Les Cahiers du GERAD ISSN: 0711–2440

Finding conjectures in graph theory with AutoGraphiX

M. Aouchiche, G. Caporossi, P. Hansen,
N. Mladenović, C. Lucas

G–2016–90

November 2016

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Finding conjectures in graph theory with AutoGraphiX

Mustapha Aouchiche [a]
Gilles Caporossi [a]
Pierre Hansen [a]
Nenad Mladenović [b]
Claire Lucas [c]

[a] *GERAD & HEC Montréal (Québec) Canada*
[b] *LAMIH, University of Valenciennes, France*
[c] *Saint Gobain Research Laboratories, France*

mustapha.aouchiche@gerad.ca
gilles.caporossi@gerad.ca
pierre.hansen@gerad.ca
nenad.mladenovic@univ-valenciennes.fr
claire.lucas@hotmail.fr

November 2016

Les Cahiers du GERAD
G–2016–90

Abstract: Graph theoretical heuristics are used extensively in many fields to solve approximately large scale optimization problems. Graph theoretical heuristics can also be used to advance graph theory *per se*, *i.e.*, to refute, find, corroborate, prove or give ideas of proof of conjectures on graph invariants, *i.e.*, numerical quantities that not depend on vertex or edge labeling. In this paper, we present and discuss the AutoGraphiX system which finds automatically or in some cases, interactively conjectures in graph theory. We consider in particular conjectures on pairs of a set of 20 graph invariants, which gives rise to 1520 cases.

Keywords: Graph, invariant, conjecture, proof, refutation, Variable neighbohood search

Résumé : Les heuristiques basées sur la théorie des graphes sont largement utilisées dans plusieurs domaines pour résoudre approximativement des problèmes d'optimisation de grandes tailles. Elles sont aussi ulilisées pour faire progresser la théorie des graphes *per se*, *i.e.* réfuter, trouver, corroborer, prouver ou donner des idées de preuve des conjectures sur des invariants graphiques, *i.e.*, des paramètres numériques indépendants de la numérotation des sommets ou des arêtes. Dans cet article, nous présentons et discutons le système AutoGraphiX qui trouve automatiquement ou, dans certains cas, interactivement des conjectures en théorie des graphes. Nous considérons en particulier des conjectures sur des paires d'invariants pris d'un ensemble de vingt.

Mots clés : Graphe, invariant, conjecture, preuve, réfutation, recherche à voisinage variable

1 Variable neighborhood search

Metaheuristics are general frameworks to build heuristics for solving combinatorial and global optimization problems. They have been the subject of intensive research since Simulated Annealing was proposed [54] as a general scheme for building heuristics which get out of local minima. Several other metaheuristics were soon proposed. For discussion of the best-known of them the reader is referred to the books of surveys [21, 41, 60]. Some of the many successful applications of metaheuristics are also mentioned there.

Variable Neighborhood Search (VNS) [49, 50, 51, 52, 56] is a metaheuristic which exploits systematically the idea of neighborhood change, both in descent to local minima and in escape from the valleys which contain them. VNS exploits systematically the following observations:

- A local minimum with respect to one neighborhood structure is not necessary so for another.
- A global minimum is a local minimum with respect to all possible neighborhood structures.
- For many problems local minima with respect to one or several neighborhoods are relatively close to each other.

Unlike many other metaheuristics, the basic schemes of VNS and its extensions are simple and require few, and sometimes no parameters. Therefore, in addition to providing very good solutions, often in simpler ways than other methods, VNS gives insight into the reasons for such a performance, which, in turn, can lead to more efficient and sophisticated implementations.

Function `VNS` (x, k_{max}, t_{max});
1 **repeat**
2     $k \leftarrow 1$;
3   **repeat**
4         $x' \leftarrow$ `Shake`(x, k);
5         $x'' \leftarrow$ `FirstImprovement`(x') ;
6   `NbhoodChange`(x, x'', k);
    until $k = k_{max}$;
7     $t \leftarrow$ `CpuTime()`;
until $t > t_{max}$;

Algorithm 1: Steps of the basic VNS

The Basic VNS (BVNS) method [56] combines deterministic and stochastic changes of neighbourhood. Its steps are given in Algorithm 1. Often successive neighbourhoods will be nested. Observe that point x' is generated at random in Step 4 in order to avoid cycling, which might occur if deterministic rules were applied. In Step 5, several neighborhoods may be used. In this case, we speak about *variable neighborhood descent* (VND), the scheme of which is given in Algorithm 2. For more details about VNS and its applications in solving problems in different domains of sciences see the recent survey [53] as well as the references therein.

Function `VND` (x, k'_{max});
repeat
    $k \leftarrow 1$;
    repeat
        $x' \leftarrow arg\min_{y \in \mathcal{N}'_k(x)} f(x)$;
        `NbhoodChange` (x, x', k);
    until $k = k'_{max}$;
until *no improvement is obtained*;

Algorithm 2: Steps of the basic VND

In all its applications, VNS is used as an optimization tool. These applications are mainly solving specific optimization problems. However, VNS can also be used in *discovery science*, *i.e.*, help in the development of theories. The first domain to be addressed in this way was graph theory. VNS is the fundamental tool exploited in the system AutoGraphiX (AGX, for short) [5, 26, 27], which is devoted to conjecture–making, and therefore to scientific discovery, in graph theory. A long series of papers (see the list in [8]) with the common title "*Variable neighborhood search for extremal graphs*" was published. Several of the papers which use AGX without being included within this series are listed in [8]. This system addresses the following problems:

- Find a graph satisfying given constraints;
- Find optimal or near optimal graphs for an invariant subject to constraints;
- Refute a conjecture;
- Suggest a conjecture (or repair or sharpen one);
- Provide a proof (in simple cases) or suggest an idea of proof.

A basic idea is then to consider all of these problems as parametric combinatorial optimization problems on the infinite set of all graphs (or in practice some smaller subset) solved with a generic heuristic. This is done by applying VNS to find extremal graphs, with a given number n of vertices (and possibly also a given number of edges). Then a VND with many neighbourhoods is used. Those neighborhoods are defined by modifications of the graphs such as the removal or addition of an edge, rotation of an edge, and so forth. Once a set of extremal graphs, parametrized by their order, is found, their properties are explored with various data mining techniques, leading to conjectures, refutations and simple proofs or ideas of proof.

2 The AutoGraphiX system

Among the first application of VNS, a computer program, called the *AutoGraphiX system* (AGX, for short) [5, 26, 27], was built for conjecture–making in graph theory. This system has been developed at GERAD, Montreal, since 1997. Conjectures obtained with AGX were proved by the present authors or by graph theorists from several countries. A graph invariant is a function of a graph G which does not depend on labeling of G's vertices or edges. Examples of graph invariants are the diameter, the radius, the average distance, the independence number and the index (definitions will be given below). Graph theory is replete with theorems involving graph invariants. They are either *algebraic*, *i.e.,* equalities or inequalities involving one or several invariants, or *structural*, *i.e.,* characterizations of the families of graphs for which an invariant takes an extremal value. Both types of results can be conjectured by AGX, in a fully automated way, or in some cases, to be carefully distinguished, in an assisted way. Let $\mathcal{G}_n$ and $\mathcal{G}_{n,m}$ denote respectively the sets of all graphs with n vertices, and with n vertices and m edges. Two basic ideas underlie the systems AGX:

- *Most problems of extremal graph theory can be viewed as problems of parametric combinatorial optimization of the form*
$$\min / \max_{G \in \mathcal{G}_n} i(G) \quad \text{or} \quad \min / \max_{G \in \mathcal{G}_{n,m}} i(G) \tag{1}$$
for some invariant $i(G)$ with parameters n and m, or the exploitation of their solutions (in practice only moderate values of n and m will be considered);
- *All problems of the form (1) can be solved approximately by a generic heuristic.*

To obtain such a heuristic, the Variable Neighborhood Search metaheuristic (VNS) is specialized. VNS exploits systematically changes in neighborhoods used in the search, both in a descent phase to obtain a locally extremal graph, and in a "shaking" phase, to get out of the corresponding valley (or away from the corresponding mountain) in order to find a better graph. Rules of VNS applied in AGX are the following:

1. Select the set of neighborhood structures $N_k,\ k = 1, \dots k_{max}$ that will be used in the search for a better locally optimal graph, and a stopping condition. Choose an initial graph G.
   Repeat until the stopping condition is met:
2. Set $k = 1$;
3. Until $k = k_{max}$, repeat the following steps:
   (a) (*shaking*) generate a graph G' from the k^{th} neighborhood of G ($G' \in N_k(G)$);
   (b) (*descent*) apply VND with G' as initial graph; denote with G'' the locally optimal graph obtained;

(c) (*improvement or continuation*) if $i(G'')$ is better than $i(G)$, the best value of i for a previously visited graph, move there, *i.e.,* replace G by G'', and continue search within $N_1(G)$; otherwise, set $k \leftarrow k+1$.

