Les Cahiers du GERAD ISSN: 0711–2440

Variable Neighborhood Search for Extremal Graphs. 27. Families of Extremal Graphs

M. Aouchiche, G. Caporossi,
P. Hansen

G–2007–87

November 2007

Les textes publiés dans la série des rapports de recherche HEC n'engagent que la responsabilité de leurs auteurs. La publication de ces rapports de recherche bénéficie d'une subvention du Fonds québécois de la recherche sur la nature et les technologies.

Variable Neighborhood Search for Extremal Graphs. 27. Families of Extremal Graphs

Mustapha Aouchiche
Gilles Caporossi
Pierre Hansen

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

mustapha.aouchiche@gerad.ca
gilles.caporossi@gerad.ca
pierre.hansen@gerad.ca

November 2007

Les Cahiers du GERAD

G–2007–87

Copyright © 2007 GERAD

Abstract

The AutoGraphiX system (AGX 1 and AGX 2) for interactive, and for several functions automated, graph theory, discovers conjectures of algebraic or structural form. In this paper, we focus on the later, *i.e.* families of extremal graphs for a series of relations between pairs of graph invariants chosen among 20 selected ones. The form of these relations generalizes that one of Nordhaus-Gaddum relations. There are 1520 cases, leading to 47 families of extremal graphs. They include many classical families but also some apparently new ones (bags, bugs, ...). Five ways to exploit these families of extremal graphs in order to enhance the performance of AutoGraphiX are studied and illustrated by examples.

Key Words: AutoGraphiX, AGX, Graph invariant, Extremal graph, Conjecture, VNS, Heuristic.

Résumé

Le système AutoGraphiX (AGX 1 et AGX 2) pour la théorie des graphes, interactif et automatisé pour plusieurs de ses fonctions, découvre des conjectures algébriques ou structurales. Dans cet article, on s'intéresse à ces dernières, *c-à-d*, aux familles de graphes extrêmes pour une série de relations entre paires d'invariants graphiques choisis parmi une sélection de 20 invariants. La forme de ces relations généralise celle des relations de Nordhaus-Gaddum. Il y a 1520 cas, menant à 47 familles de graphes extrêmes. On y trouve plusieurs familles classiques mais aussi quelques nouvelles familles (paniers, bestioles, ...). Cinq façons d'exploiter ces familles de graphes extrêmes dans le but d'accroître les performances d'AutoGraphiX sont étudiées et illustrées par des exemples.

Mots clés : AutoGraphiX, AGX, invariant graphique, graphe extrême, conjecture, VNS, heuristique.

1 Introduction

In the last section of a survey on Variable Neighborhood Search (VNS) [20], a list of desirable properties of metaheuristics is provided, spurred by a stimulating correspondence with Fred Glover. The seventh and last property listed is Innovation: the metaheuristic should suggest new types of applications, not merely new ways to solve well-known problems. The possibility of such innovation in the domain of graph theory became clear in 1997, very soon after the VNS metaheuristic was initiated [25]. Metaheuristics have long been used to solve a variety of optimization problems on graphs *e.g.* find the clique or chromatic numbers, maximum flows or multiflows and many others. Operations research problems defined on graphs have also been extensively solved using various metaheuristics. However, applications of metaheuristics to solve problems in graph theory itself, i.e. find conjectures, refutations and proofs (or ideas of proofs) in that field had not yet been explored (although man-machine theorem-proving and *a priori* generation of conjectures in graph theory to be tested on potential counter-examples had attracted some attention, see e.g. [17, 18]).

The basic idea to do so was to consider extremal problems in graph theory as possibly parametric mathematical programming problems of the form

$$\min_{G \in \mathcal{G}_n} f(G) \quad \text{or} \quad \max_{G \in \mathcal{G}_n} f(G)$$

where $f(G)$ is some formula depending on one or more invariants of G (or quantities which do not depend on the labeling of the vertices or edges of G), and $\mathcal{G}_n$ is the set of all graphs on n vertices (the sets of all graphs with m edges, or with n vertices and m edges, and others, can be considered also). Then these problems are solved by a *generic* heuristic fitting in the VNS framework. Five types of problems were considered in order to evaluate the potential of VNS in graph theory:

(a) *Find a graph satisfying given constraints.* Let $i_1(G), i_2(G), \dots i_l(G)$ denote l invariants of G and $p_1, p_2, \dots p_l$ values given to them. Consider the problem

$$\min_{G \in \mathcal{G}_n} f(G) = \sum_{k=1}^{l} |i_k(G) - p_k|.$$

Any graph such that $f(G) = 0$ satisfies these constraints. Note that constraints involving formulae on several invariants can be treated in a similar way. Note also that using min and max operators, lower and upper bounds on individual invariants as well as formulae on several invariants can be treated in a similar way.

(b) *Find optimal or near-optimal values for an invariant possibly subject to constraints.* Let $i_0(G)$ denote the objective function invariant and assume constraints expressed as above. Consider the problem

$$\min_{G \in \mathcal{G}_n} f(G) = i_0(G) + M \sum_{k=1}^{k=l} |i_k(G) - p_k|$$

where M is a constant sufficiently large to ensure that for any pair of graphs $G, G' \in \mathcal{G}_n$ such that $\sum_{k=1}^{k=l} |i_k(G) - p_k| = 0$ and $\sum_{k=1}^{k=l} |i_k(G') - p_k| > 0$, $f(G) < f(G')$. Maximum values are tackled in a similar way.

(*c*) *Refute a conjecture*: Consider a conjecture $h(G) \leq g(G)$ where $h(G)$ and $g(G)$ are formulae depending on one or more invariants of G. Then solve the problem

$$\min_{G \in \mathcal{G}_n} f(G) = g(G) - h(G).$$

If a graph G for which $f(G) < 0$ is found, the conjecture is refuted.

(*d*) *Suggest a conjecture (or sharpen one, or repair a refuted one)*: This can be done in various ways, which usually use parameterization on n or other invariants $i_1(G)$, $i_2(G), \ldots i_l(G)$. For instance, consider

$$\min_{G \in \mathcal{G}_n} f(G) = i_2(G) - i_1(G).$$

If no graph G with $f(G) < 0$ is found this suggests $i_1(G) \leq i_2(G)$. If the extremal graphs found belong to a recognizable class, it may be possible to deduce a more precise inequality in $i_1(G)$, $i2(G)$ and n. This can be done in several ways, see below.

(*e*) *Suggest a proof*: Depending on the way the extremal graphs are obtained, *i.e.*, what transformations of G are used, may suggest ways to prove conjectures by showing that all graphs admit such transformations which improve the objective function value.

As explained below, automated proofs of many simple relations between graph invariants can also be found by exploiting the fact that certain families of graphs are simultaneously extremal for several invariants.

These problems were addressed with two successive versions of the system AutoGraphiX (AGX): AXG 1 [12, 13] and AGX 2 [3].

The generic heuristic, central to AGX 1 and AGX 2, fits in the standard VNS framework [20, 25]. It first uses a variable neighborhood descent component to find a local optimum; to that effect a series of graph transformations, *e.g.* rotation, addition or removal of an edge, as well as more complex moves such as detour, short-cut, 2–opt move and the like are used in AGX 1 [12]; systematic consideration of all moves on subgraphs with up to four vertices, together with reinforcement of beneficial ones is performed in AGX 2 [3]. Once a local optimum is attained a series of nested neighborhoods is defined by removal, addition or move of $k = 1, 2, 3, \ldots$ edges. Following the VNS framework, a graph is drawn at random in the first neighborhood and a descent performed from there; the search is recentered around the new local optimum attained if it is better than the current best graph and otherwise a point is drawn from the next neighborhood (see [3, 12] for details). Over thirty papers have been written since 1998 on AGX 1, AGX 2 and their applications. Most of them, including the present one, belong to a series with the common title "Variable Neighborhood Search for Extremal Graphs".

When studying conjectures, once AGX has generated a series of (presumably) extremal graphs for the objective under study, the invariant values obtained are analyzed following one or several of three ways [13]:

(a) a *numerical method*, which exploits the mathematics of principal component analysis to find a basis of affine relations between graph invariants;
(b) a *geometric method*, which determines the convex hull of extremal graphs viewed as points in invariant space; facets of this convex hull correspond to inequalities between invariants;
(c) an *algebraic method*, which first recognizes the family of extremal graphs then substitutes to the invariants in the objective, their expression in function of one parameter (or more), usually the order n of the graph considered.

The first two approaches gives algebraic conjectures in the form of equality or inequality relations among invariants respectively. Note that these relations may be nonlinear if well-chosen expressions of invariants (e.g. square, square root, product ...) are added as new invariants. The third approach first gives a structural conjecture on the family of extremal graphs, then an algebraic conjecture after substitution of invariants and simplification.

All three methods can be completely automated. However, this implies, for the third one a database of parametric expressions for invariants on relevant families of graphs. If such a database is not available, these expressions can be computed by hand, usually in a straightforward way. References can be found in a short recent survey paper [5].

