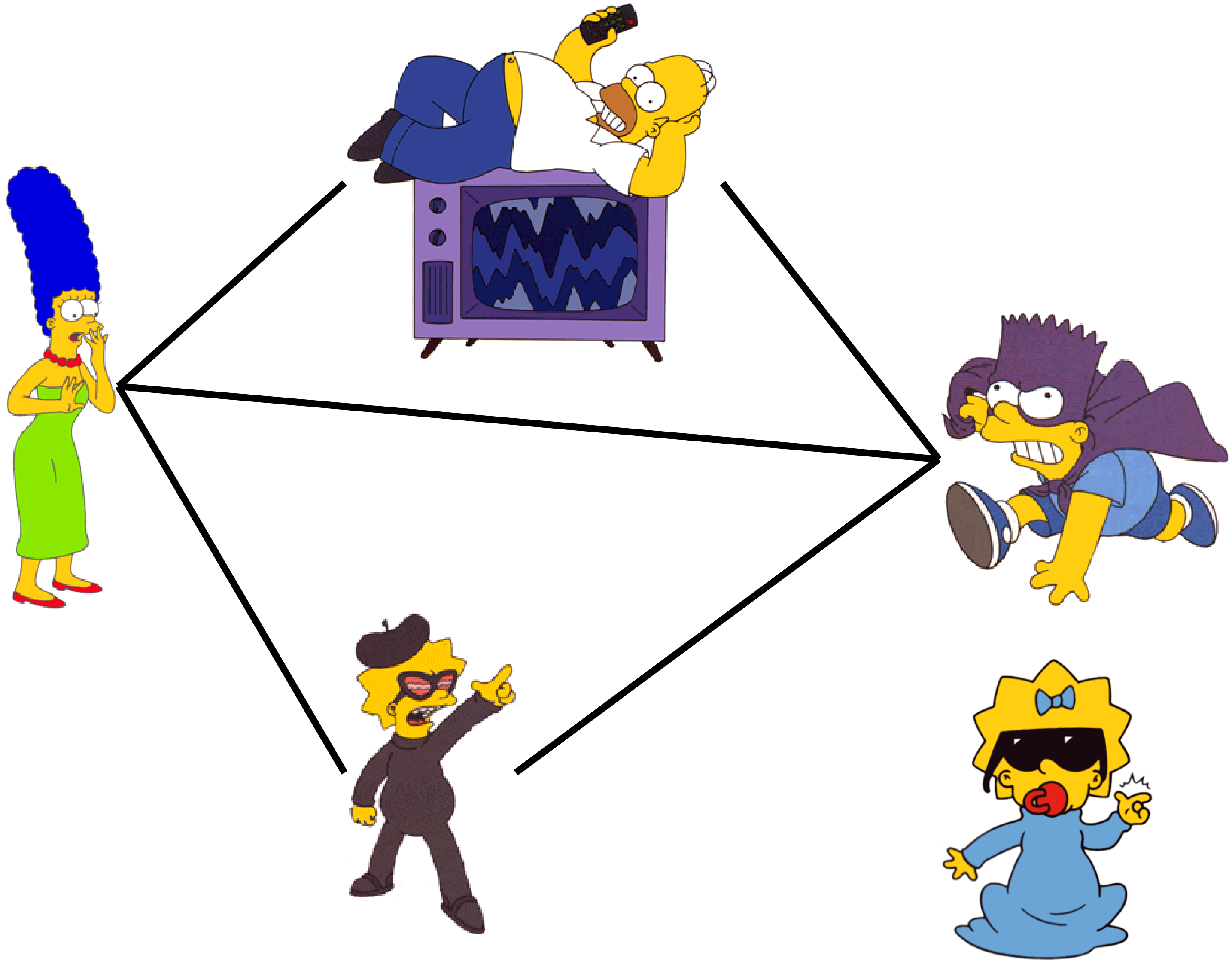
The Simpsons Social Network



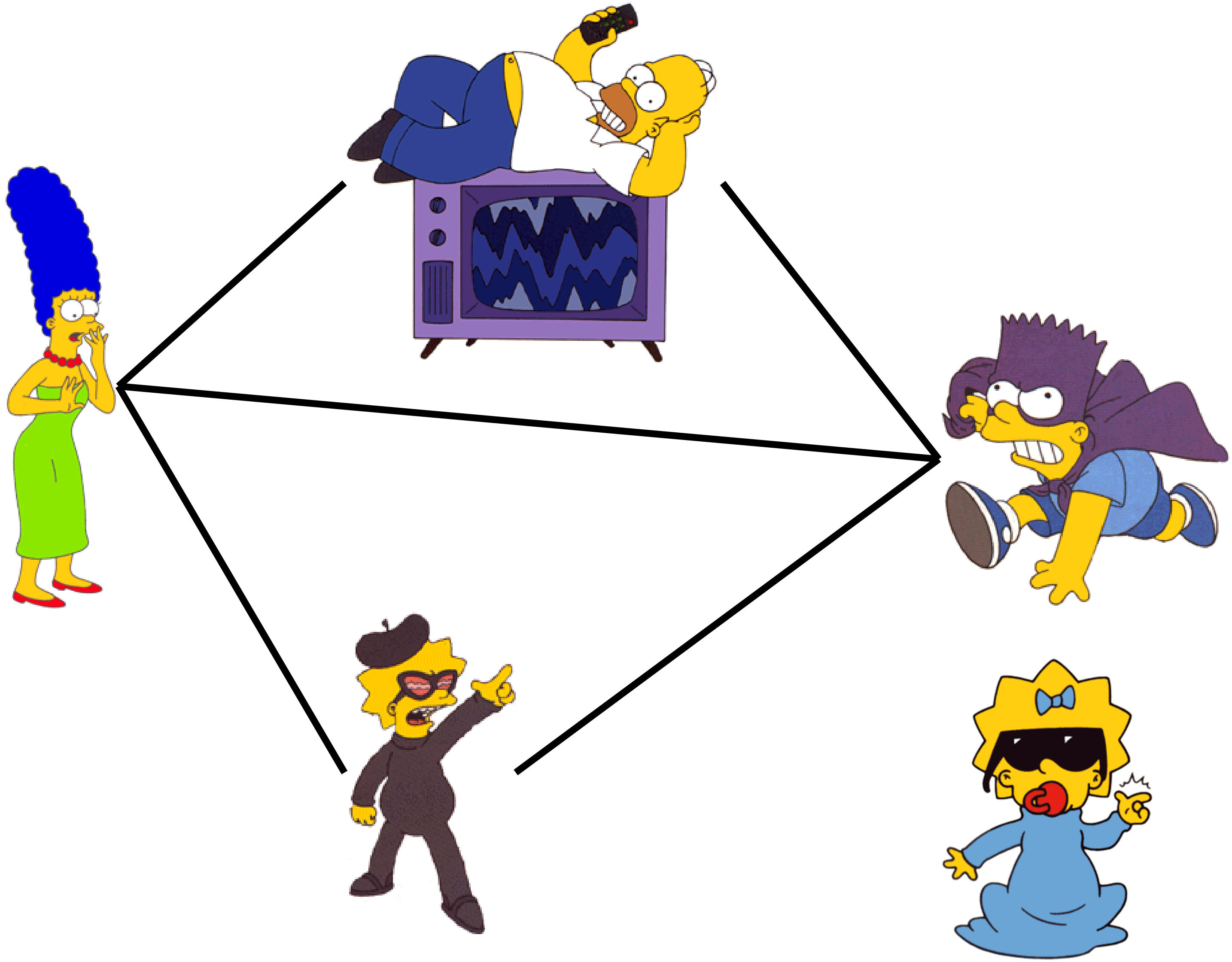
k-plexes

- Given a graph G=(V, E), a set $S \subseteq V$ is called a k-plex if every node of S has at most k-1 non-neighbors in S
- A set $S \subseteq V$ is called a co-k-plex if every node of S has at most k-1 neighbors in S
- Cliques are 1-plexes
- NP-hard to find maximum k-plex, $\omega_k(G)$, in a graph G