The stopping condition is usually a maximum computing time. The optimization routine of VNS is called *variable neighborhood descent*. It explores systematically larger and larger neighborhoods of the current graph, and performs a move whenever it is profitable (first improvement) or is also best within its neighborhood (best improvement). The neighborhoods used initially in AGX are the following: remove, add, move, detour, short cut, 2–opt, insert pending vertex, add pending vertex, and remove vertex.

In the most recent version of AGX, the VND routine is replaced by *Learning Descent* (LD), in order to keep track of which transformations are the most fruitful and to reinforce their use. The LD used in AGX is described in [24].

Once a set of (presumably) extremal graphs has been found, conjectures can be stated by one of the following 3 approaches [26]:

(i) a *numerical method* which applies the mathematics of Principle Component Analysis [25] to determine, in polynomial time, a basis of affine relations between invariants, satisfied by the extremal graphs found.

(ii) a *geometric method* which views extremal graphs as points in invariants space and applies a "gift-wrapping" algorithm to find their convex hull and linear inequality relations associated with its facets. Note that a similar approach is used in GraPHedron [28];

(iii) an *algebraic method* [3, 7, 5] which recognizes to which family (or families) of graphs the extremal graphs belong, then uses a database of formulae for invariants in function of the order of G to obtain conjectures.

3 Bounding invariants

The AGX system was built for finding extremal graphs with respect to a given invariant or an algebraic combination of invariants, *i.e.* finding graphs that minimize or maximize a given invariant function. Once the extremal graphs obtained, research is done for finding a lower bound, in the case of minimization, or an upper bound in case of maximization, on the invariant function under study. Thus, naturally, the first AGX task is bounding one invariant at a time, *i.e.*, without considering combinations of invariants. The *degree* of a vertex v in G, denoted by $d(v) = d_G(v)$ is the number of vertices adjacent to v in G. The minimum, average and *maximum degrees* in G are denoted by δ, $\overline{d}$ and Δ respectively. The distance $d(u,v) = d_G(u,v)$ between two vertices u and v in a graph G is the *length* (number of edges) of a shortest path between u and v. The *average distance* is denoted by $\bar{l}$.

The problem of upper bounding the average distance in terms of order and minimum degree was studied using AGX in [16]. Six conjectures were obtained, one of which was proved:

Theorem 1 ([16]) *Let $G=(V,E)$ be a connected graph on $n \geq 7$ vertices with average distance $\bar{l}$ and minimum degree $\delta \geq 2$. Then*

$$\bar{l} \leq \frac{n+1}{3} - \frac{8}{n} + \frac{4}{n-1}$$

with equality iff G is composed of two triangles linked by a path.

After the above result, we progressively generalized our experiments according to the value of the minimum degree: $\delta = 3$, $\delta = 4$ and $\delta = 5$.

Then, the general case, with a given lower bound on δ was considered. Among the obtained conjectures, we recall only the next two. Some graph definitions are needed.

(a) Let n and δ be integers such that $n = q(\delta+1)$ with $q \geq 2$ and $\delta \geq 3$. Consider the graph G obtained from the graph composed of q copies of $K_{\delta+1}$, say $K^i_{\delta+1}$ for $i = 1, 2, \ldots q$, by removing an edge $u^i v^i$

from each $K_{\delta+1}^i$ for $i = 2, \dots q-1$, then adding the edges $v^i u^{i+1}$, uu^2 and $v^{q-1}v$ where u is any vertex from $K_{\delta+1}^1$ and v any vertex from $K_{\delta+1}^q$. If $q = 2$, there are two copies of $K_{\delta+1}$, then we add only the edge uv. See Figure 1 for $(n, \delta) = (25, 4)$.

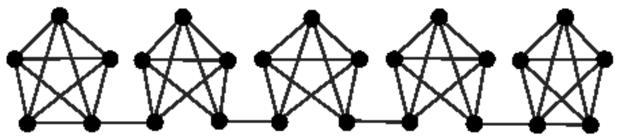

Figure 1: Presumably extremal graph for $(n, \delta) = (25, 4)$

(b) Let n and δ be integers such that $n = q(\delta + 1) + 2$ with $q \geq 2$ and $\delta \geq 3$. Consider the graph G obtained from the graph described in (a) by replacing each of $K_{\delta+1}^1$ and $K_{\delta+1}^q$ by the graph H obtained from $K_{\delta+2}$ on the set of vertices $\{w_1, w_2, \dots w_{\delta+2}\}$, by deleting the edges w_1w_2, w_1w_3 and w_iw_{i+1} for $i = 4, 6, \cdots p+1$, where $p = \delta$ if δ is even and $p = \delta + 1$ if δ is odd. The vertices u and v from the graph described in (a) correspond to w_1 from each copy of H respectively. Again, if $q = 2$, there are two copies of H, then we add only the edge uv. See Figure 2 for $(n, \delta) = (22, 4)$.

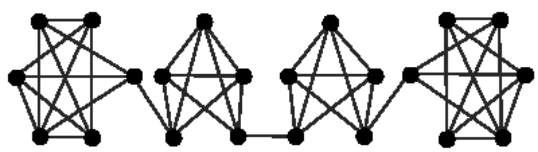

Figure 2: Presumably extremal graph for $(n, \delta) = (22, 4)$

Conjecture 1 ([16]) *Let $G = (V, E)$ be a connected graph on n vertices with minimum degree $\delta \geq 3$ where $n = (\delta + 1) \cdot k$ for some integer $k \geq 2$. Then the average distance $\bar{l}$ of G satisfies*

$$\bar{l} \leq \frac{n+1}{\delta+1} - \frac{4\delta}{n} + \frac{4\delta^2 - \delta - 2}{(\delta+1)(n-1)}$$

with equality if and only if G is obtained as described in (a).

Note that Kouider and Winkler [55] gave the extremal graphs of Conjecture 1 as extremal cases, without a proof, for the case $n = (\delta + 1)k$. However, the corresponding bound does not appear to be generalizable for all integers n and δ. If true, the next conjecture provides a global and sharp upper bound on $\bar{l}$ in terms of δ.

Conjecture 2 ([16]) *Let $G = (V, E)$ be a connected graph on n vertices with minimum degree $\delta \geq 3$. Then the average distance $\bar{l}$ of G satisfies*

$$\bar{l} \leq \frac{n+1}{\delta+1} - \frac{2\delta^2 - 14\delta + 36}{n} + \frac{12\delta^2 - 75\delta + 150}{(\delta+1)(n-1)}.$$

The bound is reachable only if $n = (\delta + 1) \cdot k + 2$ for some integer $k \geq 2$, in which case the extremal graph G is the graph obtained as described in (b).

The *adjacency matrix* A of G is a 0–1 $n \times n$–matrix indexed by the vertices of G and defined by $a_{ij} = 1$ if $ij \in E$. Denote by $(\lambda_1, \lambda_2, \dots, \lambda_n)$ the A–spectrum of G, *i.e.*, the spectrum of the adjacency matrix of G, and assume that the eigenvalues are labeled such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. The *spectral spread* of G is defined by $s(G) = \lambda_1(G) - \lambda_n(G)$. The problem of finding the maximum value of $s(G)$ among the class of connected graphs of given order n is an open problem. Experiments were done with AGX to study the problem, and the extremal graphs were found in [4] (see also [6, 14]). A conjecture was obtained, but before its statement, recall that a *complete split graph* with parameters n, q $(q \leq n)$, denoted by $CS(n, q)$, is a graph on n vertices consisting of a clique on q vertices and an independent set on the remaining $n - q$ vertices in which each vertex of the clique is adjacent to each vertex of the independent set.

Conjecture 3 ([6]) *Let G be a connected graph on $n \geq 3$ vertices. Then*

$$s(G) \leq \sqrt{4qn - 3q^2 - 2q + 1}$$

with equality iff G is the complete split graph $CS(n,q)$ with an independent set of size $n - q = \left\lceil \frac{n}{3} \right\rceil$ and a clique of size $q = \left\lfloor \frac{2n}{3} \right\rfloor$.

Note that the above conjecture did appear in [42], in terms of extremal graphs only, where it has been verified by computer for graphs up to 9 vertices, but remained unsolved.

