

Graph Theoretic Independence and Critical Independent Sets

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

C. E. Larson

August 2008

UMI Number: 3328485

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3328485

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

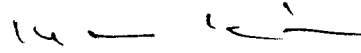
ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346

Graph Theoretic Independence and Critical Independent Sets



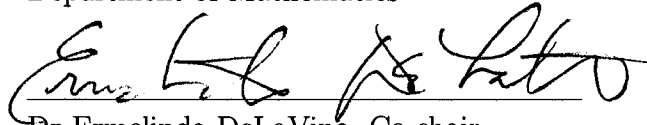
C. E. Larson

APPROVED:



Dr Klaus Kaiser, Co-chair,

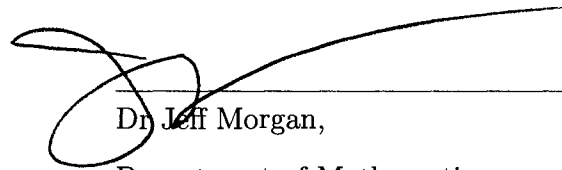
Department of Mathematics



Dr Ermelinda DeLaVina, Co-chair,

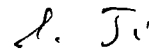
Department of Mathematics,

University of Houston-Downtown



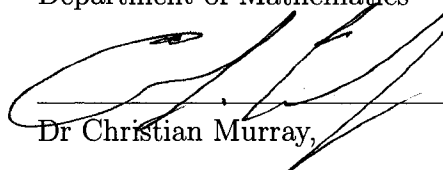
Dr Jeff Morgan,

Department of Mathematics



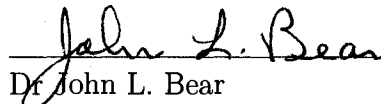
Dr Shanyu Ji,

Department of Mathematics



Dr Christian Murray,

Department of Economics



Dr John L. Bear

Dean, College of Natural Sciences
and Mathematics

ACKNOWLEDGMENTS

I have been extremely lucky to have had the support, kindness and friendship of a large number of people. This dissertation would never have been completed without their help and ongoing faith in me. This dissertation has been a long time in the making, and I would like to acknowledge all of the help I have received.

To begin I would like to thank Ermelinda DeLaVina and Klaus Kaiser. They both stood up when I was having serious difficulties; their support has been indispensable. Dr DeLaVina lent her graph theoretical expertise to read the entire dissertation, finding many mistakes, making useful suggestions, and helping me clarify some key ideas. In the last year, she spent more of her valuable time talking about mathematics with me than anyone, and I learned a lot from her. I also want to thank her for using her program, Graffiti.pc, to run conjectures for one of the invariants discussed below—leading to a nice result that is proven herein. I have known Dr Kaiser for an extremely long time, and learned about logic, algebra, lattices, orders, set theory and non-standard methods, at his feet. He always has a smile, is filled with mathematical history and interesting stories. Their acceptance of co-chairmanship of my dissertation committee was a burden that both undertook without hesitation. I am forever indebted to both of them.

I would also like to thank the remaining members of my committee. Drs Morgan and Ji, as the Department Chair and Graduate Advisor, respectively, have been incredibly supportive. I taught in the Department as a Lecturer for a number of years, and Dr Morgan was as good a boss as I could imagine: he made creative and decisive moves in advancement of undergraduate success and I have learned a lot about management watching him in action. He is a true leader; when the situation calls for it, I try to be as much like him as I can be. Dr Ji has only wanted to see me succeed; I really appreciate all of his help. Dr Murray was very kind to read a dissertation so remote from his professional interests. He has also been a source of

ready and useful advice.

Justin Leiber, now at Florida State University, has always been a great supporter. He was my advisor and mentor in my philosophy days, later my neighbor, and always a friend. He has been a role model: his breadth of knowledge, and willingness to help students, has always impressed me. I want to know what he knows, and have worked at it, with less success. His model always comes to mind when students ask me for recommendations.

My interest in graph theory is due to Siemion Fajtlowicz. The research topics of the dissertation stem from his own interests. For many years he was a great supporter and teacher. His broad interests, development of his Graffiti program, and views on artificial intelligence, have inspired me.

My parents have been constant and unwavering supporters for untold years, always believing in me and encouraging me. I first started learning mathematics as a little boy from my dad. They supported me through all of my changes of major, and focus. They have been incredible. Maybe the most important things they have taught me are to ask lots of questions, to be skeptical of authority, and to pursue knowledge and to enjoy its pursuit. Among other things, they bought me many many books, the utility of which they could never have precisely known. I am a creature made in their image.

There are many many other people that have directly or indirectly helped me and to whom I am grateful. Ryan Pepper has been my best mathematical friend—as well as my friend—and with whom I discussed many of the ideas below. Just trying to keep up with him has pushed me forward more than he realizes. William Staton, Professor of Mathematics at the University of Mississippi, was kind enough to read, comment on, and review this dissertation: it was a great favor and surely not his first choice for how to spend his time. I am also extremely lucky to have a great brother and great friends.

Lastly, none of this would have been possible without the love and support of my sweet wife Jeanine. Until I met her I was unfocused. Only with her in my life have I finally been able to achieve many of the things that I had aimed at but had stopped progressing towards. She has put up with a lot and I dedicate this to her.

Graph Theoretic Independence and Critical Independent Sets

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

C. E. Larson

August 2008

ABSTRACT

The independence number of a graph is the maximum number of pairwise non-adjacent vertices of the graph. New bounds are presented for this NP-hard invariant, new algorithms for calculating it, as well as new theoretical techniques for investigating maximum independent sets in a graph. These include:

1. A new, short proof of Graffiti's conjectured lower bound for the independence number in terms of the number of cut vertices of the graph.
2. A new proof of an upper bound for the independence number in terms of the number of cut vertices, and (together with G. Henry, R. Pepper, and D. Sexton) the characterization of those graphs where equality of this bound holds.
3. A faster version of Tarjan and Trojanowski's algorithm for finding maximum independent sets in fullerenes, together with a previously unreported computation adding to the evidence that minimizing independence is the best statistical predictor of fullerene stability.
4. A characterization of those graphs whose independence number equals its radius, which was an open problem mentioned in a 1986 paper of Fajtlowicz and Waller.
5. A new sufficient condition, following a conjecture of Graffiti, for the existence of a Hamiltonian cycle in a graph.
6. A characterization of those graphs whose independence number equals its annihilation number, which was an open problem in Pepper's 2004 dissertation.
7. A polynomial-time algorithm for finding maximum cardinality critical independent sets, which was an open problem of Butenko and Trukhanov's 2005 preprint.

on using critical independent sets in order to speed-up finding maximum independent sets.

8. The invention of the *critical independence number*, a new polynomial-time computable lower-bound for the independence number, and a polynomial-time computable characterization of the graphs where these invariants are equal.
9. A decomposition of a general graph into two unique subgraphs, such that the independence numbers of these graphs are additive, and where one of these independence number computations can be done in polynomial-time.
10. A new and simple characterization of König-Egervary graphs, resulting from a surprising conjecture of Graffiti.pc.

Contents

1	Introduction	1
1.1	Graphs and Graph Theory	2
1.2	Conjectures and Conjecture-making	7
1.3	Notation and Terminology	15
2	Graph-theoretic Independence, the Independence Number, Bounds and Applications	20
2.1	Matchings and Independence Algorithms	22
2.2	Chemical Graph Theory and Fullerenes	29
2.3	Cut Vertices, the Radius, and Independence Bounds	37
2.3.1	Cut Vertex Lower Bound	39
2.3.2	Cut Vertex Upper Bound & Characterization of Equality	40
2.3.3	Characterizing when Independence equals Radius	45
2.3.4	An application: Hamiltonicity	49
2.4	Open Problems	58
3	Critical Independent Sets, the Critical Independence Number, and Applications	60
3.1	Finding Maximum Critical Independent Sets	61
3.1.1	Weighted Critical Independent Sets	70
3.1.2	An application: Critical Independence Reductions for Fullerenes	73

3.2	An application: Characterizing when Independence equals Annihilation	74
3.3	An Independence Decomposition of a General Graph	78
3.4	An application: König-Egervary Graphs	86
3.5	Open Problems on Critical Independent Sets	88
	Bibliography	90

List of Figures

1.1	The Traveling Salesman Problem: find a minimal weight Hamiltonian cycle: starting at vertex A , visit all other vertices and return to A so that the sum of the weights on the traversed cycle is minimized. Can you do this in a number of steps which is a polynomial function of the number of vertices?	3
1.2	Tetrahedral C_{100} : the lone 100-atom fullerene isomer with isolated pentagons and tetrahedral symmetry. There are 285,914 100-atom fullerene isomers. If a 100-atom isomer appears in experiment, Fajtlowicz and this author predict it will be this model. The picture was produced by the CaGe program.	6
2.1	(a) The truncated icosahedron with pentagons highlighted and (b) with a 24-vertex maximum independent set highlighted.	21
2.2	The sets $X = \{a, b, c, d, e, f\}$ and $Y = \{a', b', c', d', e', f'\}$ form a bipartition of the graph.	24
2.3	The three pendants a , b , and c can be removed and included in any maximum independent set.	28

2.4	IP isomer $C_{100}:321(T)$ with pentagons highlighted. At most two vertices from each pentagon can be included in any maximum independent set: at most 24 of vertices covered by the pentagons can be included in a maximum independent set. The graph induced by the remaining vertices has independence number 19. Thus the independence number of this fullerene is no more than $24+19=43$	35
2.5	Cut vertices. The vertices e , x , and y are cut vertices. Removing any one of them will result in a graph with more than one component. . .	38
2.6	An odd tree. The branching points are c and e . The pendant vertices are a , b , f and g . The distance from any branching point to any pendant vertex is odd.	42
2.7	r -ciliates. The graphs are, from left to right, $C_{2 \times 1, 2-1} = C_{2,1}$, $C_{2 \times 3, 3-3} = C_{6,0}$, $C_{2 \times 2, 4-2} = C_{4,2}$. These graphs are radius critical. If a vertex is removed in any of these graphs that does not disconnect the graph, the radius will decrease.	46
2.8	In all three graphs, $r = 3$ and $R = \{v_0, v_1, \dots, v_5\}$ induces a 6-cycle, with bipartition (B, W) , where $B = \{v_0, v_2, v_4\}$ and $W = \{v_1, v_3, v_5\}$. G_1 satisfies all the conditions of Theorem 2.9 and $\alpha(G_1) = r(G_1) = 3$. $\alpha(G_2) = 4 \neq 3 = r(G_2)$ as G_2 fails condition (3). In this case, for the independent set $I' = \{v_6, v_8\}$ in $G[R']$, $ N_R(I') = 2 I' + k'$, where $k' = 0$, but the number of odd paths in $G[R \setminus N(I')] = 2$, which is greater than k' . $\alpha(G_3) = 4 \neq 3 = r(G_3)$ as G_3 fails condition (2). In this case, for the independent set $I' = \{v_8\}$ in $G[R']$, I' has fewer than $ I' $ white neighbors.	50
3.1	The vertices $\{a, b\}$ form a (maximum cardinality) critical independent set; this set of vertices can be extended to a maximum independent set of the graph.	61

- 3.2 These graphs are *totally independence reducible*. In the graph on the left, the set $I = \{a, c, e\}$ is a maximum critical independent set, and a maximum independent set. In the middle graph, the set $I_c = \{a, b, c\}$ is a maximum critical independent set, and a maximum independent set. The left and the middle graphs are bipartite (and, in fact, isomorphic); all bipartite graphs are totally independence reducible. The graph on the right is not bipartite; the set $I_c = \{a, b, c\}$ is a maximum critical independent set, and a maximum independent set. For all three of these graphs $\alpha = \alpha' = 3$ 80
- 3.3 These graphs are independence irreducible: for *any* independent set of vertices I in these graphs, $|N(I)| > |I|$ 81
- 3.4 Useful figures for following the proofs of Lemma 3.35 and Theorem 3.36. The figure on the top is a schematic of the relationship between critical independent sets I_c and J_c and their neighbors. The figure on the bottom is the same figure but shaded to identify set $I = I_c \cup J = I_c \cup J_c \setminus (I_c \cup N(I_c))$ 82
- 3.5 The vertices $I_c = \{a, b\}$ form a (maximum cardinality) critical independent set. The set $X = I_c \cup N(I_c)$ induces a decomposition of the graph into a totally independence reducible subgraph $G[X]$ and an independence irreducible subgraph $G[X^c]$, according to Theorem 3.36. 83
- 3.6 A non-bipartite König-Egervary graph. The graph has vertex set $V = \{a, b, c, d, e, f\}$ and edge set $E = \{ad, be, cf, de, ef, df\}$. So $n = |V| = 6$. The set $I = \{a, b, c\} \subseteq V$ is a maximum independent set. So $\alpha = |I| = 3$. The set $M = \{ad, be, cf\} \subseteq E$ is a maximum matching. So $\mu = |M| = 3$. Thus, $\alpha + \mu = n$ and the graph is König-Egervary. . 87

Chapter 1

Introduction

Concept formation, conjecture-making, constructing counterexamples, and theorem-proving are four of the main practices of mathematics. Among other things, this dissertation discusses a new graph theoretic invariant and its relationship to the independence number, a well-known and widely-studied graph theoretic invariant. Many of the theorems here were conjectured by this (human) author—but some were conjectured by computer programs. One of the results of Siemion Fajtlowicz’s work on his Graffiti conjecture-making program is the demonstration that computers can have mathematical intuitions as good as those of human beings.

In this chapter, the author’s mathematical results are first introduced and placed in context. Following this is a discussion of conjecture-making and the main heuristics of the Graffiti program. Mathematicians use the word “conjecture” intuitively, without any need for specifying necessary and sufficient conditions. Mathematics would continue to advance whether or not this issue was ever discussed. The issue does arise though in the context of designing and programming conjecture-making programs; ambiguity in the use of the word “conjecture” by non-mathematicians can affect the evaluation of programs. A definition is advanced below. This is of mathematical utility for two reasons. First, a mathematician should keep in mind that her

mathematical activities are part of a collective enterprise and that success is measured in terms of our pre-existing goals. Second, researchers who want to write conjecture-making programs must understand that a conjecture is not an arbitrary statement whose truth-value is unknown; it is a statement whose investigation will advance our collective mathematical enterprise. The last section of this Introduction also introduces the heuristics used in Fajtlowicz's Graffiti program. Several conjectures in this dissertation are products of either Graffiti or Graffiti.pc, a descendant written by Fajtlowicz's former student, Ermelinda DeLaVina. This author has written two papers discussing Graffiti's heuristics and automated mathematical conjecture-making [57, 55]. Mathematicians including Poincare and Hadamard have investigated how mathematics is produced, presumably with the idea that, if it were better understood, mathematicians could increase the quantity and quality of mathematical research. Understanding Graffiti's operation, the operation of a successful conjecture-making program, is relevant to this investigation.

1.1 Graphs and Graph Theory

Informally, a graph is "dots and lines"—dots representing some objects and lines between pairs of them representing the existence of some relationship between the pair. The dots could represent cities, and lines between a pair could represent that there is a regularly scheduled airline route between them. A number (or *weight*) can be associated to each line representing the length of the flight. If each pair of cities has a regularly scheduled flight between them, it can be asked what route should be taken in order to travel to every one of the cities while minimizing the total flight time. Of course, there is an obvious method of solution: consider each possible route (there are only a finite number of them) and choose the shortest one. On this approach, one would have to consider on the order of $(n - 1)!$ possible routes (if there are n

cities). If n is large, this approach is computationally infeasible. The real question is

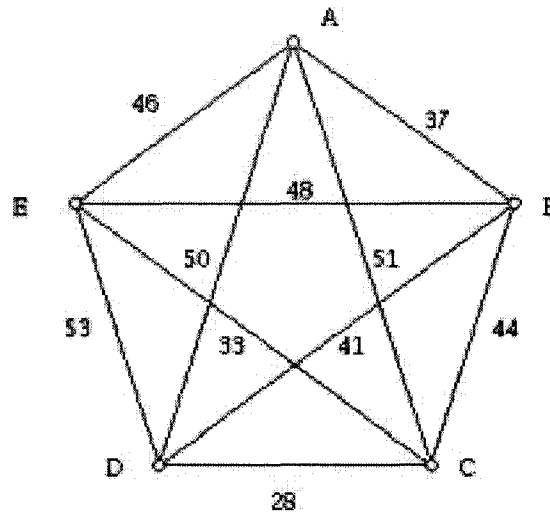


Figure 1.1: The Traveling Salesman Problem: find a minimal weight Hamiltonian cycle: starting at vertex A , visit all other vertices and return to A so that the sum of the weights on the traversed cycle is minimized. Can you do this in a number of steps which is a polynomial function of the number of vertices?

whether there is a “fast” (*polynomial-time*) algorithm for solving this problem—or a proof that a fast algorithm is impossible. This is the Traveling Salesman Problem and it is (equivalent to) the most famous unsolved problem in graph theory, whether or not $P=NP$. This problem is a Millennium Prize problem [52]. It is third on Stephen Smale’s list of “Mathematical Problems for the Next Century” [79]. He says that “its solution, partial results, or even attempts at its solution are likely to have great importance for mathematics and its development for the next century.” It has been shown that there are hundreds of problems that are computationally equivalent to the Traveling Salesman problem (see, e.g., [42]). If a polynomial-time algorithm can be found for this problem, then polynomial-time algorithms for the others would follow; if there is no polynomial-time algorithm, then there cannot be one for any of the others.

Graphs are also used to represent the bonding structure of molecules. In this case the dots represent atoms and a line between them represents a covalent bond between the atoms. In fact, J. J. Sylvester, who introduced the word *graph* to represent these objects, used graphs in exactly this way, to represent what are now called *chemical graphs* (see [5, p. 66]). Of course, a graph can only represent topological features of a molecule; the geometry, for instance, of the molecule is lost. Nevertheless, the corresponding molecular theory (the Hückel theory, see [47, 81]) has proved to be an often good approximation.

These are just two of a large number of applications of graphs; many recent applications are mentioned in [49]. Of particular recent interest and research, and of large investment by Google and other world wide web (WWW) search engine companies, is the graph of web page linkages.

The *graphs* that are considered in the following are all *finite, simple graphs*: there is at most one line between any pair of dots, and no loops. Formally these graphs can be defined as a set of *vertices* and a set of pairs of vertices, called *edges*. The dots then are vertices and the lines are edges. If v and w are vertices and $\{v, w\}$ (or simply vw) is an edge then v and w are said to be *adjacent*. A graph can be represented in a number of ways. One way, for instance, is via its *adjacency matrix*. Label the n vertices of the graph $v_1, v_2, v_3, \dots, v_n$. Then the adjacency matrix of a graph G is $A(G) = (a_{ij})$, where $a_{ij} = 1$ if vertex v_i is adjacent to vertex v_j and $a_{ij} = 0$ if these vertices are non-adjacent. Clearly there are many different adjacency matrices which can be used to represent a graph—each depending on a particular ordering or labeling of the vertices.

A graph *invariant* is a number which is associated to a graph and independent of the representation of the graph. The *order* of a graph, the number of vertices n , is a (simple) graph invariant. The *size*, the number of edges m , is another. A more complicated (and more interesting!) example is the *independence number* of a graph. A

set of vertices is an *independent set* if the vertices in the set are pairwise non-adjacent. The independence number is the cardinality of a maximum independent set of vertices. The independence number is a well-known widely-studied NP-hard invariant [42]; if a “fast” algorithm were known for computing the independence number of an arbitrary graph, it could be transformed into a fast algorithm for solving the traveling salesman problem. One avenue of investigation is finding upper and lower bounds for the independence number. This is of theoretical interest, of use in approximating the independence number, in finding the exact value of the independence number in special cases, and in constructing algorithms for finding maximum independent sets and the independence number. The results in the sequel all relate to these goals, or to applications of the independence number.

Many of the results in the sequel stem from conjectures of Siemion Fajtlowicz’s Graffiti program, or of Ermelinda DeLaVina’s Graffiti.pc program, discussed in the next section. One of Graffiti’s conjectures led to the discovery, by Fajtlowicz and this author, that the minimization of the independence number of a fullerene is the best predictor of its stability [33]. Fullerenes are the third known allotrope of carbon (along with diamond and graphite). They were discovered by R. Curl, H. Kroto, and R. Smalley at Rice University in 1985 [2]. The first to be discovered was Buckminsterfullerene, which appeared in experiment as a 60-atom carbon clump or cluster. Its structure was not known. The Fullerene Hypothesis is that the carbon atoms in these molecules each bond to three others, and that the structure forms a convex polyhedron having pentagonal and hexagonal faces (graph theoretically, a fullerene is a cubic planar graph whose faces are all pentagons and hexagons). For 60-atom carbon clusters there are 1812 mathematically possible structures (or *isomers*) satisfying the fullerene hypothesis [39]. Of these, only one has appeared in experiment. An unresolved question is both how to characterize those fullerenes which appear in experiment and to predict which isomers will appear. Statistical evidence was presented

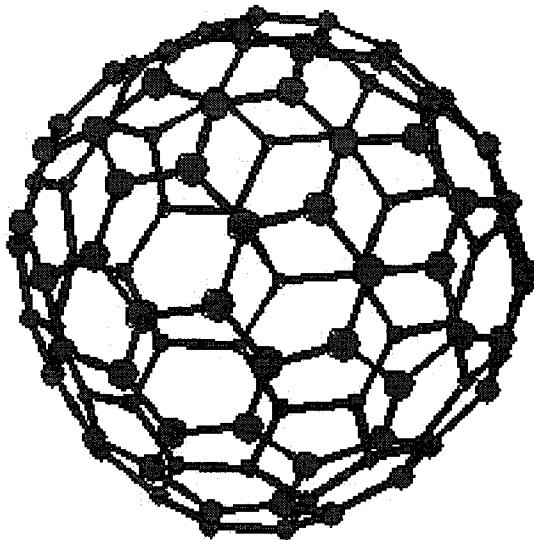


Figure 1.2: Tetrahedral C_{100} : the lone 100-atom fullerene isomer with isolated pentagons and tetrahedral symmetry. There are 285,914 100-atom fullerene isomers. If a 100-atom isomer appears in experiment, Fajtlowicz and this author predict it will be this model. The picture was produced by the CaGe program.

in [33] that the independence number of a fullerene correlates with its stability. More is presented here, together with the algorithm used to calculate the independence numbers of these graphs.