1-plexes

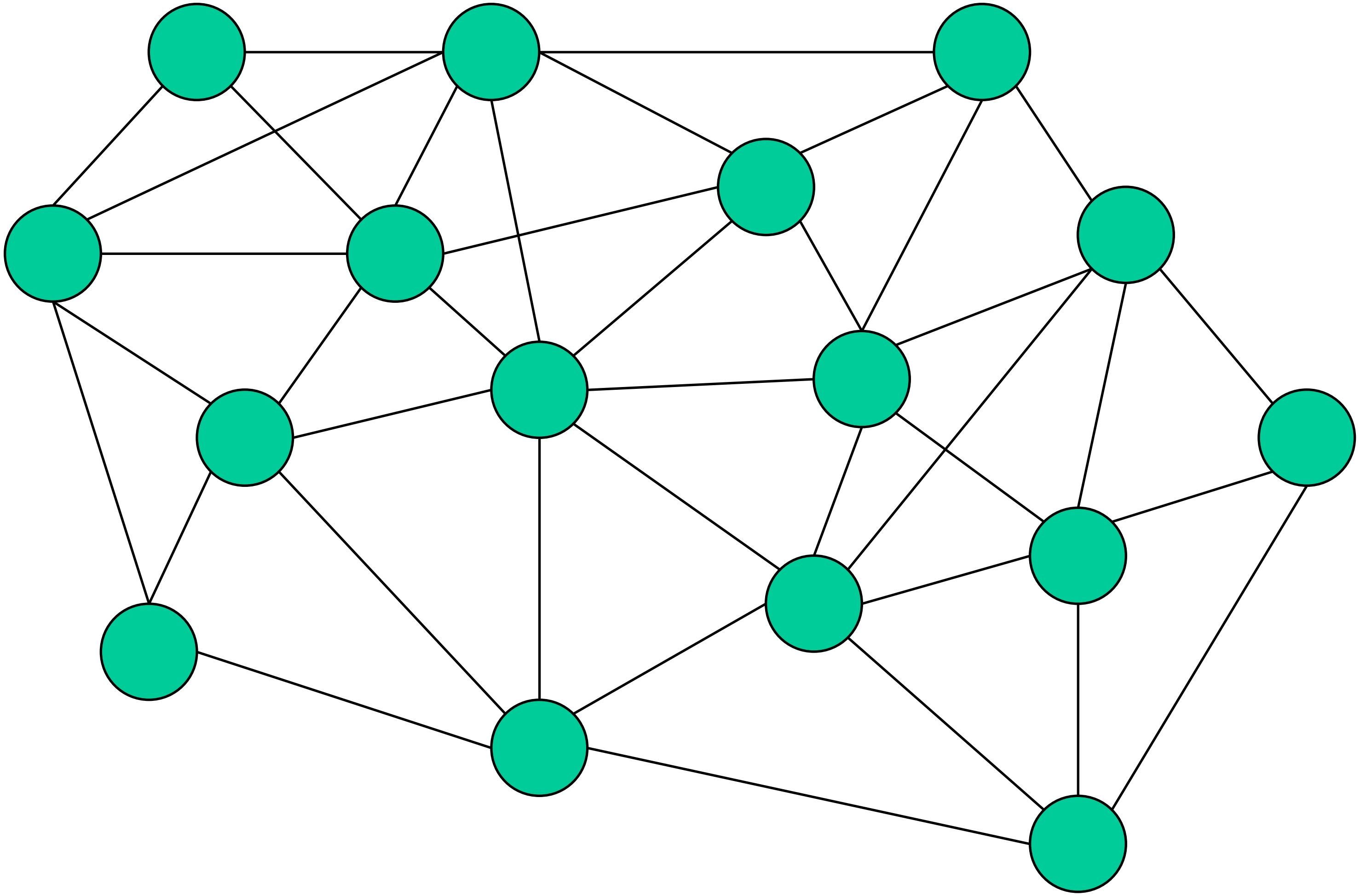


2-plexes

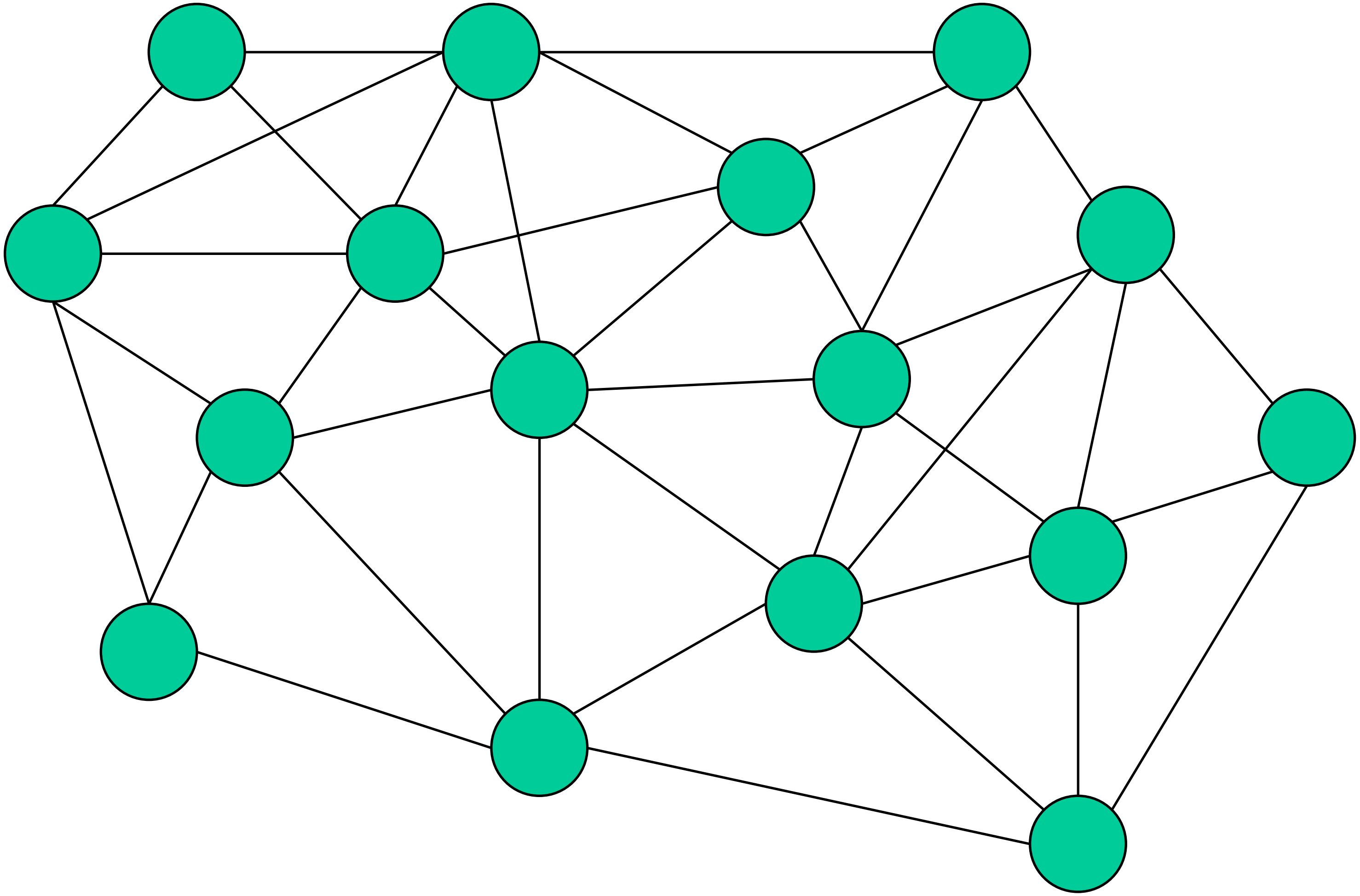


Ready for Co-*k*-plexes!!!