The *energy* $E(G)$ of a graph G, introduced in [44], is defined as the sum of the absolute values of its eigenvalues, *i.e.*

$$E(G) = \sum_{i=1}^{n} |\lambda_i(G)| = 2 \sum_{\lambda_i > 0} \lambda_i(G) = 2 \sum_{\lambda_i < 0} |\lambda_i(G)|.$$

A lollipop $Lol_{n,g}$, with $n \geq g \geq 3$, is a graph obtained from a cycle C_g and a path P_{n-g} by adding an edge between a vertex from the cycle and an endpoint from the path. $Lol_{n,n-1}$ is called the short lollipop while $Lol_{n,3}$ is the long lollipop.

In order to find lower and upper bounds on the energy, AGX was used in [23]. They found the following conjectures afterwards proved by hand.

Theorem 2 *Let G be a simple graph on n vertices and m edges with energy E. Then*

1. *$E \geq 4m/n$;*
2. *$E \geq 2\sqrt{m}$ with equality iff G is a complete bipartite graph plus possibly some isolated vertices;*
3. *if G is connected, $E \geq 2\sqrt{n-1}$ with equality iff G is the star S_n;*
4. *$E \leq 2m$ with equality iff G is composed of disjoint edges and possibly isolated vertices.*

In this study, the particular case of unicyclic graphs was considered. Some unicyclic graphs that maximize the energy are given in Figure 3. The following conjecture was stated.

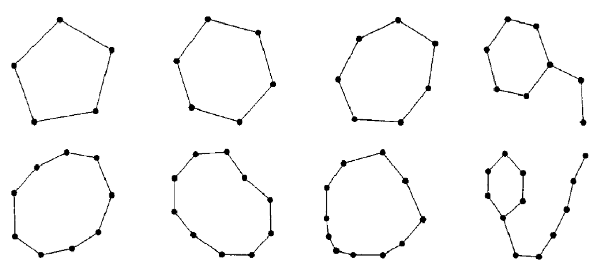

Figure 3: Unicyclic graphs with largest energy

Conjecture 4 *Among unicyclic graphs on n vertices the cycle C_n has maximal energy if $n \leq 7$ and $n = 9, 10, 11, 13$ and 15. For all other values of n the unicyclic graph with maximum energy is the lollipop $Lol_{n,6}$.*

The *Laplacian* of a graph G is the matrix defined by $L = Deg - A$, where Deg is the diagonal matrix whose diagonal entries are the vertex degrees in G and A is the adjacency matrix of G. The *Laplacian spectrum* of G is the spectrum of L and is denoted by $\mu_1, \mu_2, \ldots \mu_n$, where $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \mu_n = 0$. The second smallest Laplacian eigenvalue of a graph G is called *algebraic connectivity* of G [39] and denote $a = a(G)$. Experiments performed in [20] using AGX led to lower and upper bounds on the algebraic connectivity, *e.g.*,

Theorem 3 ([20]) *Let G be a connected graph on n vertices and m edges with algebraic connectivity a. If $G \not\cong K_n$, then*

$$a \leq \left\lfloor -1 + \sqrt{1 + 2m} \right\rfloor.$$

Moreover, the bound is sharp for all $m \geq 2$.

The *signless Laplacian* of a graph G is the matrix defined by $Q = Deg + A$, where Deg is the diagonal matrix whose diagonal entries are the vertex degrees in G and A is the adjacency matrix of G. The *signless Laplacian spectrum* of G is the spectrum of Q and is denoted by $q_1, q_2, \dots q_n$, where $q_1 \geq q_2 \geq \cdots q_n$. For more details about Q and its spectrum, see [33, 34, 35, 32]. The paper [32] reports on AGX conjectures obtained at GERAD and related to Q–spectrum of a graph.

AGX was used in [47] for studying the problem of upper bounding the largest signless Laplacian eigenvalue q_1 in terms of order n and chromatic number χ of G (the minimum number of colors that can be assigned to the vertices of a graph such that two adjacent vertices are not assigned the same color).

Theorem 4 ([47]) *Let G be a graph on n vertices with signless Laplacian index q_1, chromatic number χ. Then*

$$q_1 \leq \frac{2n(\chi - 1)}{\chi}$$

with equality iff G is is a complete regular χ-partite graph.

The above result remains true if χ is replaced by ω. The equivalent result was also obtained using AGX.

The *Albertson irregularity* $Al = Al(G)$ of a graph $G = (V, E)$, introduced in [1], is defined as the of the absolute values of the differences between the degrees of the end-vertices of the edges of G, *i.e.*,

$$Al = Al(G) = \sum_{uv \in E} |d(u) - d(v)|.$$

Note that the difference $|d(u) - d(v)|$, for an edge uv is called [1] the *imbalance* of uv. The experiments done in [48] with the use of AGX, to find an upper bound on Al, did not only conjecture a bound but also did suggest a clear idea for proving it. The extremal graphs (see Figure 4) suggested by AGX belong to the well–known family of *fanned complete split graphs*. A *fanned complete split graph* with parameters $n, q, t (n \geq q \geq t)$, denoted by $FCS(n, q, t)$, is a graph (on n vertices) obtained from a complete split graph $CS(n, q)$ by connecting a vertex from the stable set by edges to t other vertices of the stable set. The curves of the irregularity for $9 \leq n \leq 12$ and $n - 1 \leq m \leq n(n - 1)/2$ are given in Figure 5.

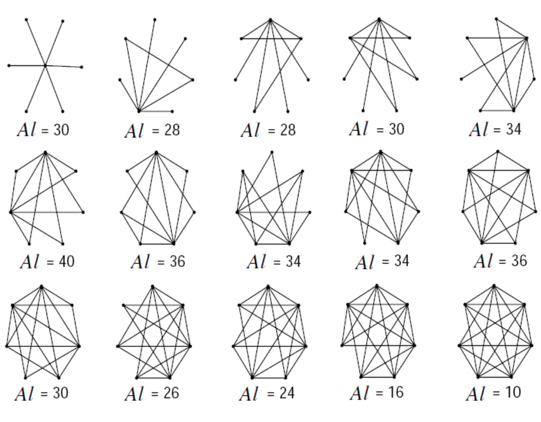

Figure 4: The extremal graphs for Al with $n = 7$ and $6 \leq m \leq 20$

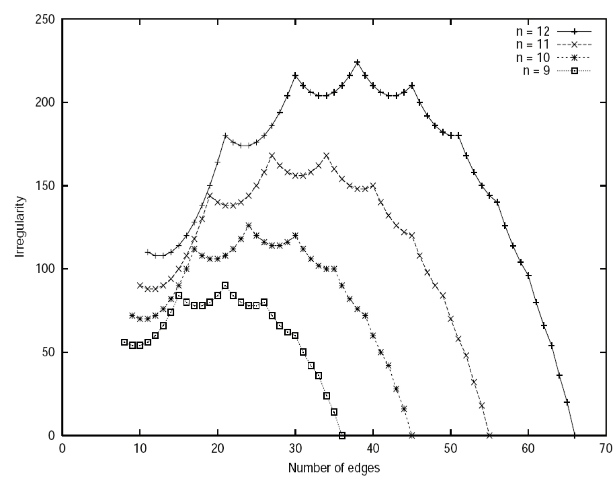

Figure 5: The curves of Al for $9 \leq n \leq 12$

The extremal graphs were obtained by AGX using a single move: *the rotation of an edge* (if $uv \in E$ and $uw \notin E$, the rotation of the uv to uw is the suppression of uv and the addition of uw). From where the proof idea: show that for any non-optimal graph, there exists an edge rotation that increases the irregularity. This proof works and the result is the following:

Theorem 5 ([48]) *If G is a graph on n vertices and m edges, then*

$$Al(G) \le s(n-s)(n-s+1) + t(t-2s-1)$$

where

$$s = \left\lfloor n - \frac{1}{2} - \sqrt{\left(n - \frac{1}{2}\right)^2 - 2m} \right\rfloor$$

and

$$t = m - s(n-s) - \frac{s(s-1)}{2}.$$

Moreover, the bound is attained iff G is a fanned complete split graph.

4 AGX Form 1

In this section, we report on a particular form of results obtained using AGX. More precisely, in our experiments, we considered a set of invariants (20 at first and then few others were added) and sought expressions of the following form (called AGX Form 1):

$$\underline{b}(n) \le i_1 \oplus i_2 \le \overline{b}(n) \tag{2}$$

where i_1 and i_2 are invariants of a graph G from the chosen set, $\oplus$ denotes one of the 4 operations $+, -, /$ and $\times$, $\underline{b}(n)$ and $\overline{b}(n)$ are, respectively, lower and upper bounding functions depending on the *order* n, or number of vertices, of G which are *best possible*, *i.e.,* such that for each value of n (except possibly very small ones, due to border effects) there is a graph G for which the bound is tight. Note that the form (2) is reminiscent of the well-known Nordhaus-Gaddum relations [12, 57]; however, it generalizes this last form in two ways:

(i) the operations $-$ and $/$ are considered in addition to $+$ and $\times$;

(ii) the invariants i_1 and i_2 are independent instead of having $i_2(G) = i_1(\overline{G})$, where $\overline{G}$ denotes the complementary graph of G, in which an edge joins vertices v_i and v_j iff there is no such edge in G.