Recognizing families of graphs, and generating them (possibly in function of some parameter(s)) are tasks required in various places of the AGX system, for several purposes. This is the subject matter of the present paper. Recognition of extremal graphs and five ways to use them are described in the next section and illustrated by examples. They bear upon initialization of the optimization, automated proof of simple results, algebraic generation of conjectures, refutation of conjectures and reparation of conjectures. Moreover, some new families of graphs, which appear to be of interest, are presented. In Section 3, an atlas of families of extremal graphs, generated in a large scale experimental study, is provided. Brief conclusions are given in Section 4. In the appendix a table is given specifying for each family of graphs considered the list of invariants for which it is extremal.

Recent applications of AGX focus upon relations of the form

$$\underline{b}_n(G) \quad \leq \quad i_1(G) \oplus i_2(G) \quad \leq \quad \overline{b}_n(G) \qquad (1)$$

where $G = (V, E)$ is a connected graph with vertex set V and edge set E, $i_1(G)$ and $i_2(G)$ are graph invariants, $\oplus$ denotes one of the four operations $+$, $-$, $\times$ and $/$, $f_1(G)$ and $f_2(G)$ are functions of $n = |V|$ which are best possible, *i.e.*, such that for each n (except possibly for very small values for which there are border effects) there exist a graph G' (respectively G") such that the lower (respectively upper) bound is attained. We call this form AGX Form 1.

Observe that it generalizes the well-known form of Nordhaus-Gaddum relations [26] in two ways: first, different invariants $i_1(G)$ and $i_2(G)$ are considered instead of the same invariant in the graph G and its complement $\bar{G}$; second, in addition to the operations $+$ and $\times$, one considers $-$ and $/$.

AGX Form 1 has several advantages:

(i) the relations obtained are sufficiently simple to be of interest;

(ii) the easiest of them can be proved automatically, either by observing that the relevant families of extremal graphs for $i_1(G)$ and $i_2(G)$ have non-empty intersection, or by simple algebraic manipulations;

(iii) they comprise some well-known results which are automatically rediscovered *e.g.* the theorem, $\bar{l}(G) \leq \alpha(G)$ where $\bar{l}(G)$ denotes average distance between pairs of vertices of G and $\alpha(G)$ the stability number, conjectured in WOW 2 [29] and proved by Chung [14];

(iv) they improve upon some well-known conjectures *e.g.* WOW 127 [29], $\delta(G)\bar{l}(G) \leq n$ where $\delta(G)$ denotes the minimum degree of G, now strengthened to $\delta(G)\bar{l}(G) \leq n-1$ which is best possible and follows from a result of Beezer, Riegsecker and Smith [10];

(v) the difficulty of proving these relations goes all the way from the very easy to the difficult and some apparently hard conjectures remain open.

In the thesis [1], AGX Form 1 has been studied for a series of 20 graph invariants. They are defined in Table 1 and listed in Table 2 together with lower and upper bound on their values as functions of n and the corresponding extremal graphs. This gives rise to 1520 cases. Results are presented in the 210 pages long Appendix to [1]. Quantitative results, *i.e.*, number of cases in which a bound is found and proved automatically, or proved by hand, or found interactively, etc. …is briefly presented in [4]. This paper complement this last one by presenting structural results, *i.e.*, results one the families of extremal graphs obtained.

2 Extremal Graphs in AGX

AGX 1 and AGX 2 do not only find extremal graphs, but also takes advantage of them in various ways, including the five discussed in this section.

2.1 Initialization of the optimization

A study on the optimization process sensitivity, in AGX 1 and AGX 2, showed that, in many cases, the choice of the initial graph influences considerably the results obtained. This phenomenon appears, mainly, when the invariants considered in the objective functions are integers, such as the minimum and maximum degree, the radius, the diameter and the chromatic number. When the variation of some or all of the invariants under study is sensitive to any change in the graph, as in the case of average degree, average distance and index, the choice of the initial graph influences less the optimization process.

Among strategies adopted to circumvent this difficulty, the two following ones exploit the extremal graphs:

- The first strategy is interactive. The user chooses a graph among those defined in AGX as an initial graph in the optimization process.
- The second strategy is automated. The system evaluates the objective function for each graph in the list of extremal graphs available in AGX and selects as an initial graph the best one.

A graph is defined in AutoGraphiX, if a routine to generate it is implemented. Presently, about forty families, most of which are known or defined in Table 6, can be generated. Some of them, such as a path, a star, a cycle and a complete graph, use one parameter

Table 1: Definitions of the 20 selected invariants

Δ	The maximum degree. The degree of a vertex is the number of its neighbors.
δ	The minimum degree.
$\overline{d}$	The average degree.
$\overline{l}$	The average distance between all pairs of vertices. The distance between two vertices is the length of a shortest path between them.
D	The diameter, the largest distance between pairs of vertices.
r	The radius, the minimum eccentricity. The eccentricity of a vertex is the maximum distance from this vertex to another one.
g	The girth, the length of the smallest cycle in a graph.
ecc	The average eccentricity.
π	The proximity is the minimum normalized transmission. The transmission of a vertex in a graph, is the sum of all distances between the vertex and other ones. It is said to be normalized if divided by $n-1$, where n is the order of the graph.
ρ	The remoteness is the maximum normalized transmission.
λ_1	The index (the spectral radius) is the largest eigenvalue of the adjacency matrix of a graph.
R	The Randić index is defined by $R = \sum 1/\sqrt{d_i d_j}$ where the summation is on the set of all edges ij and d_i denotes the degree of the vertex i.
a	The algebraic connectivity, the second smallest eigenvalue of the Laplacian matrix L, that is defined by: $L_{ii} = -d_i$ and $L_{ij} = a_{ij}(i \neq j)$ where d_i is the degree of the vertex i and $A = (a_{ij})$ is the adjacency matrix of the graph.
ν	The vertex connectivity, minimum number of vertices of a connected graph the removal of which disconnects the graph or reduces it to a single vertex.
κ	The edge connectivity, minimum number of vertices of a connected graph the removal of which disconnects the graph.
α	The independence number, maximum cardinality of an independent set, *i.e.*, a set of pairwise non adjacent vertices.
β	The domination number, minimum cardinality of a dominant set, *i.e.*, a set of vertices such that each vertex is in the set or adjacent to a vertex of the set.
ω	The number of vertices in the largest clique in the graph.
χ	The chromatic number, minimum number of colors to assign to the vertices such that two adjacent vertices have different colors.
μ	The matching number, the maximum number of disjoint edges.

Table 2: Selected invariants together with lower and upper bounds for all connected graphs with at least 3 vertices

Inv.	Name	Lower bound $\underline{b}(n)$	Extremal graphs for $\underline{b}(n)$	Upper bound $\overline{b}(n)$	Extremal graphs for $\overline{b}(n)$
Δ	Maximum degree	2	P_n, C_n	$n-1$	G with a dominating vertex (K_n, S_n, ...)
δ	Minimum degree	1	G with a pending vertex (Tree, P_n, S_n, ...)	$n-1$	K_n
$\overline{d}$	Average degree	$2-\frac{2}{n}$	Tree (P_n, S_n, ...)	$n-1$	K_n
$\bar{l}$	Average distance	1	K_n	$\frac{n+1}{3}$	P_n
D	Diameter	1	K_n	$n-1$	P_n
r	Radius	1	G with a dominating vertex (K_n, S_n, ...)	$\left\lfloor\frac{n}{2}\right\rfloor$	P_n, C_n, ...
g	Girth	3	G with a triangle (K_n, ...)	n	C_n
ecc	Average eccentricity	1	K_n	$\frac{3n+1}{4}\cdot\frac{n-1}{n}$ if n is odd $\frac{3n-2}{4}$ if n is even	P_n
π	Proximity	1	G with a dominating vertex (K_n, S_n, ...)	$\frac{n+1}{4}$ if n is odd $\frac{n}{4}+\frac{n}{4n-4}$ if n is even	P_n and C_n
ρ	Remoteness	1	K_n	$\frac{n}{2}$	P_n
λ_1	Index	$2cos\frac{\pi}{n+1}$	P_n	$n-1$	K_n
R	Randić index	$\sqrt{n-1}$	S_n	$\frac{n}{2}$	G regular (K_n, C_n, ...)
a	Algebraic connectivity	$2-2cos\frac{\pi}{n}$	P_n	n	K_n
ν	Node connectivity	1	G with a cut vertex	$n-1$	K_n
κ	Edge connectivity	1	G with a cut edge	$n-1$	K_n
α	Independence number	1	K_n	$n-1$	S_n
β	Domination number	1	G with a dominating vertex (K_n, S_n, ...)	$\left\lfloor\frac{n}{2}\right\rfloor$	$K_{\lceil\frac{n}{2}\rceil}$ + $\left\lfloor\frac{n}{2}\right\rfloor$ disjoint pending edges
ω	Clique number	2	C_n, P_n, Tree, ...	n	K_n
χ	Chromatic number	2	G bipartite, (Tree, P_n, ...)	n	K_n
μ	Matching number	1	S_n	$\left\lfloor\frac{n}{2}\right\rfloor$	K_n, P_n, C_n, ...

