An Open Database of Interesting Graphs: the House of Graphs

CSD 6

Portorož, 21 – 25 may 2012

Gunnar Brinkmann[1], Kris Coolsaet[1], Jan Goedgebeur[1] and Hadrien Mélot[2]

[1] Vakgroep Toegepaste Wiskunde en Informatica, UGent, Belgium

[2] Service d'Algorithmique, UMONS, Belgium

Remarks

- What this talk does not contain : new theoretical results
- But... hopefully it can help you to get new results

House of Graphs

- open database of special graphs
- needed and discussed since the first editions of CSD (name "House of Graphs" used by Hansen in CSD2)
- prototype first presented at CSD5 (Sheffield, 2010)
- since, it has been redesigned and... is open for your submissions !
- presentation paper in press (available online 16 Aug 2012, *Discrete Appl. Math.*)

Objectives

The House of Graphs (HoG):

- complete lists and generators of some graph classes
- searchable database of special graphs :
  - interesting and relevant in the study of graph theoretic problems
  - counterexamples to conjectures

Facts

- Number of graphs grows very fast

OEIS A001349 (Number of connected graphs with n vertices)

n		n		n	
2	1	8	11117	14	29003487462848061
3	2	9	261080	15	31397381142761241960
4	6	10	11716571	16	63969560113225176176277
5	21	11	1006700565	17	245871831682084026519528568
6	112	12	164059830476	18	1787331725248899088890200576580
7	853	13	5033590786921	19	24636021429399867655322650759681644

Facts

- Number of graphs grows very fast
- It grows exponentially fast even if one restricts class of graphs (regular graphs, cubic graphs, trees, etc.)

Facts

- Number of graphs grows very fast
- It grows exponentially fast even if one restricts class of graphs (regular graphs, cubic graphs, trees, etc.)
- Lists of graphs are useful for tests and scientific discovery (conjectures, proofs, tight examples, etc.)

Facts

- Number of graphs grows very fast
- It grows exponentially fast even if one restricts class of graphs (regular graphs, cubic graphs, trees, etc.)
- Lists of graphs are useful for tests and scientific discovery (conjectures, proofs, tight examples, etc.)
- Some graphs (e.g., the Petersen graph) or graph classes (e.g., snarks) appear repeatedly in the literature

Facts

- Number of graphs grows very fast
- It grows exponentially fast even if one restricts class of graphs (regular graphs, cubic graphs, trees, etc.)
- Lists of graphs are useful for tests and scientific discovery (conjectures, proofs, tight examples, etc.)
- Some graphs (e.g., the Petersen graph) or graph classes (e.g., snarks) appear repeatedly in the literature
- Others are probably just part of the huge mass

If Orwell had been a mathematician he could have said. . .

> *"All graphs are interesting, but some graphs are more interesting than others."*

Our goals

- Offer a searchable database
  - $\Rightarrow$ relatively small size
- that still gives you a good chance to be useful
  - to find counterexamples and/or
  - obtain results that allow generalization
- gather (links to) already existing lists of graphs

What make a graph relevant of interesting?

We will not try to give an exact definition or a definitive answer !

- Depends on the problem one wants to study
- Which graph could be eligible?
  - appears in the literature;
  - contained in some lists (e.g., websites);
  - is pointed out by a conjecture-making system;
  - was used by you and other graph theorists during any step of the discovery process.

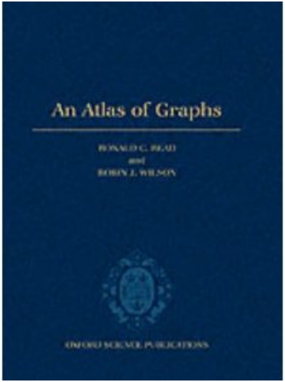
What make a graph relevant of interesting?

We will not try to give an exact definition or a definitive answer !