Table 1: The 20 invariants considered in [3] for the AGX Form 1

Δ	The maximum degree.
δ	The minimum degree.
$\overline{d}$	The average degree.
$\overline{l}$	The average distance between all pairs of vertices.
D	The diameter.
r	The radius.
g	The girth, the length of the smallest cycle in a graph.
ecc	The average eccentricity.
π	The proximity or minimum normalized transmission.
ρ	The remoteness is maximum normalized transmission.
λ_1	The index or spectral radius.
Ra	The Randić index.
a	The algebraic connectivity or second smallest Laplacian eigenvalue.
ν	The vertex connectivity.
κ	The edge connectivity.
α	The independence number.
β	The domination number.
ω	The clique number.
χ	The chromatic number.
μ	The matching number.

In the thesis [3] expressions of AGX Form 1 were systematically studied for all pairs of invariants among a list of 20, given in Table 1. This amounts to 1520 cases. Results are summarized in [8]. For each case, we give the formulae for the lower and upper bounds together with the status of the conjecture: known (K), trivial (T), open (O), assisted open (AO), structural open (SO), refuted (R). For a proved automated, assisted or structural conjecture, we refer to the paper where it is proved, and we indicate that no result is obtained (NR) whenever it is the case. Statistics on the numbers of cases which fall in these categories are given in Table 2. It appears that:

(i) cases in which no result was obtained (because the graphs obtained by AGX do not present sufficient regularity) are rare (3.62%);

(ii) known results rediscovered by AGX are also rare (2.43%);

(iii) complete results, *i.e.*, algebraic formulae and extremal graphs, are frequent (82.89%). They comprise obvious results, usually proved automatically by AGX (55.59%), and non trivial results proved by hand either at GERAD or by graph theorists of various countries (23.75%), in such cases references to the proofs are given;

(iv) in some other cases only structural conjectures, *i.e.*, only families of extremal graphs are obtained (11.06%), in some cases formulas are obtained by hand (5.67%);

(v) cases where AGX conjectures were refuted are rare (3.62%);

(vi) there remains a consequent number of open conjectures (8.42%). This is due to the fact that our systematic effort done to prove some families of conjectures was not enough or that some invariants appearing there are hard to handle or that some conjectures appear to be hard.

Table 2: Summary of results

Known results reproduced	37	(2.43 %)
Obvious results	845	(55.59 %)
Complete results proved by hand	361	(23.75 %)
Proved structural results and formulae by hand	46	(3.03 %)
Proved structural results only	21	(1.38 %)
Open complete results	33	(2.17 %)
Open structural results and formulae by hand	34	(2.24 %)
Open structural results only	61	(4.01 %)
Refuted complete results	21	(1.38 %)
Refuted structural results and formulae by hand	6	(0.40 %)
Refuted structural results only	0	(0.00 %)
No results	55	(3.62 %)
Total	1520	(100 %)

Results for a pair of invariants can be *complete*, *i.e.,* consist of both conjectured best possible functions $\underline{b}(n)$ and $\overline{b}(n)$ and the corresponding characterizations of the extremal graphs, or *structural*, *i.e.,* consist of the characterizations of extremal graphs only. This last case occurs when algebraic expressions for $\underline{b}(n)$ and $\overline{b}(n)$ are too difficult for AGX to obtain, or when such expressions do not exist, *e.g.* because they correspond to solutions of an equation of degree 5 or more.

In some fairly frequent cases, complete results are simple and can be proved by AGX in a fully automated way; we then refer to them as *observations*. If results are structural, algebraic expressions for $\underline{b}(n)$ and $\overline{b}(n)$ can sometimes be deduced, in an assisted way, from the characterization of extremal graphs. In some fairly rare cases the graphs obtained by AGX and conjectured to be extremal present very little or no regularity and no results are obtained.

In each case, *i.e.*, each bound, graphs with 5 to 20 vertices were considered. Computing time on Intel Xeon with 2.66 GHz and 2 Gb RAM, at that moment, varied from less than 1 second in the frequent case in which a bound could be obtained automatically, without using VNS, up to 75 seconds per graph in the most complex cases, whether results were obtained or not. Trying longer computing times did not give better results.

Among all bounds conjectured in [3], 128 remain open, and among all possible cases, AGX failed to find a conjecture or a false conjecture in 82 cases. Under the assumption that these open conjectures are difficult to prove and that AGX failed when the cases are difficult to handle, we tried to figure out the reasons of these difficulties and we gathered the statistics summarized in Table 4 regarding the operations and Table 5 regarding the bounds. In these tables, we use O, AO and SO for open, assisted open and structural open conjecture, respectively, and R and AR for refuted conjecture and refuted assisted conjecture. NR is used to say that no result is obtained in the corresponding case. T–O and T–R are used for the total over open conjectures and cases with no result or with refuted conjectures, respectively. Total indicates the sum of T–O and T–R.

According to the statistics, the most difficult invariant to handle is the domination number β with a total of 46 occurrences over 420 (10.95 %). The second most difficult invariant seems to be the Randić index Ra with 39 occurrences (9.51 %). Then comes a set of three invariants with 35 occurrences each (8.33 %). Two of these three invariants are eigenvalues, the index λ_1 and the algebraic connectivity a, and the third is a metric invariant, namely, the remoteness ρ.

After that, we can find three sets each containing two invariants with almost the same occurrences: the average eccentricity ecc and the average distance $\bar{l}$ with 30 and 29 occurrences (7.14 % and 6.91 %), respectively, the proximity π and the independence number α with 25 occurrences each (5.95 %), and the radius r and the maximum degree Δ with 20 and 19 occurrences, respectively. The remaining nine invariants can be split into three sets each with three invariants with almost the same number of occurrences: the average degree $\overline{d}$, the diameter D (the maximum distance in a graph) and the chromatic number χ, with 14, 13 and 13 occurrences, respectively, form the first set, the minimum degree δ, the edge connectivity κ and the clique number ω, with 9, 9 and 8 occurrences, respectively, form another set, and finally, the set of the less frequent invariants is composed of the matching number μ, the girth g and the vertex connectivity ν with 6, 5 and 5 occurrences, respectively.

Table 3: Difficulties regarding the invariants

Inv.	O	AO	SO	T–O	NR	R	AR	T–R	Tot.
β	9	11	12	32	12	0	2	14	46
Ra	9	12	6	27	8	4	0	12	39
λ_1	5	1	11	17	10	7	1	18	35
a	4	6	19	29	6	0	0	6	35
ρ	3	8	17	28	6	1	0	7	35
ecc	5	6	9	20	6	4	0	10	30
$\bar{l}$	3	5	9	17	9	2	1	12	29
π	6	3	8	17	8	0	0	8	25
α	5	4	7	16	7	2	0	9	25
r	5	4	2	11	8	1	0	9	20
Δ	0	2	3	5	12	1	1	14	19
$\overline{d}$	2	1	6	9	1	1	3	5	14
D	3	0	2	5	3	4	1	8	13
χ	0	2	6	8	2	2	1	5	13
δ	2	0	2	4	1	4	0	5	9
κ	2	1	0	3	2	4	0	6	9
ω	0	0	2	2	3	1	2	6	8
μ	1	1	0	2	3	1	0	4	6
g	1	0	1	2	1	2	0	3	5
ν	1	1	0	2	2	1	0	3	5

If we consider the difficulty with respect to the operations, it is easy to see that the product is the most difficult combination to handle. It occurs 79 times over 210 (37.62 %). The other three operations appear to present the same degree of difficulty: 41 occurrences (19.52 %) for the addition, 43 occurrences (20.48 %) for the subtraction and 47 occurrences (22.38 %) for the division.

If we distinguish between lower and upper bound, it is almost the same degree of difficulty in both cases even if the upper bounds seems to be slightly more difficult than the lower bound with 117 (55.71 %) cases among 210.

Among the bounds considered in the thesis [3], some were already known in the graph theory literature, *e.g.*, $\delta \leq \overline{d} \leq \lambda_1 \leq \Delta$; $\bar{l} \leq \alpha$ [29]; $\chi \leq \lambda_1 + 1$ [64]; and $a \leq \frac{n\delta}{n-1}$ [39].