(the order); others, such as a complete bipartite, double-star, a comet, a path-complete graph and a bag, use two parameters (order and another invariant; size, maximum clique number, independence number, maximum degree ...); still others, such as a double-comet and a bug, use three parameters. Some routines generate random graphs such as bipartite ones, trees and graphs with given density.

Sorting graphs, by the value they give to the objective function, enables to start the optimization with a graph near to the optimum. Moreover, as the number of observed families of extremal graphs is moderate (at least for AGX Form 1), it is fairly frequent that the selected graph is optimal for the chosen objective.

Example 1 When solving the problem of finding the upper bound and the extremal graphs for the sum of average distance and average degree, $\bar{l} + \bar{d}$, the selected initial graph, for any value of the parameter n, is the complete graph. AGX 2 never improves this initial solution and automatically states the conjecture.

Conjecture 1 *For any connected graph on $n \geq 3$ vertices we have*

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

with equality if and only if the graph is complete.

This is proved in [8].

Example 2 The complete graph was also found to be extremal for the upper bounds on ω/R and χ/R, where ω, χ and R denote the clique number, the chromatic number and the Randić index respectively. The corresponding results are as follows.

Conjecture 2 *For any connected graph G on $n \geq 2$ with chromatic number χ and Randić index R,*

$$\chi \leq 2R$$

with equality if and only if G is the complete graph K_n.

This conjecture is proved in [22]. The proof is five pages long and relies on the following preliminary result: if v is a vertex of minimum degree in a graph G,

$$R(G) - R(G - v) \geq \frac{1}{2}\sqrt{\frac{\delta}{\Delta}}$$

which appears to be of interest in its own right.

Conjecture 3 *For any connected graph G on $n \geq 2$ vertices with clique number ω and Randić index R,*

$$\omega \leq 2R$$

with equality if and only if G is the complete graph K_n.

Since $\omega \leq \chi$ for any graph, this conjecture is a corollary of Conjecture 2.

Example 3 In a similar way, AutoGraphiX found that the star S_n is the extremal graph for the upper bound on R/μ, where R and μ denote the Randić index and the matching number respectively. The corresponding result is the following.

Conjecture 4 *For any connected graph on $n \geq 2$ vertices with Randić index R and matching number μ,*

$$R \leq \mu\sqrt{n-1}$$

with equality if and only if G is the star S_n.

This conjecture is proved in [9].

2.2 Automated proof of simple results

AGX 2 contains a knowledge base of invariants. Each of its items represents a graph invariant and contains the invariant name in the system, its name in the (graph theory) literature, the upper and the lower bounding functions in terms of the order of the graph, and a list of families of extremal graphs relevant to each bound. For example the objects that represent the average degree and the diameter are summarized in Table 3 and Table 4, respectively. The first three rows in each of these tables contain, respectively, the name, the invariant symbol in AGX system and the most used notation of the invariant in the graph theory literature. The fourth row contains, respectively the lower and upper bounds on the invariant (average degree and diameter), functions of the order n of a graph. The last row contains list of extremal graphs for which the (lower and upper) bounds are reached, respectively.

Table 3: The average degree in the knowledge base

The average degree	
AVDEG	
$\overline{d}$	
$2-\frac{2}{n}$.	$n-1$.
Tree: path, star, comet,...	Complete graph, pineapple ...

Table 4: The diameter in the knowledge base

The diameter	
DIAMETER	
D	
1.	$n-1$.
Complete graph	Path

Let $i_1(G)$ and $i_2(G)$ denote two graph invariants, $\oplus$ (resp. $\ominus$) be one of the two operations $+$ and $\times$ (resp.$-$ and $/$). It is easy to see that

(*i*) $Max\ (Min)\ \ i_1(G) \oplus i_2(G) = Max\ (Min)\ \ i_1(G) \oplus Max\ (Min)\ \ i_2(G)$ if and only if both $i_1(G)$ and $i_2(G)$ are maximum (minimum) for a same (extremal) graph.

(*ii*) $Max\ (Min)\ \ i_1(G) \ominus i_2(G) = Max\ (Min)\ \ i_1(G) \ominus Min\ (Max)\ \ i_2(G)$ if and only if both $i_1(G)$ is maximum (minimum) and $i_2(G)$ is minimum (maximum) for a same (extremal) graph.

Using (*i*) and (*ii*) as rules and taking two graph invariants chosen by the user, AGX 2 operates on the knowledge base and tests if the intersection of the respective sets of families of extremal graphs is empty or not. If it is empty, the optimization component is automatically called to find the "extremal" graphs. If the intersection is not empty, AGX 2 generates the corresponding result together with its proof. The result take the form of an algebraic formula, that is a combination of the corresponding bounds, together with the associated extremal graphs.

Example 4 Consider the invariants given in tables 3 and 4. The diameter reaches its minimum for a complete graph with $D = 1$, and its maximum for a path with $D = n - 1$; the average degree reaches its minimum for a tree with $\bar{d} = 2 - \frac{2}{n}$, and its maximum for a complete graph. In this case, AGX 2 applies the rule (*ii*) and generates the following results together with their proofs.

Observation 1 *For all connected graphs with at least 2 vertices with diameter D and average degree $\bar{d}$,*

$$3 - n - \frac{2}{n} \leq \bar{d} - D \leq n - 2.$$

Moreover, the lower (resp. upper) bound is reached if and only if the graph is a path (resp. complete graph).

Observation 2 *For all connected graphs with at least 2 vertices with diameter D and average degree $\bar{d}$,*

$$(2 - \frac{2}{n})/(n - 1) \leq \bar{d}/D \leq n - 1.$$

Moreover, the lower (resp. upper) bound is reached if and only if the graph is a path (resp. a complete graph).

For the pair of chosen invariants (D and $\bar{d}$), the intersections of the sets of families of extremal graphs for the sum and the product are empty. Thus AGX 2 uses its optimization component (based on the VNS heuristic), and after recognizing the extremal graphs generates automatically the following conjectures:

Conjecture 5 *For all connected graphs with at least 2 vertices with diameter D and average degree $\bar{d}$,*

$$4 - \frac{2}{n} \leq \bar{d} + D \leq n + 1 - \frac{2}{n}$$

Moreover, the lower (resp. upper) bound is reached if and only if the graph is a star (resp. a path or a complete graph minus an edge, $K_n - e$).

Conjecture 6 *For all connected graphs with at least $n \geq 2$ vertices with diameter D and average degree $\bar{d}$,*

$$D \cdot \bar{d} \geq 4 - \frac{4}{n}$$

with equality if and only if the graph is a star. Moreover, the upper bound is reached if and only if the graph is a bug.

Both Conjecture 5 and Conjecture 6 are true.

Proof of Conjecture 5

For the lower bound, if $D = 1$ the graph is complete and $D + \bar{d} = n$. If not, we have $D \geq 2$ and $\bar{d} \geq 2 - \frac{2}{n}$. Thus $D + \bar{d} \geq 4 - \frac{2}{n}$ with equality if and only if $D = 2$ and $\bar{d} = 2 - \frac{2}{n}$, *i.e.,* the graph is a star.

For the upper bound, according to Harary [23], if the order n and the size m are fixed, the diameter D is maximum for, among others, a path-complete graph $PK_{n,m}$, composed of a clique a disjoint path and one or more additional edges joining one endpoint of the path each to a distinct vertex of the clique. We denote the diameter of $PK_{n,m}$ by $D_{n,m}$. So for any graph of order n, size m and diameter D, we have

$$D + \bar{d} = D + \frac{2m}{n} \leq D_{n,m} + \frac{2m}{n}.$$

For $PK_{n,m}$, let k denotes the number of edges between the path and the clique. Then

$$\begin{aligned} D + \bar{d} &\leq D_{n,m} + \tfrac{2m}{n} \\ &\leq D_{n,m} + \tfrac{2}{n}(D_{n,m} - 2 + k + \tfrac{1}{2}(n + 1 - D_{n,m})(n - D_{n,m})) \\ &\leq n + 1 + (D_{n,m}^2 - (n-1)D_{n,m} + 2(k-2))/n \end{aligned}$$

This last expression is an increasing function in k, so it reaches its maximum for $k = n - D$. Thus

$$\begin{aligned} D + \bar{d} &\leq n + 1 + \tfrac{1}{n}(D_{n,m}^2 - (n-1)D_{n,m} + 2(n - D_{n,m} - 2)) \\ &\leq n + 1 + \tfrac{1}{n}(D_{n,m}^2 - (n+1)D_{n,m} + 2(n-2)). \end{aligned}$$

This bound, as a function in D, reaches its maximum for $D = 2$ or $D = n - 1$. For both these values, the expression is equal to $n + 1 - \frac{2}{n}$. It is then easy to see that this bound is attained only for two graphs:

- a path P_n which corresponds to the case $D = n - 1$;
- a complete graph minus an edge $K_n - e$ which corresponds to the case $D = 2$. □

Proof of Conjecture 6
The lower bound, as well as the characterization of the associated extremal graphs, on $D + \bar{d}$ is proved exactly like the lower bound of Conjecture 5. The structural part of the conjecture follows from Theorem 4 of [23]. □

2.3 Algebraic generation of conjectures

The first step in generating conjectures with AGX is to find extremal graphs, using the VNS metaheuristic, and then try to find the algebraic relation (formula), between selected invariants, that is satisfied by all graphs obtained. AutoGraphiX can find easily the formula if it is linear, using algebraic properties of Principal Component Analysis [12, 13]. Sometimes we can get a relation even in the frequent case where it is not an affine one, by using some device such as introducing a function of n (square, square root, inverse, ...). Another way to circumvent this difficulty is to use the properties of the extremal graphs, after their recognition by AGX 2. Usually, when a family of extremal graphs is well-defined in function of some parameters, we can compute the relevant expressions for several invariants as functions of n. By substitution of the invariants in the objective function, we can get the algebraic formula. To illustrate, consider the following conjecture.