In 1990 Zhang introduced the concept of a *critical set*, a set C such that the difference between the cardinalities of it and its neighbor set is maximized [84]. His introduction of this concept was motivated by a relation to pre-existing concepts (independent sets, and the binding and isoperimetric numbers) which measured certain ratios between cardinalities of certain vertex sets and their neighbors. He also introduced the concept of a *critical independent set*, which is an independent set that is

also a critical set. Zhang also showed that these sets can be found in polynomial-time.

Much later, Butenko and Trukhanov showed that the identification of critical independent sets is related to the problem of finding a maximum independent set (MIS) in a graph: they showed that any critical independent set can be extended to a maximum independent set [10]. The consequence of this is that the problem of finding a MIS can be *reduced*. In order to find a MIS, first find a critical independent set I_c . Then there must be a MIS I such that $I_c \subseteq I$. I_c and its neighbors can be removed from the graph; all that remains to do is to find a MIS J of the graph induced on the remaining vertices. $I = I_c \cup J$. Computationally this reduces the problem of solving an NP-hard problem on a graph to one on a (possibly smaller) subgraph. The reduction can be found in polynomial-time.

A graph can have critical independent sets of different cardinalities. In fact, for some graphs, the empty set and a set containing as many as half of the vertices are both critical independent sets. Identifying the empty set will yield no reduction (and hence no algorithmic speed-up) of the MIS problem. Identifying the larger critical independent set may result in a complete reduction of the problem. In their paper, Butenko and Trukhanov mentioned the open problem of finding maximum cardinality critical independent sets. The solution is presented in the sequel. The definition of the *critical independence number* as the cardinality of a maximum independent set led to a conjecture of DelaVina's Graffiti.pc program giving a simple new characterization of König-Egervary graphs. The conjecture and proof are in Section 3.4.

1.2 Conjectures and Conjecture-making

Mathematicians invent new concepts, conjecture and prove theorems, and find counterexamples to conjectures. The importance and centrality of conjecture-making in the process of mathematical research is often overlooked—but one cannot prove a the-

orem without first having a conjecture. Some rare mathematicians, including Paul Erdős, are as well-known for their conjectures as for their theorems.

Research on artificial intelligence began in the 1950's. Researchers were interested in what mental processes could be mechanized. Mathematics, due to its formality, was an early and obvious area to investigate. Most of the work that has been done has been in automated theorem-proving. Automated conjecture-making has been largely overlooked. Hao Wang did some initial work on this in the late 1950's and early 1960's [83] and raised the fundamental question: from a mass of statements with unknown truth-values, which should be investigated, which are "significant"?

What is a conjecture? Any mathematician has a pretty good idea what a conjecture is, and there has not been any real effort towards specifying precise necessary and sufficient conditions. Nevertheless, if one wants to write or evaluate a conjecture-making program some clarity about what is meant is required.

In the 1970's Douglas Lenat wrote a program AM that produced known mathematical statements and theorems including Goldbach's Conjecture [58, 67, 63, 64, 62, 66, 65, 60, 61, 59, 76]. Nevertheless, it did not produce any statements that a working mathematician would call a conjecture or that would initiate any new mathematical investigation. An attempt is made here to introduce some clarity. A main consequence of this investigation is an intrinsic definition of a "conjecture" modeled after Einstein's definition of "simultaneity." It is an important observation that our mathematical judgments are in the context of our mathematical practice—which typically begins with problems and interests of our teachers and the broader mathematical community.

Wang and Lenat raised the issue of how to program a machine to produce conjectures. Siemion Fajtlowicz solved this problem in the 1980's when he initially wrote his Graffiti program. This program has produced, among other things, conjectured bounds for the independence number of a graph. Graffiti uses few examples, imitating

what seems to be true of human conjecture-makers. Its heuristics, described below, are almost certainly of wider applicability and will be the basis of future experimentation by this author.

The following material is adapted from this author's [56].

What counts as a "conjecture" and, thus, success or failure of various programs that might be called "conjecture-making programs" is partly a terminological question. The word "conjecture" is used in various ways: a teacher might call a student's proposal for trisecting an angle with ruler-and-compass a "conjecture" even though he knows such a construction is impossible; a mathematician who proposes some non-novel proposition may be credited with having made a "conjecture," for instance if evidence suggests that it was put forth independently.

That a statement is "interesting" is a plausible necessary condition for mathematical conjecture-hood. When is a statement interesting? What are the criteria for this? A statement might be defined to be interesting if it inspired mathematical research resulting in publication. That a statement led to a mathematical publication may not be due to any fact about the statement—it could be for non-mathematical reasons. A researcher may have investigated the truth of a statement simply because it struck her fancy, was entertaining, in the way a crossword puzzle is entertaining. The pursuit of most mathematical research may involve this element—but the pursuit of various human whims can hardly provide the criteria for distinguishing what to count as a conjecture.

What epitomizes mathematical research is that it contributes to the advancement of mathematics. The clearest advances occur when our existing mathematical questions are answered. The determination of the truth-value of an "interesting" mathematical statement, however interesting-ness is defined, may or may not advance or answer any of our existing mathematical questions. Hence, the clearest way to define what to count as a "conjecture" is not in terms of concepts which may be

extrinsic to our mathematical goals, but to define it intrinsically, directly in terms of our mathematical goals. A conjecture should be defined to be *a new mathematical proposition the determination of whose truth would be relevant to the advancement of our existing mathematical questions*. G. H. Hardy famously claimed that a mathematician's product should be judged by its "beauty" and "seriousness." He described the seriousness of a mathematical theorem

in the *significance* of the mathematical ideas which it connects. We may say, roughly, that a mathematical idea is "significant" if it can be connected in a natural and illuminating way, with a large complex of other mathematical ideas. Thus a serious mathematical theorem, a theorem which connects significant ideas, is likely to lead to important advances in mathematics itself and even in other sciences. [50, p. 89]

Applied to novel mathematical statements, a conjecture, as here defined, would have some degree of "seriousness" (as Hardy meant the word).

The mere novelty of a statement cannot be a sufficient condition for conjecturehood—it is trivial to produce new statements (or to write a program to produce them). Being new, though, is a necessary condition. If a human or program today conjectured that there are infinitely many twin primes (pairs of primes of the form p and $p + 2$), it would not be a contribution to the advancement of mathematics as this conjecture is already known, discussed and researched—and we would not credit the human or program with having made the conjecture. A new formula for the exact number of primes up to n , a new formula for the exact number or even a bound for any mathematical quantity for which formulas are sought, a proposition that would imply the existence or non-existence of odd perfect numbers, would all count as conjectures on this definition. The definition of "conjecture" given here provides an unambiguous criterion, removed from the vagaries of our psychology and sociology, and explained immediately with reference to actual mathematical practice.

Among statements which count as conjectures under this definition, there is a continuum of relevance: these conjectures are certainly not of equal value—knowledge of the truth-values of some of them will answer or advance more of our mathematical questions than others. (Similarly, with Hardy’s criterion for the “seriousness” of a mathematical theorem, there is a continuum of seriousness—some theorems connect more mathematical ideas than others.)

Graffiti, a program conceived by Fajtlowicz at the University of Houston (and developed, from 1990 to 1993, with Ermelinda DeLaVina), was the first program to have actually made (research) conjectures. The statements Graffiti has produced are largely novel. Since the 1980s, Fajtlowicz has maintained a list, *Written on the Wall* (WoW), of hundreds of these statements [30]. They have inspired research by numerous mathematicians. There are numerous papers, theses, and dissertations in which these statements (or weaker or stronger variants) are proved or disproved.¹ Graffiti’s collaborators include such well-known graph theorists as Noga Alon, Bela Bollobas, Fan Chung, Paul Erdős, Jerry Griggs, Daniel Kleitman, Laszlo Lovasz, Paul Seymour and Joel Spencer [7, 11, 20, 21, 44, 53, among others].

Many of Graffiti’s statements are explicit bounds for quantities for which bounds are desired; others imply bounds. Mathematicians have published numerous papers on bounds for various graph-theoretic quantities (invariants), for instance, the independence number of a graph; Graffiti’s statements have provided new bounds for many of these quantities and thus these statements are conjectures.

Graffiti’s first conjectures were in the field of graph theory. Its underlying ideas, as described in Fajtlowicz’s papers [23, 25, 27, in particular], apply not just to graphs: Graffiti has also made conjectures in geometry, number theory, and chemistry—conjectures about the structure of fullerenes (as represented by their graphs) have already led to papers by, among others, the fullerene expert Patrick Fowler [37, 31, 33].

¹A partial list can be found on the WWW at: cms.dt.uh.edu/faculty/delavinae/wowref.html.

One of Graffiti's conjectures led to the discovery, by Fajtlowicz and this author, that the minimization of the independence number of a fullerene is at least as good a predictor of stability as several predictors which have been used by chemists in predicting the stability of these molecules [33].

Suppose conjectures about objects of a given type (for example, graphs) are desired, and that representations of some number of these objects, O_1, O_2, \dots, O_n , are stored in the computer's memory. An invariant of these objects is a function which associates a number to each of the objects (in the case of graphs, the independence number of the graph is an invariant). Let $\alpha_1, \alpha_2, \dots, \alpha_r$ be computable invariants: for an object O , $\alpha_i = \alpha_i(O)$. Let f_1, f_2, \dots, f_s represent operations of some algebraic system (these might include, for instance, the ordinary arithmetic operators "plus," "times," &c., or any other unary, binary or n-ary operators.) Any term, for instance, $f(\alpha_1, \alpha_2)$, represents a new numerical invariant. (If "plus" is an operation, then $\alpha_1 + \alpha_2$ is a term—representing a new invariant). Statements can then be formed from relations of these terms. If t and s are two such terms, the expression $t \leq s$ —which should be interpreted as the statement, "For every object O (of the type of object under consideration), $t(O) \leq s(O)$ "—is a candidate for a conjecture.

Graffiti's main heuristic for culling the stream of candidate conjectures is the *Dalmation* heuristic [27, pp. 370-371]. Given a statement of the form $t \leq s$ and a (possibly empty) database of pre-existing conjectures of similar form, $t \leq u_1, t \leq u_2, \dots, t \leq u_l$ —the program checks if the statement " $s(O) < u_1(O)$ and $s(O) < u_2(O)$ and ... and $s(O) < u_l(O)$ " is true for at least one of the objects O from the set O_1, \dots, O_n . That is, it checks if there is an object for which the value of the candidate upper bound is less than the minimum value of the existing conjectured upper bounds. If it is, then, with respect to this object, the relation says something "stronger" than all the stored conjectures and, with respect to the objects stored in memory, the relation $t \leq s$ says something informative—that is, the relation says something that

was not implied by the totality of the previous conjectures of that form that had been kept in the program's database—so the relation remains a candidate for Graffiti to add to the database of conjectures. Otherwise, Graffiti rejects the relation as a possible conjecture—with respect to the databases of objects and pre-existing conjectures it is uninformative.

The second heuristic, applied to those relations which survive the Dalmatian heuristic, is to test for the truth of the relation with respect to the stored objects. If the relation is true of all of these objects then it is added to the database of conjectures; and if the relation is false for any of these objects then the general statement that the relation of term functions (the relation of invariants) represents is false—and the relation is not accepted as a conjecture. These first two heuristics are the heart of the program and express the following principle of Fajtlowicz: *make the strongest conjecture for which no counterexample is known.*

Another heuristic used in Graffiti is applicable only when objects of a proper superclass of objects are already stored in the computers memory, the *Echo* heuristic. [24, p. 190] Suppose the database of objects includes $O_1, \dots, O_m, O_{m+1}, \dots, O_n$ of a type A and conjectures are desired of a type B , a subclass of A . Suppose the objects of type B are the objects O_1, \dots, O_m . The Echo heuristic is used to cull those possible conjectures which are true of each of the objects O_{m+1}, \dots, O_n : conjectures which could be true of all objects of type A —when what is desired are conjectures about its proper subclass B —are not specific enough and are rejected. In Fajtlowicz's papers he often calls this superclass, the "background." Thus, if conjectures about fullerenes (fullerene graphs) are desired, Graffiti can be directed to cull all conjectures about graphs in general: here graphs are the superclass and fullerenes are the proper subclass; the resulting conjectures will be true of the fullerenes in the database of graphs but false for at least one of the non-fullerene graphs.

Graffiti's conjectures may, naturally, be false. These can be removed by inform-

ing the program of a counterexample, that is, by adding a new representation to the program's database of objects. Counterexamples can be found automatically, by producing representations of objects of the given type and testing the stored conjectures against them, or counterexamples can be provided by another intelligent agent—whether human or another computer.

Graffiti's operation is sped along by keeping its databases of objects and conjectures relatively small. The program only stores (representations of) objects which it has found "informative," that is, which have served as a counterexample to some previously made conjecture. Graffiti's database of conjectures is kept relatively small by eliminating conjectures that are no longer informative. Whenever a new relation is added to the database of conjectures, it is possible that one or more pre-existing relations are no longer informative (with respect to the objects stored in the computer): if there is an object for which a bound is better than those given by all the other conjectured bounds then this bound is kept, otherwise it is removed (as, with respect to the stored objects, it does not improve on the other existing bounds). This implies that the number of conjectured bounds of a given form (for instance, upper bounds for a given invariant), stored at a given moment, can never exceed the number of objects stored at that moment.

Graffiti's operation may be seen as mimicking the brain's operation in the sense that our brains do not and cannot store records of all the specific objects and relations holding between those objects that we encounter and experience; only some of those objects and relations make an "impression." Furthermore, just as we humans do, Graffiti can fall back on previously accepted but superseded beliefs when its present beliefs are proved wrong.

1.3 Notation and Terminology

The terminology and notation in graph theory is not consistent. In particular the terminology in chemical graph theory is quite a bit different than that used in well-known graph theory texts like [9]. Chemists, for instance, say that a bipartite graph is *alternant*. Outside of graph theory, graphs are often called *networks*. The notation used is largely consistent with [9], which should also be consulted if any terminology is not defined here.

Most, if not all, of the following definitions appear in the sequel where the concept is first introduced. This glossary was primarily conceived as a reference for the convenience of the reader who has forgotten a definition that was introduced earlier in the text. Thus, the definitions are presented in alphabetical order.

- For a graph G with vertices $V = \{v_1, v_2, \dots, v_n\}$, having degrees $d_i = d(v_i)$, with $d_1 \leq d_2 \leq \dots \leq d_n$, and having e edges, the **annihilation number** $a = a(G)$ is defined to be the largest index such that $\sum_{i=1}^a d_i \leq e$. Pepper originally defined the annihilation number of a graph in terms of a reduction process on the degree sequence of the graph (akin to the Havel-Hakimi process; see, for example, [44]). The definition here is due to Fajtlowicz. The definitions are proved to be equivalent in [73].
- A graph G is **bipartite** if there are independent sets A and B such that $V(G) = A \cup B$. A bipartite graph can be colored with two colors such that adjacent vertices are colored different colors. The independence number of any bipartite graph can be computed in polynomial-time.
- The **chromatic number** χ of a graph G is the minimum number of colors that are needed to color the vertices of G such that no pair of adjacent vertices has the same color.

- A **clique** in a graph G is a set of vertices C such that, for each pair of vertices $x, y \in V(G)$, xy is an edge in G ; that is the set C induces a complete subgraph in G . The **clique number** is the cardinality of a maximum clique.
- A graph is **complete** if, for every pair of vertices $x, y \in V(G)$, xy is an edge in G .
- The **complement** \bar{G} of a graph G is the graph with vertex set $V(G)$ and, for vertices $x, y \in V(G)$, edge $xy \in E(\bar{G})$ if, and only if, $xy \notin E(G)$.
- A **component** of a graph G is a maximal connected subgraph.
- A graph is **connected** if there is a path between any pair of its vertices.
- A vertex v of a graph is a **connector** if v is not a cut vertex.
- The **critical difference** of a graph G is the maximum value of $|S| - |N(S)|$ for any (possibly empty) subset S of $V(G)$. Zhang showed that this is the same as the maximum value of $|I| - |N(I)|$ for any (possibly empty) independent subset I of $V(G)$ [84].
- A **critical set** in a graph G is any set of vertices S such that $|S| - |N(S)|$ equals the critical difference of G . Zhang showed that if S is a critical set, then $S \setminus N(S)$ is a critical set which is independent [84]. A **critical independent set** is such a set. The **critical independence number** α' is the cardinality of a maximum critical independent set. It is a lower bound for the independence number of the graph.
- A graph is **cubic** if every vertex of the graph has degree three, that is if the graph is regular of degree three. Cubic graphs are also called *trivalent*.
- A vertex v of a graph G is a **cut vertex** if the number of components of $G - v$ is greater than the number of components of G .

- A **cycle** in a graph G is a path v_1, v_2, \dots, v_k , together with the edge $\{v_1, v_k\}$ (assuming vertices v_1 and v_k are adjacent).
- The **degree** $d(v)$ of a vertex v in a graph is the number of edges incident to v . In a simple graph this number is the same number as the number of vertices adjacent to v .
- The **distance** $d(v, w)$ between vertices v and w in the same component of a graph G is the length of a shortest path between the vertices. If $v = v_0, v_1, v_2, \dots, v_k$ is a path which minimizes the number of vertices—a shortest path—then $d(v, w) = k$.
- A subset S of the vertex set is a **dominating set** in the graph if every vertex is either in S or adjacent to a vertex in S . The **dominance number** γ is the cardinality of a minimum dominating set.
- The **eccentricity** of a vertex v in a graph G is the maximum distance $d(v, w)$, for every vertex w in the component of G containing v .
- A **fullerene** is a cubic, planar graph whose faces are either hexagons or pentagons.
- A **Hamiltonian cycle** (or **spanning cycle**) in a graph is a cycle which includes all of the vertices of the graph. A **Hamiltonian path** is a path which includes all of the vertices. A **Hamiltonian graph** is one having a Hamiltonian cycle. Determining whether a graph is Hamiltonian is an NP-hard problem.
- An **independent set** of vertices in a graph is a set of vertices which are pairwise non-adjacent. The **independence number** α of a graph is the cardinality of a maximum independent set. Many books and papers use the symbol β_0 for the independence number.

- If S is a subset of vertices of a graph G , then the graph which consists of the vertices S together with all those edges in G incident to a pair of vertices in S is the **induced subgraph** $G[S]$.
- A **König-Egervary graph** is a graph where the sum of the independence and matching numbers equals the number of vertices of the graph ($\alpha + \mu = n$). Bipartite graphs are König-Egervary graphs.
- A **matching** in a graph is a set of edges such that no pair of the edges are incident to the same vertex, that is, where no two of the edges have a common endpoint; a matching is an independent set of edges. The **matching number** μ of a graph is the cardinality of a maximum matching. A **perfect matching** is a matching that covers all of the vertices of the graph (thus, $n = 2\mu$). A graph has a **pseudo-perfect matching** if the graph formed by deleting any single vertex has a perfect matching.
- The **neighbors** $N(S)$ of a set of vertices S in a graph G are all those vertices in $V(G)$ adjacent to any vertex $v \in S$.
- A **path** is a sequence of distinct vertices v_1, v_2, \dots, v_k such that each vertex in the sequence is adjacent to its successor.
- The **path covering number** ρ of a graph G is the minimum number of vertex disjoint paths that cover (or contain) all of the vertices of G . If a graph has a Hamiltonian path, then $\rho = 1$.
- A **pendant vertex** (or simply, a **pendant**) is a vertex of degree one, necessarily incident to exactly one edge.
- A **planar graph** is one that can be embedded in the plane without any edge crossings.

- The **radius** of a connected graph is the minimum eccentricity of all of its vertices.
- The **Randić index** R of a graph G is the sum, over all the edges of G , of the reciprocal of the products of the degree weights of the two vertices incident to the edge. That is, $R = \sum_{xy \in E(G)} \frac{1}{d(v)d(w)}$.
- A graph is **regular** if the degrees of its vertices are the same.
- A set of edges M of a graph **saturates** a vertex v if v is incident to at least one edge in M . A vertex which is not incident to any of these edges is **unsaturated** (with respect to M).
- A **vertex cover** of a graph is a set of vertices such that every edge of the graph is incident to at least one of the vertices. The **vertex covering number** τ of a graph is the cardinality of a minimum vertex cover. Lovasz and Plummer [70] use this notation. Bondy and Murty [9] use β .
- The **Wiener index** W of a connected graph G is the sum, over all pairs of vertices of G , of the distances between those pairs. That is, $W = \sum_{\{v,w\} \subseteq V} d(v,w)$.

Chapter 2

Graph-theoretic Independence, the Independence Number, Bounds and Applications

A set of vertices in a graph is *independent* if the vertices in the set are pairwise non-adjacent or, equivalently, if the graph has no edge incident to two of the vertices in the set. The *independence number* of a graph is the cardinality of a maximum independent set.

Some of the basic facts about the independence number of a graph include the following: For the complete graph on n vertices, K_n , $\alpha(K_n) = 1$. For the cycle on n vertices, C_n , $\alpha(C_n) = \lfloor \frac{n}{2} \rfloor$. For the path on n vertices, P_n , $\alpha(P_n) = \lfloor \frac{n}{2} \rfloor$. For a bipartite graph G with bipartition (X, Y) , $\alpha(G) \geq \max\{|X|, |Y|\}$.

The independence number is directly and simply related to many other graph invariants, including the *chromatic number* χ , the *clique number* c , the *vertex covering number* τ , the *domination number* γ and the *matching number* μ . For any graph G , $\chi(G)\alpha(G) \geq n(G)$, $c(G) = \alpha(\bar{G})$ and, conversely, $c(\bar{G}) = \alpha(G)$. For any graph G , $\alpha(G) + \tau(G) = n(G)$ (this is one of the Gallai Identities, see [70]). For any graph G ,

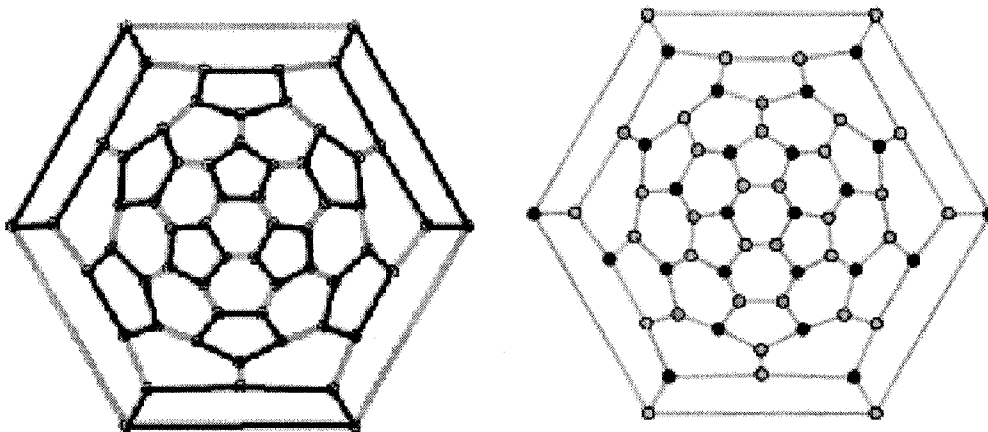


Figure 2.1: (a) The truncated icosahedron with pentagons highlighted and (b) with a 24-vertex maximum independent set highlighted.

if I is a maximum independent set then every vertex in $V(G) \setminus I$ must be adjacent to some vertex in I . Thus, I is a dominating set and $\gamma(G) \leq \alpha(G)$. A necessary condition relating the matching and independence numbers is: $\alpha \geq n - 2\mu$. This is explained in the next subsection.