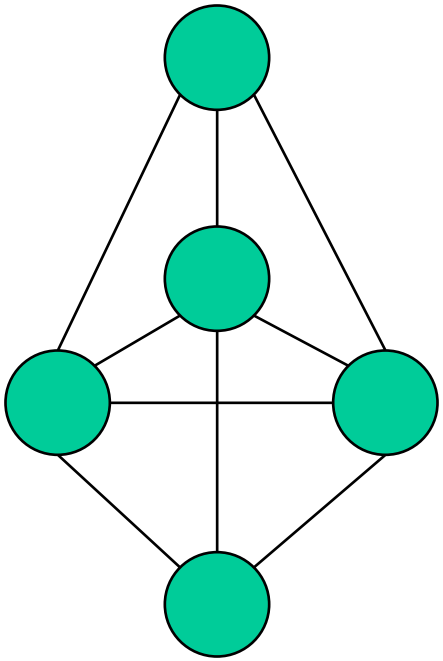
Another Example: Retail Location



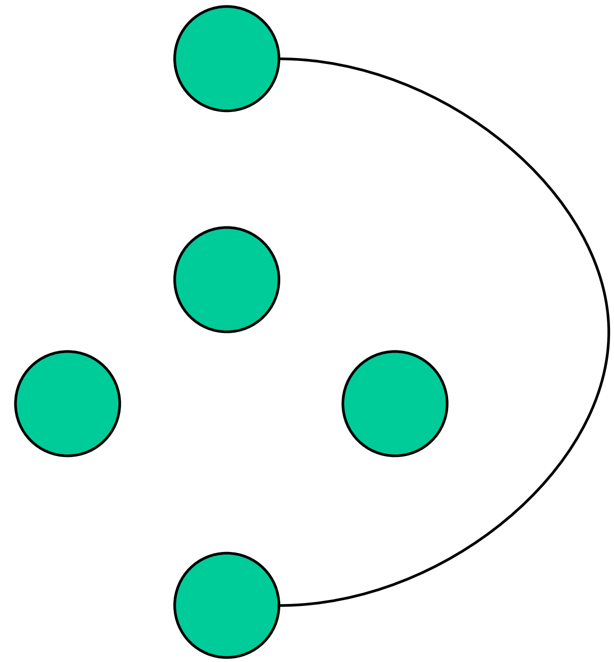
Another Example: Retail Location



k-plexes and co-*k*-plexes

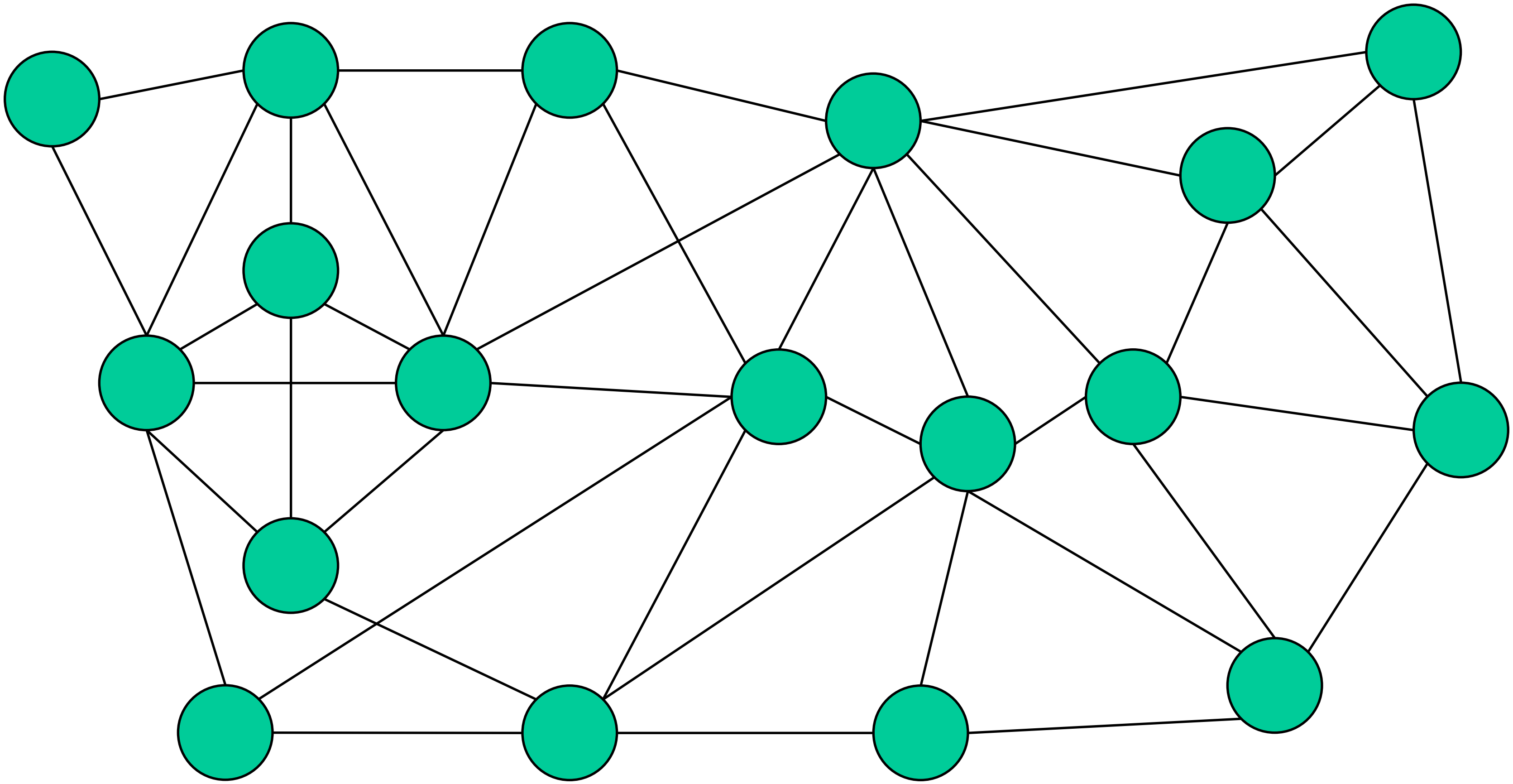


G



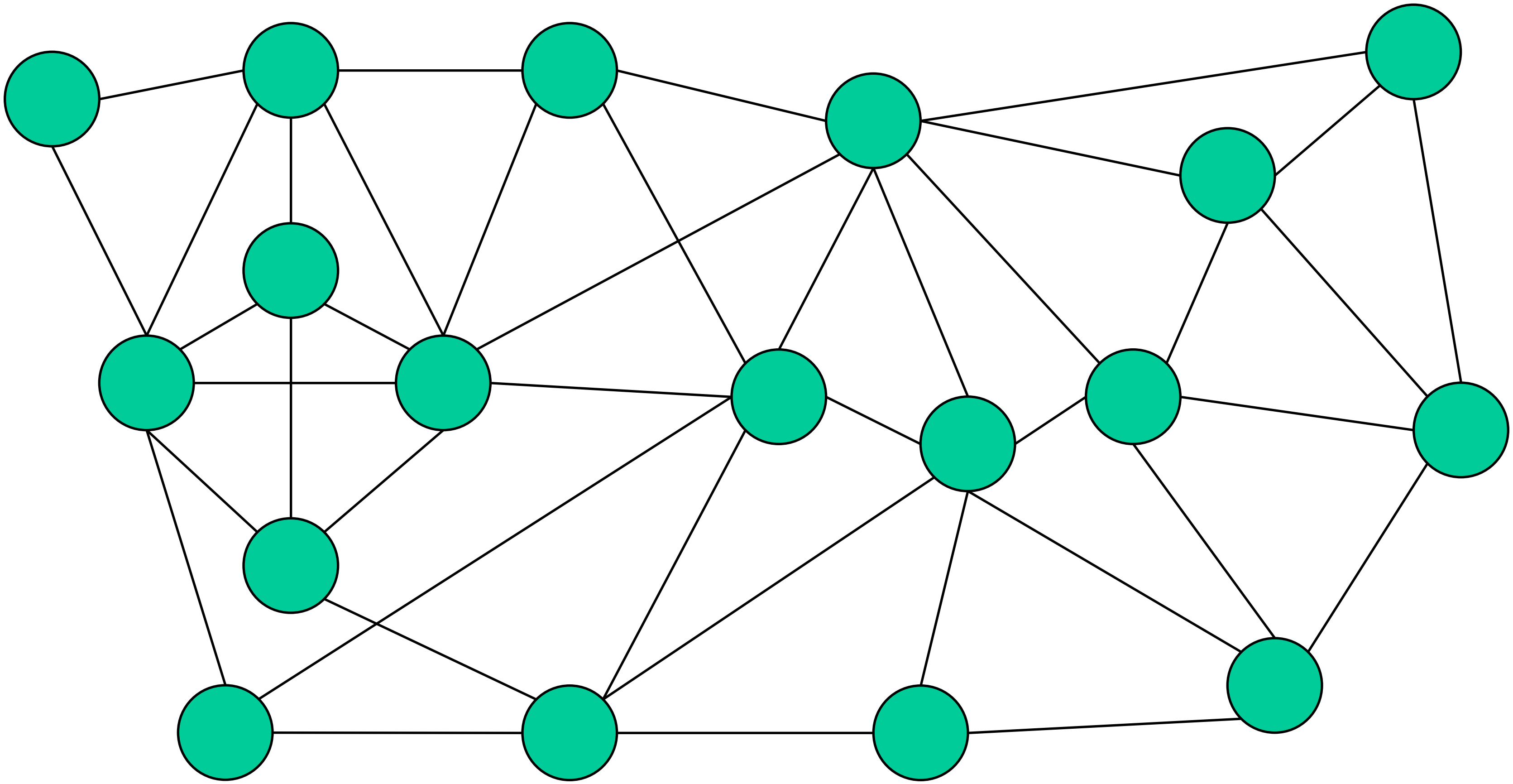
G^C

Graph Coloring



$$\omega(G) \leq \chi(G)$$

Co-*k*-plex Coloring



$$\omega_k(G) \leq \chi_k(G)$$

Polyhedral Approach

- Let N[v] denote the closed neighborhood of vertex v
- Let d(v) denote $|V \setminus \mathrm{N}[v]|$