Table 4: Difficulties regarding the operations

Op.	O	AO	SO	T–O	NR	R	AR	T–R	Tot.
$-$	11	9	11	31	6	4	2	12	43
$+$	5	7	8	20	14	6	1	21	41
$/$	5	10	18	33	12	1	1	14	47
$\times$	12	8	24	44	23	10	2	35	79

Table 5: Difficulties regarding the bounds

Bound	O	AO	SO	T–O	NR	R	AR	T–R	Tot.
Lower	11	18	22	51	31	9	2	42	93
Upper	22	16	39	77	24	12	4	40	117

Note that some of the above listed inequalities were obtained twice. For instance, the inequality $\lambda_1 \leq \Delta$ was obtained as $\lambda_1 - \Delta \leq 0$ and $\lambda_1/\Delta \leq 1$.

Some of the bounds are naturally easy to obtain. When both invariants considered come from the same vector or matrix, say S, by taking its minimum ($m = \min S$), average $\overline{s} = \frac{1}{|S|}\sum_{s\in S} s$) or maximum value ($M = \max S$), it is obvious that

$$m \leq \overline{s} \leq M$$

with equality iff the entries of S are equal. Immediate consequences of this double inequality are

$$M - \overline{s} \geq 0; \overline{s} - m \geq 0; M - m \geq 0; \frac{M}{\overline{s}} \geq 1; \frac{\overline{s}}{m} \geq 1; \frac{M}{m} \geq 1.$$

For example, for all connected graphs G with $n \geq 2$ vertices, diameter D (the maximum among all the distances in G) and average distance $\bar{l}$,

$$D - \bar{l} \geq 0 \quad \text{and} \quad \frac{D}{\bar{l}} \geq 1$$

with the equalities iff G is a complete graph.

There exists another kind of bounds easy to obtain. Actually, when the relevant families of extremal graphs for the invariants i_1 and i_2 are considered and if they have a non-empty intersection a proved and best possible bounding function is obtained. For our next example, we need the following definitions. The *eccentricity* $ecc(v)$ of a vertex v in G is the maximum among the distances from v to all other vertices in G. The *radius* $r = r(G)$ of a graph G is the maximum over the eccentricities of its vertices. The *Randić index* $Ra(G)$ of a graph $G = (V, E)$, introduced in [59], is defined by

$$Ra = Ra(G) = \sum_{uv\in E} \frac{1}{\sqrt{d_u d_v}},$$

where d_u and d_v denote the degree of the vertices u and v, respectively. It is well known that, on the one hand, the Randić index Ra is minimum for the star S_n, which is among the graphs that minimize the radius r, and on the other hand, Ra is maximum for any regular graph, and among the regular graph the cycle C_n maximizes r. Thus the following bounds are immediately obtained.

$$1 + \sqrt{n-1} \leq Ra + r \leq \tfrac{n}{2} + \left\lfloor \tfrac{n}{2} \right\rfloor \quad \text{and} \quad \sqrt{n-1} \leq Ra \cdot r \leq \tfrac{n}{2} \cdot \left\lfloor \tfrac{n}{2} \right\rfloor$$

with equality in both lower (resp. upper) bounds iff G is the star S_n (resp. the cycle C_n). Another example: the average distance $\bar{l}$ is minimum (resp. maximum) for the complete graph K_n (resp. path P_n) with $\bar{l}(K_n) = 1$ (resp. $\bar{l}(P_n) = (n+1)/3$, while the maximum degree Δ is maximum for K_n, with $\Delta(K_n) = n-1$, and minimum for P_n, with $\Delta(P_n) = 2$. Thus

$$2 - n \leq \bar{l} - \Delta \leq \frac{n-5}{3} \quad \text{and} \quad \frac{1}{n-1} \leq \frac{\bar{l}}{\Delta} \leq \frac{n+1}{6}$$

with equality in both lower (resp. upper) bounds iff G is the complete graph K_n (resp. path P_n).

The other results were obtained as conjectures and can be divided into three types. A common step for all the three types is the VNS optimization. At that step, the optimization component of AGX is executed and presumably extremal graphs are obtained. Then, a component, aimed for finding (linear) relations between selected invariants, is executed. In case of success, we obtain a formula: a lower bound for a minimizing problem or an upper bound for a maximizing problem. Thus, we get a conjecture containing a bound with corresponding extremal graphs and we speak about *complete conjectures*, that constitutes the first type of results. Among such results, we cite the following theorems and conjectures.

Theorem 6 ([17]) *Let G be a connected graph on $n \geq 3$ vertices with index λ_1 and average distance $\bar{l}$. Then*

$$\lambda_1 + \bar{l} \leq n$$

with equality iff G is the complete graph K_n.

Conjecture 5 ([46]) *Let G be a connected graph on $n \geq 6$ vertices with signless Laplacian spectral radius q_1 and chromatic number χ. Then*

$$q_1 - \chi \leq \frac{3n-8}{2}$$

with equality iff G is the $\lfloor n/2 \rfloor$–partite graph $K_{p,2,2,\ldots,2}$, where $p = 2 + n \bmod(2)$.

A relation between q_1 and Δ, obtained by AGX, is proved in [32].

Theorem 7 ([32]) *Let G be a connected graph on $n \geq$ vertices with signless Laplacian index q_1 and maximum degree Δ. Then*

$$q_1 - \Delta \geq 1$$

with equality iff G is the star S_n.

The *girth* $g = g(G)$ of a connected graph G on $n \geq 3$ vertices with at least n edges, is the length (number of edges) of its smallest cycle. The next theorem, proved in [18], was first conjectured by AGX.

Theorem 8 ([18]) *Let G be a connected graph on $n \geq 3$ vertices and $m \leq n$ edges with girth g and average distance $\bar{l}$. Then*

$$\frac{\bar{l}}{g} \geq \begin{cases} \frac{n}{4(n-1)} & \text{if } n \text{ is even,} \\ \frac{n+1}{4n} & \text{if } n \text{ is odd.} \end{cases}$$

Moreover, the bound is reached for cycles.

The *matching number* $\mu = \mu(G)$ of a graph is the maximum number of independent (pairwise non-incident) edges in G. The following result was conjectured using AGX and then proved in [63].

Theorem 9 ([63]) *Let G be a connected graph, $G \ncong K_3$, on $n \geq 3$ vertices with adjacency index λ_1 and matching number μ. Then*

$$\lambda_1 - \mu \leq n - 1 - \left\lfloor \frac{n}{2} \right\rfloor$$

with equality iff G is the complete graph K_n. Also,

$$\frac{\lambda_1}{\mu} \leq \sqrt{n-1}$$

with equalities iff G is the star S_n.

Note that there exists an infinite family of counterexamples [63] for the relation $\lambda_1 + \mu \geq \sqrt{n-1} + 1$ first conjectured by AGX.

When AGX could not provide a complete conjecture, an interactive procedure for recognizing the extremal graphs was launched. If the extremal graph are recognized and the corresponding formulas of the invariants under study are available in the database, substitutions are done and then bounds are obtained. The results so obtained are called *assisted conjectures*. First, recall that the *vertex connectivity* $\nu = \nu(G)$ of a connected graph G is the minimum number of vertices whose removal disconnects G.

Theorem 10 ([37, 65]) *Let G be a connected graph on $n \geq 3$ vertices with index λ_1 and vertex connectivity ν. Then*

$$\lambda_1 - \nu \leq n - 3 + t; \quad \frac{\lambda_1}{\nu} \leq n - 2 + t,$$

where t is such that $0 < t < 1$ and $t^3 + (2n-3)t^2 + (n^2 - 3n + 1)t - 1 = 0$. Moreover, equalities hold iff G is the kite $Ki_{n,n-1}$.

Finding the bound in the above theorem in an automated way was not possible since it contains a factor that uses an implicit solution of a difficult to solve equation.

Another example with a complicated bound is the following theorem proved in [9].

Theorem 11 ([9]) *Let $G = (V, E)$ be a connected graph of order n with independence number α and maximum degree Δ. Then*

$$\alpha - \Delta \quad \leq \quad \max\left\{ \left\lfloor n - \frac{n-1}{\lceil\sqrt{n-1}\rceil} \right\rfloor - \lceil\sqrt{n-1}\rceil, \ \left\lfloor n - \frac{n-1}{\lfloor\sqrt{n-1}\rfloor} \right\rfloor - \lfloor\sqrt{n-1}\rfloor \right\}.$$

The bound is reached for every n.

For the above theorem, the difficulty is in the fact that the bound is an integer that implies a combination of fractions and square roots of integers.

Finally, when the recognition of the extremal graphs succeeded, but no formulae were found, we state a conjecture about the structure of the extremal graphs. In this case, we speak about *structural conjectures*.