Bounds for the independence number are of theoretical and practical interest. Several, including the Caro-Wei lower bound, can be found in Gutin's survey [46] in the Handbook of Graph Theory. One very good upper bound, Cvetkovic's spectral bound, is not included there. For graph G with adjacency matrix $A(G)$, the *spectrum* of G is the set $\mathbf{Sp}(G)$ of eigenvalues of $A(G)$. While the adjacency matrix of G is not uniquely defined, the spectrum is—it is same for each adjacency matrix. Let p_0 , p_+ and p_- be the number of zero, positive and negative eigenvalues, respectively. Then Cvetkovic's Theorem [13, p. 88] is that, for any graph, $\alpha \leq p_0 + \min\{p_+, p_-\}$.

Several bounds, with very simple forms, were conjectured by Fajtlowicz's Graffiti program. These include the radius, average distance, and residue lower bounds. The first follows immediately from the Induced Path Lemma: that every connected graph must contain an induced path of length at least $2r - 1$ [19, 22]. The second bound

was proved by Chung [11] and published in a 1988 paper which is perhaps the first paper due to a conjecture of a computer program. The third bound was originally proved by Favaron, Mahéo and Saclé. Various simpler proofs have been published (including, for instance, one by Griggs and Kleitman [44]).

Other bounds for the independence number, in terms of the number of cut vertices in a graph and its critical independence number, are discussed and proved in the sequel.

2.1 Matchings and Independence Algorithms

A *matching* in a graph is a set of edges, no pair of which are incident to the same vertex; in other words, a matching is an independent set of edges. The *matching number* μ of a graph is the cardinality of a maximum matching. The independence number is related to matchings and the matching number via the fact that, if M is a maximum matching of a graph then the vertices not covered by the edges in the matching must be independent. Since the number of vertices covered by M is 2μ it follows that $\alpha \geq n - 2\mu$.

A *bipartite graph* is one whose vertex set can be divided into two independent subsets. For any bipartite graph G , it is known that $\alpha + \mu = n$ (this is the König-Egervary Theorem, see e.g. [70]). An efficient (that is, polynomial time) algorithm for finding a maximum matching (and hence the matching number μ) has been known since at least the 1950's. Kuhn called it the Hungarian method (see, e.g., [9, pp. 82-83]). It follows then that the independence number of a bipartite graph can be found in polynomial time. This is an important fact that the analysis of the MCIS algorithm, discussed in Chapter 3, depends on. It is a result of Edmonds that a maximum matching of a general graph can be found in polynomial time [70]. The author's own programs use another algorithm, one which finds a maximum flow on a

bipartite graph formed by adding a “source” vertex adjacent to all the vertices in one subset of the graph and a “sink” vertex adjacent to all of the vertices in the other subset (see [54, pp. 120-122]). An algorithm for producing a maximum independent set in a bipartite graph, rather than simply the independence number, is discussed below.

The following theorem was conjectured by DeLaVina’s Graffiti.pc program. It illustrates the fact that computers can often have better intuitions than humans. In particular, conjectures of Graffiti and Graffiti.pc often involve seemingly unrelated invariants.

Theorem 2.1. (*Graffiti.pc #113*) *For connected graphs with more than one vertex, $\mu \geq \lceil \frac{g}{2} \rceil$, where g is the number of distinct degrees of the graph.*

Proof. Let G be a connected graph with more than one vertex. Let $S = \{v_1, v_2, \dots, v_g\}$ be a set of g vertices having distinct degrees $d_1 < d_2 < \dots < d_g$.

Case I. Assume that g is even, and let $k = \frac{g}{2}$. Let $V = \{v_{k+1}, v_{k+2}, \dots, v_g\}$. Since $d_{k+1} \geq k + 1$ and $|V| = k$, v_{k+1} must be adjacent to some vertex u_1 not in V . Since $d_{k+2} \geq k + 2$ and $|V \cup \{u_1\}| = k + 1$, v_{k+2} must be adjacent to some vertex u_2 not in $V \cup \{u_1\}$. In general, since $d_{k+i} \geq k + i$ and $|V \cup \{u_1, u_2, \dots, u_{i-1}\}| = k + (i - 1)$, v_{k+i} must be adjacent to some vertex u_i not in $V \cup \{u_1, u_2, \dots, u_{i-1}\}$. Thus, the set $\{(v_{k+1}, u_1), (v_{k+2}, u_2), \dots, (v_{k+k}, u_k)\}$ is a set of k independent edges and $m \geq \frac{g}{2}$.

Case II. When g is odd the proof is similar, but slightly more delicate. Assume $g = 2k + 1$. So it must be shown that $\mu \geq \lceil \frac{2k+1}{2} \rceil = k + 1$. Let $V = \{v_{k+1}, v_{k+2}, \dots, v_g\}$. Since it is assumed that G is finite and simple, v_{k+1} must be adjacent to d_{k+1} vertices besides itself. Since $d_{k+1} \geq k + 1$, $|V \setminus \{v_{k+1}\}| = k$, and v_{k+1} is not adjacent to itself, it follows that v_{k+1} must be adjacent to some vertex u_1 not in V . Since $d_{k+2} \geq k + 2$, $|V \cup \{u_1\}| = k + 2$, and v_{k+2} is not adjacent to itself, v_{k+2} must be adjacent to some vertex u_2 not in $V \cup \{u_1\}$. In general, since $d_{k+i} \geq k + i$, $|V \cup \{u_1, u_2, \dots, u_{i-1}\}| = k + (i - 1)$, and v_{k+i} ($1 \leq i \leq k = 1$) is not adjacent to

itself, v_{k+i} must be adjacent to some vertex u_i not in $V \cup \{u_1, u_2, \dots, u_{i-1}\}$. Thus, the set $\{(v_{k+1}, u_1), (v_{k+2}, u_2), \dots, (v_{k+(k+1)}, u_{k+1})\}$ is a set of $k + 1$ independent edges and $m \geq k + 1 = \frac{g}{2}$. \square

A simple-minded, but computationally inefficient, approach for finding a maximum independent set in a graph (and, hence, its independence number α) is to list every subset of the vertex set of the graph, test each set to determine whether it is an independent set, and then choose one of these of maximum cardinality. For a graph G with $n = n(G) = |V(G)|$ vertices, 2^n subsets would have to be checked.

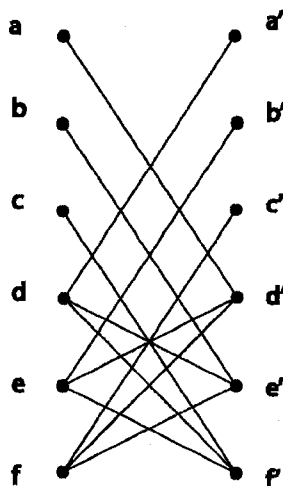


Figure 2.2: The sets $X = \{a, b, c, d, e, f\}$ and $Y = \{a', b', c', d', e', f'\}$ form a bipartition of the graph.

Given a maximum matching of a bipartite graph, it will be shown how to find a maximum independent set of the graph. This is a key step in the algorithm (discussed in Section 3.1) for finding a critical independent set in a graph. Once the matching number μ of a bipartite graph is known, the independence number α follows immediately from the König-Egervary Theorem. Finding a maximum independent set is not difficult but less obvious. The following is the author's own algorithm, used in his programs, for finding a maximum independent set in a bipartite graph, given a

maximum matching, together with a proof of the correctness of this algorithm.

Assume G is a bipartite graph with bipartition (X, Y) , and maximum matching M . Assume furthermore that G is connected: if G is not connected then the following algorithm can be run on each component of G . For each edge $xy \in M$, at most one of x or y can belong to a maximum independent set I of G . Since $\alpha(G) + \mu(G) = n(G)$, and since there must be $n - 2\mu$ vertices which are not incident to any edge in M (they are unsaturated by the edges of M), it follows that I must consist of these unsaturated vertices together with exactly one vertex from each edge in M .

Let $V(M)$ be the vertices saturated by M . Let $I_0 = V(G) \setminus V(M)$ be the unsaturated vertices. I_0 is an independent set: if I_0 were not independent, then M would not be maximal. Since X and Y are independent sets, each edge in M must be incident to one vertex in X and one in Y . If M is a perfect matching then there are no unsaturated vertices; moreover both X and Y are maximum independent sets.

The main idea of the algorithm presented below is simple: start with the set I_0 . Find its neighbors J_0 . These cannot be in a maximum independent set. These are saturated by the matching M . Hence, the vertices I_1 matched to J_0 by M must be in the maximum independent set; continue this process. Find the new neighbors J_1 of I_1 (that is, those which did not appear in the previously defined sets). These cannot be in a maximum independent set. These are saturated by the matching M . Hence, the vertices I_2 matched to J_1 by M must be in the maximum independent set; iterate this process until J_i is empty. It must terminate for finite graphs. Any remaining vertices cannot be adjacent to any of the I 's by construction. The remaining vertices are perfectly matched by edges in M . So X intersected with the remaining vertices must be an independent set containing one vertex from each of these remaining edges. These vertices together with the I 's must form a maximum independent set of the graph.

1. Let $i = 0$, $I'_0 = I_0 = V_0 = V \setminus V(M)$, $\bar{V}_0 = V \setminus V_0$, and $J_0 = N(I_0) \setminus V_0$.

2. If $J_i = \emptyset$. Stop.
3. $i := i + 1$. Let I_i be the vertices matched to J_{i-1} under M . $I'_i = I'_{i-1} \cup I_i$.
 $V_i = V_{i-1} \cup J_{i-1} \cup I_i$, $\bar{V}_i = V \setminus V_i$, and $J_i = N(I_i) \setminus V_i$. Return to Step 2.

After this algorithm stops, let $I = I'_i \cup (X \cap \bar{V}_i)$. It will be shown that I is a maximum independent set of G . It follows from the preceding observations that what must be shown is that (1) I is independent, that (2) it contains all vertices not covered by M , and that (3) it contains exactly one vertex from each edge of M . Since $I_0 \subseteq I'_i \subseteq I$ is the set of uncovered vertices, (2) is immediate.

In order to show (1) and (3), it will be shown that, following each iteration j , I'_j is independent and that the edges in M incident to vertices in I'_j cover all the vertices in $V_i \setminus V_{i-1}$. This means that these vertices are perfectly matched by M and, thus, that the vertices in $G[\bar{V}_j]$ are perfectly matched by M . Since $X \cap \bar{V}_j$ is a maximal independent set in $G[\bar{V}_j]$ and since, by construction, following the termination of the algorithm, no vertex in I'_j is adjacent to any vertex in \bar{V}_j , it follows that $I = I'_i \cup (X \cap \bar{V}_i)$ is a maximum independent set in G .

$I_0 = V_0$ is independent by construction. The vertices in I_0 are uncovered by M and must be in every independent set. J_0 is the set of neighbors of I_0 . These vertices cannot be in any independent set. I_1 is the set of vertices matched to J_0 under M . $V_1 = V_0 \cup J_0 \cup I_1$. Since no vertex in J_0 is contained in any maximum independent set and since a maximum independent set must contain one vertex from each edge of M , it follows that I_1 is independent, that I_1 must be contained in every maximum independent set of G , that $I'_1 = I_0 \cup I_1$ is an independent set, and that I'_1 must be contained in every maximum independent set of G . Also note that, for every edge in M that is incident to any vertex in V_1 , that edge is incident to a vertex in I'_1 . This is by construction: it was true for every edge in M incident to a vertex in $V_0 = I_0$, and the only edges in M incident to vertices in $V_1 \setminus V_0$ are incident to $I_1 \subseteq V_1 \setminus V_0$.

Assume that after j iterations of this algorithm that I'_j is independent, contained

in every maximum independent set of G and that each edge in M incident to any vertex in V_j is incident to a vertex in I'_j . If $J_j = \emptyset$, the algorithm terminates and we have shown what we set out to show. Otherwise, the algorithm continues. I_{j+1} is the set of vertices matched to J_j under M . $V_{j+1} = V_j \cup J_j \cup I_{j+1}$. Since no vertex in J_j is contained in any maximum independent set and since a maximum independent set must contain one vertex from each edge of M , it follows that I_{j+1} is independent, that I_{j+1} must be contained in every maximum independent set of G , that $I'_{j+1} = I_j \cup I_{j+1}$ is an independent set, and that I'_{j+1} must be contained in every maximum independent set of G . Also note that, for every edge in M that is incident to any vertex in V_{j+1} , that edge is incident to a vertex in I'_{j+1} . This is by construction: it was true for every edge in M incident to a vertex in V_j , and the only edges in M incident to vertices in $V_{j+1} \setminus V_j$ are incident to $I_{j+1} \subseteq V_{j+1} \setminus V_j$. Thus, truth of the claim holds for each step of the algorithm, proving the claim above.

For general graphs, the fundamental idea used in the fastest algorithms for finding maximum independent sets is due to Tarjan and Trojanowski [80]. Their idea is that, for a graph G and any vertex $v \in V(G)$, either v belongs to some maximum independent set of G or it does not. Then the problem of finding a maximum independent set can be divided into two smaller subproblems. If v is in some maximum independent set, then it is enough to find a maximum independent set in $G - v - N(v)$ and adding v to it; since $G - v - N(v)$ does not contain any neighbor of v , v cannot be adjacent to any vertex in any (independent) set of $G - v - N(v)$. If v is not in any maximum independent set then it is enough to find a maximum independent set in $G - v$. It also leads to a recursive solution to the problem of finding a maximum independent set. This leads to a solution of the problem of finding the independence number of G which can be formalized as follows: For any graph G and vertex $v \in V(G)$, $\alpha(G) = \max\{\alpha(G - v), 1 + \alpha(G - v - N(v))\}$.

An additional idea that can reduce branching in the recursion tree resulting from

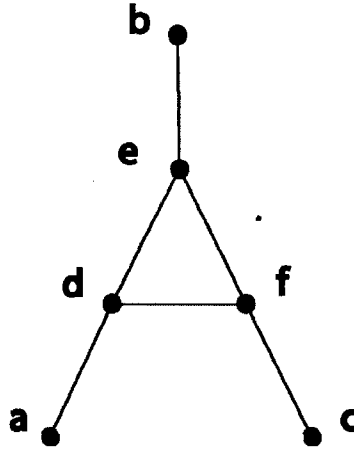


Figure 2.3: The three pendants a , b , and c can be removed and included in any maximum independent set.

the preceding procedure is to identify pendant vertices (vertices adjacent to exactly one other). If v is a pendant vertex of a graph G , then it can be included in a maximum independent set of G . If I is a maximum independent set of G , w is the unique vertex adjacent to v , and $w \in I$, then $I \setminus \{w\} \cup \{v\}$ is an independent set with the same cardinality. Thus, if G has a pendant vertex v , v can be included in any maximum independent set, and v and its neighbor can be removed without introducing any new branching: that is, $\alpha(G) = 1 + \alpha(G - v - N(v))$, and only the single subgraph $G - v - N(v)$ need be considered, rather than two subgraphs.

In the following section, a slight modification of the Tarjan-Trojanowski algorithm for finding a maximum independent set in a recursive graph is discussed. If a connected graph has adjacent pendant vertices then the graph must be the complete graph K_2 on two vertices. It clearly follows from the discussion above that for any connected graph G other than K_2 , having set P of pendant vertices, that $\alpha(G) = |P| + \alpha(G - P - N(P))$. In Chapter 3 a generalization of this idea is discussed.

2.2 Chemical Graph Theory and Fullerenes

Graphs were originally used to represent the structures of molecules in the 18th century. In the 19th century, graphs were used to represent their bonding structure, that is with an edge representing a chemical bond between two atoms. These structural drawings were called *graphical notation*. As mentioned in the Introduction, Sylvester initiated the use of the word *graph* as a shortening of *graphical notation*. Crum Brown may have been the first to use graphs representing molecular bonding structure as they are now used in chemistry. Two molecules with the same chemical composition can have different chemical properties: the molecules may have different bonding structures. They are different *isomers*. The problem of isomerism was identified in the 19th century and led to work on counting numbers of isomers, most famously Cayley's work in counting the numbers of alkane isomers [5, 78]. This marks the beginning of *chemical graph theory*. Graphs were used not only to represent molecular structure, but now theorems about graphs were conjectured and proved, motivated by chemical problems.

Isomer enumeration is one of four key areas that Rouvray [78] identifies in the development of chemical graph theory. Another is the development of topological indices. What chemists call "topological indices" are the same as what graph theorists call "invariants." The aims in chemical graph theory are to find indices or invariants that correlate with chemical properties such as boiling points, and also indices that distinguish between isomers, indices that give different values for the molecules in an isomer class. Such an index is said to be *non-degenerate*. In 1947, Wiener defined an index, now called the *Wiener index*, which was designed to measure, in some sense, the "connectivity" of a molecular graph and showed that it correlated with several physical properties [78]. Wiener is taken to have initiated this kind of investigation. Another well-known widely-studied index is the *Randić index*, introduced by Randić in 1975, designed to measure, in some sense, the "branching-ness" of a molecular

graph.

The historical development of these and other topological indices is also discussed in [4]. The independence number of the graph of a molecule corresponds to the largest set of atoms of the molecule, no pair of which are bonded. In their history Balaban and Ivanciuc do not give any examples of the use of the independence number α as a topological index. They do give an example of the use of a related invariant: Merrifield and Simmons investigated the number σ of independent sets in a graph. In [71] they give a recursive formula (similar to Tarjan and Trojanowski's recursive formula for α) for computing σ , and show that for the lower alkanes this invariant correlates both with the alkane heats of formation and their boiling points.

The use of the independence number as a topological index in chemical graph theory seems to have been initiated by a conjecture of Fajtlowicz's Graffiti program. Fajtlowicz and this author showed (in [33]) that minimizing this invariant appears to be a useful selector in identifying stable fullerene isomers. The experimentally characterized isomers with 60, 70 and 76 atoms uniquely minimize this number among the classes of possible structures with, respectively, 60, 70 and 76 atoms. Other experimentally characterized isomers also rank extremely low with respect to this invariant. These findings were initiated by a conjecture of the computer program Graffiti.

Fullerenes with a wide range of numbers of carbon atoms have been produced in experiment. Isomers with 60, 70, 76, 78, and 84 atoms have been produced in sufficient quantity to be characterized by NMR spectroscopy. The term "stable" is ambiguous, is used to refer alternately to thermodynamic and kinetic stability and, less formally, it is applied to those fullerenes that have actually been observed. These uses of the term are presumably related. For the present purposes fullerenes that have been produced in bulk quantity (or have been isolated) are referred to as *stable* fullerenes. These include at least $C_{60}(I_h)$, $C_{70}(D_{5h})$, $C_{76}(D_2)$, $C_{78}(D_3)$, $C_{78}(C_{2v})$ (2

kinds), $C_{84}(D_2)$ and $C_{84}(D_{2d})$ [11]. (In the numbering scheme of [39] these are $C_{60} : 1$, $C_{70} : 1$, $C_{76} : 1$, $C_{78} : 1$, $C_{78} : 2$, $C_{78} : 3$, $C_{84} : 22$ and $C_{84} : 23$.)

The problem for chemists is two-fold: characterizing those fullerenes that have been produced in experiment and predicting, for a given isomer class, which fullerenes are most likely to appear in future experiments. In each instance, the number of mathematically possible structures satisfying the fullerene hypothesis—that their carbon framework forms a trivalent polyhedron whose faces are either hexagons or pentagons—is enormous. Various rules-of-thumb have been proposed for reducing these numbers of possible isomers; the IPR hypothesis—that stable fullerenes have pentagonal faces which are isolated—is the most commonly used. Other rules-of-thumb include the maximum value of HOMO-LUMO [69], Raghavachari’s uniform hexagon environments criteria [74], and Fowler’s qualitative version of the same, the second moment of the hexagon neighbor signature [39]. The stability-independence hypothesis is that stable fullerenes tend to minimize their independence numbers.

Statistical evidence for the utility of this new rule-of-thumb was presented in [33], and is reproduced in Table 2.1. Since then we have learned that known facts about benzenoid stability support this rule-of-thumb. Benzenoids are finite regions of the infinite hexagonal lattice having carbon molecules at the vertices of this lattice. On one theory of fullerene formation, fullerenes are formed by folding up finite graphite sheets, that is, by folding up benzenoids. Stable benzenoids are known to have Kekule structures; mathematically, this means that their graphs have perfect matchings. Since benzenoids are alternant hydrocarbons (mathematically, that their graphs are bipartite), this implies that stable benzenoids with a given number of vertices n are ones that minimize their independence numbers. This is an important fact that has been overlooked, for instance in [40], which claims to refute the stability-independence hypothesis but which does not propose any better predictor of fullerene stability.