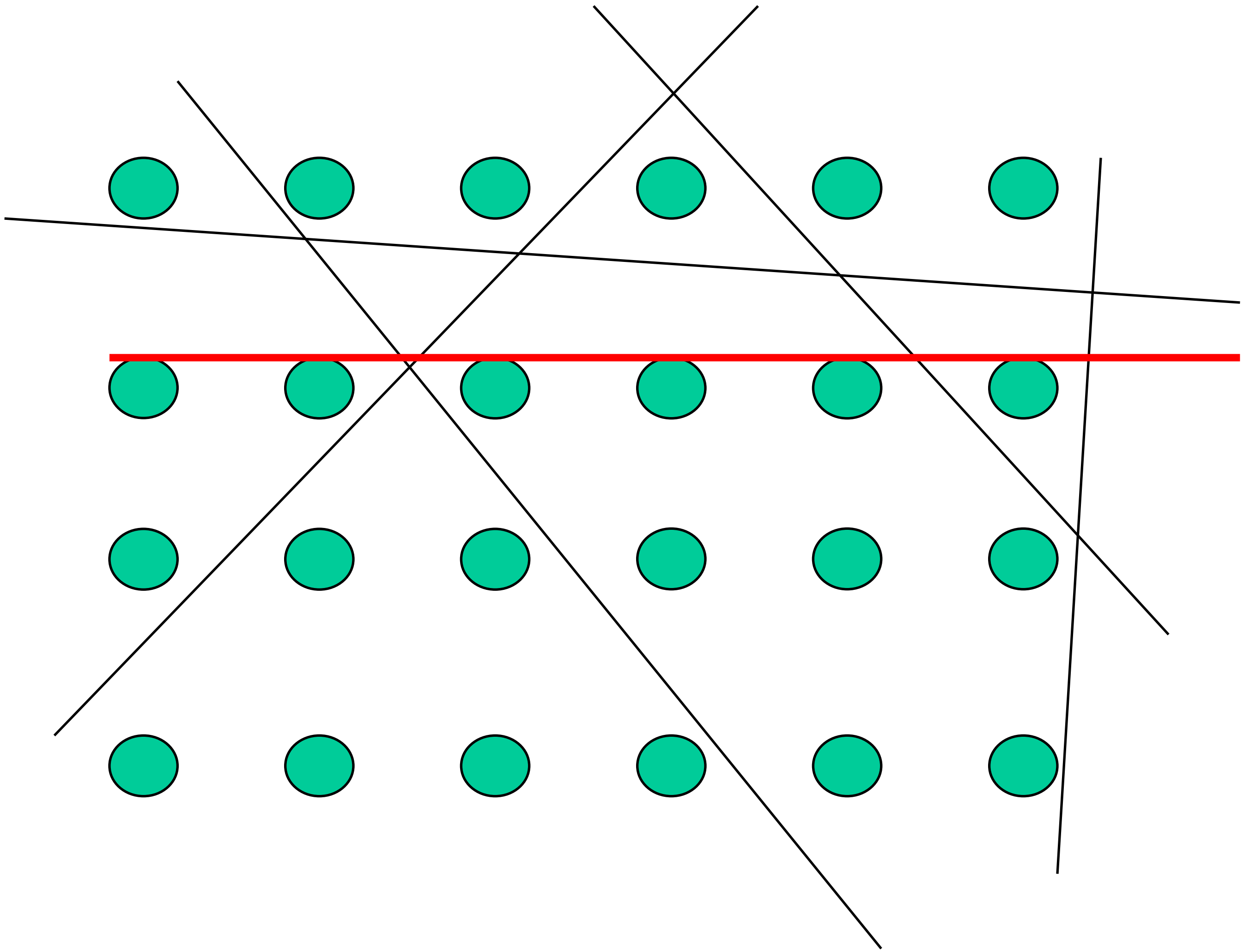
$$\text{Max} \sum_{v \in V} x_v$$

$$st.$$

$$\sum_{u \in V \setminus \mathrm{N}[v]} x_u \leq (k - 1)x_v + \mathrm{d}(v)(1 - x_v) \;\; \forall v \in V$$

$$x_v \in \{0, 1\} \;\; \forall v \in V$$

Polyhedral Approach



Wrap-Up

- Social Networks
- *k*-plexes & co-*k*-plexes
- Co-*k*-plex coloring

Acknowledgments

- My collaborator: Ben McClosky, Ph. D.
- NSF
  - DMI 0521209
  - DMS 0611723
  - CMMI 0926618

Any Questions?