Conjecture 7 *For all connected graphs with at least $n \geq 3$ vertices with index λ_1 and average degree $\bar{d}$,*

$$1 \leq \lambda_1/\bar{d} \leq \frac{n}{2\sqrt{n-1}}.$$

The lower bound is attained if and only if the graph is regular, and the upper bound if and only if the graph is a star.

The lower bound is obtained automatically by AGX 2; it is true and well-known [16]. The upper bound is obtained interactively. After the optimization, one can easily see that the extremal graphs for the upper bound are stars. Thus, using the fact that for a star of order n, $\lambda_1 = \sqrt{n-1}$ and $\bar{d} = 2 - \frac{2}{n}$, we can get the formula of the upper bound. It is also true.

Proof of Conjecture 7
The lower bound is known [16].
The upper bound: In [24], Hong showed that $\lambda_1 \leq \sqrt{2m-n+1}$, therefore

$$\lambda_1/\bar{d} = n\lambda_1/(2m) \leq n\sqrt{2m-n+1}/(2m) = f(m)$$

Considering the last expression as a function of m we have $\dot{f}(m) = n(n-1-m)/(2m^2\sqrt{2m-n+1}) < 0$ for $m > n-1$. Then $max_{m \geq n-1} f(m) = f(n-1) = n/(2\sqrt{n-1})$ and this maximum is attained if and only if $m = n-1$, *i.e.*, if the graph is a tree. Finally it is easy to see that among all trees $\lambda_1/\bar{d} = n/(2\sqrt{n-1})$ if and only if $\lambda_1 = \sqrt{n-1}$. It is well known [16] that, for any tree T, $\lambda_1(T) \leq \sqrt{n-1}$ with equality if and only if T is the star S_n. Thus the result follows. □

Conjecture 8 *For any connected graph with at least $n \geq 3$ vertices with algebraic connectivity a and clique number ω,*

$$2 - n \leq a - \omega \leq \lfloor (1 - \frac{1}{\lfloor \sqrt{n} \rfloor}) n \rfloor - \lfloor \sqrt{n} \rfloor$$

The lower bound is attained if and only if the graph is a short kite $Kite_{n,n-1}$. The upper bound is attained if and only the graph is a complement of a Turan graph in which (the complement) the size of each clique is almost the same and almost equal to the number of cliques.

The lower bound is obtained automatically while the upper one is obtained by observing the extremal graphs (multipartite graphs) and then using their properties to compute the algebraic expression. This conjecture is proved in [27].

In order to render the algebraic method efficient three tasks have to be addressed:

(*i*) *Recognize* the family or families of extremal graphs to which the extremal graphs belong.

(*ii*) *Substitute* formulae for the value of the invariants in the objective function on those families.

(*iii*) *Simplify.*

The first task is easy and automated in AGX 2 for most families of extremal graphs. Indeed, characterization of these families in terms of graph invariants, which AGX 2 computes very rapidly, are readily available. To illustrate, $D = 1$ characterizes complete graphs; $m = n - 1$ and $D \leq n$ characterizes trees; these last conditions and $r = 1$ characterizes stars; etc. ...

The second task is also easy provided a database giving formulae for invariant values on various families of graphs as functions of n are available. Such a database is partly constructed and used in AGX 2. More precisely, formulae for almost all of the selected invariants have been computed for the best known families of graphs such as complete ones, paths, stars, cycles, etc. ... Similar tables for new families of graphs (see below) have also been obtained. AGX 2 itself can be used for that purpose as it computes very quickly the numerical values of the selected invariants. However, for some invariants, such as the index or the algebraic connectivity, no formulae can be obtained. Table 5 presents formulae for the 20 selected invariants and several families of graphs.

The third task, *i.e.* simplification of algebraic formulae can be done by or with the help of computer systems such as Matlab or Mathematica.

2.4 Refutation of conjectures

As explained in the introduction, a conjecture of the form $h(G) \leq g(G)$ can be refuted by minimizing the function $f(G) = g(G) - h(G)$, if a graph G such that $f(G) < 0$ is found. AGX 2 can often find a counter-example provided there is one with a moderate number of

Table 5: Expressions of values of selected invariants on some families of graphs as functions of n

	K_n	S_n	C_n	$PA_{n,q}$	$Bag_{p,q}$	Bug_{p,q_1,q_2}
Δ	$n-1$	$n-1$	2	$n-1$	$p-1$	$p-1$
δ	$n-1$	1	2	1	2	1
$\overline{d}$	$n-1$	$2-\frac{2}{n}$	2	$2+\frac{q(q-3)}{n}$	$2+\frac{p(p-3)}{p+q-2}$	$2+\frac{p(p-3)-2}{p+q_1+q_2-2}$
$\overline{l}$	1	$2-\frac{2}{n}$	$\frac{n+1}{4}$ if n odd $\frac{n^2}{4n-4}$ if n even	$2-\frac{2n+q(q-3)}{n(n-1)}$	$\frac{2p(2p-9)+q(q^2-3q-9)+2pq(q+2)+19}{4(p+q-2)(p+q-3)}$ if q odd $\frac{(q+2)(q(q-5)+2pq)+4(p-2)(p-3)}{4(p+q-2)(p+q-3)}$ if q even	$\frac{8q_1^3-6q_1^2-14q_1+6pq_1^2+6pq_1+3p^2-15p+18}{3(p+q_1+q_2-2)(p+q_1+q_2-3)}$ if $q_1=q_2$ $\frac{8q_1^3-18q_1^2-2q_1+6pq_1^2+3p^2-15p+18}{3(p+q_1+q_2-2)(p+q_1+q_2-3)}$ if $q_1=q_2+1$
D	1	2	$\lfloor\frac{n}{2}\rfloor$	2	$\lceil\frac{q}{2}\rceil$	q_1+q_2
r	1	1	$\lfloor\frac{n}{2}\rfloor$	1	$\lceil\frac{q}{2}\rceil$	$\lceil\frac{q_1+q_2}{2}\rceil$
g	3	∞	n	3	3	3
ecc	1	$2-\frac{1}{n}$	$\lfloor\frac{n}{2}\rfloor$	$2-\frac{1}{n}$	$\lceil\frac{q}{2}\rceil$	$\frac{3q_1^2+(p-1)q_1}{p+2q_1-2}$ if $q_1=q_2$ $\frac{3q_1q_2+(p-1)q_1}{p+q_1+q_2-2}$ if $q_1=q_2+1$
π	1	1	$\frac{n+1}{4}$ if n odd $\frac{n^2}{4n-4}$ if n even	1	$\frac{4p+2q+q^2-11}{4p+4q-12}$	$\frac{p+q_1^2+q_1-3}{p+2q_1-3}$ if $q_1=q_2$ $\frac{p+q_1^2-3}{p+2q_1-3}$ if $q_1=q_2+1$
ρ	1	$2-\frac{1}{n-1}$	$\frac{n+1}{4}$ if n odd $\frac{n^2}{4n-4}$ if n even	$2-\frac{1}{n-1}$	$\frac{q^2+2(p-2)(q+1)-1}{4(p+q-3)}$ if q odd $\frac{q(q^2+2pq-4p)}{4(p+q-3)}$ if q even	$\frac{(q_1+q_2)^2+q_1+q_2+2q_1(p-3)}{2(p+q_1+q_2-3)}$
λ_1	$n-1$	$\sqrt{n-1}$	2	Not known	Not known	Not known
R	$\frac{n}{2}$	$\sqrt{n-1}$	$\frac{n}{2}$	$\frac{q-2}{2}+\frac{n-q+\sqrt{q-1}}{\sqrt{n-1}}$	$\frac{p+q-3}{2}+\frac{\sqrt{2(p-1)}-1}{p-1}$	$\frac{p+q_1+q_2-6+2\sqrt{2}}{2}+\frac{\sqrt{2(p-1)}-1}{p-1}$
a	n	1	$2-2cos\frac{2\pi}{n}$	1	Not known	Not known
ν	$n-1$	1	2	1	2	1
κ	$n-1$	1	2	1	2	1
α	1	$n-1$	$\lfloor\frac{n}{2}\rfloor$	$n-q+1$	$\lceil\frac{q}{2}\rceil$	$\lceil\frac{q_1+q_2+1}{2}\rceil$
β	1	1	$\lceil\frac{n}{3}\rceil$	1	$\lceil\frac{q+1}{3}\rceil$	$\lceil\frac{q_1+q_2+1}{3}\rceil$
ω	n	2	2	q	$p-1$	$p-1$
χ	n	2	3 if n odd 2 if n even	q	$p-1$	$p-1$
μ	$\lfloor\frac{n}{2}\rfloor$	1	$\lfloor\frac{n}{2}\rfloor$	$\lceil\frac{q}{2}\rceil$	$\lfloor\frac{p+q-2}{2}\rfloor$	$\lfloor\frac{p+q_1+q_2-2}{2}\rfloor$

vertices (say $n \leq 20$). Otherwise one can test if the conjecture is refuted or not by one of the families of graphs for which formulae giving the values of its invariants are available. There is no need in this case to build these graphs explicitly. One only needs to check for all values of n up to some large limit (say n=10000) if the conjecture holds or not. The first procedure is illustrated by the following reformulation of Conjecture 834 of Graffiti [29].