- Depends on the problem one wants to study
- Which graph could be eligible?
  - appears in the literature;
  - contained in some lists (e.g., websites);
  - is pointed out by a conjecture-making system;
  - was used by you and other graph theorists during any step of the discovery process.
- House of Graphs allows users to add new graphs

Graphs in the literature (papers or books)

Counterexamples; tight graphs; extremal graphs; classes of graphs; lists of graphs; etc.



Examples of books:

- Brandstädt, Le and Spinrad, Graph classes: a survey (1999)
- Capobianco and Molluzzo, Examples and Counterexamples in Graph Theory (1978)
- Read and Wilson, An atlas of graphs (1999)

$\Rightarrow$ unusable for coding tests

Graphs in websites (static lists)

Connected regular graphs

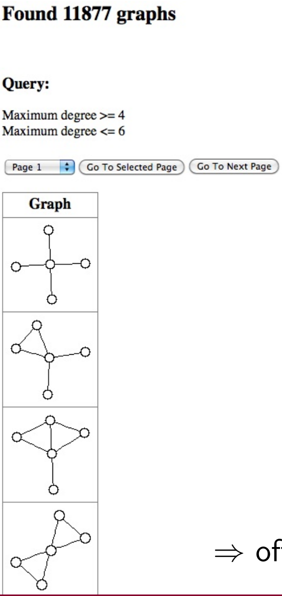
The following table contains numbers of connected regular graphs with given number of vertices a degree. For the empty fields the number is not yet known (to me). The latest numbers (for *n*=19,*k*= *n*=16,*k*=6; *n*=16,*k*=7) have been contributed by Jason Kimberley (University of Newcastle, Austra 2009), who ran GENREG on up to 250 cores.

Vertices	Degree 3	Degree 4	Degree 5	Degree 6	Degree 7
4	1	0	0	0	0
5	0	1	0	0	0
6	2	1	1	0	0
7	0	2	0	1	0
8	5	6	3	1	1
9	0	16	0	4	0
10	19	59	60	21	5
11	0	265	0	266	0
12	85	1544	7848	7849	1547
13	0	10778	0	367860	0
14	509	88168	3459383	21609300	21609301
15	0	805491	0	1470293675	0
16	4060	8037418	2585136675	113314233808	733351105934
17	0	86221634	0		0
18	41301	985870522			
19	0	11946487647			
20	510489				
22	7319447				
24	117940535				
26	2094480864				

- Brendan McKay
- Markus Meringer
- Gordon Royle
- Ted Spence
- ...

$\Rightarrow$ static and non searchable

Graphs in websites (databases)



- AutoGraphiX, `www.gerad.ca/~agx` : DB of AGX conjectures + extremal graphs
- Jason Grout, the "small graph database" (graphs up to 8 vertices)
- ChemSpider, specialized in chemical graphs
- GReGAS Atlas
- etc.

⇒ often very specialized ...

Conjecture-making systems

Input: a particular problem (*e.g.*, conjecture, set of invariants, etc.)
Output: graphs relevant for this problem

Examples:

- AutoGraphiX 1 & 2 [Caporossi, Hansen]: extremal graphs;
- Graffiti [Fajtlowicz] & Graffiti.PC [DeLa Viña]: counterexamples;
- GraPHedron [Mélot]: vertex-graphs (= "conglomerates", see later);
- newGraph [Stevanović et al.]: small predefined set of graphs;
- GrInvIn [Brinkmann et al.]: counterexamples.

Initial list of special graphs

House of Graphs = meta-directory (lists & generators) + list of special graphs

- Meta-directory contains $> 10^8$ graphs

Initial list of special graphs

- Extremal graphs found by GraPHedron (explained after)
- Named graphs (from MathWorld and GraphData in Mathematica)
- Snarks which occurred as couterexamples
- Ramsey graphs
- etc.