Relevant Literature

- Seidman & Foster (1978)
  - Introduced *k*-plexes in context of social network analysis
- Balasundaram, Butenko, Hicks, and Sachdeva (2006)
  - IP formulation for maximum *k*-plex problem
  - NP-complete complexity result
- McClosky & Hicks (2007)
  - Co-2-plex polytope
- McClosky & Hicks (2008)
  - Graph algorithm to compute k-plexes

Co-k-plexes

- Given a graph $G=(V, E)$, a set $S \subseteq V$ is called a co-k-plex if $\Delta(G[S]) \leq k - 1$, where Δ denotes maximum degree
- Stable sets are co-1-plexes and co-k-plexes form independence systems
- NP-hard to find maximum co-k-plex, $\alpha_k(G)$ in a graph G
- Co-2-plexes correspond to vertex induced subgraphs of isolated nodes and matched pairs

Co-k-plex Polytope

- Given graph G, let $\mathscr{I}^k$ be the set of co-k-plexes in G
- For all $S \in \mathscr{I}^k$, let x^S be the incidence vector for S.
- Define $P_k(G) = \text{conv}(\{x^S : S \in \mathscr{I}^k\})$
- $P_2(G)$ shares many properties with $P_1(G)$

Co-2-plex analogs

- Padberg (1973)
  - Clique and odd hole inequalities
- Trotter (1975)
  - Web inequalities
- Minty (1980)
  - claw-free graphs

2-plex Inequalities

- Theorem (Padberg): If K is a maximal clique in G, then $\sum_{v \in K} x_v \leq 1$ is a facet for P_1(G).

- Theorem (M & H, B et al.): If K is a maximal 2-plex in G such that $|K| > 2$, then $\sum_{v \in K} x_v \leq 2$ is a facet for P_2(G)

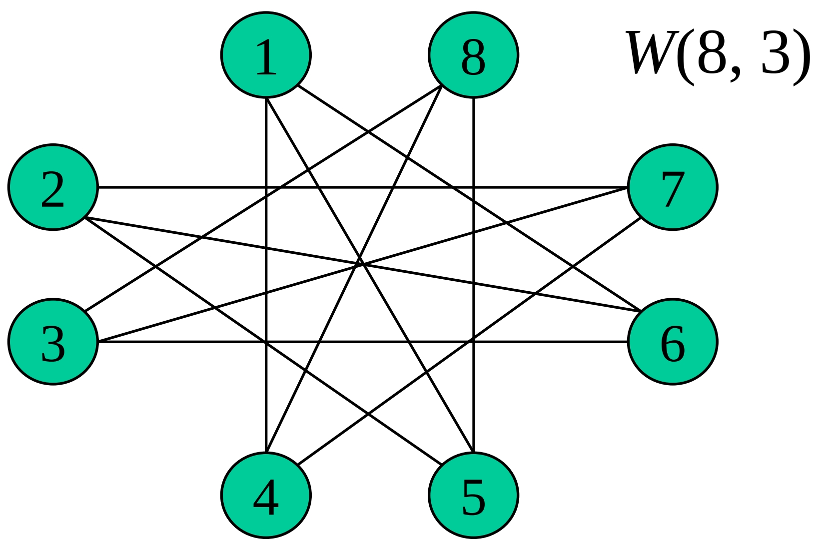
Odd-mod Hole Inequalities

- Theorem (Padberg): If C is an n-chordless cycle such that $n > 3$ is odd, then $\sum_{v \in V(C)} x_v \leq \lfloor n/2 \rfloor$ is a facet for P_1(C).

- Theorem (M & H): If C is an n-chordless cycle such that $n > 2$ and $n \neq 0 \bmod 3$, then $\sum_{v \in V(C)} x_v \leq \lfloor 2n/3 \rfloor$ is a facet for P_2(C)

Webs

- For fixed integers $n \geq 1$and p such that $1\leq p \leq \lfloor n/2 \rfloor$, the web $W(n, p)$ has n vertices and edges $E=\{(i, j)\text{: } j=i+p, \ldots, i+n -p;\ \forall \text{ vertices } i\}$



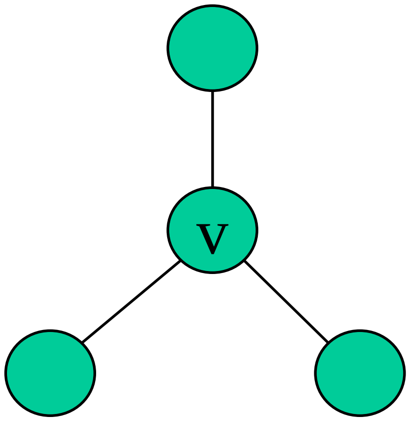
Web Inequalities

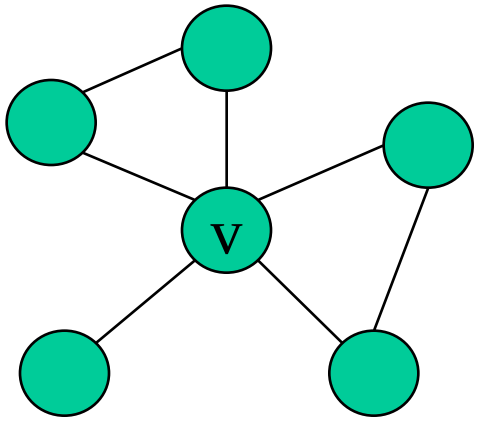
- Theorem (Trotter): If gcd(n, p) = 1, then $\sum_{v \in V(W(n,p))} x_v \leq p$ is a facet for $P_1(W(n, p))$.

- Theorem (M & H): If gcd(n, p + 1) = 1, then $\sum_{v \in V(W(n,p))} x_v \leq p + 1$ is a facet for $P_2(W(n, p))$.

k-claws

- Given an integer $k \geq 1$, the graph G is a k-claw if there exists a vertex v of G such that $V(G)$=N[v], N(v) is a co-k-plex, and $|N(v)| \geq \max\{3, k\}$