Conjecture 9 *For any connected graph on $n \geq 2$ vertices with minimum degree δ and average distance $\bar{l}$*

$$(1+\delta) \cdot \bar{l} \leq n.$$

This conjecture was refuted [11] using the AutoGraphiX system. The smallest counter-example found is a graph composed by two triangles linked by a path with seven edges. This graph belongs to a family of extremal graphs presented below.

The second procedure is illustrated by the study of the upper bound on the product of the index λ_1 and the girth g, using AutoGraphiX led to the following conjecture.

Conjecture 10 *For any connected graph on $n \geq 3$ vertices with index λ_1 and finite girth g*

$$\lambda_1 \cdot g \leq 3(n-1)$$

with equality if and only if G is the complete graph K_n.

Using a series of graph families to test the above conjecture, it turned out that it is refuted by the turnips (graphs composed of a cycle and one or more pendent edges all incident with the same vertex of the cycle) with large order. The smallest counter-example is the turnip $T_{n,g}$ on $n = 52$ vertices and girth $g = 36$. This conjecture is discussed in [1, 7].

2.5 Reparation of conjectures

In the previous subsection, the use of graph families for conjecture refutation was discussed. At the end of this process and if the best counter-examples belong to a well defined family of graphs one can often state a new conjecture. This is what is called conjecture reparation. Note that the new conjecture, so obtained, may be structural as it is in the following example.

The AutoGraphiX system generated the following conjecture.

Conjecture 11 *For any connected graph G on $n \geq 3$ vertices with average distance $\bar{l}$ and finite girth g,*

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

with equality if and only if G is the cycle C_n.

The optimization was over graphs of order from 3 to 20. Using a series of graph families to test this conjecture for graphs of order $n \geq 21$, counter-examples were found for $n \geq 31$ and the corresponding graphs are again turnips. Using these new results Conjecture is adjusted to

Conjecture 11' *Let $G = (V, E)$ be a connected graph on $n \geq 3$ vertices with average distance $\bar{l}$ and finite girth g.*

(i) If $n \leq 30$, then

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

with equality if and only if G is the cycle C_n.

(ii) If $n \geq 31$, then $\bar{l}/g$ is minimum for some turnip $T_{n,g}$.

The above conjecture is discussed (but remains open) in [6]. The following table gives some presumably optimal values for the girth g for $31 \leq n \leq 190$.

n	$[31, 38]$	$[39, 56]$	$[57, 74]$	$[75, 93]$	$[94, 115]$	$[116, 138]$	$[139, 163]$	$[164, 190]$
g	15	17	19	21	23	25	27	29

The next conjecture was also tested using the families of extremal graphs.

Conjecture 12 *For any connected graph on $n \geq 3$ vertices with clique number ω and minimum degree δ,*

$$\omega - \delta \geq 2 - \left\lfloor \frac{n}{2} \right\rfloor$$

with equality if and only if G is the balanced complete bipartite graph $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$.

This conjecture was refuted and then repaired giving the following one, which is proved in [27].

Conjecture 12' *For any connected graph on $n \geq 3$ vertices with clique number ω and minimum degree δ,*

$$\omega - \delta \geq \lfloor \sqrt{n} \rfloor - \left\lfloor \left(1 - \frac{1}{\lfloor \sqrt{n} \rfloor}\right) n \right\rfloor = \lceil \sqrt{n} \rceil - \left\lfloor \left(1 - \frac{1}{\lceil \sqrt{n} \rceil}\right) n \right\rfloor$$

The lower bound is best possible as shown by balanced complete multipartite graph $T_{\lfloor \sqrt{n} \rfloor}(n)$ or $T_{\lceil \sqrt{n} \rceil}(n)$. The upper bound is attained if and only if G is the short kite $KT_{n,n-1}$ or K_3.

2.6 Finding new families of graphs

As a by-product of the study of relations of AGX Form 1 between the 20 selected graph invariants, several apparently new families of extremal graphs were obtained (It is likely however that some graphs belonging to these families have appeared somewhere in the vaste graph theory literature. We will be glad to learn about the first occurrences of any of them).

Example 5 The difference between the index of a graph G and its average degree $\lambda_1 - \overline{d}$, introduced in [15], is known as the *irregularity* of G (so called as $\lambda_1 - \overline{d} = 0$ if and only if G is a regular graph). The upper bound on the irregularity is studied using AGX 2. Extremal graphs, *i.e* the most irregular ones, are those composed of a clique together with some pending edges all incident to the same vertex from the clique. They are now called pineapples and their family is added to the AGX 2 database of extremal graphs. The associated conjecture, discussed in [2] and which remains open, is the following.

Conjecture 13 *The most irregular connected graph on n ($n \geq 10$) vertices is a pineapple $PA(n,q)$ in which the clique number was $\omega = q = \lceil n/2 \rceil + 1$.*

Example 6 A bag $Bag_{n,q}$ is a graph obtained from a complete graph K_{n-q} by replacing an edge uv by a path P_q. A bag is odd if q is odd, otherwise it is even. The bags, which are now available in the system's database, were obtained by AutoGraphiX as extremal graphs for the lower bound on the ratio $\overline{l}/r$ and for the upper bounds on $\overline{d} \cdot r$ and $\lambda_1 \cdot r$, where $\overline{l}$, r, $\overline{d}$ and λ_1 denote respectively the average distance, the radius, the average degree and the index of a graph. Note that these bounds remain open conjectures. Besides these conjectures, the following theorem is proved in [21].

Theorem 1 *Among all connected graphs on n vertices with radius $r \geq 3$, the odd bag $Bag_{n,2r+3}$ maximizes the index λ_1.*

3 Some Families of Extremal Graphs

The families of extremal graphs issued of the automated comparison of graph invariants [1] are listed in Table 6 below. The first column contains the notation (used in the literature or suggested in this paper) for the graph, the second is for the name of the graph if any, and in the last column the definition of the corresponding graph is given. Finally, in the last column, the absolute frequency with which the graphs considered among the 1520 cases studied are extremal is given. Note that the sum of values exceeds 1520 as for many cases extremal graphs are not unique.

We observe that:

(i) The number of families is moderate, *i.e.* only 47.

(ii) Frequencies vary largely from 599 for the complete graph to 2 for graphs that contain a triangle and with maximum degree 3.

(*iii*) By far the most frequent families appear to be the classical simplest ones of graph theory: complete graph (frequency = 599), path (324), star (281), cycle (148).

(*iv*) Some such graphs to which an edge has been added also appear to be frequent: short kite (157), star with an additional edge (108).

(*v*) Another simple graph to which several pending edges are added is also frequent: pineapple (100).