The energy predicted by the Hückel model is another predictor of fullerene sta-

Atoms	Isomer	# of Isomers	Independence Number	Rank	Max	Min
60	$C_{60}:1 (I_h)$	1812	24	1	28	24
70	$C_{70}:1 (D_{5h})$	8149	29	1	33	29
76	$C_{76}:1 (D_2)$	19151	32	1	36	32
78	$C_{78}:1 (D_3)$	24109	33	1 (3)	37	33
	$C_{78}:3 (C_{2v})$		34	2		
	$C_{78}:2 (C_{2v})$		33	1 (3)		
84	$C_{84}:22 (D_2)$	51592	36	1 (17)	40	36
	$C_{84}:23 (D_{2d})$		36	1 (17)		

Table 2.1: *Independence Number* data for experimentally produced fullerenes. *Rank* is by smallest value of *Independence Number*. *Max* and *Min* are the largest and smallest values within the corresponding class. The numbers in parentheses record the number of isomers that share the corresponding rank or value.

bility: stable fullerenes should minimize their energy. The relative Hückel energy (the relative molecular energy predicted by the Hückel theory) is a purely topological invariant. It is shown here that this invariant compares unfavorably with the independence number as a predictor of fullerene stability. While the Hückel model is a simplification that does not take into consideration π -electron interactions, Gutman and Polansky note that “the HMO total π -electron energy is in a perfect linear correlation with the kinetic energy of π -electrons as calculated by rather accurate (STO-3G) ab initio methods” [47, p. 136]. The ordering of the molecules in an isomer class with respect to this energy can be found by computing the following topological invariant E_h , where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of the n -atom molecule. In the case where n is even,

$$E_h = \sum_{i=1}^{\frac{n}{2}} \lambda_i$$

Atoms	Isomer	# of Isomers	Hückel Energy	Rank	Min	Max
60	$C_{60}:1 (I_h)$	1812	93.1616	24	93.0745	93.4768
70	$C_{70}:1 (D_{5h})$	8149	108.8136	1	108.8136	109.2523
76	$C_{76}:1 (D_2)$	19151	118.3267	21	118.2711	118.6952
78	$C_{78}:1 (D_3)$	24109	121.5358	282	121.3839	121.8600
	$C_{78}:3 (C_{2v})$		121.5909	2289		
	$C_{78}:2 (C_{2v})$		121.5576	657		
84	$C_{84}:22 (D_2)$	51592	130.9796	564	130.7431	131.3079
	$C_{84}:23 (D_{2d})$		131.0442	5356		

Table 2.2: *Hückel energy* data for experimentally produced fullerenes. *Rank* is by smallest value of Hückel energy. *Min* and *Max* are the smallest and largest values within the corresponding class.

and if n is odd,

$$E_h = \sum_{i=1}^{\frac{n+1}{2}} \lambda_i$$

The computed rankings are presented in Table 2.2.

Computing the independence number of a graph is, in general, a computationally intractable problem. It is known that computing whether the independence number of a cubic planar graph is less than a given number is NP-complete [42], it is not known whether this is the case for fullerenes, which are cubic and planar but which have additional structure. For the relatively low-order fullerene IPRs, it is often easy to compute the independence number. An upper bound for this number may be computed by noting that no more than two vertices of any pentagonal face can belong to a maximum independent set—so the pentagonal faces may contribute no more than 24 vertices to a maximum independent set. Removing the pentagonal faces, it may then be easy to compute the independence number of the remaining subgraph. The independence number of the fullerene can be no more than 24 plus the sum of the

independence numbers for these components. In Figure 2.1, the highlighted set of vertices demonstrated that a 24 element independent set of vertices exists for C_{60} and thus that the independence number is 24.

Ramras proved that, for any planar cubic graph whose vertices can be covered by disjoint pentagons the independence number is $\frac{2n}{5}$ [75]. Thus, the independence number for the graph of Buckminsterfullerene is 24—which Ramras showed, using this graph as an illustration even before the molecule had ever been isolated. One underlying idea is that an upper bound for the independence number of a graph is the sum of the independence numbers of any collection of disjoint induced subgraphs of a graph that contain all of the vertices of that graph. This idea can be used to find the independence numbers for each of the eight stable isomers highlighted above [41].

In the case of $C_{100}:321(T)$ (Fig. 2.4), the vertices not belonging to the pentagons form a connected subgraph in the shape of a “peace sign”—there is an outer cycle, together with a center point from which three paths extend and meet this cycle. It is not difficult to show that the independence number of this subgraph is 19. Thus, the independence number of the graph of this fullerene is no more than $19 + 24 = 43$. It is also not difficult to find an independent set which realizes this upper bound.

Given that our graphs represent fullerenes and thus are cubic (or trivalent, that is, that all vertices have degree three), after an initial vertex v is chosen and the subgraphs formed by removing it and removing both it and its neighbors are considered, these subgraphs as well as every subsequent subgraph considered in the recursive algorithm must contain an isolated vertex, a pendant vertex, or one joined to exactly two others (a vertex of degree two). After the first step of the algorithm, and the removal and inclusion of isolated and pendant vertices in a largest independent set, either the algorithm terminates or the remaining non-empty subgraphs contain a vertex w of degree two. In the latter case, consider a neighbor u of w . A modification of the stan-

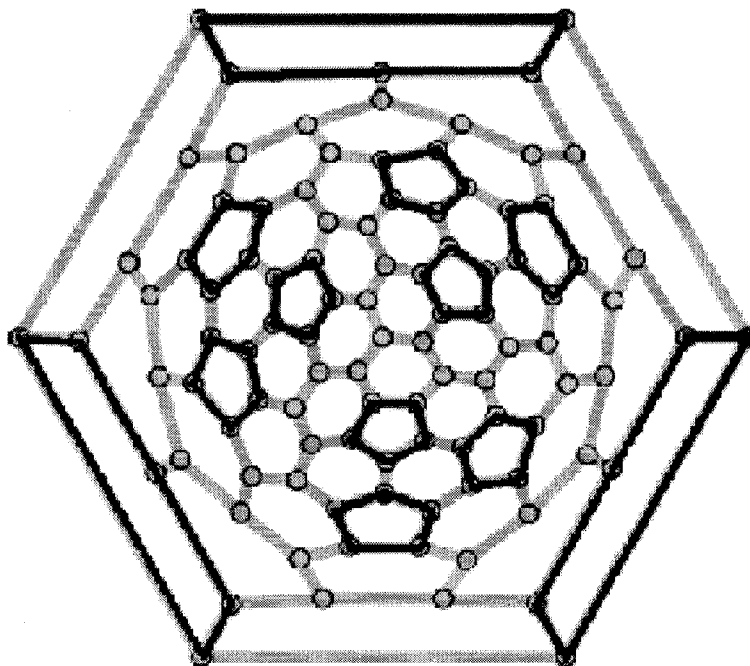


Figure 2.4: IP isomer $C_{100}:321(T)$ with pentagons highlighted. At most two vertices from each pentagon can be included in any maximum independent set: at most 24 of vertices covered by the pentagons can be included in a maximum independent set. The graph induced by the remaining vertices has independence number 19. Thus the independence number of this fullerene is no more than $24+19=43$.

standard algorithm is to reduce the problem to the problem of finding the independence number of the graph formed by removing u and the one formed by removing u and its neighbors. In the first of these graphs, w is a pendant vertex and can be removed immediately. The size of the problem is reduced by two vertices instead of one. This algorithm made it possible to compute independence numbers of fullerenes with up to 100 atoms in a reasonable amount of time. The algorithm follows.

1. Let v be any vertex of G . Find $\alpha(G - v)$ and $\alpha(G - v - N(v))$. Then $\alpha(G) = \max\{\alpha(G - v), 1 + \alpha(G - v - N(v))\}$.
2. If G is the empty graph, stop.
3. If G has an isolated vertex v , find $\alpha(G - v)$. $\alpha(G) = 1 + \alpha(G - v)$.
4. If G has a pendant vertex v , find $\alpha(G - v - N(v))$. $\alpha(G) = 1 + \alpha(G - v - N(v))$.
Return to Step 2.
5. If G has a vertex v of degree two, and w is a neighbor of v , find $\alpha(G - w)$ and $\alpha(G - w - N(w))$. Then $\alpha(G) = \max\{\alpha(G - w), 1 + \alpha(G - w - N(w))\}$. Return to Step 2.

As mentioned the only difference between this algorithm and the standard one is to search for a degree two vertex in the branching step (Step 5): for a connected cubic graph, after the initial branching step, each subgraph is guaranteed to have a vertex of degree zero, one or two. The independence numbers for fullerenes were computed both with the unmodified recursive algorithm and the cubic graph modification discussed in Section 2.1. In all cases the computations agreed. In empirical testing on fullerenes the modified algorithm was roughly twice as fast as the basic recursive algorithm. The computed independence numbers were checked against a large number of special and common graphs whose independence numbers are known, including Paley graphs up to 200 vertices, PR[n] graphs (defined in [30]) up to 200 vertices; and they

agree with the independence numbers for the stable fullerenes calculated by hand, as outlined above. Furthermore, that the isomers we have identified as minimizing their independence numbers in a given isomer class do minimize their independence numbers was checked in two ways. Firstly, the program implementing the algorithm above produced sets of vertices of cardinality corresponding to the computed independence number—and the independence of these vertices in the graph of the corresponding isomer was confirmed. Secondly, this was checked by using other algorithms—either a greedy algorithm or the unmodified recursive algorithm—to confirm that the other isomers had larger independence numbers. These computations also agree with the original computations of this author, utilizing the unmodified independence algorithm, whose results were confirmed for selected fullerenes by programs of Brendan McKay and Wendy Myrvold.

2.3 Cut Vertices, the Radius, and Independence Bounds

In this section upper and lower bounds for the independence number in terms of the number of cut vertices of a graph are proved. A vertex is a *cut vertex* if its removal from the graph (together with the edges incident to it) results in an increase in the number of components; that is, more formally, v is a cut vertex of G if the number of components of $G - v$ is greater than the number of components of G . Fajtlowicz's Graffiti program conjectured that $\frac{C}{2} + k \leq \alpha$, where C is the number of cut vertices, k is the number of components, and α is the independence number. This conjecture led to the following result:

$$\frac{C}{2} + 1 \leq \alpha \leq n - \frac{C}{2} - \frac{1}{2}.$$

n is the number of vertices of the graph. The proof of the lower bound is originally due to Greg Henry and Ryan Pepper and first appears in Pepper's dissertation [73].

Graffiti's original conjecture follows immediately from this lower bound. A new, short and illuminating proof is given here. The proof of the upper bound is due to Fajtlowicz and independently to this author together with Pepper; the proof given here is new and used in giving a characterization of when the cut vertices upper bound equals the independence number.

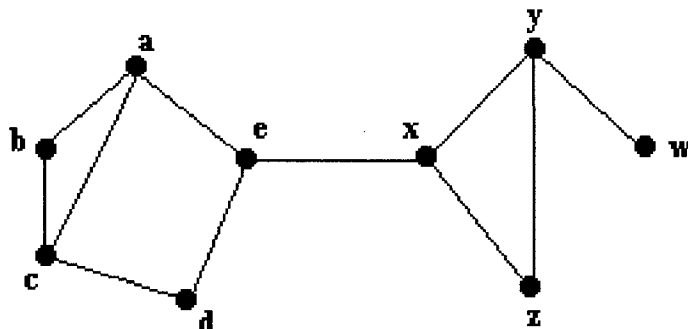


Figure 2.5: Cut vertices. The vertices e , x , and y are cut vertices. Removing any one of them will result in a graph with more than one component.

It is of theoretical and possibly practical interest to characterize those graphs where the equality of a bound holds. Several proofs of Graffiti's early conjecture that the radius of a graph is no more than its independence number have been found. The *eccentricity* of a vertex of a connected graph is the maximum of the distances between that vertex and each of the other vertices of the graph. The *radius* of a connected graph is the minimum eccentricity of any of its vertices. Characterizing those connected graphs where the radius r of the graph is equal to its independence number α was mentioned as an open problem in [35]; a solution is given below.

Graffiti made a conjecture which implied that graphs where $r = \alpha$ have Hamiltonian paths. A graph has a *Hamiltonian path* if there is a path which includes every vertex. This was known for graphs where the radius is no more than three: the case where $r = 2$ is a consequence of a theorem of Chvatal and Erdős [12], where the case

when $r = 3$ was proved by Fajtlowicz (in an unpublished note). The general case was proved by DeLaVina, Pepper and Waller in [17]. This author conjectured that if $r = \alpha$ as well as contained a $2r$ -cycle then the graph is Hamiltonian. A graph is *Hamiltonian* if it contains a spanning cycle. The proof of this conjecture, building on the characterization of those graphs where $r = \alpha$, for the case where $r > 4$, is below.

2.3.1 Cut Vertex Lower Bound

The main idea of the following new proof of the cut vertex lower bound is that, starting from any non-cut vertex v , and traveling to any cut-vertex, there must be at least one vertex on the “other side” of this cut-vertex. For each cut-vertex choose one of these. For the set of cut vertices at an even distance from v , this set of vertices on the other side will be independent; similarly, this will be true for the set of vertices at an odd distance from v . One of these sets will be at least half the size of the set of cut vertices. Adding v to this set gives an independent set at least as large as the lower bound.

Theorem 2.2. (*Henry and Pepper*) *If G is a connected graph then $\alpha \geq \frac{C}{2} + 1$.*

Let a *connector* be a vertex which is not a cut-vertex. Let $C = C(G)$ be the number of cut-vertices of a graph G . Every graph contains a connector. Every connected graph with more than one vertex contains at least two connectors: if v and w are vertices at maximum distance from each other in a graph, then they are both connectors. We say vertices v and u are *separated by w* if vertices v and u are in different components of $G - w$.

If v is a vertex of a connected graph G , and w is any vertex, let

$$D_v(w) = \{u \in N(w) \mid u \text{ and } v \text{ are separated by } w \text{ and } d(v, u) = d(v, w) + 1\}.$$

If w is a cut vertex, then the set $D_v(w)$ contains the neighbors u of w where every path from v to u must pass through w . (These sets were originally defined as singleton

sets, each containing a choice of a vertex from the present sets. Fajtlowicz suggested that consideration of the present sets were of more general interest). Let $G_{v,w}$ be the graph induced on $D_v(w)$. So $G_{v,w} = G[D_v(w)]$.

$$\mathcal{O}_v = \{D_v(w) | w \text{ is a cut vertex of } G \text{ and } d(v, w) \text{ is odd} \}$$

$$\mathcal{E}_v = \{D_v(w) | w \text{ is a cut vertex of } G \text{ and } d(v, w) \text{ is even} \}$$

Note that the number C of cut vertices of G equals $|\mathcal{O}_v| + |\mathcal{E}_v|$, and that $\max\{|\mathcal{O}_v|, |\mathcal{E}_v|\} \geq \frac{C}{2}$.

Having described the necessary preliminaries, the new proof can now be given:

Proof. Let v be a connector. For each cut vertex w , let w' be a vertex in $D_v(w)$. Let O'_v be a choice of a vertex from each set in \mathcal{O}_v . O'_v is an independent set and $|O'_v| = |\mathcal{O}_v|$. Similarly let E'_v be a choice of a vertex from each set in \mathcal{E}_v . E'_v is an independent set and $|E'_v| = |\mathcal{E}_v|$. No element in O'_v or E'_v is adjacent to v .

Thus,

$$\alpha \geq \max\{|O'_v \cup \{v\}|, |E'_v \cup \{v\}|\} = \max\{|\mathcal{O}_v| + 1, |\mathcal{E}_v| + 1\} \geq \frac{C}{2} + 1.$$

□

2.3.2 Cut Vertex Upper Bound & Characterization of Equality

The upper bound was proved independently by several different people. A new proof, due to this author and presented below, begins by noting some necessary facts and drawing out their consequences. These ideas are then used in characterizing those graphs where the cut vertices upper bound equals the independence number of the graph.

Lemma 2.3. *For every tree with at least two vertices,*

$$\alpha \leq n - \frac{C}{2} - \frac{1}{2}.$$

Proof. The only tree in which two pendant vertices are adjacent is the path on two vertices—for which the theorem holds. Assume T is a tree with more than two vertices. Let I be a maximum independent set containing all pendant vertices. Let P be the set of pendant vertices and let $p = |P|$. Note that any vertex in $I \setminus P$ has degree two or more. Thus, the number e' of edges incident to vertices in I is at least

$$2|I \setminus P| + |P| = 2(\alpha - p) + p = 2\alpha - p.$$

Since, for any tree, $p = n - C$,

$$n - 1 = e \geq e' \geq 2\alpha - p = 2\alpha - n + C.$$

Rearranging gives the desired inequality. □

Theorem 2.4. *For every connected graph with at least two vertices,*

$$\alpha \leq n - \frac{C}{2} - \frac{1}{2}.$$

Proof. The theorem is true for the path on two vertices. Let G be a connected simple graph with more than two vertices. Let T be any spanning tree of G . Since, $n(G) = n(T)$, $C(G) \leq C(T)$, and $\alpha(G) \leq \alpha(T)$, by the previous lemma we have

$$n(G) - \frac{C(G)}{2} - \frac{1}{2} \geq n(T) - \frac{C(T)}{2} - \frac{1}{2} \geq \alpha(T) \geq \alpha(G).$$

□

Fajtlowicz found another proof of the upper bound by noting its equivalence to the following conjecture of Graffiti:

Conjecture 2.5. *For any connected graph, $L \geq n + 1 - 2\mu$.*

Here, L is the maximum number of pendants of any spanning tree of G and μ is the matching number of G , the maximum cardinality of a set of non-incident edges of G . Fajtlowicz's proof of this conjecture can be found in Ermelinda DeLaVina's list "Written on the Wall II" of conjectures of Graffiti.¹

Equality holds for, and only for, odd trees.

Definition 2.6. A branching point in a tree is a vertex of degree greater than or equal to three. An odd tree is an odd path (a path with an odd number of vertices) or a non-path tree where the distance from any branching point to each pendant vertex is odd.

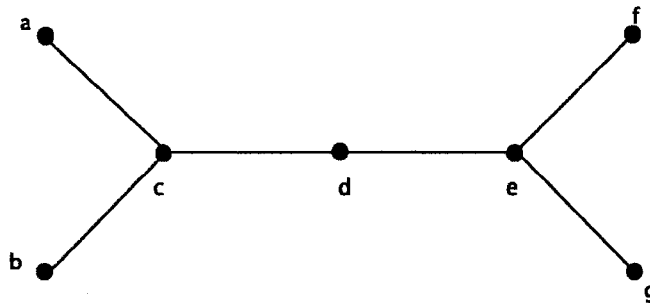


Figure 2.6: An odd tree. The branching points are c and e . The pendant vertices are a , b , f and g . The distance from any branching point to any pendant vertex is odd.

Clearly, the distance between any two pendant vertices in an odd tree is even.

Lemma 2.7. For every tree T , $\alpha = n - \frac{C}{2} - \frac{1}{2}$ if, and only if, T is an odd tree.

Proof. Suppose the theorem is true for all trees with no more than k vertices. Let T be a tree with $k + 1$ vertices. The theorem is true for paths, so it will be assumed that T contains a branching point. Let v be a pendant vertex such that the distance from v to its nearest branching point w is maximized.

¹This list can be found on the WWW at: <http://cms.dt.uh.edu/faculty/delavinae/research/wowII/>

If v and w are adjacent, then every pendant must be adjacent to a branching point. Since any tree has more pendant vertices than branching points, there must be a branching point adjacent to two or more pendant vertices. Assume that w is such a branching point. Let $T' = T - v$. In this case, $C(T') = C(T)$, $\alpha(T') = \alpha(T) - 1$, and $n(T') = n(T) - 1 = k$.

Suppose that, $\alpha(T) = n(T) - \frac{C(T)}{2} - \frac{1}{2}$. By substitution, it follows that, $\alpha(T') = n(T') - \frac{C(T')}{2} - \frac{1}{2}$ and, by the inductive assumption, that T' is an odd tree. Since T' is an odd tree, and T is formed by adding a pendant vertex to a branching point that is adjacent to a pendant vertex in T' , T is also an odd tree.

Now suppose that T is an odd tree. Then T' is also an odd tree, since T' was formed by removing a pendant vertex adjacent to a branching point of T . By assumption, $\alpha(T') = n(T') - \frac{C(T')}{2} - \frac{1}{2}$, and the result follows by substitution.

Now consider the case where v and w are not adjacent. Let v' be the unique (degree two) neighbor of v . Let $T' = T - \{v, v'\}$. In this case, $\alpha(T') = \alpha(T) - 1$, $n(T') = n(T) - 2$ and $C(T') = C(T) - 1$, in the case where v' is adjacent to w or, otherwise, $C(T') = C(T) - 2$.

Suppose that $\alpha(T) = n(T) - \frac{C(T)}{2} - \frac{1}{2}$. If $C(T') = C(T) - 1$, by substitution it follows that $\alpha(T') = n(T') - \frac{C(T')}{2}$. But, by Theorem 2.4, $\alpha(T') \leq n(T') - \frac{C(T')}{2} - \frac{1}{2}$. So $C(T') = C(T) - 2$, and $\alpha(T') = n(T') - \frac{C(T')}{2} - \frac{1}{2}$. By the inductive assumption, T' is an odd tree and, since T' was formed by removing a pendant vertex and its degree two neighbor v' from T , T is an odd tree.

Now suppose that T is an odd tree. Since the distance between v and w is odd, and v is not adjacent to w , it follows that T' is an odd tree and that $C(T') = C(T) - 2$. By the inductive assumption, $\alpha(T') = n(T') - \frac{C(T')}{2} - \frac{1}{2}$. The desired result follows immediately by substitution. \square

Theorem 2.8. *For every connected graph G , $\alpha = n - \frac{C}{2} - \frac{1}{2}$ if, and only if, G is an odd tree.*

Proof. Assume the theorem is true for graphs with no more than k vertices. Let G be a connected graph with $k + 1$ vertices. One direction is an immediate consequence of Lemma 2.7. Assume then that $\alpha(G) = n(G) - \frac{C(G)}{2} - \frac{1}{2}$.

Let T be a spanning tree of G . Since $\alpha(G) \leq \alpha(T)$, $n(G) = n(T)$, and $C(G) \leq C(T)$,

$$\alpha(G) \leq \alpha(T) \leq n(T) - \frac{C(T)}{2} - \frac{1}{2} \leq n(G) - \frac{C(G)}{2} - \frac{1}{2}.$$

The second inequality follows from Theorem 2.4. Since the first and last terms are assumed to be equal, all the terms are equal and $\alpha(G) = \alpha(T)$ and $C(G) = C(T)$. Since every cut-vertex in G is a cut-vertex in T , it follows that the sets of cut-vertices in G and T are identical. Finally, since $\alpha(T) = n(T) - \frac{C(T)}{2} - \frac{1}{2}$, Lemma 2.7 implies that T is an odd tree. This argument also shows that *any* spanning tree of T must be an odd tree.

If G is a path, then the result follows from Lemma 2.7. Otherwise, there is a spanning tree of G with a branching point. Assume T has a branching point. Since there are more pendants than branching points in any tree, there is a branching point v and pendant vertices u and w , such that the unique (odd-length) paths from u to v and from v to w contain no branching points.