The well–known result, in spectral graph theory, $\lambda_1(G) \geq \overline{d}(G)$ with equality iff G is a regular graph, was proved in [30]. Then, they proposed to consider the difference between the index and the average degree as a measure of the *irregularity* of a graph (other definitions of irregularity in graphs have been proposed, see [1, 19], and for a comparison between them see [43]). Thus the irregularity of a graph G is defined by $Irr(G) = \lambda_1(G) - \overline{d}(G)$. The problem of finding an upper bound on the irregularity and characterizing the most irregular graphs remains open. The following conjecture related to the irregularity of a graph have been formulated after some experiments with the system AGX. First, we need the following definition. A *pineapple* with parameters n, q $(q \leq n)$, denoted by $PA(n, q)$, is a graph on n vertices consisting of a clique (a set of pairwise adjacent vertices) on q vertices and an independent set (a set of pairwise non-adjacent vertices) on the remaining $n-q$ vertices in which each vertex of the independent set is adjacent to a unique and the same vertex of the clique.

Conjecture 6 ([4, 6]) *The most irregular connected graph on n $(n \geq 10)$ vertices is a pineapple $PA(n, q)$ in which the clique size q is equal to $\lceil \frac{n}{2} \rceil + 1$.*

The issue in the above theorem is the difficulty to get an explicit formulae of the index for some classes of graphs.

The well–known result, in spectral graph theory, $\lambda_1(G) \geq \overline{d}(G)$ with equality if and only if G is a regular graph, was proved by Collatz and Sinogowitz [30] in 1957. Then, they proposed to consider the difference between the index and the average degree as a measure of the *irregularity* of a graph (other definitions of irregularity in graphs have been proposed, see [1, 19], and for a comparison between them see [43]). Thus the irregularity of a graph G is defined by $Irr(G) = \lambda_1(G) - \overline{d}(G)$. The problem of finding an upper bound on the irregularity and characterizing the most irregular graphs remains open. The following conjecture related to the irregularity of a graph have been formulated after some experiments with the system AGX. First, we need the following definition. A *pineapple* with parameters n, q $(q \leq n)$, denoted by $PA(n, q)$, is a graph on n vertices consisting of a clique (a set of pairwise adjacent vertices) on q vertices and an independent set (a set of pairwise non-adjacent vertices) on the remaining $n-q$ vertices in which each vertex of the independent set is adjacent to a unique and the same vertex of the clique. Some pineapples are illustrated in Figure 6.

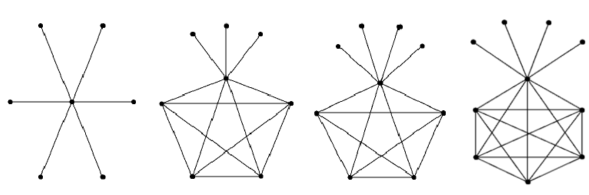

Figure 6: Presumably most irregular graphs for $n = 7, 8, 9, 10$

Conjecture 7 ([4, 6]) *The most irregular connected graph on n $(n \geq 10)$ vertices is a pineapple $PA(n, q)$ in which the clique size q is equal to $\lceil \frac{n}{2} \rceil + 1$.*

The issue in the above theorem, as well as in the next, is the difficulty to get an explicit formulae of the index for some classes of graphs.

Theorem 12 ([10]) *Over all connected graphs on $n \geq 4$ vertices and $m \geq n$ edges with girth g and index λ_1, $g + \lambda_1$ is maximum for the kite $Ki_{n,3}$ (see Figure 7 for $Ki_{9,3}$). Moreover, for each $t > 0$, there exists an integer n_t such that for all $n \geq n_t$,*

$$3 + \sqrt{5} - t < g(Ki_{n,3}) + \lambda_1(Ki_{n,3}) < 3 + \sqrt{5}.$$

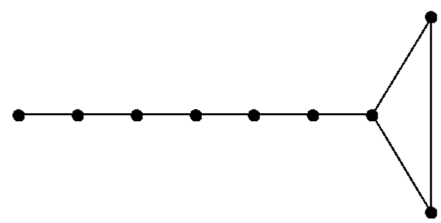

Figure 7: $Ki_{9,3}$**: an extremal graph in Theorem 12**

A study similar to that of [3] was done in [46] where the signless Laplacian spectral radius q_1 is compared to 19 other graph invariants. The results, to which belongs Conjecture 5, are summarized in [8].

A first generalization of AGX Form 1 to AGX Form 2 was introduced in [62]:

$$\underline{b}(m) \leq i_1 \oplus i_2 \leq \overline{b}(m) \tag{3}$$

in which the lower and upper bounding functions $\underline{b}(m)$ and $\overline{b}(m)$ depend on the size m (or number of edges) of the graph instead of its order. Otherwise the symbols have the same meaning and assumptions are the same. Among the AGX Form 2 results, we give the following theorem.

Theorem 13 ([62]) *Let G be a connected graph with size $m \geq 1$, radius r and minimum degree δ. Let k and l be integers such that $m = k(k-1)/2 - l$, where $0 \leq l < k - 1$. Then*

$$\left.\begin{array}{ll} \text{if } l = 0, & 2-k \\ \text{if } 0 < l < k/2, & 3-k \\ \text{if } k/2 \leq l \leq k-1, & 4-k \end{array}\right\} \leq r - \delta \leq \left\lfloor \frac{m-1}{2} \right\rfloor; \tag{4}$$

$$\left.\begin{array}{ll} \text{if } l = 0, & 1/(1-k) \\ \text{if } 0 < l < k/2, & 1/(2-k) \\ \text{if } k/2 \leq l \leq k-1, & 1/(3-k) \end{array}\right\} \leq \frac{r}{\delta} \leq \left\lfloor \frac{m+1}{2} \right\rfloor; \tag{5}$$

$$2 \leq r + \delta \leq \left\lfloor \frac{m}{2} \right\rfloor + 2; \tag{6}$$

$$1 \leq r + \delta \leq \left\lfloor \frac{m}{2} \right\rfloor. \tag{7}$$

The lower bounds for (4) and (5) are attained by the complete graph K_k if $l = 0$, by $K_k \setminus M_l$, where M_l is a matching containing l edges, if $0 < l \leq k/2$ and by $K_k \setminus C_l$, where C_l is a cycle containing l edges, if $k/2 < l < k - 1$. The lower bounds for (6) and (7) are attained by the star S_{m+1}. The upper bounds for (4) and (5) are attained by the path P_{m+1}. The upper bounds for (6) and (7) are attained by the cycle C_m.

4.1 Other forms

Besides bounding invariants and bounds of AGX Form 1, several results of different forms were studied using AutoGraphiX. In this section, we report on relations that do not belong to those described in the two previous sections. As a first example, we give relations involving more than two graph invariants, in addition to the order n. Such relationships are rare in the graph theory literature. A second example is a result about one invariant, in which we consider the behavior of the invariant instead of its minimum or maximum values. Other examples are given below and more can be found in [4, 6, 13, 26, 32, 33, 34, 35, 36, 48].

Any tree is a bipartite graph and therefore its vertex set can be partitioned into two independent subsets. Let a be the number of vertices in one subset and b in the other. In this case, we speak about an (a, b)–partition. Assume, without loss of generality that $a \geq b$ and let $\mathcal{T}_{a,b}$ be the class of all trees that can be partitioned into an (a, b)–partition. In [36], the authors considered the problem of finding extremal trees $T \in \mathcal{T}_{a,b}$ with respect to the adjacency index $\lambda_1(T)$, *i.e.*, solving the problems $\min\{\lambda_1(T) : T \in \mathcal{T}_{a,b}\}$ and $\max\{\lambda_1(T) : T \in \mathcal{T}_{a,b}\}$ for given a and b. Among their results, we recall the following two theorems and conjecture.

Theorem 14 ([22, 36]) *For fixed order $n = a + b$ and for $T \in \mathcal{T}_{a,b}$, the minimal value of $\lambda_1(T)$ increases monotonously with $a - b$.*

Conjecture 8 ([22, 36]) *A vertex from the subset with a vertices in a minimal tree over the class $\mathcal{T}_{a,b}$, with respect to λ_1, has degree 1 or 2.*

For the statement of the next conjecture, we need the following definition. A *comet* $Co_{n,\Delta}$ is the tree obtained from a star S_Δ by inserting $n - \Delta$ vertices (of degree 2) into the same edge.

Theorem 15 ([22, 36]) *For $a = b + 2$ and $n = a + b \geq 6$, trees $T^* \in \mathcal{T}_{a,b}$ with minimal λ_1 are comets $Co_{n,4}$. Moreover*

$$\lim_{n\to+\infty} \lambda_1(T^*) = 2.$$

In [26], after experiments using AutoGraphiX on trees in $\mathcal{T}_{a,b}$ with fixed a and b, the authors obtained the following unexpected conjecture involving five invariants.