Table 6: The list of the extremal graphs

Symbol	**Name**	**Definition**	**Freq.**
K_n	Complete graph	A graph on n vertices that contains all possible edges	599
P_n	Path	A graph with vertex set $\{v_1, v_2, \cdots v_n\}$ and edge set $\{v_i v_{i+1}, i = 1, \cdots n-1\}$	324
C_n	Cycle	A graph with vertex set $\{v_1, v_2, \cdots v_n\}$ and edge set $\{v_1 v_n, v_i v_{i+1}, i = 1, \cdots n-1\}$	148
S_n	Star	A graph with vertex set $\{v_1, v_2, \cdots v_n\}$ and edge set $\{v_1 v_i, i = 2, \cdots n\}$	281
$K_{p,q}$	Complete bipartite graph	A graph with vertex set $\{u_1, u_2, \cdots u_p, v_1, v_2, \cdots v_q\}$ and edge set $\{u_i v_j, 1 \leq i \leq p, 1 \leq j \leq q\}$	51
$K_{p_1, \cdots p_k}$	Complete k-partite graph		12
$SK_{n,\alpha}$	Complete split graph	A graph with vertex set $\{u_1, u_2, \cdots u_\alpha, v_1, v_2, \cdots v_{n-\alpha}\}$ and edge set $\{u_i v_j, 1 \leq i \leq p, 1 \leq j \leq q\} \cup \{v_i v_j : 1 \leq i, j \leq n-\alpha\}$	46
$PK_{n,m}$	Path complete or Soltes graph	For fixed integers n and m such that $n-1 \leq m \leq \frac{n(n-1)}{2}$, let k and l the only integers such that $m = n-k+l+\frac{k(k-1)}{2}$ and $0 \leq l \leq k-2$. $PK_{n,m}$ is obtained from a clique K_k and a path P_{n-k} by adding $l+1$ edges between K_k and one end point of P_{n-k}	58
$PK_{n,n}$	Long lollipop	A path complete graph with $m = n$	57
$K_n - e$		Complete graph from which we delete an edge	46
$K_{n-1} + e$	Short kite	A clique on $n-1$ vertices with an additional pending vertex	157
$S_n + e$		A star to which we add an edge	108
$C_{n-1} + e$	Short lollipop	A cycle on $n-1$ vertices together with a pending edge	36
TC_n	Triangulated cycle	A cycle on n vertices plus an edge between two vertices that have a common neighbor	17
$CO_{n,\Delta}$	Comet	A star S_Δ together with a path joined to a pending vertex from the star	56
DC_{n,Δ_1,Δ_2}	Double-comet	Two stars S_{Δ_1} and S_{Δ_2} joined by a path	48
$PC_{n,g}$	Path cycle or lollipop	A cycle C_g together with an appended path	12
UT_n	Triangulated unicyclic graph	A connected graph that contains a triangle as its unique cycle	21
RG	Regular graph	A graph where all the vertices have the same degree	12
TPT		A graph composed of two triangles joined by a path	9
LC	Long claw	A graph composed of three paths with a common extremity	39
$G(n, n-1)$		Graph with at least a dominating vertex (*i.e.* a vertex which is adjacent to all other vertices	17

$G(n, n-1, 1)$		Graph with exactly one dominating vertex	22
$G(n, n-1, 1, 1)$		Graph with at least a dominating and a pending vertex	76
$TG(n, n-1)$		Graph with at least a dominating vertex and contains at least a triangle	28
$TG3$		Graph that contains a triangle and with maximum degree 3	2
$Tnp_{n,g}$	Turnip	A graph composed of a cycle C_g and $n-g$ pending edges all incident to the same vertex of C_g	14
Ur_n	Urchin	A clique on $\lceil \frac{n}{2} \rceil$ vertices and $\lfloor \frac{n}{2} \rfloor$ vertices adjacent to the clique's ones in a one-to-one mapping	47
PG		A graph that contains at least one pending vertex	12
PTG		A triangulated graph containing at least one pending vertex	14
$KeK_{n,p}$		A graph composed of two cliques of almost equal size joined by an edge	19
$KK_{n,p}$		A graph composed of two cliques of almost equal size with a common vertex	66
$KPK_{n,p,q}$		A graph composed of two cliques joined by a path	14
$Bag_{p,q}$	Bag	Graph obtained from a complete graph K_p by replacing an edge uv with a path P_q $(n = p + q - 2)$	13
Bug_{p,q_1,q_2}	Bug	A graph obtained from a complete graph K_p by deleting an edge uv and attaching paths P_{q_1} and P_{q_2} at u and v, respectively $(n = p + q_1 + q_2)$	26
RG	Regular graph	A regular bipartite graph	20
TRG	Triangulated regular graph	A triangulated regular graph	13
$ComE_{n,p}$		A complement of a set of disjoint edges	12
$ComE_{n,\lfloor \frac{n}{2} \rfloor}$		A complement of a matching with $\lfloor \frac{n}{2} \rfloor$ edges	23
$ComMC$		A complement of a minimum (edge-vertex) covering	13
$CTn_{n,n_1,\cdots n_p}$	Connected Turan graph	A set of p disjoint cliques, on $n_1, \cdots n_p$ vertices respectively, together with $p-1$ disjoint edges joining between the cliques making the graph connected, also known as a connected Turan graph	5
$KT_{n,p}$		A clique on $n-p$ vertices together with $p-1$ edges connecting the p vertices that not belong to the clique	18
Cat_n	Caterpillar	A path with additional pending edges	56
Dr_n	Dragon	The graph obtained from a bag $Bag_{4,3}$ and a path P_{n-5} by adding an edge between an end point of P_{n-5} and the vertex of degree 2 in $Bag_{4,3}$	18
$PA_{n,q}$	Pineapple	The graph obtained from a clique K_q by attaching $n-q$ vertices to one vertex in the clique	100
T	Tree	Connected graph without cycle	38
$Kite_{n,q}$	Kite	The graph obtained from a clique K_q and a path P_{n-q} by adding an edge between an endpoint of the path and a vertex from the clique	43

Figure 1 contains the extremal graphs obtained when the automated comparison of graph invariants was done, and which can be defined in terms of their order (number of vertices) only. Figure 2 contains those can be defined using the order and one other parameter, such as clique number (*e.g.* the kite), the size (*e.g.* the path complete graph or maximum degree (*e.g.* the comet).