There are three cases to consider: (1) the case where both u and w are adjacent to v , (2) the case where both u and w are adjacent to degree-two neighbors, and (3) the case where exactly one of u or w is adjacent to a degree-two neighbor.

If u is adjacent to v in T , let $G' = G - u$. Then $C(G') = C(G)$. If u is not adjacent to v , then there is a unique degree-two neighbor of u in T . Call this vertex u' . In this case let $G' = G - \{u, u'\}$. Here, $C(G') = C(G) - 2$. Because u is in every maximum independent set of T and $\alpha(G) = \alpha(T)$, u is in every maximum independent set of G . So, in either case, $\alpha(G') = \alpha(G) - 1$, and it follows that $\alpha(G') = n(G') - \frac{C(G')}{2} - \frac{1}{2}$. The inductive hypothesis implies that G' is an odd tree. Note that w is a pendant in G' . Thus the only vertices w can be adjacent to in G are its unique neighbor in

T , u and, if u has a degree-two neighbor in T , u' . Arguing symmetrically, the only vertices that u can be adjacent to in G are its unique neighbor in T , w and, if w has a degree-two neighbor in T , that neighbor (call it w').

In all three cases there can be no edge in G between u and w . If u and w were adjacent, it is easy to modify T to form a spanning tree of G which is not an odd tree, contradicting the fact that any spanning tree of T is an odd tree. In case (2) there cannot be edges in G between u and w' or between w and u' for the same reason. In case (3), assume that u is adjacent in G to v , that w is adjacent to a degree-two vertex w' and that u is adjacent to w' . Then there is at least one degree-two vertex on the path between w' and v : this vertex is a cut vertex in T but not in G , contradicting the fact that the set of cut vertices is the same in both graphs, proving the result.

Thus, G has no more edges than its spanning tree T , is identical to it and is an odd tree, proving the result.

□

2.3.3 Characterizing when Independence equals Radius

The *eccentricity* of a vertex of a connected graph is the maximum distance from the vertex to any other vertex of the graph. The *radius* of the graph is the minimum eccentricity of the vertices of the graph. Let $r = r(G)$ be the radius. The computer program Graffiti conjectured [30] that, for any connected graph, $\alpha \geq r$. This conjecture follows immediately from the Induced Path Theorem, whose proof in a paper of Erdős, Saks and Sos, [19, Thm. 2.1] is credited to Fan Chung. In [22] Fajtlowicz mentions four different proofs of this conjecture as of 1988. Fajtlowicz and Waller provide one proof in [35] and remark that characterizing the case of equality remained open.

An *r-ciliate* is a cycle with $2q$ ($q \geq 1$) vertices and appended to each of these vertices is a path with $r - q$ vertices. They are denoted $C_{2q, r-q}$ (see Figure 2.7). It

is easy to see that r -ciliates are bipartite and that $\alpha(C_{2q,r-q}) = \frac{|V(G)|}{2} = qr - q^2 + q$. In the case where $q = 1$, the cycle is degenerate and identical to the path on two vertices. Clearly, $r \geq q$. In the extreme cases, where $q = 1$ and $r = q$, the r -ciliate is a path or a cycle, respectively.

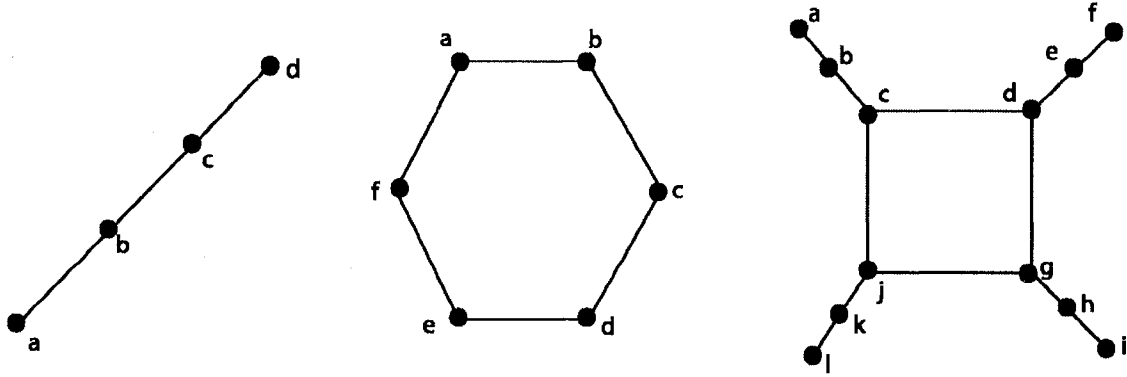


Figure 2.7: r -ciliates. The graphs are, from left to right, $C_{2 \times 1, 2-1} = C_{2,1}$, $C_{2 \times 3, 3-3} = C_{6,0}$, $C_{2 \times 2, 4-2} = C_{4,2}$. These graphs are radius critical. If a vertex is removed in any of these graphs that does not disconnect the graph, the radius will decrease.

A connected graph G is *radius-critical* if v is any non-cut vertex, then $G - v$, the subgraph formed by deleting v and the edges incident to it, has radius $r - 1$. Fajtlowicz proved that a graph is radius critical if, and only if, it is a ciliate [22]. The radius of an r -ciliate is r . It follows that every connected graph with radius r contains an induced r -ciliate. Fajtlowicz noted that a connected graph with radius r , independence number α , with $\alpha = r$, necessarily contains an induced path with $2r$ vertices (a $2r$ -path) or an induced cycle with $2r$ vertices (a $2r$ -cycle) [22]. This result is the foundation for the characterization of those connected graphs whose independence number equals its radius.

If G is a connected graph with radius r , a radius-critical graph having radius r can be found by removing vertices until it is no longer possible to remove a vertex without either disconnecting the graph or without decreasing its radius: at some point

all remaining vertices are either cut vertices or will result in a decrease of the radius. By Fajtlowicz's theorem, this process yields an r -ciliate. Let R be the set of vertices of an induced r -ciliate. Let $R' = V(G) \setminus R$. Since r -ciliates are bipartite, let (B, W) be the bipartition of the vertices (so $R = B \cup W$, and $|B| = |W|$); a vertex in B is *black*, while a vertex in W is *white*.

The following notation is used here: the vertex set of a graph G is $V(G)$, the set of neighbors of a set $S \subseteq V(G)$ is $N(S)$, the set of neighbors of S that are in set T is $N_T(S) = N(S) \cap T$, and the graph induced on a set $S \subseteq V(G)$ is $G[S]$. All graphs are assumed to be finite and simple.

The following characterization focuses on properties of sets of independent "external" vertices. Another characterization, long in progress, due to DeLaVina, Pepper, Waller and this author, focusing on the relationships between individual external vertices, has just been completed.

Theorem 2.9. *If G is a connected graph, with independence number $\alpha = \alpha(G)$, radius $r = r(G)$, where R is the vertex set of an induced r -ciliate, with bipartition (B, W) of $G[R]$ (vertices in B are black and vertices in W are white), and $R' = V(G) \setminus R$ then, $\alpha = r$ if, and only if, G satisfies the following conditions:*

1. *G contains an induced $2r$ -path or an induced $2r$ -cycle.*
2. *Every independent set I' in $G[R']$ has at least $|I'|$ white neighbors and at least $|I'|$ black neighbors (and $|N_R(I')| \geq 2|I'|$).*
3. *For every independent set I' of $G[R']$, with $|N_R(I')| = 2|I'| + k'$, the number of odd components in $G[R \setminus N(I')]$ is no more than k' .*

Proof. Assume that G is a connected graph G , with independence number $\alpha = \alpha(G)$, and radius $r = r(G)$, where R is the vertex set of an induced r -ciliate and $R' = V(G) \setminus R$.

We first show that, if $\alpha(G) = r(G)$, then conditions 1, 2, and 3 necessarily follow.

That G must satisfy condition 1 is an observation of Fajtlowicz. Indeed, every graph has an induced r -ciliate of the form $C_{2q, r-q}$. It is easy to see that r -ciliates are bipartite and that $\alpha(C_{2q, r-q}) = qr - q^2 + q$. So, $\alpha(G) \geq qr - q^2 + q$. $qr - q^2 + q = r$ if, and only if, $q = 1$ or $q = r$. In the first case, G has an induced $2r$ -path, while in the second case, G has an induced $2r$ -cycle.

Suppose I' is an independent set of $G[R']$. Suppose, that I' has fewer than $|I'|$ white neighbors, that is, $|N(I') \cap W| < |I'|$. Let $N_W = N(I') \cap W$. Let $I = (W \setminus N_W) \cup I'$. Then I is an independent set and $\alpha(G) \geq |I| = |W| - |N_W| + |I'| > |W| - |N_W| + |N_W| = |W| = r(G)$ which contradicts the assumption that $\alpha(G) = r(G)$. Analogous reasoning shows that I' has at least $|I'|$ black neighbors. So, G must also satisfy condition 2.

It is worth noting that condition 2 implies that every vertex $v \in R'$ is adjacent to at least one black vertex in R and at least one white one.

Suppose I' is an independent set in $G[R']$. By condition 2, $|N_R(I')| = 2|I'| + k'$, for some integer $k' \geq 0$. (If $I' = \emptyset$, then $k' = 0$.) So, $|R \setminus N(I')| = |R| - 2|I'| - k' = 2r - 2|I'| - k'$. If $G[R]$ is a $2r$ -path, then the components of $G[R \setminus N(I')]$ are paths. If $G[R]$ is a $2r$ -cycle then either $N_R(I') = \emptyset$ or $N_R(I') \neq \emptyset$. If $N_R(I') = \emptyset$ then $I' = \emptyset$, and $\alpha(G) = \alpha(G[R]) = r(G)$. If $N_R(I') \neq \emptyset$ then the components of $G[R \setminus N(I')]$ are paths. If H is a graph which is a union of paths, k of which have an odd number of vertices, then $\alpha(H) = \frac{|V(H)| - k}{2} + k$. Let k be the number of components of $G[R \setminus N(I')]$ with an odd number of vertices.

Suppose the number k of odd paths in $G[R \setminus N(I')]$ is greater than k' (that is $k > k'$). Then,

$$\begin{aligned} \alpha(G[R \setminus N(I')]) &= \frac{|R \setminus N(I')| - k}{2} + k > \frac{|R \setminus N(I')| - k'}{2} + k' \\ &= \frac{|R| - |N_R(I')| - k'}{2} + k' = \frac{2r - (2|I'| + k') - k'}{2} + k' = r - |I'| + k'. \end{aligned}$$

Let J be a maximum independent set of $G[R \setminus N(I')]$. So $|J| = \alpha(G[R \setminus N(I')]) > r - |I'| + k' \geq r - |I'|$. Now, $J \cup I'$ is independent, so $\alpha(G) \geq |J \cup I'| = |J| + |I'| > (r - |I'|) + |I'| = r$. This contradicts the assumption that $\alpha(G) = r(G)$. It then follows that $k \leq k'$, and condition 3 is satisfied.

We now show that, if conditions 1, 2 and 3 are satisfied, then $\alpha(G) = r(G)$.

By condition 1, G contains an induced $2r$ -path or an induced $2r$ -cycle. Let R be the vertex set of this induced r -ciliate, with bipartition (B, W) and $R' = V(G) \setminus R$.

Let I be a maximum independent set of G and $I' = I \cap R'$. Condition 2 implies that $|N_R(I')| = 2|I'| + k'$, for some $k' \geq 0$. (If $I' = \emptyset$, then $k' = 0$.) So, $|R \setminus N(I')| = |R| - (2|I'| + k') = 2r - 2|I'| - k'$. If $G[R]$ is a $2r$ -path, then the components of $G[R \setminus N(I')]$ are paths. If $G[R]$ is a $2r$ -cycle and $I' \neq \emptyset$, then the components of $G[R \setminus N(I')]$ are also paths. Let k be the number of components of $G[R \setminus N(I')]$ with an odd number of vertices. (If $I' = \emptyset$, then $k = 0$.) So,

$$\begin{aligned} |I \cap R| &\leq \alpha(G[R \setminus N(I')]) = \frac{|R \setminus N(I')| - k}{2} + k \\ &= \frac{2r - (2|I'| + k') - k}{2} + k = r - |I'| - \frac{k'}{2} + \frac{k}{2} \leq r - |I'|. \end{aligned}$$

The last inequality follows as $k \leq k'$ since condition 3 is assumed to be satisfied and, thus, $|I \cap R| \leq r - |I'|$. So, $\alpha(G) = |I'| + |I \cap R| \leq |I'| + (r - |I'|) = r$. But, by the Induced Path Theorem (cited earlier in this section), $\alpha(G) \geq r(G)$. Thus, $\alpha(G) = r(G)$. \square

2.3.4 An application: Hamiltonicity

The previous characterization led the author to a conjecture and partial results for a new sufficient condition for the Hamiltonicity of a graph. The condition was known to be true for graphs of radius one and two, and believed to be true for graphs of any radius. A proof was submitted to the *Journal of Graph Theory*. The referee found

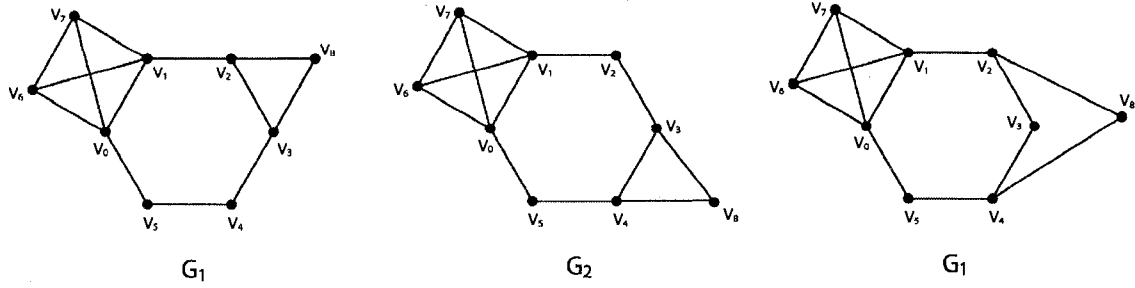


Figure 2.8: In all three graphs, $r = 3$ and $R = \{v_0, v_1, \dots, v_5\}$ induces a 6-cycle, with bipartition (B, W) , where $B = \{v_0, v_2, v_4\}$ and $W = \{v_1, v_3, v_5\}$. G_1 satisfies all the conditions of Theorem 2.9 and $\alpha(G_1) = r(G_1) = 3$. $\alpha(G_2) = 4 \neq 3 = r(G_2)$ as G_2 fails condition (3). In this case, for the independent set $I' = \{v_6, v_8\}$ in $G[R']$, $|N_R(I')| = 2|I'| + k'$, where $k' = 0$, but the number of odd paths in $G[R \setminus N(I')]$ is 2, which is greater than k' . $\alpha(G_3) = 4 \neq 3 = r(G_3)$ as G_3 fails condition (2). In this case, for the independent set $I' = \{v_8\}$ in $G[R']$, I' has fewer than $|I'|$ white neighbors.

errors with the cases where the radius equals three and four. These cases have not yet been remedied. The result for the case where the radius is greater than four stands and is included here. Finding sufficient conditions for graphs having Hamilton paths or cycles is a well-researched problem [43].

The distance between vertices v and w in a graph G is the length of a shortest path between v and w in G and is denoted $d_G(v, w)$. If n, n' and m are integers, with $0 \leq n < m$ and $0 \leq n' < m$, let $d(n, n') \pmod{m} = \min\{n - n' \pmod{m}, n' - n \pmod{m}\}$. If the modulus referred to is clear, then $d(v, w)$ may also be used.

Definition 2.10. For a graph with a distinguished set of vertices $R = \{v_0, v_1, \dots, v_{2r-1}\}$, let $N_{i,j}$ be the set of vertices adjacent to both v_i and v_j but not adjacent to any other vertex in R . If the graph $G[N_{i,j}]$ induced on these vertices is complete, let $P_{i,j}$ be a Hamilton path in $G[N_{i,j}]$. Let $N_{i,j,k}$ be the set of vertices adjacent to v_i, v_j and v_k in R , but not adjacent to any other vertex in R . If the graph $G[N_{i,j,k}]$ induced on these vertices is complete, let $P_{i,j,k}$ be a Hamilton path in $G[N_{i,j,k}]$.

Theorem 2.11. *For every connected graph where the independence number equals its radius r , having an induced path or cycle, $v_0, v_1, \dots, v_{2r-1}$ (with v_0 adjacent to v_{2r-1} in the latter case), on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$, then*

1. *the graph $G[N_{i,j}]$ is complete,*
2. *the graph $G[N_{i,j,k}]$ is complete,*
3. *every vertex in $N_{i,j}$ is adjacent to every vertex in $N_{j,k}$, and*
4. *every vertex in $N_{i,j}$ is adjacent to every vertex in $N_{i,j,k}$.*

Proof. Suppose G is a graph with independence number $\alpha = \alpha(G)$, radius $r = r(G)$, $\alpha = r$, and has an induced $2r$ -path, $v_0, v_1, \dots, v_{2r-1}$ (with v_0 adjacent to v_{2r-1} in the case that G has an induced $2r$ -cycle), on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$. Let $R' = V(G) \setminus R$. Let (B, W) be a bipartition of $G[R]$. The vertices in B are *black* and the vertices in W are *white*.

We first show (1) that $G[N_{i,j}]$ is complete. Suppose $v, w \in N_{i,j}$. Suppose they are not adjacent. Let $I = \{v, w\}$ be the set containing these vertices; I is an independent set in $G[R']$. Theorem 2.9 implies that I has at least $|I| = 2$ white neighbors and $|I| = 2$ black neighbors. But, by definition, the only neighbors of I in R are v_i and v_j , contradicting the theorem. Thus, v and w must be adjacent and $G[N_{i,j}]$ is complete.

The proof of (2) is similar. $G[N_{i,j,k}]$ is complete: if two vertices in $N_{i,j,k}$ are independent they must have at least four neighbors in R but, by definition, there are only three.

We will now show (3) that every vertex in $N_{i,j}$ is adjacent to every vertex in $N_{j,k}$. Suppose $v \in N_{i,j}$ and $w \in N_{j,k}$ and $I = \{v, w\} \subseteq R'$ is an independent set in $G[R']$. Then Theorem 2.9 implies that I has at least four neighbors in R but, by definition, the only neighbors of I in R are v_i, v_j , and v_k , which is a contradiction. Thus, v and w must be adjacent.

The proof of case (4) is analogous to the proof of case (3).

□

Theorem 2.12. *Every connected graph where the independence number equals its radius r , having an induced $2r$ -cycle, $v_0, v_1, \dots, v_{2r-1}$, on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$, is 2-connected.*

Proof. Assume G is a connected graph where the independence number equals its radius r , having an induced $2r$ -cycle, $v_0, v_1, \dots, v_{2r-1}$, on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$. Let $R' = V \setminus R$.

Let v , x and y be any vertices of G . It will be shown that there is a path from x to y in $G - v$. Theorem 2.9 implies that every vertex in R' is adjacent to at least two vertices in R . Thus, if x or y are in R' then it must be adjacent to at least one vertex in $R - v$. Since R induces a cycle in G , the graph induced on $R - v$ must be connected. Then either x and y belong to this connected induced subgraph or are adjacent to one of the vertices of this subgraph. Thus, the deletion of at least two vertices is required in order to disconnect G . □

Theorem 2.13. *For every connected graph where the independence number equals its radius r , $r > 3$, having an induced $2r$ -cycle, $v_0, v_1, \dots, v_{2r-1}$, on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$, if $v \in R' = V \setminus R$ then $v \in N_{i,i+1}$ or $v \in N_{i,i+1,i+2}$ (where $i \in \{0, 1, \dots, 2r - 1\}$ and all indices are assumed to be mod $2r$).*

Proof. Let G be a connected graph where the independence number equals its radius r , having an induced $2r$ -cycle, $v_0, v_1, \dots, v_{2r-1}$, on the set $R = \{v_0, v_1, \dots, v_{2r-1}\}$. Note that for $i, j \in \{0, 1, \dots, 2r - 1\}$, $d_G(v_i, v_j) \leq d(i, j) \leq r$. Let $R' = V \setminus R$ and let $v \in R'$. Let $W = \{v_0, v_2, \dots, v_{2r-2}\}$ and $B = \{v_1, v_3, \dots, v_{2r-1}\}$. (B, W) is a bipartition of the graph induced by R . A vertex in B is *black*, while a vertex in W is *white*.

Theorem 2.9 implies that v is adjacent to at least one white and one black vertex. Suppose that v is adjacent to v_i and v_j . It will be shown that $d(i, j) \leq 2$. Suppose it is not; that is, suppose $d(i, j) \geq 3$.

First consider the case where $3 < d(i, j) < r$. Assume $j = i + d(i, j) \pmod{2r}$. It will be shown that the eccentricity of v_i is no more than $r - 1$. Let $R_1 = \{v_i, v_{i+1}, \dots, v_{i+(d(i,j)-1)} = v_{j-1}\}$, $R_2 = \{v_{i-1}, v_{i-2}, \dots, v_{i-(r-2)} = v_{i+r+2}\}$, and let $R_3 = \{v_j, v_{j+1}, v_{j+2}, \dots, v_{j+r+1-d(i,j)} = v_{i+r+1}\}$. $R = R_1 \cup R_2 \cup R_3$. If $w \in R_1$, $d_G(v_i, w) \leq d(i, j) - 1 < r - 1$. So, $d_G(v_i, w) \leq r - 2$. If $w \in R_2$, $d_G(v_i, w) \leq r - 2$. If $w \in R_3$, $d_G(v_i, w) \leq d_G(v_i, v_j) + d_G(v_j, v_{j+r+1-d(i,j)}) \leq 2 + (r + 1 - d(i, j)) \leq r + 3 - d(i, j) \leq r - 1$. So if $w \in R$, $d_G(v_i, w) \leq r - 1$ and, for $w \in R - v_{i+r+1}$, $d_G(v_i, w) \leq r - 2$. If $w \in R'$ then Theorem 2.9 implies that it must be adjacent to at least two vertices in R and, in particular, to some vertex besides v_{i+r+1} . Assume $w \in R'$ and that w is adjacent to $v_l \in R - v_{i+r+1}$. Then $d_G(v_i, w) \leq d_G(v_i, v_l) + d_G(v_l, w) \leq (r - 2) + 1 = r - 1$. So the eccentricity of v_i is no more than $r - 1$ and $r(G) < r$.