Conjecture 9 ([22, 26]) *For fixed integers a and b, let $T \in \mathcal{T}_{a,b}$ with size m, independence number α, diameter D, radius r and n_1 pendent edges. Then*

$$2\alpha - m - n_1 + 2r - D = 0.$$

The above conjecture is not valid for the class of trees in general. Experiments done in [22, 26] with AutoGraphiX led to the following theorem, first obtained as a conjecture.

Theorem 16 ([22, 26]) *Let T be a tree on n vertices and m edges with independence number α, diameter D, radius r and n_1 pendent edges. Then*

$$m + n_1 + D - 2r - \left\lfloor \frac{n-2}{2} \right\rfloor \leq 2\alpha \leq m + n_1 + D - 2r$$

In 1956, Nordhaus and Gaddum [57] proved that

$$2\sqrt{n} \leq \chi(G) + \chi(\bar{G}) \leq n + 1 \text{ and } n \leq \chi(G) \cdot \chi(\bar{G}) \leq \frac{(n+1)^2}{4},$$

where χ is the chromatic number of a graph. Finck [40] showed that these bounds were sharp (taking floors and ceilings if necessary) and characterized extremal graphs. Similar bounds, there after called Nordhaus-Gaddum relations, were obtained for a large number of graph invariants by a variety of authors. For an extensive survey of such relations see [12] and over 350 references therein. Here, we are interested in Nordhaus–Gaddum relations only for the index. Nosal [58] and Amin and Hakimi [2] independently proved that

$$n - 1 \leq \lambda_1(G) + \lambda_1(\bar{G}) \leq \sqrt{2}(n - 1).$$

The best upper bound known up to now is proved by Csikvári [31] in 2009:

$$\lambda_1(G) + \lambda_1(\overline{G}) \leq \frac{1 + \sqrt{3}}{2} n - 1.$$

The problem of finding an upper bound for the index of the Nordhaus–Gaddum type was studied using AGX [4, 6]. The AutoGraphiX conjecture about the upper bound is as follows.

Conjecture 10 ([4, 12]) *For any simple graph G, with complement $\overline{G}$, index $\lambda_1(G)$ and n vertices we have*

$$\lambda_1(G) + \lambda_1(\overline{G}) \leq \frac{4}{3}n - \frac{5}{3} - \begin{cases} f_1(n) & if\ n \equiv 1[3] \\ 0 & if\ n \equiv 2[3] \\ f_2(n) & if\ n \equiv 0[3] \end{cases}$$

where $f_1(n) = 3n - 2 - \sqrt{9n^2 - 12n + 12}/6$ *and* $f_2(n) = 3n - 1 - \sqrt{9n^2 - 6n + 9}/6$.

This bound is sharp and attained if and only if G or $\overline{G}$ is a complete split graph with an independent set on $\lfloor n/3 \rfloor$ vertices (and also on $\lceil n/3 \rceil$ vertices if n mod$(3) = 2$).

The process that led to the above conjecture is described in [4, 6], where some partial results can be found.

The problem of finding Nordhaus–Gaddum inequalities was also considered with AutoGraphiX for the two other invariants. The *transmission* $t(v)$ of a vertex v in a connected graph G, is the sum of the distances from v to all other vertices in G. It is said to be *normalized*, and then denoted $\tilde{t}(v)$, when divided by $n-1$. The *proximity* $\pi = \pi(G)$ and *remoteness* $\rho = \rho(G)$ [3, 7] of G are, respectively, the minimum and the maximum normalized transmission in G. That is

$$\pi = \min_{v \in V} \tilde{t}(v) \qquad \text{and} \qquad \rho = \max_{v \in V} \tilde{t}(v).$$

Some properties of proximity and remoteness are studied in [3, 7, 15, 11, 61]. In [13], the authors derived and proved Nordhaus–Gaddum type inequalities for π and for ρ. The results are stated below.

Theorem 17 ([13]) *For any connected graph G on $n \geq 5$ vertices for which $\overline{G}$ is connected*

$$\frac{2n}{n-1} \leq \pi + \overline{\pi} \leq \begin{cases} \frac{n+1}{4} + \frac{n+1}{n-1} & \text{if } n \text{ is odd,} \\ \frac{n}{4} + \frac{n}{4(n-1)} + \frac{n+1}{n-1} & \text{if } n \text{ is even.} \end{cases}$$

The lower bound is attained if and only if $\Delta(G) = \Delta(\overline{G}) = n - 2$. The upper bound is attained if and only if either G or $\overline{G}$ is the cycle C_n;

$$\frac{n^2}{(n-1)^2} \leq \pi \cdot \overline{\pi} \leq \begin{cases} \frac{(n+1)^2}{4(n-1)} & \text{if } n \text{ is odd,} \\ \frac{n(n+1)}{4(n-1)} + \frac{n(n+1)}{4(n-1)^2} & \text{if } n \text{ is even.} \end{cases}$$

The lower bound is attained if and only if $\Delta(G) = \Delta(\overline{G}) = n - 2$. The upper bound is attained if and only if either G or $\overline{G}$ is the cycle C_n.

Similar results involving ρ are also obtained using AGX and proved in [13]

5 Conclusions

- Development and use of AGX by many researchers showed that this system can generate many conjectures in graph theory which range from the obvious to the very difficult.
- Mostly relations of AGX form 1 have been studied up to now. There are many other possibilities as shown in the survey [45]. The AGX Form 2 uses the size m in stead of n in AGX Form 1 (see [62]).
- Forbiden subgraphs characterizations of graphs are used in a way similar to AGX [38].

References

[1] M. O. Alberston, The irregularity of a graph, Ars Combinatoria, 46 (1997) 215–225.

[2] A. T. Amin and S. L. Hakimi, Upper bounds on the order of a clique of a graph. SIAM J. Appl. Math. 22 (1972) 569–573.

[3] M. Aouchiche. Comparaison automatisée d'invariants en théorie des graphes. PhD Thesis, École Polytechnique de Montréal, February 2006.

[4] M. Aouchiche, F. K. Bell, D. Cvetković, P. Hansen, P. Rowlinson, S. Simić and D. Stevanović, Variable neighborhood search for extremal graphs. 16. Some conjectures related to the largest eigenvalue of a graph. European J. Oper. Res. 191 (2008) 661–676.

[5] M. Aouchiche, J.-M. Bonnefoy, A. Fidahoussen, G. Caporossi, P. Hansen, L. Hiesse, J. Lacheré and A. Monhait. Variable neighborhood search for extremal graphs. 14. The AutoGraphiX 2 system. In L. Liberti and N. Maculan (editors), Global Optimization: From Theory to Implementation, Springer (2006) 281–310.

[6] M. Aouchiche, G. Caporossi and P. Hansen, Open problems on graph eigenvalues studied with AutoGraphiX. EURO J Comput Optim, 1 (2013) 181–199.

[7] M. Aouchiche, G. Caporossi and P. Hansen, Variable neighborhood search for extremal graphs 20. Automated comparison of graph invariants. Match Commun. Math. Comput. Chem. 58 (2007) 365–384.

[8] Aouchiche M, Caporossi G, Hansen P, Lucas C (2013) Variable Neighborhood Search for Extremal Graphs. 28: AutoGraphiX After Fifteen Years Les Cahiers du GERAD, G-2013-12.

[9] M. Aouchiche, O. Favaron and P. Hansen, Variable neighborhood search for extremal graphs. 22. Extending bounds for independence to upper irredundance. Discrete Appl. Math. 157 (2009) 3497–3510.

[10] M. Aouchiche, O. Favaron and P. Hansen, Recherche à voisinage variable de graphes extrémaux. 26. Nouveaux résultats sur la maille. (French) RAIRO Oper. Res. 43 (2009) 339–358.

[11] M. Aouchiche and P. Hansen, Proximity, remoteness and girth in graphs. Submitted.

[12] M. Aouchiche and P. Hansen, A survey of Nordhaus–Gaddum type relations. Discrete Appl. Math. 161 (2013) 466-546

[13] M. Aouchiche and P. Hansen, Nordhaus–Gaddum relations for proximity and remoteness in graphs. Computers & Mathematics with Applications 59 (2010) 2827–2835.

[14] M. Aouchiche and P. Hansen, A survey of automated conjectures in spectral graph theory. Linear Algebra Appl. 432 (2010) 2293–2322.

[15] M. Aouchiche and P. Hansen, Proximity and remoteness in graphs: results and conjectures. NETWORKS 58 (2011) 95–102.

[16] M. Aouchiche and P. Hansen, Bounding average distance using order and Minimum degree. Graph Theory Notes of New-York 56 (2009) 21–29.