$\Rightarrow$ Currently contains > 1300 graphs

GraPHedron

Conjecture-making system GraPHedron [Mélot, Discrete Appl. Math. (2008)]

- Computer assisted and automated conjectures
- Use a polyhedral approach
- led to several theoretical results (connected variant of the famous Turán theorem, tight bounds for the Merrifield-Simons index, etc.)
- Used here to feed HoG by running many problems

GraPHedron's type of problems

Problem

What are all the best linear inequalities among a set of invariants, valid for all graphs of order n that belong to a given class of graphs?

Example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

Polyhedral approach – an example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

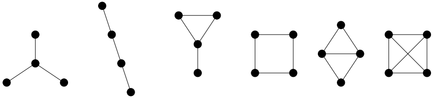
For a given value of the order n :   Example: $n = 4$

Polyhedral approach – an example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

For a given value of the order n : Example: $n = 4$

1. Generate graphs

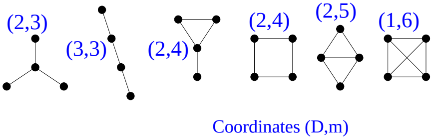


Polyhedral approach – an example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

For a given value of the order n : Example: $n = 4$

1. Generate graphs
2. Compute invariants



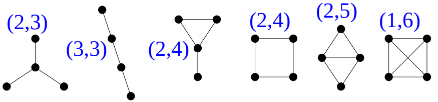
Polyhedral approach – an example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

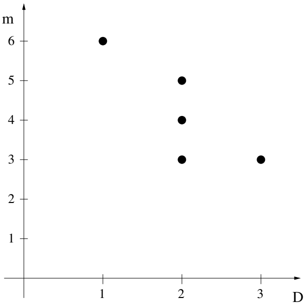
For a given value of the order n :

1. Generate graphs
2. Compute invariants
3. Consider graphs as points in the space

Example: $n = 4$



Coordinates (D,m)



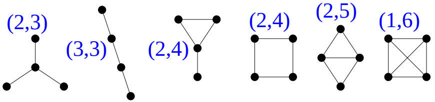
Polyhedral approach – an example

What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

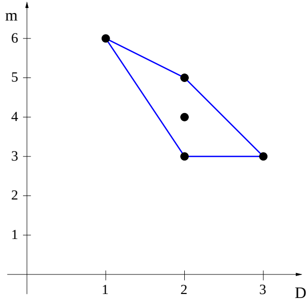
For a given value of the order n :

1. Generate graphs
2. Compute invariants
3. Consider graphs as points in the space
4. Compute the convex hull (polytope $\mathcal{P}$)

Example: $n = 4$



Coordinates (D,m)



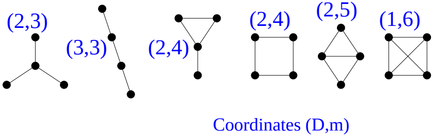
Polyhedral approach – an example

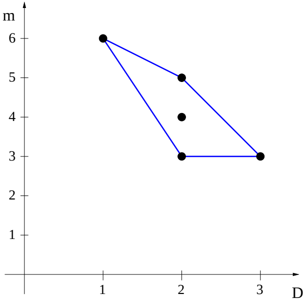
What are all the best linear inequalities among the diameter D and the number of edges m of connected graphs?

For a given value of the order n :

1. Generate graphs
2. Compute invariants
3. Consider graphs as points in the space
4. Compute the convex hull (polytope $\mathcal{P}$)
5. Analyze facets and vertices of $\mathcal{P}$

Example: $n = 4$

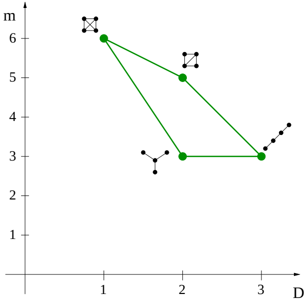




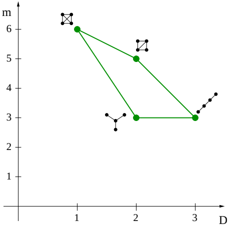
Polyhedral approach

Information outputted :

- facets
- vertices (and corresponding extremal graphs)