Secondly, consider the case where $d(i, j) = r$, that is $j = i + r$, and $r > 3$. It will be shown that the eccentricity of v_i is no more than $r - 1$. Let $R_1 = \{v_i, v_{i+1}, \dots, v_{i+(r-2)}\}$, $R_2 = \{v_{i-1}, v_{i-2}, \dots, v_{i-(r-2)}\}$, and let $R_3 = \{v_{i+(r-1)}, v_{i+r} = v_j, v_{i+(r+1)} = v_{i-(r-1)}\}$. $R = R_1 \cup R_2 \cup R_3$. If $w \in R_1$, $d_G(v_i, w) \leq d_G(v_i, v_{i+(r-2)}) \leq (i + r - 2) - i \leq r - 2$. If $w \in R_2$, $d_G(v_i, w) \leq d_G(v_i, v_{i-(r-2)}) \leq r - 2$. If $w \in R_3$, $d_G(v_i, w) \leq 3 \leq r - 1$ (since $r > 3$). So if $w \in R$, $d_G(v_i, w) \leq r - 1$ and, for $w \in R \setminus \{v_{i+(r-1)}, v_{i-(r-1)}\}$, $d_G(v_i, w) \leq r - 2$. If $w \in R'$ then Theorem 2.9 implies that it must be adjacent to at least one white vertex and one black vertex. Since $v_{i+(r-1)}$ and $v_{i-(r-1)}$ are either both black or both white, w is adjacent to some vertex in $R \setminus \{v_{i+(r-1)}, v_{i-(r-1)}\}$. Assume $w \in R'$ and that w is adjacent to $v_l \in R \setminus \{v_{i+(r-1)}, v_{i-(r-1)}\}$. Then $d_G(v_i, w) \leq d_G(v_i, v_l) + d_G(v_l, w) \leq (r - 2) + 1 = r - 1$. So the eccentricity of v_i is no more than $r - 1$ and $r(G) < r$.

Lastly, consider the case where $d(i, j) = 3$. So, $v_j = v_{i+3}$. Let $R_1 = \{v_i, v_{i+1}, v_{i+2}\}$.

Let $R_2 = \{v_i, v_{i-1}, \dots, v_{i-(r-2)} = v_{i+r+2}\}$. Let $R_3 = \{v_{i+3}, v_{i+4}, \dots, v_{(i+3)+(r-4)} = v_{i+r-1}\}$. $R_1 \cup R_2 \cup R_3 = R \setminus \{v_{i+r}, v_{i+r+1}\}$. If $w \in R_1$, $d_G(v_i, w) \leq d_G(v_i, v_{i+2}) = 2 \leq r - 2$. If $w \in R_2$, $d_G(v_i, w) \leq d_G(v_i, v_{i-(r-2)}) = r - 2$. If $w \in R_3$, $d_G(v_i, w) \leq d_G(v_i, v) + d_G(v, v_{i+3}) + d_G(v_{i+3}, v_{(i+3)+(r-4)}) = 2 + (r - 4) = r - 2$. If $w \in R'$, Theorem 2.9 implies that it is adjacent to at least two vertices in R . If $w \in R'$ is adjacent to a vertex $v_l \in R_1 \cup R_2 \cup R_3$ then $d_G(v_i, w) \leq d_G(v_i, v_l) + d_G(v_l, w) \leq (r - 2) + 1 = r - 1$. So if no vertex in R' is adjacent only to vertices v_{i+r} and v_{i+r+1} , then the eccentricity of v_i is no more than $r - 1$ and $r(G) \leq r - 1$, contradicting the fact that $r(G) = r$. Assume, then, that there is a vertex $w_i \in R'$ adjacent only to vertices v_{i+r} and v_{i+r+1} .

By a parallel argument it can be shown that the eccentricity of v_{i+4} is no more than $r - 1$ unless there is a vertex $w_{i+4} \in R'$ adjacent only to vertices $v_{(i+4)+(r-1)}$ and $v_{(i+4)+r}$. So assume that there is a vertex $w_{i+4} \in R'$ adjacent only to these vertices. Here we have $d_G(v_{i+4}, v_{(i+4)+(r-1)}) \leq d_G(v_{i+4}, v_{i+r}) + d_G(v_{i+r}, w_i) + d_G(w_i, w_{i+4}) + d_G(w_{i+4}, w_i) = (r - 4) + 1 + 1 + 1 = r - 1$. So if w_i is adjacent to w_{i+4} then there is a path of length less than r from v_{i+4} to w_{i+4} , and the eccentricity of v_{i+4} is less than r and $r(G) < r$, contradicting the fact that $r(G) = r$. Thus, $I' = \{w_i, w_{i+4}\}$ is an independent set in R' . Then Theorem 2.9 implies that $|N_R(I')| = 2|I'| + k'$, for some non-negative integer k' . Here $N_R(I') = \{v_{i+r}, v_{i+r+1}, v_{(i+4)+(r-1)}, v_{(i+4)+r}\}$. So $|N_R(I')| = 4$ and $k' = 0$. But Theorem 2.9 also implies that the number of paths with an odd number of vertices in $G[R \setminus N(I')]$ is no more than $k' = 0$. But the number of paths with an odd number of vertices in $G[R \setminus N(I')]$ is two. It is enough to show that there is a single path in $G[R \setminus N(I')]$ which has an odd number of vertices. $v_{i+r+2} \in R \setminus N(I')$ but $v_{i+r+1}, v_{i+r+3} \in N(I')$. Thus, the path induced on the single vertex $\{v_{i+r+2}\}$ is an odd component of $G[R \setminus N(I')]$. So the assumption that $d(i, j) = 3$ yields a contradiction.

In each case the assumption that v is adjacent to vertices $v_i, v_j \in R$ where $d(i, j) \geq 3$ leads to a contradiction. So $d(i, j) \leq 2$. If $v \in R'$, Theorem 2.9 implies that it is

adjacent to at least two vertices in R , at least one of which is white and one of which is black. If v_i is white then the only possibilities for a black vertex are either v_{i-1} or v_{i+1} . v may be adjacent to at most one more vertex in R : if v is adjacent to both v_i and v_{i+1} then v may also be adjacent to either v_{i-1} or v_{i+2} , and if v is adjacent to both v_i and v_{i-1} then v may also be adjacent to either v_{i-2} or v_{i+1} . So, if $v \in R'$, adjacent to $v_i \in R$, then v must be in $N_{i,i+1}$, $N_{i-1,i}$, $N_{i,i+1,i+2}$, $N_{i-1,i,i+1}$, or $N_{i-2,i-1,i}$. All of these sets have one of two forms: either there are two indices which are consecutive integers, or there are three indices which are consecutive integers. Thus, if $v \in R'$ then $v \in N_{j,j+1}$ or $v \in N_{j,j+1,j+2}$, for some integer $j \in \{0, 1, \dots, 2r - 1\}$, with all indices assumed to be mod $2r$. \square

Fajtlowicz's Graffiti program made the following conjecture:

Conjecture 2.14. *For any connected graph with independence number α , radius r , and path covering number ρ , $\alpha \geq r + \frac{\rho-1}{2}$ [16].*

The *path covering number* of a graph is the minimum number of vertex disjoint paths so that each vertex of the graph is included in (exactly) one of the paths. If the graph has a Hamiltonian path then the path covering number of the graph is one. So the conjecture implies that, if $\alpha = r$, then $\rho = 1$, and the graph has a Hamilton path. This conjecture is a generalization of Graffiti's early and well-known conjecture, mentioned above, that for a connected graph $\alpha \geq r$. DeLaVina, Fajtlowicz, and Waller have proved the conjecture for trees [16]. DeLaVina's Graffiti.pc has made a number of conjectures of sufficient conditions for a connected graph to have a Hamilton path including: if the independence number of a connected graph equals its radius, then the graph has a Hamilton path [14]. Fajtlowicz has proved this conjecture for graphs with radius no more than three [27, 28]. DeLaVina, Pepper and Waller proved the general case in [17].

It was noted above that Fajtlowicz showed that connected graphs whose independence number equals its radius r have either an induced $2r$ -path or an induced

$2r$ -cycle. The preceding characterization of connected graphs whose independence number equals its radius and structural consequences can now be used to show that those graphs having induced $2r$ -cycles are Hamiltonian.

Conjecture 2.15. *Every connected graph whose independence number equals its radius r , having an induced $2r$ -cycle, is Hamiltonian.*

The case where $r = 1$ is pathological in the sense that the definition of a cycle standardly requires at least three vertices, while the conditions of the theorem require an induced 2-vertex cycle. If $r = 1$ and the graph does have at least three vertices, then it is indeed Hamiltonian. In this case $R = \{v_0, v_1\}$, R' is non-empty, and any independent set in $G[R']$ can have at most one vertex and $G[R']$ is complete. Let P be a Hamilton path in $G[R']$. Theorem 2.9 implies that every vertex in R' is adjacent to both v_0 and v_1 . So $v_0 P v_1$ is a Hamilton cycle.

If $r = 2$, the theorem holds as a consequence of Chvatal and Erdős' Theorem: If a graph with at least three vertices is s -connected and $\alpha \leq s$, then the graph has a Hamilton cycle [12]. In this case $\alpha(G) = 2$ and Theorem 2.12 implies that G is 2-connected; thus, the Chvatal-Erdős Theorem applies and G has a Hamilton cycle.

The cases where $r = 3$ and $r = 4$ remain open.

Theorem 2.16. *Every connected graph whose independence number equals its radius r , $r > 4$, having an induced $2r$ -cycle, is Hamiltonian.*

Proof. Let G be a connected graph with independence number $\alpha = \alpha(G)$, radius r , $r > 3$, and $\alpha = r$. Let $v_0, v_1, \dots, v_{2r-1}$ be an induced $2r$ -cycle. Let $R = \{v_0, v_1, \dots, v_{2r-1}\}$, and let $R' = V(G) \setminus R$.

Assume then that $r > 4$ (and $|R| > 8$). It can also be assumed that R' is not empty. If R' is empty then G is the cycle induced on R and thus has a Hamilton cycle.

Theorem 2.13 implies that, if $v \in R'$, then $v \in N_{i,i+1}$ or $v \in N_{i,i+1,i+2}$ (for some integer $i \in \{0, 1, \dots, 2r-1\}$, and where the indices are assumed to be mod $2r$). Thus,

$$R' = \bigcup_{i \in \{0,1,\dots,2r-1\}} (N_{i,i+1} \cup N_{i,i+1,i+2}) =$$

$$N_{0,1} \cup N_{0,1,2} \cup N_{1,2} \cup N_{1,2,3} \cup \dots \cup N_{2r-2,2r-1} \cup N_{2r-2,2r-1,0} \cup N_{2r-1,0} \cup N_{2r-1,0,1},$$

where any of the sets $N_{i,i+1}$ or $N_{i,i+1,i+2}$ may be empty.

Theorem 2.11 implies that the graphs induced on each of the sets $N_{i,i+1}$ and $N_{i,i+1,i+2}$ are complete. Complete graphs have Hamilton paths. Represent these paths as $P_{i,i+1}$ and $P_{i,i+1,i+2}$, respectively. Then the path represented

$$\mathbf{v}_0 P_{0,1} P_{0,1,2} \mathbf{v}_1 P_{1,2} P_{1,2,3} \mathbf{v}_2 \dots \mathbf{v}_{2r-2} P_{2r-2,2r-1} P_{2r-2,2r-1,0} \mathbf{v}_{2r-1} P_{2r-1,0} P_{2r-1,0,1}$$

is a Hamilton path in G .

v_0 , by definition, is adjacent to every vertex in $N_{0,1}$ and, in particular, to the first vertex in the Hamilton path $P_{0,1}$ in $G[N_{0,1}]$. Theorem 2.11 implies that every vertex in $N_{0,1}$ is adjacent to every vertex in $N_{0,1,2}$ and, in particular, the last vertex of the Hamilton path $P_{0,1}$ is adjacent to the first vertex of the Hamilton path $P_{0,1,2}$ in $G[N_{0,1,2}]$. Every vertex in $N_{0,1,2}$ is adjacent, by definition, to v_1 and, in particular, the last vertex of $P_{0,1,2}$ is adjacent to v_1 .

If either or both of $N_{0,1}$ or $N_{0,1,2}$ is empty the given path still works. If, for instance, $N_{0,1}$ is empty then, since v_0 is adjacent by definition to every vertex in $N_{0,1,2}$, v_0 is adjacent to the first vertex of $P_{0,1,2}$. If both $N_{0,1}$ and $N_{0,1,2}$ are empty then, since v_0 is adjacent to v_1 , the path schematized above remains well-defined.

A similar explanation can be given for every sequence from v_i to v_{i+1} . v_i , by definition, is adjacent to every vertex in $N_{i,i+1}$ and, in particular, to the first vertex in the Hamilton path $P_{i,i+1}$ in $G[N_{i,i+1}]$. Theorem 2.11 implies that every vertex in $N_{i,i+1}$ is adjacent to every vertex in $N_{i,i+1,i+2}$ and, in particular, the last vertex of the Hamilton path $P_{i,i+1}$ is adjacent to the first vertex of the Hamilton path $P_{i,i+1,i+2}$

in $G[N_{i,i+1,i+2}]$. Every vertex in $N_{i,i+1,i+2}$ is adjacent, by definition, to v_{i+1} and, in particular, the last vertex of $P_{i,i+1,i+2}$ is adjacent to v_{i+1} .

In general, if either or both of $N_{i,i+1}$ or $N_{i,i+1,i+2}$ is empty the given path still works. If, for instance, $N_{i,i+1}$ is empty then, since v_i is adjacent by definition to every vertex in $N_{i,i+1,i+2}$, v_i is adjacent to the first vertex of $P_{i,i+1,i+2}$. If both $N_{i,i+1}$ and $N_{i,i+1,i+2}$ are empty then, since v_i is adjacent to v_{i+1} , the path schematized above remains well-defined.

If $N_{2r-1,0,1}$ is not empty then the last vertex in $P_{2r-1,0,1}$ is adjacent to v_0 and this Hamilton path is actually a Hamilton cycle. In case $N_{2r-1,0,1}$ is empty then, since v_{2r-1} is adjacent to v_0 , the path above is still actually a cycle, and G is Hamiltonian.

□

2.4 Open Problems

Included here are open problems raised by the research presented above, or related to this research.

1. Fullerenes

- (a) Ryan Pepper has conjectured that for any fullerene other than C_{24} , the independence number is at least two-fifths the number of vertices. Computations verify this conjecture for fullerenes with up to 100 vertices.
- (b) Is it possible to find the independence number of a fullerene in polynomial-time? Fullerenes belong to the class of cubic planar graphs. For these graphs the decision problem for independence number is known to be NP-complete. But fullerenes have additional structure.

2. Radius

- (a) Is it possible to determine if the independence number of a graph equals its radius in polynomial-time?
- (b) Characterizing the graphs where $\alpha = qr - q^2 + q$.
- (c) Graffiti's conjecture: $\alpha \geq r + \frac{r-1}{2}$.

3. Hamiltonicity

- (a) Does the sufficient condition of Theorem 2.16 hold in the cases where $r = 3$ and $r = 4$?

Chapter 3

Critical Independent Sets, the Critical Independence Number, and Applications

When trying to find a maximum independent set (MIS) in a graph, the pendants of the graph may always be included. These vertices and their neighbors can then be removed, reducing the problem to that of finding a MIS on the remaining subgraph. A maximum critical independent set is a generalization of the set of pendant vertices: it may be included in a maximum independent set of the graph. It and its neighbors may be removed, similarly reducing the problem of finding an MIS to a subgraph. An algorithm for finding these sets in polynomial-time is presented, along with a proof that the algorithm works. Following this is a theoretical result about decomposing a graph into two subgraphs dividing the problem of finding the independence number into “easy” and “hard” parts. Two applications are also included.

The work in the first section appeared in [57].

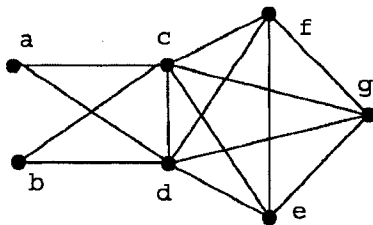


Figure 3.1: The vertices $\{a, b\}$ form a (maximum cardinality) critical independent set; this set of vertices can be extended to a maximum independent set of the graph.

3.1 Finding Maximum Critical Independent Sets

Finding a maximum independent set (MIS) in a graph is a well-known widely-studied NP-hard problem [42]. A polynomial-time algorithm for reducing this problem to the MIS problem on a subgraph is described here.

An independent set of vertices I is a *critical independent set* if $|I| - |N(I)|$ is maximized. Butenko and Trukhanov proved that any critical independent set is contained in a maximum independent set [10]. This can lead to a speed-up of the problem of finding a maximum independent set (MIS) and the independence number of a graph: if I is a critical independent set of a graph G , then the problem of finding a MIS can be reduced to finding one for $G \setminus (I \cup N(I))$. In fact, Butenko and Trukhanov demonstrate that the speed-up from this reduction can be dramatic.

The algorithm Butenko and Trukhanov use for finding a critical independent set does not always result in a non-empty critical independent set in cases where there is, in fact, such a set, and thus does not always result in a reduction of the problem of finding a MIS to a smaller graph. A criterion is given here for when a non-empty critical independent set exists as well as an algorithm for finding one in polynomial-time.

Butenko and Trukhanov ask “how to find the largest critical independent set in a graph?” This question is answered here. The specified algorithm can be extended to

yield a maximum-cardinality critical independent set.

Definition 3.1. $C \subseteq V(G)$ is a critical set of a graph G if $|C| - |N(C)| \geq |U| - |N(U)|$ for every $U \subseteq V(G)$.

Definition 3.2. $I \subseteq V(G)$ is a critical independent set of a graph G if I is an independent set of vertices and $|I| - |N(I)| \geq |U| - |N(U)|$ for every independent set $U \subseteq V(G)$.

A graph may contain critical independent sets of different cardinalities. A graph consisting of a single edge (K_2 , the complete graph on two vertices) has critical independent sets of cardinalities 0 and 1. A graph may not contain a non-empty critical independent set. For instance, the empty set is the unique critical independent set of K_3 . In fact, for any graph with a perfect matching (which is, in a well-defined sense, almost every graph with an even number of vertices [6, p. 178]), the empty set is a critical independent set. Finding a critical independent set using Ageev's algorithm may yield no reduction in these cases. There are, though, graphs with perfect matchings which have non-empty critical independent sets: K_2 is an example.

We now define the *bi-double graph* of a given graph. This graph is utilized in a proof of Zhang [84], and referred to in the papers of Ageev and Butenko and Trukhanov, and is a more generally useful proof technique (see, for instance, [3]).

Definition 3.3. For a graph G , the bi-double graph $B(G)$ has vertex set $V \cup V'$, where V' is a copy of V . If $V = \{v_1, v_2, \dots, v_n\}$, let $V' = \{v'_1, v'_2, \dots, v'_n\}$. Then, $(x, y') \in E(B(G))$ if, and only if, $(x, y) \in E(G)$.

Theorem 3.4. (Zhang [84, p. 437]) If C is a critical set then the isolated points in $G[C]$, the graph induced on C , is a critical independent set.

Theorem 3.5. (Ageev [1, p. 294]) For a graph G , if I is a maximum independent set in the bi-double graph $B(G)$, then $U = V(G) \cap I$ is a critical set for G .

The preceding two theorems imply that the following algorithm results in a critical independent set I_c in a graph G :

1. Construct the bi-double graph $B(G)$ of G .
2. Find a maximum independent set J in $B(G)$.
3. Let $J' = V \cap J$.
4. Let I_c be the set of isolated points in $G[J']$.

Since a maximum independent set in a bipartite graph can be found in polynomial-time, this algorithm yields a critical independent set in polynomial time. (Zhang was the first to prove the existence of such an algorithm).

Butenko and Trukhanov showed that identifying a non-empty critical independent set gives a polynomial-time reduction of the problem of finding a maximum independent set to a proper subgraph. The following lemma identifies a fact about the structure of critical independent sets, and leads to a new proof (below) of Butenko and Trukhanov's theorem.

Definition 3.6. *For disjoint subsets X and Y of the vertices of a graph G , there is a matching of X into Y if there is a set of disjoint edges having one endpoint in X and the other in Y and that saturates all of the vertices in X .*

Lemma 3.7. *(The Matching Lemma) If I_c is a critical independent set, then there is a matching from $N(I_c)$ into I_c .*

Proof. Note that I_c and $N(I_c)$ are disjoint. Let B be the subgraph of G whose vertices are $I_c \cup N(I_c)$, with $(x, y) \in E(B)$ if, and only if, $x \in I_c$, $y \in N(I_c)$, and $(x, y) \in E(G)$. Clearly, it is enough to prove the claim for this subgraph. Let $J \subseteq N(I_c)$. Suppose, $|J| > |N(J)|$. Let $X = I_c - N(J)$. Then $N(X) \subseteq N(I_c) - J$ (so $|N(I_c)| \geq |N(X)| + |J|$), and

$$|X| - |N(X)| > |X| - |N(X)| - (|J| - |N(J)|) =$$

$$(|X| + |N(J)|) - (|N(X)| + |J|) \geq |I_c| - |N(I_c)|.$$

This contradicts the fact that $|I_c|$ is a critical independent set. So, it must be that $|N(J)| \geq |J|$. Since this is true for every subset $J \subset N(I_c)$, Hall's Theorem (see, for instance, [9]) implies the claim. \square

What follows is a corollary that will be needed in the sequel.

Corollary 3.8. *If I_c is a critical independent set in a graph G , and M is a maximum matching of G , then all the vertices of $N(I_c)$ are saturated by M .*

Proof. Let I_c is a critical independent set in a graph G , and M is a maximum matching of G . Let N be a matching from $N(I_c)$ into I_c . So $|N| = |N(I_c)|$. Such a matching is guaranteed by Lemma 3.7. Let M' be the matching formed by removing all edges from M which are incident to vertices in $N(I_c)$. Thus $|M'| \geq |M| - |N(I_c)|$. By construction, the edges in N and M' must be independent. Let $N' = N \cup M'$. N' is a matching and $|N'| = |N| + |M'| \geq |N(I_c)| + (|M| - |N(I_c)|) = |M|$. Hence N' is a maximum matching which saturates the vertices in $N(I_c)$. If M does not saturate the vertices of $N(I_c)$, then N' is a matching with greater cardinality than M , contradicting the assumption that it is maximum. \square

The following proof of Butenko and Trukhanov's central theorem is new.