[17] M. Aouchiche, P. Hansen and D. Stevanović, Variable neighborhood search for extremal graphs 17. Further conjectures and results about the index. Discuss. Math. Graph Theory 29 (2009) 15–37.

[18] S. Bekkai and M. Kouider, On mean distance and girth. Discrete Appl. Math. 158 (2010) 1888–1893.

[19] F.K. Bell, A note on the irregularity of graphs, Linear Algebra Appl. 161 (1992) 45–54.

[20] S. Belhaiza, N. M. M. Abreu, P. Hansen and C. S. Oliveira, Variable neighborhood search for extremal graphs. 11. Bounds on algebraic connectivity. In Graph Theory and Combinatorial Optimization, D. Avis, A. Hertz and O. Marcotte (eds.), Springer, 1–16, 2005.

[21] E. K. Burke and G. Kendall, Search methodologies. Introductory tutorials in optimization and decision support techniques. Springer, Berlin, 2005.

[22] G. Caporossi, Découverte par Ordinateur en Théorie des Graphes. PhD Thesis, École Polythechnique de Montréal, 2000.

[23] G. Caporossi, D. Cvetković, I. Gutman and P. Hansen, Variable neighborhood search for extremal graphs 2. Finding graphs with extremal energy. J. Chem. Inf. Comput. Sci. 39 (1999) 984–996.

[24] G. Caporossi and P. Hansen, A learning optimization algorithm in Graph Theory. Versatile Search for extremal graphs using a learning algorithm. Lecture Notes in Computer Science 7219 (2012) 16–30.

[25] G. Caporossi and P. Hansen, Finding relations in polynomial time. In Proceedings of the XVIth International Joint Conference on Artificial Intelligence, Stockholm 2 (1999) 580–585.

[26] G. Caporossi and P. Hansen. Variable neighborhood search for extremal graphs: V. Three ways to automate finding conjectures. Disc. Math, 276 (2004) 81–94.

[27] G. Caporossi and P. Hansen, Variable neighborhood search for extremal graphs: I. The AutoGraphiX system. Disc. Math. 212 (2000) 29–44.

[28] J. Christophe, S. Dewez, J.-P. Doignon, G. Fasbender, P. Grégoire, D. Huygens, M. Labbé, S. Elloumi, H. Mélot and H. Yaman, Linear inequalities among graph invariants: using GraPHedron to uncover optimal relationships. Networks 52 (2008) 287–298.

[29] F. Chung. The Average distance is not more than the independence number. J. Graph Theory, 12 (1988) 229–235.

[30] L. Collatz, U. Sinogowitz, Spektren endlicher Grafen. Abh. Math. Sem. Univ. Hamburg, 21(1957) 63–77.

[31] P. Csikvári. On a conjecture of V. Nikiforov. Discrete Math., 309 (2009) 4522–4526.

[32] D. Cvetković, P. Rowlinson and S. Simić, Eigenvalue bounds for the signless Laplacian. Publ. Inst. Math. (Beograd), 81(95) (2007) 11–27.

[33] D. Cvetković and S. Simić, Towards a spectral theory of graphs based on the signless Laplacian, I. Publ. Inst. Math. (Beograd) 85(99) (2009) 19–33.

[34] D. Cvetković and S. K. Simić, Towards a spectral theory of graphs based on the signless laplacian. II. Linear Algebra Appl. 432 (2010) 2257–2272.

[35] D. Cvetković and S. K. Simić, Towards a spectral theory of graphs based on the signless Laplacian. III. Appl. Anal. Discrete Math. 4 (2010) 156–166.

[36] D. Cvetković, S. Simić, G. Caporossi and P. Hansen, Variable neighborhood search for extremal graphs. 3. On the largest eigenvalue of color-constrained trees. Linear and Multilinear Algebra 49 (2001) 143–160.

[37] K.C. Das, Conjectures on index and algebraic connectivity of graphs. Linear Algebra Appl. 433 (2010) 1666–1673.

[38] C Desrosiers, P. Galinier, P. Hansen, A. Hertz, Automated generation of conjectures on forbidden subgraph characterization. Discrete Appl. Math. 162 (2014) 177–194.

[39] M. Fiedler, Algebraic connectivity of graphs. Czechoslovak Math. J. 23 (1973) 298–305.

[40] H.-J. Finck, On the chromatic numbers of a graph and its complement, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 99–113.

[41] F. Glover and G. Kochenberger, (Eds.), Handbook of metaheuristics. Kluwer, Amsterdam, 2003.

[42] A.D. Gregory, D. Hershkovitz, S.J. Kirkland, The spread of the spectrum of a graph. Linear Algebra Appl. 332-334 (2001) 23–35.

[43] I. Gutman, P. Hansen and H. Mélot, Variable neighborhood search for extremal graphs. 10. Comparison of irregularity indices for chemical trees. Journal of Chemical Information and Modeling 45 (2005) 222–230.

[44] I. Gutman, The energy of a graph. Ber. Math.–Statist. Sekt. Forsch–ungsz. Graz 103 (1978) 1–22.

[45] P. Hansen, M. Aouchiche, G. Caporossi, H. Mélot and D. Stevanović, What Forms Do Interesting Conjectures Have in Graph Theory? In Graphs and Discovery, S. Fajtlowicz, P.W. Fowler, P. Hansen, M.F. Janowitz, F.S. Roberts (eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 69, AMS, Providence, 231–251, 2005.

[46] P. Hansen and C. Lucas, Bounds and conjectures for the signless Laplacian index of graphs. Linear Algebra Appl. 432 (2010) 3319–3336.

[47] P. Hansen and C. Lucas, An inequality for the signless Laplacian index of a graph using the chromatic number. Graph Theory Notes of New-York 57 (2009) 39–42.

[48] P. Hansen and H. Mélot, Variable neighborhood search for extremal graphs 9. Bounding the irregularity of a graph. In Graphs and Discovery, S. Fajtlowicz, P.W. Fowler, P. Hansen, M.F. Janowitz, F.S. Roberts (eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 69, AMS, Providence, 253–264, 2005.

[49] P. Hansen and N. Mladenović, Variable neighborhood search. In [41], 145–184.

[50] P. Hansen and N. Mladenović, Variable neighborhood search: principles and applications. European Journal of Operational Research, 130 (2001) 449–467.

[51] P. Hansen and N. Mladenović, Developments of variable neighborhood search. In C. Ribeiro and P. Hansen (Eds.), Essays, surveys in metaheuristics. Kluwer, Amsterdam, 2001, 415–440.

[52] P. Hansen and N. Mladenović, An introduction to variable neighborhood search. In S. Voss *et al.* (Eds.), Metaheuristics, advances, trends in local search paradigms for optimization. Kluwer, Amsterdam, 1999, 433–458.

[53] P. Hansen, N. Mladenović and J. A. Moreno Pérez, Variable neighborhood search: methods and applications. Ann. Oper. Res. 175 (2010) 367–407.

[54] S. C. D. Kirkpatrick, Jr. Gellatt and P. Vecchi, Optimization by simulated annealing. Science 220 (1983) 671–680.

[55] M. Kouider and P. Winkler, Mean Distance and Minimum Degree. J. Graph Theory 25 (1997) 95–99.

[56] N. Mladenović and P. Hansen, Variable neighborhood search. Computers and Operations Research, 24 (1997) 1097–1100.

[57] E. A. Nordhaus and J. W. Gaddum. On Complementary Graphs. Amer. Math. Monthly, 63 (1956) 175–177.

[58] E. Nosal, Eigenvalues of Graphs, Master Thesis, University of Calgary, 1970.

[59] M. Randić, On Characterization of Molecular Branching. J. Am. Chem. Soc. 97 (1975), 6609–6615

[60] C. R. Reeves, (Ed.), Modern heuristic techniques for combinatorial problems. Blackwell Scientific, Oxford, 1993.

[61] J. Sedlar, Remoteness, proximity and few other distance invariants in graphs. Filomat 27:8 (2013) 1425–1435.

[62] J. Sedlar, D. Vukicević and P. Hansen, Using size for bounding expressions of graph invariants. Ann. Oper. Res. 188 (2011) 415–427.

[63] D. Stevanović, Resolution of AutoGraphiX conjectures relating the index and matching number of graphs. Linear Algebra Appl. 433 (2010) 1674–1677.

[64] H. S. Wilf, The eigenvalues of a graph and its chromatic number. J. London Math. Soc. 42 (1967) 330–332.

[65] M.–L. Ye, Y.–Z. Fan and H.–F. Wang, Maximizing signless Laplacian or adjacency spectral radius of graphs subject to fixed connectivity. Linear Algebra Appl. 433 (2010) 1180–1186.