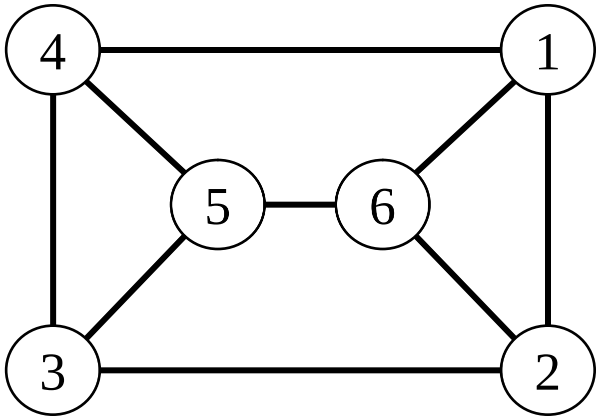
2-claw free graphs

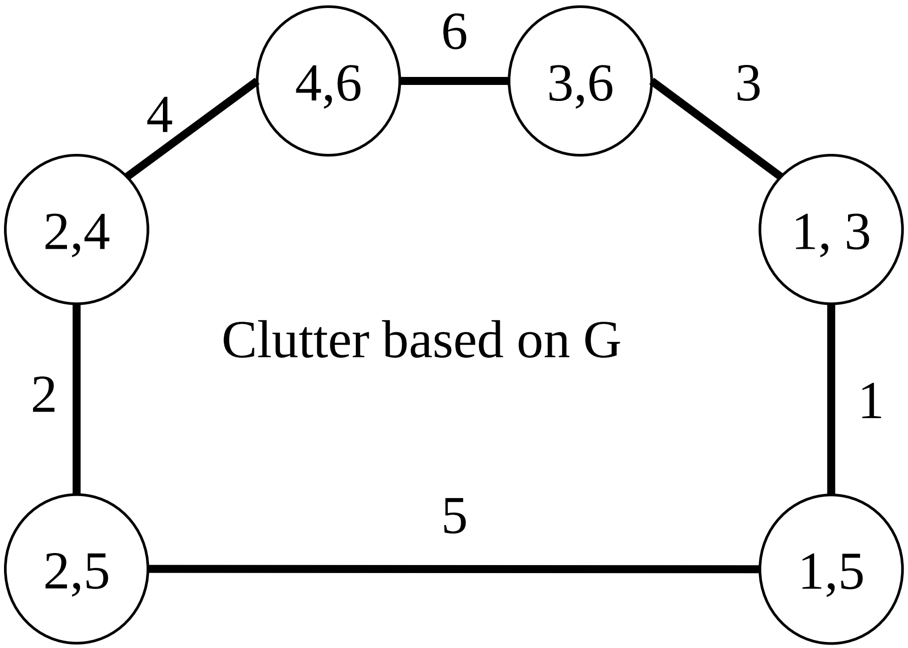
- Theorem (B & H): A graph G is 2-claw free if and only if $\Delta(G) \leq 2$ or G is 2-plex.
- This theorem will be used to describe a class of 0-1matrices A for which the polytope $P=\{x \in R^n_+ : Ax \leq 2, x \leq 1\}$ is integral.

Clutters

- A clutter is a pair (V, E) where V is a finite set and E is a family of subsets of V none of which is included in another.