Theorem 3.9. *(Butenko & Trukhanov) If I_c is a critical independent set of a graph G , then there exists a maximum independent set I of G , such that $I_c \subseteq I$.*

Proof. Let I_c be a critical independent set of G . Let J be a maximum independent set of G , and let $J_N = J \cap N(I_c)$. Let $J'_N \subseteq I_c$ be the vertices matched with vertices of J_N by the matching from $N(I_c)$ to I_c given by Lemma 3.7. Let $J' = (J \setminus J_N) \cup J'_N$. Clearly, $J \cap J'_N$ is empty, $|J_N| = |J'_N|$, J' is an independent set in G , and $|J| = |J'|$. So J' is a maximum independent set. Now $I_c \cup J'$ is an independent set and, since J' is a maximum independent set, it follows that $|I_c \cup J'| = |J'|$ and, thus, $I_c \subseteq J'$, proving the theorem. \square

If there is *any* non-empty independent set I such that $|I| \geq |N(I)|$, then there is a non-empty critical independent set. If there is a pendant for instance, there is a non-empty critical independent set. In this sense, and in the sense that identification and removal of these sets reduces the problem of finding maximum independent sets, a critical independent set can be viewed as a generalization of a pendant.

The critical independent set algorithm given above yields a critical independent set—but this set may be empty in the case where non-empty independent sets exist. The following results imply an algorithm which yields a non-empty critical independent set when one exists.

Lemma 3.10. *If I_c is a critical independent set of the graph G then $I = I_c \cup (V' - N(I_c))$ is a maximum independent set of the bi-double graph $B(G)$.*

Proof. Clearly I is independent. Suppose there is a maximum independent set $J \subseteq V(B(G))$ such that $|J| > |I|$. Let $J_V = J \cap V$ and $J_{V'} = J \cap V'$. So, $|J| = |J_V| + |J_{V'}|$ and $|I| = |I_c| + |V' \setminus N(I_c)|$. Since $J_{V'} = V' \setminus N(J_V)$, $|V' \setminus N(J_V)| = |V| - |N(J_V)|$, and $|V' \setminus N(I_c)| = |V| - |N(I_c)|$, we have the following equations:

$$|J| = |J_V| + |J_{V'}| = |J_V| + |V' \setminus N(J_V)| = |J_V| + |V| - |N(J_V)|,$$

$$|I| = |I_c| + |V' \setminus N(I_c)| = |I_c| + |V| - |N(I_c)|.$$

Since $|J| > |I|$, it follows that $|J_V| + |V| - |N(J_V)| > |I_c| + |V| - |N(I_c)|$, and $|J_V| - |N(J_V)| > |I_c| - |N(I_c)|$. Since J_V is independent, I_c is not a critical independent set, according to the definition of a critical independent set, contradicting our assumption that it was. Thus, I is a maximum independent set of $B(G)$. \square

Theorem 3.11. *A graph G contains a non-empty critical independent set if, and only if, there is a maximum independent set of the bi-double graph $B(G)$ containing both v and v' , for some vertex $v \in V(G)$, and its copy $v' \in V'$.*

Proof. If I_c is a non-empty critical independent set of G then, by Lemma 3.10, $I = I_c \cup (V' - N(I_c))$ is a maximum independent set of $B(G)$. For any vertex $v \in I_c$, v is not adjacent to v' in $B(G)$. Thus, $v' \in V' - N(I_c)$. Thus, v and v' are in I .

Suppose I is a maximum independent set of $B(G)$ containing both v and v' . Then $v \in J = I \cap V(G)$. By Theorem 3.5, J is a critical set in G , and by Theorem 3.4, the isolated points in $G[J]$ are a critical independent set in G . Suppose v is not an isolated point in $G[J]$. Then there is a vertex $w \in J$ such that v is adjacent to w in G . This implies that w is adjacent to v' in $B(G)$. So I contains v' and w and I is independent, contradicting the assumption that v is not isolated in $G[J]$. Thus, the set of isolated points of $G[J]$ is non-empty. \square

Corollary 3.12. *A graph G contains a non-empty critical independent set if, and only if, there is a vertex $v \in V(G)$ such that $\alpha(B(G)) = \alpha(B(G) - \{v, v'\}) - N(\{v, v'\}) + 2$.*

Proof. This follows immediately from Theorem 3.11. \square

Corollary 3.12 suggests a polynomial-time algorithm for finding a non-empty critical independent set in a graph if one exists:

1. Construct graph $B(G)$.
2. Set BOOL=false.
3. For $i = 1, \dots, n = |V(G)|$, set BOOL=true if $\alpha(B(G)) = \alpha(B(G) - \{v_i, v'_i\}) - N(\{v_i, v'_i\}) + 2$. If BOOL=true, break.
4. If BOOL=true,
 - (a) Find a maximum independent set J in $B(G) - \{v_i, v'_i\} - N(\{v_i, v'_i\})$.
 - (b) Let $J' = J \cap V$.
 - (c) Let I be the set of isolated points in $G[J']$ together with v .

If $\text{BOOL}=\text{true}$, I is a non-empty critical independent set. If $\text{BOOL}=\text{false}$, then no non-empty critical independent set exists in G .

Definition 3.13. *A critical independent set is maximal if there is no critical independent set which properly contains it. It is maximum if there is no critical independent set with larger cardinality.*

Butenko and Trukhanov raised the question of how to identify maximum critical independent sets. These sets will result in a maximum reduction in the problem of finding a MIS. The following theorem justifies an algorithm that yields these sets.

Theorem 3.14. *Any critical independent set is contained in a maximum critical independent set.*

Proof. Suppose I_c is a critical independent set and J_c is a maximum critical independent set. Let $I = I_c \cup J$, where $J = J_c \setminus (I_c \cup N(I_c))$. It is enough to show that I is a maximum critical independent set. Clearly I is independent.

We will first show that I is a critical (independent) set; in particular, that $|I| - |N(I)| \geq |I_c| - |N(I_c)|$. Since I_c and J are disjoint, $|I| = |I_c| + |J|$. $N(I) \subseteq N(I_c) \cup [N(J_c) \setminus (I_c \cup N(I_c))]$ and $|N(I)| \leq |N(I_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|$. So,

$$|I| - |N(I)| \geq |I_c| - |N(I_c)| + |J| - |N(J_c) \setminus (I_c \cup N(I_c))|.$$

It is enough then to show, $|J| \geq |N(J_c) \setminus (I_c \cup N(I_c))|$.

Now, $J_c = J \cup (I_c \cap J_c) \cup (N(I_c) \cap J_c)$. By definition, J , $I_c \cap J_c$, and $N(I_c) \cap J_c$ are mutually disjoint. So,

$$|J_c| = |J| + |I_c \cap J_c| + |N(I_c) \cap J_c|.$$

Also, $N(J_c) = (I_c \cap N(J_c)) \cup (N(I_c) \cap N(J_c)) \cup (N(J_c) \setminus (I_c \cup N(I_c)))$. Clearly, $I_c \cap N(J_c)$, $N(I_c) \cap N(J_c)$ and $N(J_c) \setminus (I_c \cup N(I_c))$ are disjoint. So,

$$|N(J_c)| = |I_c \cap N(J_c)| + |N(I_c) \cap N(J_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|.$$

Then,

$$|J_c| - |N(J_c)| = \tag{3.1}$$

$$(|N(I_c) \cap J_c| - |N(J_c) \cap I_c|) + |I_c \cap J_c| + |J| - (|N(I_c) \cap N(J_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|).$$

Now, Theorem 3.7 guarantees that there is a matching from $N(J_c)$ to J_c and from $N(I_c)$ to I_c . Since the vertices in $I_c \cap N(J_c) \subseteq N(J_c)$ must be matched to vertices in $N(I_c) \cap J_c$, and the vertices in $N(I_c) \cap J_c \subseteq N(I_c)$ must be matched to vertices in $I_c \cap N(J_c)$, it follows that $|I_c \cap N(J_c)| = |J_c \cap N(I_c)|$ and $|N(I_c) \cap J_c| - |N(J_c) \cap I_c|$, the first term in Equation 3.1, is 0.

Assume that $|N(J_c) \setminus (I_c \cup N(I_c))| > |J|$. Note that $N(I_c \cap J_c) \subseteq N(I_c) \cap N(J_c)$ and $|N(I_c \cap J_c)| \leq |N(I_c) \cap N(J_c)|$. Then Equation 3.1 gives,

$$|J_c| - |N(J_c)| = |I_c \cap J_c| + |J| - (|N(I_c) \cap N(J_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|)$$

$$< |I_c \cap J_c| - |N(I_c) \cap N(J_c)| \leq |I_c \cap J_c| - |N(I_c \cap J_c)|,$$

contradicting the fact that J_c is a critical (independent) set. Thus, $|N(J_c) \setminus (I_c \cup N(I_c))| \leq |J|$ and I is a critical independent set.

Lastly, we show that I is maximum; in particular that $|I| = |J_c|$. It was noted above that $J_c = (J_c \cap I_c) \cup (J_c \cap N(I_c)) \cup J$, $|J_c| = |J_c \cap I_c| + |J_c \cap N(I_c)| + |J|$, and $|I| = |I_c| + |J|$. Since I_c is a critical independent set, by the Matching Lemma 3.7 there is a matching from $N(I_c)$ to I_c . Let $J'_N \subseteq I_c$ be the vertices matched to $(J_c \cap N(I_c)) \subseteq N(I_c)$ under this matching. So $|J'_N| = |J_c \cap N(I_c)|$. Clearly J'_N and $J_c \cap I_c$ are disjoint. $|I| = |I_c| + |J| = |J'_N| + |I_c \setminus J'_N| + |J| \geq |J_c \cap N(I_c)| + |J_c \cap I_c| + |J| = |J_c|$. Note that $(J_c \cap I_c) \subseteq (I_c \setminus J'_N)$. So $|I| \geq |J_c|$ and, since J_c is a maximum critical independent set, $|I| = |J_c|$. Thus, I is a maximum critical independent set. □

Corollary 3.15. *If a vertex v of a graph G is contained in some critical independent set, then there is a maximum critical independent set which contains v .*

Corollary 3.16. *A maximal critical independent set is a maximum critical independent set.*

The idea of the following algorithm is to find a maximal critical independent set I by choosing a vertex, testing if it is contained in a critical independent set and, if it is, adding it to I and removing it and its neighbors. In either case the process is repeated on the graph induced on the remaining vertices.

The Maximal Critical Independent Set (MCIS) algorithm

1. Construct the bi-double graph $B(G)$ of G .
2. $i := 1, I_c := \emptyset$;
3. If $i > |V(G)|$, return I_c .
4. If $v_i \notin V(B(G))$, $i := i + 1$, and return to Step 3.
5. If $\alpha(B(G)) = \alpha(B(G) - \{v_i, v'_i\} - N(\{v_i, v'_i\})) + 2$, $\text{BOOL} := \text{true}$. Else, $\text{BOOL} := \text{false}$.
6. If $\text{BOOL} = \text{true}$, $I_c := I_c \cup \{v_i\}$, $V(B(G)) := V(B(G)) \setminus (\{v_i, v'_i\} \cup N(\{v_i, v'_i\}))$,
 $i := i + 1$. Return to Step 3.
7. If $\text{BOOL} = \text{false}$, $i := i + 1$. Return to Step 3.

Theorem 3.17. *The MCIS algorithm yields a maximum critical independent set.*

Proof. By Corollary 3.16 it is enough to show that this algorithm produces a maximal critical independent set. For graphs with a single vertex, the MCIS algorithm returns a one-element set. This set is clearly a maximal critical independent set.

Suppose the MCIS algorithm produces a maximal critical independent set for graphs with n or fewer vertices. Suppose G has $n + 1$ vertices. If the only critical independent set of G is the empty set then, by Corollary 3.12, the test in Step 6 will be negative and set I will remain empty after each loop. I is a maximal critical independent set.

Suppose G contains a non-empty critical independent set. Let i be the first index so that v_i belongs to a critical independent set. Corollary 3.12 guarantees that the test in Step 6 will be negative. The MCIS algorithm then sets $I := \{v_i\}$ and continues on the graph G' induced on $V(G) - \{v_i\} - N(\{v_i\})$. This graph has n or fewer vertices. By assumption, the MCIS algorithm (then) yields a maximal critical independent set J for G' . So $I = J \cup \{v_i\}$ is a maximal critical independent set for G . \square

Theorem 3.18. *If I is a maximum critical independent set of G , then the only critical independent set of $G - I - N(I)$ is the empty set.*

This means that any further repetition of the MCIS algorithm will not yield any further reduction.

3.1.1 Weighted Critical Independent Sets

The results of Ageev, Butenko and Trukhanov all have analogues for weighted graphs. An anonymous referee for the *Bulletin of the ICA* asked “whether these results can be extended to the weighted version of the MIS problem as well.” The answer is *yes and maybe*: a criterion can be specified for the existence of a critical weighted independent set which exactly parallels the criterion for existence in the non-weighted case, but it is an open question whether any critical weighted independent set is contained in a maximum weight independent set.

Definition 3.19. *A weighted graph is a graph where a nonnegative number w_i (called a weight) is associated to each vertex v_i . The weight of a set $S \subseteq V$ is $w(S) = \sum_{v_i \in S} w_i$. S is a critical weighted set if*

$$w(S) - w(N(S)) \geq w(T) - w(N(T))$$

for any set $T \subseteq V$. A set $I \subseteq V$ is a critical weighted independent set if I is independent and a critical weighted set. I is a maximum weight independent set

if I is independent and $w(I) \geq w(J)$ for any other independent set J . A critical weighted independent set is maximum if there is no critical weighted independent set with larger weight.

The weighted extensions of Zhang and Ageev's theorems are reproduced here as they are required for the proof of Theorem 3.23.

Theorem 3.20. (Zhang, [84]) *If C is a critical weighted set then the isolated points in $G[C]$, the graph induced on C , is a critical weighted independent set.*

Theorem 3.21. (Ageev, [1]) *For a graph G , if I is a maximum weighted independent set in the bi-double graph $B(G)$, then $U = V(G) \cap I$ is a critical weighted set for G .*

The criterion for determining if a graph has a non-empty critical independent set can be extended to the weighted case: there is a polynomial-time criterion for determining if a weighted graph has a non-empty critical weighted independent set.

Lemma 3.22. *If I_c is a critical weighted independent set of the graph G then $I = I_c \cup (V' - N(I_c))$ is a maximum weight independent set of the bi-double graph $B(G)$.*

Proof. Clearly I is independent. Suppose there is a maximum weight independent set $J \subseteq V(B(G))$ such that $w(J) > w(I)$. Let $J_V = J \cap V$ and $J_{V'} = J \cap V'$. So, $w(J_V) + w(J_{V'}) = w(J) > w(I) = w(I_c) + w(V' \setminus N(I_c))$. Since $J_{V'} = V' \setminus N(J_V)$, $w(V' \setminus N(J_V)) = w(V) - w(N(J_V))$, and $w(V' \setminus N(I_c)) = w(V) - w(N(I_c))$, it follows that $w(J_V) + w(V) - w(N(J_V)) > w(I_c) + w(V) - w(N(I_c))$. By Theorem 3.21, J_V is a critical weighted set in G and, thus, that I_c is not a critical weighted set (nor a critical weighted independent set).

Thus, since critical weighted independent sets are critical weighted sets, I_c is not a critical weighted set. Thus, I is a maximum weighted independent set of $B(G)$. \square

Theorem 3.23. *A graph G contains a non-empty critical weighted independent set if, and only if, there is a maximum weight independent set of the bi-double graph $B(G)$ containing both v and v' , for some vertex $v \in V(G)$, and its copy $v' \in V'$.*

Proof. If I_c is a non-empty critical weighted independent set of G then, by Lemma 3.22, $I = I_c \cup (V' - N(I_c))$ is a maximum weighted independent set of $B(G)$. For any vertex $v \in I_c$, v is not adjacent to v' in $B(G)$. Thus, $v' \in V' - N(I_c)$. Thus, v and v' are in I .

Suppose I is a maximum weight independent set of $B(G)$ containing both v and v' . Then $v \in J = I \cap V(G)$. By Theorem 3.21, J is a critical weighted set in G , and by Theorem 3.20, the isolated points in $G[J]$ are a critical weighted independent set in G . Suppose v is not an isolated point in $G[J]$. Then there is a vertex $w \in J$ such that v is adjacent to w in G . This implies that w is adjacent to v' in $B(G)$. So I contains v' and w and I is independent, contradicting the assumption that v is not isolated in $G[J]$. Thus, the set of isolated points of $G[J]$ is non-empty. \square

If I is a maximum weight independent set of a graph G , let the weighted independence number of G be $\alpha_w(G) = w(I)$. Since every graph has a maximum weight independent set, $\alpha_w(G)$ is well-defined.

Corollary 3.24. *A graph G contains a non-empty critical weighted independent set if, and only if, there is a vertex $v \in V(G)$ such that $\alpha_w(B(G)) = \alpha_w(B(G) - \{v, v'\}) - N(\{v, v'\}) + 2w(v)$.*

Proof. This follows immediately from Theorem 3.23. \square

Whether Theorem 3.14 is extendable to the weighted case is an open question. Is every critical weighted independent set contained in a maximum critical weighted independent set? The proof that every critical independent set is contained in a maximum critical independent set made use of Theorem 3.7, which cannot be extended: it is not true that, if I is a critical weighted independent set, then there is a matching from $N(I)$ into I . Consider the complete graph K_3 with three vertices. Let the vertices be $V = \{v_1, v_2, v_3\}$, having weights $w_1 = 1$, $w_2 = \frac{1}{2}$, and $w_3 = \frac{1}{2}$. Let $I = \{v_1\}$.

It is easy to verify that I is a critical weighted independent set of K_3 . But there is no matching from $N(I) = \{v_2, v_3\}$ to I .

3.1.2 An application: Critical Independence Reductions for Fullerenes

This investigation was inspired by an attempt, using Butenko and Trukhanov's theorem and the Zhang/Ageev algorithm, to reduce the problem of finding maximum independent sets in fullerene graphs—there is strong statistical evidence that the independence number of a fullerene is a predictor of its stability [33]. No reduction was found and, in fact, no reduction is possible.

Definition 3.25. *A fullerene or fullerene graph is a connected, cubic, planar graph whose faces are either pentagons or hexagons.*

Theorem 3.26. *The empty set is the only critical independent set in a fullerene. (Equivalently, for every non-empty independent set I of a fullerene, $|N(I)| > |I|$.)*

Proof. Suppose G is a fullerene and G contains a non-empty independent set I such that $|N(I)| \leq |I|$. Since G is cubic, there are $3|I|$ edges incident to set I . There are at least $3|I|$ edges incident to $N(I)$. Thus, $3|N(I)| \geq 3|I|$ and $|N(I)| \geq |I|$ (and $|N(I)| = |I|$). Thus, the graph $G[I \cup N(I)]$, induced on $I \cup N(I)$, is bipartite. Since G is connected, $G = G[I \cup N(I)]$. So G is bipartite, which contradicts the fact that fullerenes are not bipartite (it is a simple consequence of Euler's Theorem that they have twelve pentagonal faces and, thus, odd cycles). \square

The details of the proof actually give the following stronger theorem.

Theorem 3.27. *If G is a connected, non-bipartite, regular graph then the empty set is the only critical independent set.*

3.2 An application: Characterizing when Independence equals Annihilation

Pepper proved the annihilation number a of a graph is an upper bound for the independence number α of a graph and identified several classes of graphs where equality holds. In this section the critical independence number of a graph is used to provide a characterization of all graphs where equality holds. Furthermore, a polynomial-time algorithm is provided for determining if these invariants are equal. The independence number is a well-known NP-hard invariant. For these graphs, the independence number is computable in polynomial-time.

Pepper originally defined the *annihilation number* of a graph in terms of a reduction process on the degree sequence of the graph (akin to the Havel-Hakimi process; see, for example, [44]). Fajtlowicz later defined a still-unnamed invariant which he conjectured was a better upper bound for the independence number of a graph than Pepper's invariant. Pepper proved that the invariants are, in fact, the same [73]. The following definition is due to Fajtlowicz.

Definition 3.28. *For a graph G with vertices $V = \{v_1, v_2, \dots, v_n\}$, having degrees $d_i = d(v_i)$, with $d_1 \leq d_2 \leq \dots \leq d_n$, and having e edges, the annihilation number $a = a(G)$ is defined to be the largest index such that $\sum_{i=1}^a d_i \leq e$.*

Theorem 3.29. *(Pepper [73]) For any graph G , $\alpha(G) \leq a(G)$.*

Pepper originally proved this theorem using his original definition of annihilation number; Fajtlowicz later found a shorter proof using the definition above (both proofs are in [73]). Pepper also provided examples of graphs where the annihilation number of a graph is a better upper bound than any of several others, including the minimum of the numbers of non-negative and non-positive eigenvalues (Cvetkovic's bound [13, Thm 3.14]).

Lemma 3.30. (Pepper [73]) For any graph G , $a(G) \geq \lfloor \frac{n(G)}{2} \rfloor$.

Pepper's theorem and Butenko and Trukhanov's Theorem 3.9 are required in the proof of the characterization.

Lemma 3.31. For a graph G and vertex v , $a(G - v) \leq a(G)$.

Proof. Let G be a graph and $v \in V(G)$. Let $a = a(G)$ and $a' = a(G - v)$. It will be shown that $a' \leq a$. Let $d_G(w)$ be the degree of a vertex w in G . For a set $A \subseteq V$ of vertices of G , let $d_G(A)$ be the sum of the degrees in G of the vertices in A . Thus, $d_G(A) = \sum_{w \in A} d_G(w)$. Let $e = e(G)$ be the size of G and $e' = e(G - v) = e(G) - d_G(v)$.

Suppose the annihilation number of $G - v$ is at least $a + 1$. Then there is a set $A \subseteq V(G - v)$ of $|A| = a + 1$ vertices such that $d_{G-v}(A) \leq e'$. Then

$$d_G(A) \leq d_{G-v}(A) + d_G(v) \leq e' + d_G(v) = e.$$

That is, there is a set of $a + 1$ vertices in G where the sum of their degrees is less than the number of edges of G , and the annihilation number of G is at least $a + 1$, contradicting the fact that the annihilation number of G is a . So the assumption that $a(G - v) > a(G)$ is false. \square

Theorem 3.32. For a graph G , the independence number α equals its annihilation number a if, and only if, either (1) $a \geq \frac{n}{2}$ and $\alpha' = a$, or (2) $a < \frac{n}{2}$ and there is a vertex $v \in V(G)$ such that $\alpha'(G - v) = a(G)$.