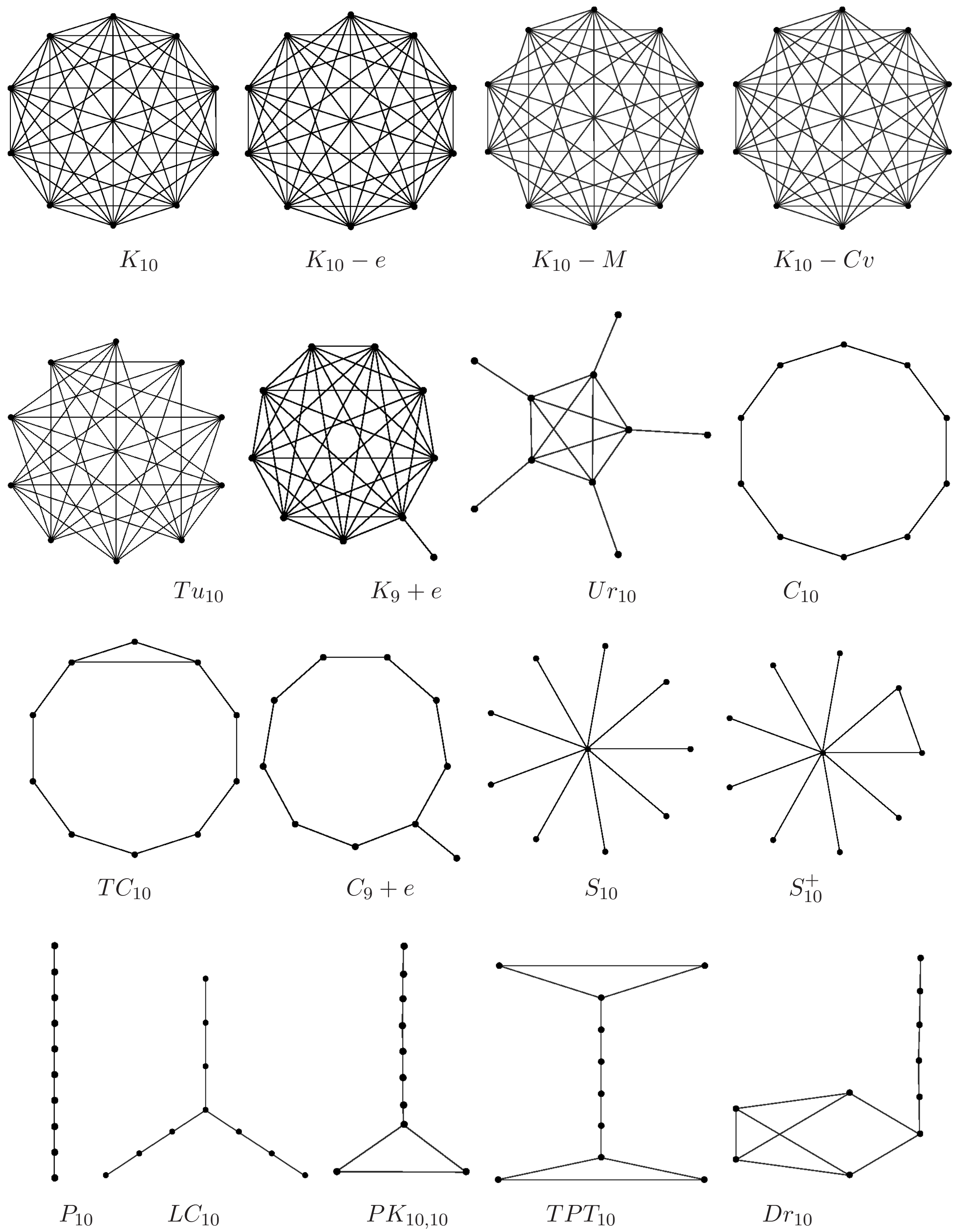

Figure 1: Graphs defined using only their order

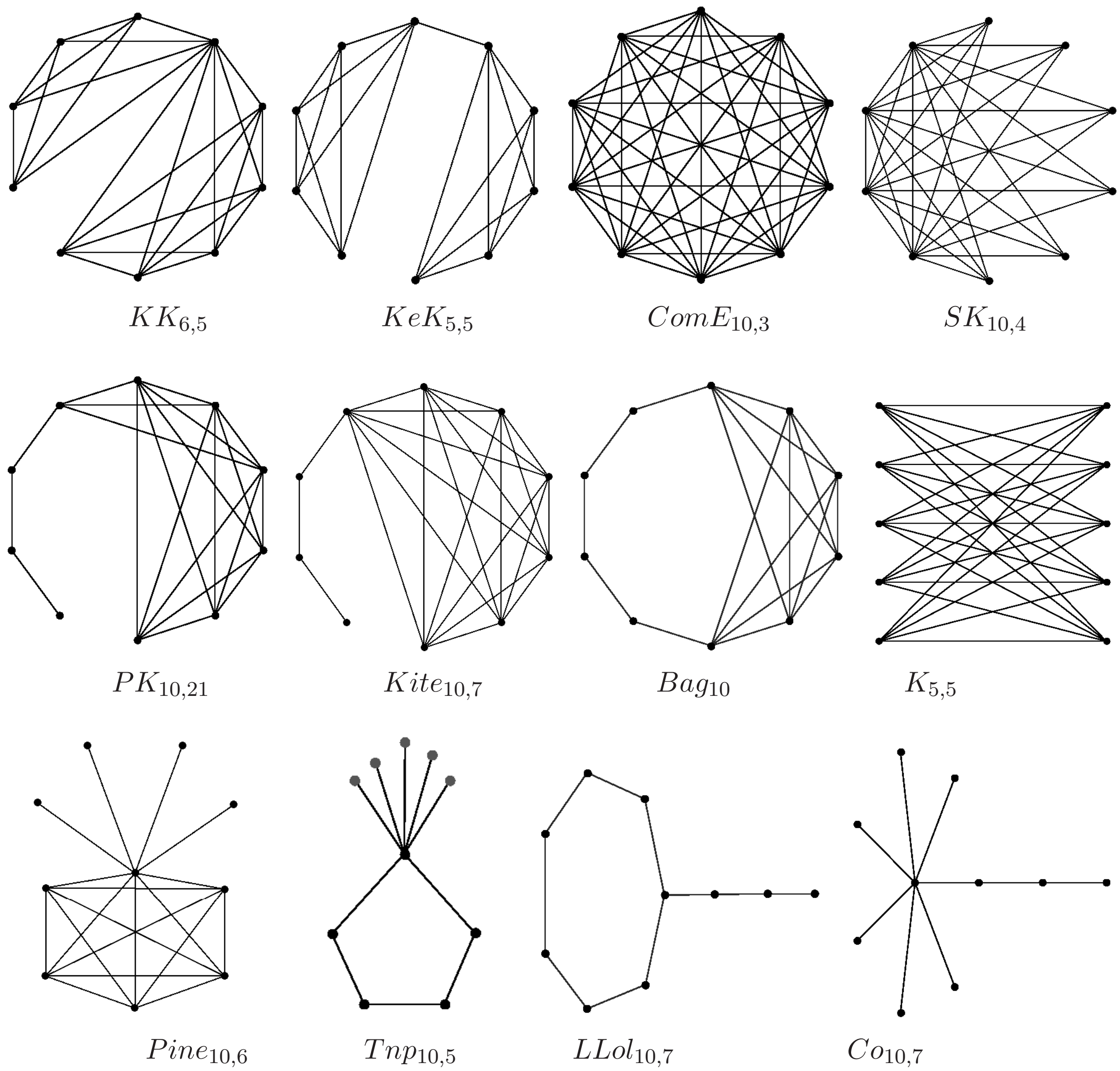

Figure 2: Graphs defined using two parameters

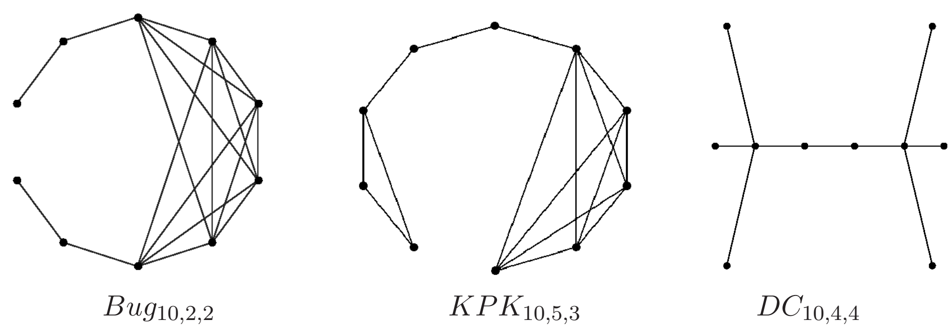

Figure 3: Graphs defined using more three parameters

Figure 3 contains the extremal graphs that need 3 parameters to be defined: the order together with 2 other invariants. The bag_{n,p_1,p_2} is defined by its order n and the lengths of its pending paths p_1 and p_2 respectively.

4 Conclusion

The AutoGraphiX system (AGX 1 and AGX 2) is a sophisticated tool for interactive graph theory, many functions of which are already fully automated. In particular, it can be used to obtain conjectures. These take the form of algebraic conjectures, *i.e.* algebraic inequality or sometimes equality relations between graph invariants, and structural conjectures, *i.e.* conjectures about which families of graphs are extremal. Families of graphs which happen to be extremal for various relations of graph invariants are of great interest. In this paper, we report on families of extremal graphs obtained when studying relations of AGX Form 1 between 20 selected invariants. We show how the use of these families can enhance in 5 different ways the performance of AutoGraphiX system: initialization of the optimization, automated proof of simple results, algebraic generation of conjectures, refutation of conjectures, and reparation of conjectures. Finally, we describe several apparently new families of graphs.

Appendix

In the following tables, which represent the interaction between the extremal graphs and the invariants, we put m (resp. M) when the graph minimizes (resp. maximizes) the invariant and s (resp. S) when the value of the invariant for the graph is at the second smallest (resp. largest) possible value. The symbol s/m (resp. S/M) means that for the corresponding graph the value of the invariant is the second smallest or the smallest (resp. second largest or largest) value.