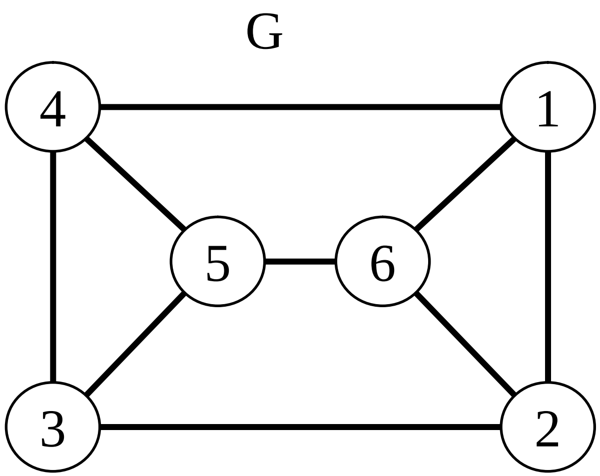
G

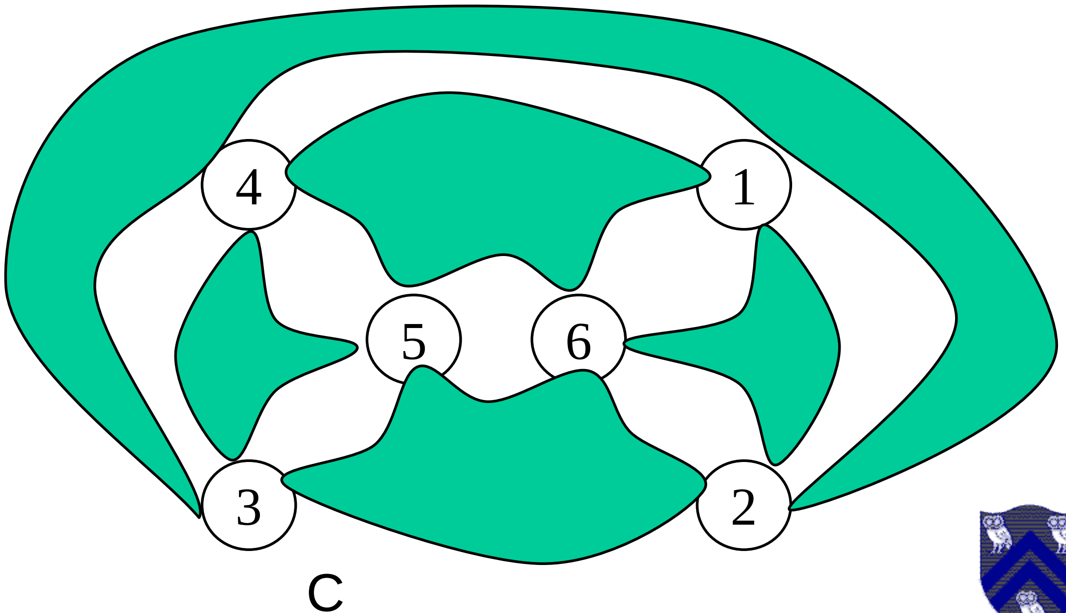




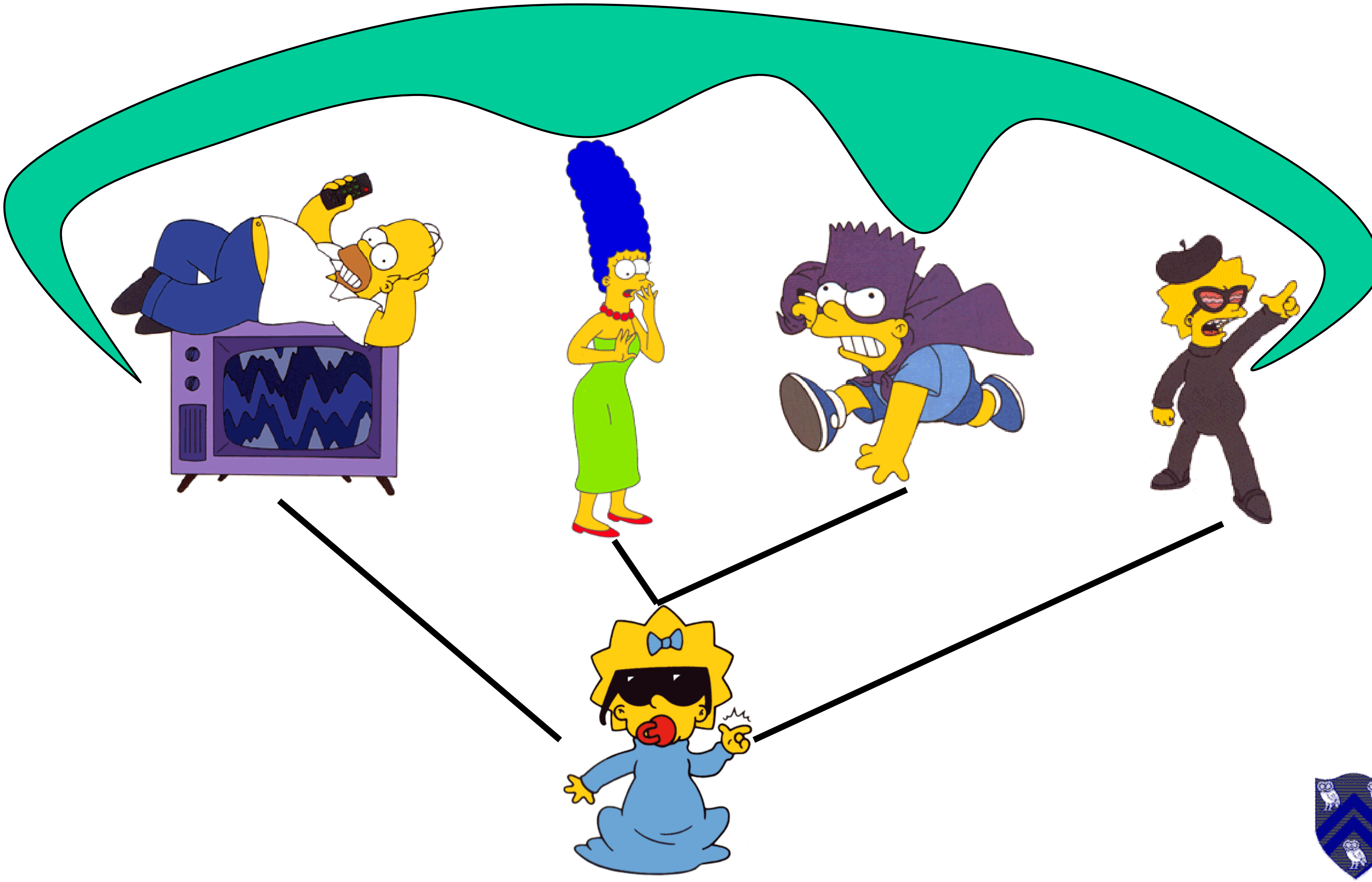
Clutters of Maximal 2-plexes

- Given a graph *G*, let **C** be the clutter whose vertices are V(*G*) and whose edges are maximal 2-plexes of *G*.





Clutters of Maximal 2-plexes



2-plex Clutter Matrices

Let A be the edge-vertex incidence matrix of ***C***.

- Theorem (M & H): Let A be the 2-plex clutter matrix of G. The polytope $P=\{x \in R^n_+ : Ax \leq 2, x \leq 1\}$ is integral if and only if the components of G are 2-plexes, co-2-plexes, paths, or 0 mod 3 chordless cycles.

- Corrollary (M & H): Given a 2-plex clutter matrix A, there is a polynomial-time algorithm to determine if $P=\{x \in R^n_+ : Ax \leq 2, x \leq 1\}$ is integral.

Future Work

- Combinatorial algorithm to compute maximum k-plexes (involves *k*-plex coloring)
- Find facets of P_k(G) for $k > 2$.
- Can k-plex clutter matrices give insight in polyhedra defined as $P=\{x \in R^n_+ : Ax \leq k,\ x \leq 1\}$?

Other inequalities

- Stable Sets
  - $\sum_{v \in I} x_v \leq k \quad \forall$ stable sets I s.t. $|I| \geq k+1$
- Holes
  - $\sum_{v \in H} x_v \leq k + 1 \quad \forall$ holes H s.t. $|H| \geq k+3$