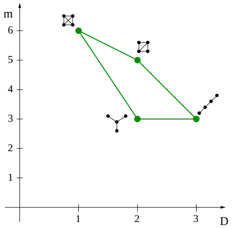
Polyhedral approach



- Facets = "all the best linear inequalities" among the selected set of invariants :
  - cannot be deduced from other valid inequalities
  - constitute a minimal system describing the polytope

$\Longrightarrow$ useful for conjecture-making

Polyhedral approach



- Facets = "all the best linear inequalities" among the selected set of invariants :
  - cannot be deduced from other valid inequalities
  - constitute a minimal system describing the polytope

  $\Longrightarrow$ useful for conjecture-making
- Vertices = extremal graphs (called "vertex-graphs")

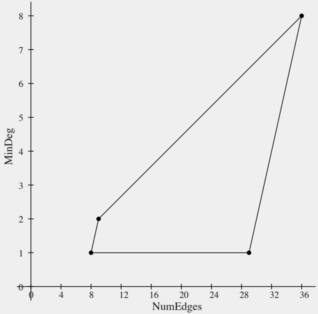
  $\Longrightarrow$ useful for conjecture-making but also in the HoG framework

Conglomerates

Possible problem: there can be a lot of graphs sharing the same pair of coordinates (a "conglomerate").

Example

Consider the minimum degree δ and the number of edges m of connected graphs ($n = 9$).



If T is a tree, then its minimum degree δ is 1 and its number of edges m is $n - 1$.

$\Longrightarrow$ all 47 trees with 9 vertices share coordinates $(8, 1)$

Representation of Conglomerates

Thus, such conglomerates can be problematic (recall that we want a searchable – small – database).

However

- graphs in a conglomerate can be considered as similar for a given problem (they share some properties);
- only one graph of a conglomerate is sufficient to represent its properties;
- for example, using a star (among all trees) is sufficient to represent the fact that some graphs are connected and acyclic.

How the graphs were selected for HoG?

Given a list of invariants $\Rightarrow$ many possible problems ($=$ pairs of invariants)

Let C be a set of conglomerates found for all problems. Then, the desired set of graphs (induced by C) is the minimum set of graphs that cover all conglomerates of C.

- finding the minimum set of graphs is equivalent to the Minimum Set Cover problem (NP-hard)
- no hope to have an (efficient) exact algorithm
- a greedy heuristic is known as the best-possible (polynomial) approximation algorithm for this problem

Computational results

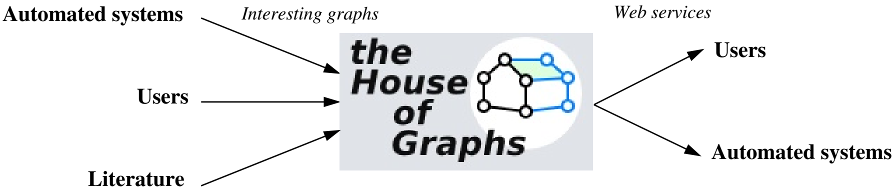
Total: 389 problems (for each value of $n = 4, 5, \ldots, 10$).

n	# gr.	# pol. vert.	# cong.	# int. gr.	pc.
4	11	1402	63	11	100.00%
5	34	1602	126	25	73.53%
6	156	1751	176	46	29.49%
7	1044	1932	236	73	6.99%
8	12346	2039	242	89	0.72%
9	274668	2253	320	127	0.05%
10	12005168	2338	323	168	0.001%

- # gr.: number of non isomorphic graphs
- # pol. vert.: total number of vertices for all polytopes
- # cong.: number of distinct conglomerates
- # int. gr.: number of interesting graphs (approx. by greedy heuristic)
- pc.: percent of interesting graphs

Demonstration

Perspectives



How can you help?

- By using it.
- By uploading special graphs.
- Mail us about problems, suggestions, ideas, etc.

Thank you for your attention!

http://hog.grinvin.org