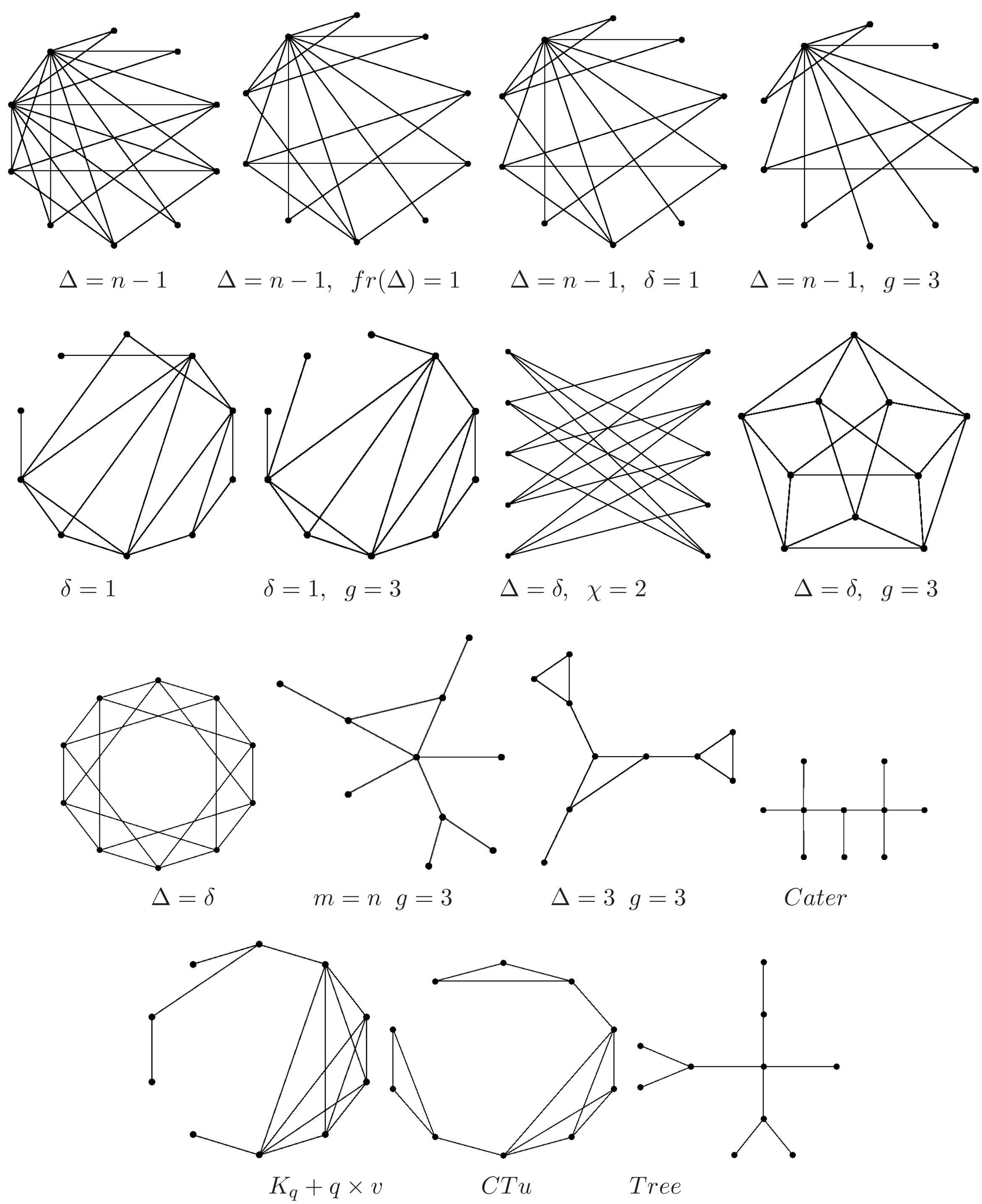

Figure 4: Graphs defined using their properties

Table 7: Interaction between extremal graphs and invariants

	Δ	δ	$\bar{d}$	$\bar{l}$	D	r	g	ecc	π	ρ	λ_1	R	a	ν	κ	α	β	ω	χ	μ
K_n	M	M	M	m	m	m	m	m	m	m	M	M	M	M	M	m	m	M	M	M
P_n	m	m	m	M	M	M		M	M	M	m		m	m	m			m	m	M
C_n	m	s	s			M	M		M			M		s	s			m	m/s	M
S_n	M	m	m		s	m			m			m		m	m	M	m	m	m	m
$K_{p,q}$					s	s	s											m	m	
$K_{p_1,\cdots p_k}$					s	s	m													
$SK_{n,\alpha}$	M				s	m	m		m								m			M
$PK_{n,m}$																				M
$PK_{n,n}$	s	m	s		S	S/M	m							m	m			s	s	M
$K_n - e$	M	S	S	s	s	m	m	s	m	s	S		S	S	S	s	m	S	S	M
$K_{n-1} + e$	M	m			s	m	m		m					m	m	s	m	S	S	M
$S_n + e$	M	m	s		s	m	m		m					m	m	S	m	s	s	s
$C_{n-1} + e$	s	m	s			M	S							m	m			m	s/m	M
TC_n	s	s				S/M	m							s	s			s	s	M
$CO_{n,\Delta}$		m	m											m	m			m	m	
DC_{n,Δ_1,Δ_2}		m	m											m	m			m	m	
$PC_{n,g}$	s	m	s											m	m			m/s	m/s	M
UT_n		m	s				m							m	m			s	s	
RG												M								
TPT	s	s					m							m	m			s	s	
LC	s	m	m											m	m			m	m	S/M
$\Delta = n-1$	M				s	m			m								m			
$!\Delta = n-1$	M				s	m			m								m			
$\Delta = n-1, \delta = 1$	M	m			s	m			m					m	m		m			

Table 8: Interaction between extremal graphs and invariants

	Δ	δ	$\overline{d}$	$\overline{l}$	D	r	g	ecc	π	ρ	λ_1	R	a	ν	κ	α	β	ω	χ	μ
$\Delta = n-1, g = 3$	M				s	m	m		m								m			
$\Delta = g = 3$	s						m													
$Tnp_{n,g}$		m	s											m	m					
Ur_n		m				s	m							m	m					M
$\delta = 1$		m												m	m					
$\delta = 1, g = 3$		m					m							m	m					
K_peK_q							m							m	m					M
K_pK_q	M					m	m		m					m						M
$KpPk_q$							m							m	m					M
Bag							m													M
Bug							m													M
$\Delta = \delta, \chi = 2$												M						m	m	
$\Delta = \delta, g = 3$							m					M								
E	S/M	S					m							S	S					M
M	S/M	S					m							S	S					M
$\overline{Cover}$	S/M						m													M
$K_{p_1}K_{p_2}\cdots$							m							m	m					
$K_{n-p} + p \times e$		m					m							m	m					
Cat		m	m											m	m			m	m	
Dr_n	s	m					m							m	m			s	s	M
$PA_{n,q}$	M	m				m	m		m					m	m		m			
Tr_n		m	m											m	m			m	m	
$Kite_n$		m					m							m	m					M

References

[1] M. Aouchiche, *Comparaison Automatisée d'invariants en Théorie des Graphes. (French)* Ph.D. thesis, École Polytechnique de Montréal, February 2006, available at "www.gerad.ca/∼agx".

[2] M. Aouchiche, F. K. Bell, D. Cvetkovic, P. Hansen, P. Rowlinson, S. Simic and D. Stevanovic, *Variable Neighborhood Search for Extremal Graphs. 16. Some Conjectures Related to the Largest Eigenvalue of a Graph.* To appear in European Journal of Operations Research.

[3] 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.

[4] 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 (2006) 365–384.

[5] M. Aouchiche, G. Caporossi, P. Hansen and M. Laffay, *AutoGraphiX: a survey.* Electronic Notes in Discrete Mathematics, 22 (2005) 515–520.

[6] M. Aouchiche, O. Favaron and P. Hansen, *Recherche à Voisinage Variable de Graphes Extrémaux 26. Nouveaux Résultats sur la Maille.* (French) Submitted.

[7] M. Aouchiche and P. Hansen, *Recherche à Voisinage Variable de Graphes Extrémaux 13. À propos de la Maille.* (French) RAIRO Oper. Res., 39 (2005) 275–293.

[8] M. Aouchiche and P. Hansen, *Automated Results and Conjectures on Average Distance in Graphs.* In A. Bondy, J. Fonlupt, J.-L. Fouquet, J.-C. Fournier, J.-L. Ramírez Alfonsín (eds), Graph Theory in Paris, Trends in Mathematics VI (2007) 21-36.

[9] M. Aouchiche, P. Hansen and M. Zheng, *Variable neighborhood search for extremal graphs. 18. Conjectures and results about the Randić index.* MATCH Commun. Math. Comput. Chem., 56 (2006) 541–550.

[10] R. A Beezer, J. E. Riegsecker and B. A. Smith, *Using Minimum Degree to Bound Average Distance.* Discrete Math., 226 (2001) 365–371

[11] G. Caporossi, *Découverte par Ordinateur en Théorie des Graphes. (French)* Ph.D. thesis, École Polytechnique de Montréal, August 2000, available at "www.gerad.ca/∼gillesc".

[12] G. Caporossi, P. Hansen, *Variable Neighborhood Search for Extremal Graphs. I. The AutoGraphiX System.* Discrete Math., 212 (2000) 29–44.

[13] G. Caporossi and P. Hansen, *Variable Neighborhood Search for Extremal Graphs. V. Three Ways to Automate Finding Conjectures.* Discrete Math., 276 (2004) 81–94.

[14] F. R. K. Chung, *The Average Distance and the Independence Number.* J. Graph Theory, 12 (1988) 229–235.

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

[16] D. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs: Theory and Application.* Academic Press, New York, 1995.

[17] D. Cvetkovic and I. Pevac, *Man-Machine Theorem-Proving in Graph-Theory* Artificial Intelligence, 35 (1988) 1–23.

[18] S. Fajtlowicz, *On conjectures of Graffiti.* Discrete Math., 72 (1988) 113–118.

[19] P. Hansen, *Computers in Graph Theory.* Graph Theory Notes of New York, 43 (2002) 20–34.

[20] P. Hansen and N. Mladenović, *Variable Neighborhood Search: Principles and Applications.* European J. Oper. Res., 130 (2001) 449–467.

[21] P. Hansen and D. Stevanovic, *On bags and bugs.* Electron. Notes Discrete Math., 19 (2005) 111–116.

[22] P. Hansen and D. Vukicevic, *Variable Neighborhood Search for Extremal Graphs. 23. On the Randic Index and the Chromatic Number.* Submitted.

[23] F. Harary, *The Maximum Connectivity of a Graph.* Proc. Nat. Acad. Sci. U.S., 48 (1962) 1142–1146.

[24] Y. Hong. A Bound of the Spectral Radius of Graphs. *Linear Algebra Appl.*, 108 (1988) 135–139.

[25] N. Mladenović and P. Hansen, *Variable Neighborhood Search.* Comput. Oper. Res. 24 (1997) 1097–1100.

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

[27] J. Sedlar, D. Vukičević, M. Aouchiche and P. Hansen, *Variable Neighborhood Search for Extremal Graphs. 24. Conjectures and Results about the Clique Number.* Submitted.

[28] L. Soltès, *Transmission in Graphs: a Bound and Vertex Removing.* Math. Slovaca, 41 (1991) 11–16.

[29] *Written on the Wall.* Electronic file available from http://math.uh.edu/∼clarson/, 1999.