- Co-k-plexes
  - $\sum_{v \in S} x_v \leq \omega_k(S) \quad \forall$ co-k-plexes S

2-plex Computational Results

G	n	m	density	ω(G)	BIS	UB	Time (sec)
c.200.1	200	1534	.077	12	12	12	57.3
c.200.2	200	3235	.163	24	24	24	46.9
c.200.5	200	8473	.426	58	58	58	40.2
h.6.2	64	1824	.905	32	32	32	.47
h.6.4	64	704	.349	4	6	6	4.4
h.8.2	256	31616	.969	128	128	130	>86000
h.8.4	256	20864	.639	16	16	46	>86000
j.8.2.4	28	210	.556	4	5	5	3.6
j.8.4.4	70	1855	.768	14	14	14	7424
j.16.2.4	120	5460	.765	8	10	14	>86000
k.4	171	9435	.649	11	15	26	>86000
m.a9	45	918	.927	16	26	26	2.3

Pop Quiz: Question #1

Who was the first African-American to receive a PhD in Mathematics?

Elbert F. Cox

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Dissertation: Polynomial Solutions of Difference Equations
Ph.D. Cornell University, 1925
Advisor: William Lloyd Garrison

Pop Quiz: Question #2

Who was the first African-American to receive a PhD in Mathematics at Rice University?

Raymond Johnson

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Dissertation: A Priori Estimates and Unique Continuation
Theorems for Second Order Parabolic Equations
Ph.D. Rice University, 1969
Advisor: Jim Douglass Jr.