Proof. The theorem can be easily verified for all graphs with three or fewer vertices. Assume that it is true for all graphs with fewer than n vertices. Let G be a graph with n vertices.

Suppose $a(G) \geq \frac{n(G)}{2}$ and $\alpha'(G) = a(G)$. Since $\alpha' \leq \alpha \leq a$ for any graph, it follows that $\alpha(G) = a(G)$.

Alternately, suppose that $a(G) < \frac{n(G)}{2}$ and there is a vertex v such that $\alpha'(G - v) = a(G)$. Since $\alpha'(G - v) \leq \alpha(G - v) \leq \alpha(G) \leq a(G)$ and the first and last terms are equal, every term must be equal and, thus, $\alpha(G) = a(G)$.

Suppose now that $\alpha(G) = a(G)$. It will be shown that either (1) $a(G) \geq \frac{n(G)}{2}$ and $\alpha'(G) = a(G)$, or (2) $a(G) < \frac{n(G)}{2}$ and there is a vertex $v \in V(G)$ such that $\alpha'(G - v) = a(G)$.

Suppose $a(G) < \frac{n(G)}{2}$. G must have an edge; otherwise, $a(G) = n(G) \geq \frac{n(G)}{2}$. If G has an edge then there is a vertex v which is not in every maximum independent set. So $\alpha(G - v) = \alpha(G) = a(G)$. Suppose that $a(G - v) < a(G)$. Then $a(G - v) < \alpha(G - v)$, which contradicts the fact that $\alpha \leq a$ for any graph. Since Lemma 3.31 implies that $a(G - v) \leq a(G)$, it follows that $a(G - v) = a(G)$ and, furthermore, that $\alpha(G - v) = a(G - v)$.

Since $a(G - v) = a(G) \geq \lfloor \frac{n(G)}{2} \rfloor = \lfloor \frac{n(G-v)+1}{2} \rfloor$, it follows that $a(G - v) \geq \frac{n(G-v)}{2}$ and thus, by the inductive assumption, that $\alpha'(G - v) = a(G - v)$. Then,

$$\alpha'(G - v) = \alpha(G - v) = \alpha(G) = a(G - v) = a(G).$$

Thus, it is shown that there is a vertex v , such that $\alpha'(G - v) = a(G)$.

The remaining case to consider is when $a(G) \geq \frac{n(G)}{2}$. In this case it must be shown that $\alpha'(G) = a(G)$. Let J_c be a maximum critical independent set of G . Since it was assumed that $\alpha(G) = a(G)$, $\alpha(G) \geq \frac{n(G)}{2}$ and it follows that G has a non-empty critical independent set. Thus, there is a vertex $u \in J_c$.

Suppose $N(J_c) = \emptyset$. So J_c is a discrete set of vertices and $\alpha'(G - J_c) = 0$ and $\alpha'(G - J_c + u) = 1$. Also $\alpha(G - J_c) = \alpha(G) - |J_c|$ and $a(G - J_c) = a(G) - |J_c|$, and $\alpha(G - J_c) = a(G - J_c)$. If $a(G - J_c) \geq \frac{n(G-J_c)}{2}$ then, the inductive assumption implies, $\alpha'(G - J_c) = a(G - J_c)$. It then follows that $\alpha(G - J_c) = 0$. So J_c is also a maximum independent set and G is a graph with no edges. If $a(G - J_c) < \frac{n(G-J_c)}{2}$ then $a(G - J_c) = \frac{n(G-J_c)-1}{2}$ and $a(G - J_c + u) = \frac{n(G-J_c+u)+1}{2} \geq \frac{n(G-J_c+u)}{2}$ and, the inductive assumption implies, $\alpha'(G - J_c + u) = a(G - J_c + u)$. It then follows that $\alpha(G - J_c + u) = 1$. Since u has no neighbors, this implies that $\alpha(G - J_c) = 0$. It again follows that J_c is a maximum independent set and G is a graph with no edges. It was noted earlier that the theorem follows for these graphs.

So it can be assumed that $N(J_c) \neq \emptyset$. Let w be a vertex in $N(J_c)$. It follows from Butenko and Trukhanov's Theorem 3.9 that J_c is contained in a maximum independent set I . Since $w \notin I$, w is not in every maximum independent set of G . Thus, by the previous reasoning, it follows that

$$\alpha'(G - w) = \alpha(G - w) = \alpha(G) = a(G - w) = a(G).$$

It is enough to show then that $\alpha'(G) = \alpha'(G - w)$.

Let $N_H(Y)$ be the neighbors of the set of vertices Y in the graph H . Since J_c is a critical independent set of G , $|J_c| - |N_G(J_c)| \geq |X| - |N_G(X)|$ for any set of vertices X in G . Note that $|J_c| - |N_G(J_c)| = |J_c| - (|N_G(J_c) - w| + 1)$. Let J'_c be a maximum critical independent set in $G' = G - w$. So $\alpha'(G - w) = |J'_c|$. Note that, since $w \in N_G(J_c)$, $J_c \subseteq V(G')$. Also, $|N_{G'}(J_c)| = |N_G(J_c)| - 1$. Since J'_c is a critical independent set, $|J'_c| - |N_{G'}(J'_c)| \geq |J_c| - |N_{G'}(J_c)| = |J_c| - (|N_G(J_c)| - 1)$.

Since $J'_c \subseteq V(G')$, $w \notin J'_c$. There are two cases to consider: (1) the case where $w \in N_G(J'_c)$, and (2) the case where $w \notin N_G(J'_c)$. If $w \in N_G(J'_c)$, then $N_G(J'_c) = N_{G'}(J'_c) \cup \{w\}$, and $|N_G(J'_c)| = |N_{G'}(J'_c)| + 1$. Since $|J'_c| - |N_{G'}(J'_c)| \geq |J_c| - |N_G(J_c)| + 1$, it follows that $|J'_c| - (|N_G(J'_c)| - 1) \geq |J_c| - |N_G(J_c)| + 1$ and, thus, $|J'_c| - |N_G(J'_c)| \geq |J_c| - |N_G(J_c)|$. So J'_c is a critical independent set of G and $\alpha(G) \geq \alpha'(G) \geq |J'_c| = \alpha'(G - w) = \alpha(G - w) = \alpha(G)$. It follows that $\alpha'(G) = \alpha'(G - w)$, which was to be shown.

If $w \notin N_G(J'_c)$, then $N_G(J'_c) = N_{G'}(J'_c)$ and $|N_G(J'_c)| = |N_{G'}(J'_c)|$. So $|J'_c| - |N_{G'}(J'_c)| \geq |J_c| - |N_{G'}(J_c)| = |J_c| - |N_G(J_c)|$. Thus, J'_c is a critical independent set of G and $\alpha(G) \geq \alpha'(G) \geq |J'_c| = \alpha'(G - w) = \alpha(G - w) = \alpha(G)$, proving in this case too that $\alpha'(G) = \alpha'(G - w)$.

□

Let G be a graph. The steps to determine whether $a(G) = \alpha(G)$ are as follows.

1. Calculate $a(G)$.

2. If $a(G) \geq \frac{n}{2}$, calculate $\alpha'(G)$. If $\alpha'(G) = a(G)$, then Theorem 3.32 implies that $\alpha(G) = a(G)$. If $\alpha'(G) \neq a(G)$, then Theorem 3.32 implies that $\alpha(G) \neq a(G)$.
3. If $a(G) < \frac{n}{2}$, choose an edge vw . It cannot be that both vertex v and vertex w are in every maximum independent set of G . Calculate $\alpha'(G - v)$ and $\alpha'(G - w)$. If $a(G) = \alpha'(G - v)$ or $a(G) = \alpha'(G - w)$ then Theorem 3.32 implies that $\alpha(G) = a(G)$.

The proof of Theorem 3.32 actually shows that if (a) $\alpha(G) = a(G)$, (b) $a(G) < \frac{n(G)}{2}$, and (c) vertex u is not in every maximum independent set of G , then (d) $\alpha'(G - u) = a(G)$. Since either v or w is not in every maximum independent set of G then, if $a(G) \neq \alpha'(G - v)$ and $a(G) \neq \alpha'(G - w)$, it follows that $\alpha(G) \neq a(G)$.

Since a and α' can be calculated in polynomial-time, and since the preceding algorithm will terminate after at most three calculations of these invariants, determining whether $a(G) = \alpha(G)$ can be done in polynomial-time.

3.3 An Independence Decomposition of a General Graph

The *critical independence number* of a graph G , denoted $\alpha' = \alpha'(G)$, is the cardinality of a maximum critical independent set. If I_c is a maximum critical independent set, and so $\alpha'(G) = |I_c|$, then clearly $\alpha' \leq \alpha$. It was shown above in Section 3.1 (and by this author in [57]) that the critical independence number of a graph can be computed in polynomial-time.

A graph is *independence irreducible* if $\alpha' = 0$. For these graphs the number of neighbors of any independent set of vertices is greater than the number of vertices in the set; fullerene graphs, for instance, are independence irreducible [57]. A graph is

independence reducible if $\alpha' > 0$. A graph is *totally independence reducible* if $\alpha' = \alpha$; K_2 is an example. It will be shown that, for any graph G , there is a set $X \subseteq V(G)$ such that $G[X]$ is totally independence reducible, $G[X^c]$ is independence irreducible, and $\alpha(G[X]) + \alpha(G[X^c]) = \alpha(G)$.

Zhang [84] gave a different definition of the critical independence number of a graph. He defined it to be the quantity $|I_c| - |N(I_c)|$ for a critical independent set I_c . This quantity though is not an “independence number”; that is, it is not the cardinality of an independent set of vertices. A better name for Zhang’s invariant would be the *critical difference* of the graph.

An obvious and inefficient algorithm for computing α' of a graph G is to find every independent set I of vertices of the graph and compute $|I| - |N(I)|$. Then α' is the cardinality of the largest of the sets that maximizes this difference. A polynomial-time algorithm for finding a Maximum Critical Independent Set (MCIS) and, thus, computing α' is given above (Section 3.1).

Finding a maximum independent set in a graph and its independence number are NP-hard problems. When attacking these problems it would be useful to be able to decompose the problem into finding maximum independent sets and independence numbers for subgraphs whose vertex sets are disjoint and whose union is the vertex set of the original graph. For any graph it is shown that there is such a decomposition into two subgraphs, where the independence number of one can be computed in polynomial-time and where the critical independence number of the other is zero.

The following terms were defined above.

Definition 3.33. *A graph is independence irreducible if $\alpha' = 0$. A graph is independence reducible if $\alpha' > 0$. A graph is totally independence reducible if $\alpha' = \alpha$.*

All bipartite graphs, for instance, are totally independence reducible. This will be proved in Section 3.4. Other examples are in Figure 3.3.

It will be shown that any graph can be decomposed into totally independence

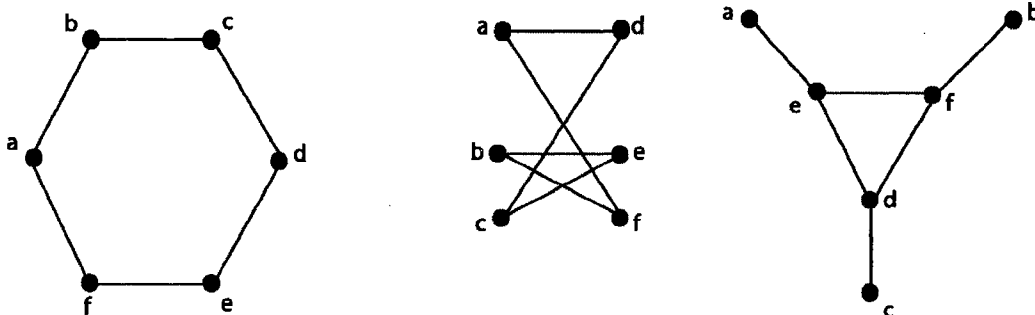


Figure 3.2: These graphs are *totally independence reducible*. In the graph on the left, the set $I = \{a, c, e\}$ is a maximum critical independent set, and a maximum independent set. In the middle graph, the set $I_c = \{a, b, c\}$ is a maximum critical independent set, and a maximum independent set. The left and the middle graphs are bipartite (and, in fact, isomorphic); all bipartite graphs are totally independence reducible. The graph on the right is not bipartite; the set $I_c = \{a, b, c\}$ is a maximum critical independent set, and a maximum independent set. For all three of these graphs $\alpha = \alpha' = 3$.

reducible and independence irreducible subgraphs whose independence numbers sum to the independence number of the parent graph.

The following characterization of graphs whose independence numbers equal their critical independence numbers will be needed in the proof of the main result.

Theorem 3.34. *For any graph G , $\alpha = \alpha'$ if, and only if, there is a maximum critical independent set I_c such that $I_c \cup N(I_c) = V(G)$.*

Proof. Let G be a graph. Suppose first that $\alpha(G) = \alpha'(G)$. Let I_c be a maximum critical independent set. It is also a maximum independent set by assumption and, thus, it and its neighbors must exhaust the vertices of G .

Suppose now that I_c is a maximum critical independent set and $I_c \cup N(I_c) = V(G)$. Theorem 3.14 implies that there is a maximum independent set I of G which contains I_c . If there is a vertex $v \in I \setminus I_c$ then, by assumption, $v \in N(I_c)$. But then v is adjacent to some vertex in I_c and I is not independent. So $I = I_c$ and $\alpha = \alpha'$. \square

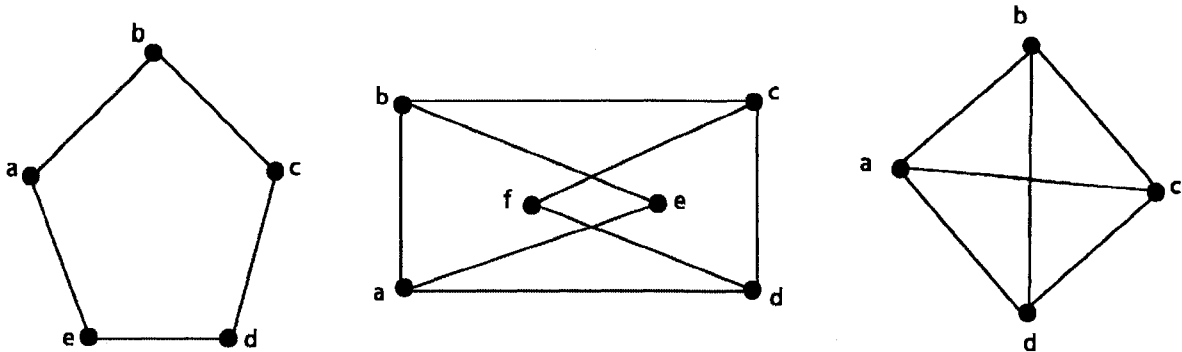


Figure 3.3: These graphs are independence irreducible: for *any* independent set of vertices I in these graphs, $|N(I)| > |I|$.

Since a maximum critical independent set of a graph can be found in polynomial-time, Theorem 3.34 implies that determining whether a graph is totally independence reducible can be determined in polynomial-time.

Lemma 3.35. *If G is a graph with critical independent sets I_c and J_c , where $J = J_c \setminus (I_c \cup N(I_c))$, and $I = I_c \cup J$ then,*

1. $|I_c \cap N(J_c)| = |J_c \cap N(I_c)|$,
2. $|J| \geq |N(J_c) \setminus (I_c \cup N(I_c))|$, and
3. I is a critical independent set.

Proof. The Matching Lemma 3.7 guarantees that there is a matching from the vertices in $N(J_c)$ to (a subset of) the vertices in J_c and from the vertices in $N(I_c)$ to (a subset of) the vertices in I_c . Since the vertices in $I_c \cap N(J_c) \subseteq N(J_c)$ must be matched to vertices in $N(I_c) \cap J_c$, and the vertices in $N(I_c) \cap J_c \subseteq N(I_c)$ must be matched to vertices in $I_c \cap N(J_c)$, it follows that $|I_c \cap N(J_c)| = |J_c \cap N(I_c)|$, proving (1).

Applying the Matching Lemma again, we have that $N(J_c)$ is matched into J_c , that is, every vertex in $N(J_c)$ can be paired with a distinct adjacent vertex in J_c . Notice

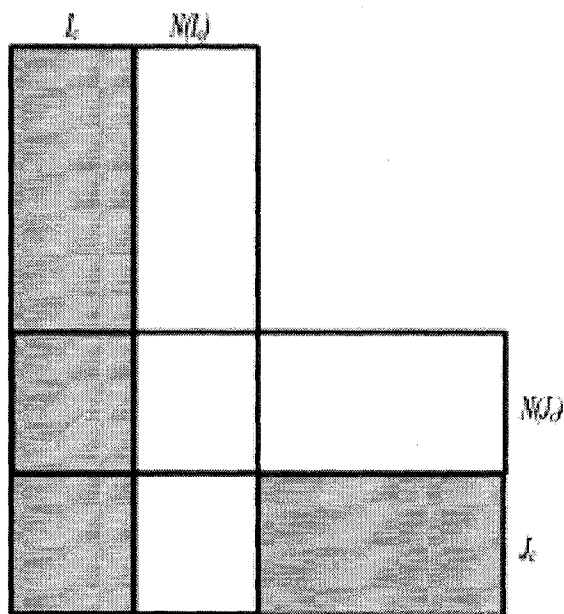
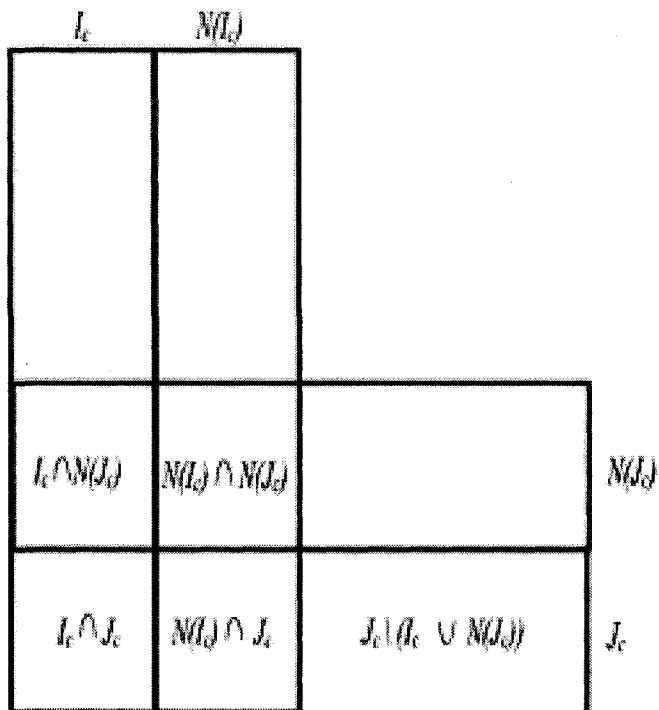


Figure 3.4: Useful figures for following the proofs of Lemma 3.35 and Theorem 3.36. The figure on the top is a schematic of the relationship between critical independent sets I_c and J_c and their neighbors. The figure on the bottom is the same figure but shaded to identify set $I = I_c \cup J = I_c \cup J_c \setminus (I_c \cup N(I_c))$.

that a vertex v in $N(J_c) \setminus (I_c \cup N(I_c))$ cannot be matched to a vertex in $J_c \cap N(I_c)$ under any matching, as the proof of (1) guarantees that these are only matched to vertices in $I_c \cap N(J_c)$. Furthermore, a vertex v in $N(J_c) \setminus (I_c \cup N(I_c))$ cannot be matched to a vertex w in $I_c \cap J_c$. If it were, then since $w \in I_c$ and v is adjacent to w , it follows that $v \in N(I_c)$, contradicting the fact that $v \notin N(I_c)$. Thus vertices in $N(J_c) \setminus (I_c \cup N(I_c))$ can only be matched to vertices in $J_c \setminus (I_c \cup N(I_c))$. Since every vertex in $N(J_c) \setminus (I_c \cup N(I_c))$ is matched to a vertex in $J_c \setminus (I_c \cup N(I_c))$, it follows that $|J| = |J_c \setminus (I_c \cup N(I_c))| \geq |N(J_c) \setminus (I_c \cup N(I_c))|$, proving (2).

$I = I_c \cup J$. Since I_c and J are independent, and $J = J_c \setminus (I_c \cup N(I_c))$, I is independent. Since I_c and J are disjoint, $|I| = |I_c| + |J|$. $N(I) \subseteq N(I_c) \cup [N(J_c) \setminus (I_c \cup N(I_c))]$ and $|N(I)| \leq |N(I_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|$. So, $|I| - |N(I)| \geq (|I_c| + |J|) - (|N(I_c)| + |N(J_c) \setminus (I_c \cup N(I_c))|) = (|I_c| - |N(I_c)|) + (|J| - |N(J_c) \setminus (I_c \cup N(I_c))|)$. Since (2) implies that the last term is non-negative, it follows that $|I| - |N(I)| \geq |I_c| - |N(I_c)|$ and, thus, that I is a critical independent set, proving (3). \square

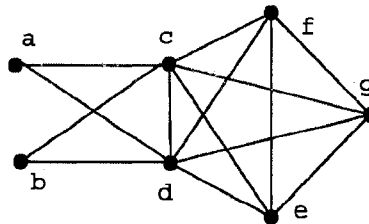


Figure 3.5: The vertices $I_c = \{a, b\}$ form a (maximum cardinality) critical independent set. The set $X = I_c \cup N(I_c)$ induces a decomposition of the graph into a totally independence reducible subgraph $G[X]$ and an independence irreducible subgraph $G[X^c]$, according to Theorem 3.36.

Theorem 3.36. *For any graph G , there is a unique set $X \subseteq V(G)$ such that*

1. $\alpha(G) = \alpha(G[X]) + \alpha(G[X^c])$,
2. $G[X]$ is totally independence reducible,

3. $G[X^c]$ is independence irreducible, and

4. for every maximum critical independent set J_c of G , $X = J_c \cup N(J_c)$.

Proof. Let I_c be a maximum critical independent set of G . Let $X = I_c \cup N(I_c)$. I_c is an independent set in $G[X]$.

Suppose I_c is not a maximum independent set in $G[X]$. Let Y be an independent set of $G[X]$ such that $|Y| > |I_c|$. Let $Y_I = Y \cap I_c$ and $Y_N = Y \cap N(I_c)$. So $Y = Y_I \cup Y_N$, $|Y_I| + |Y_N| = |Y|$, and $|Y_I| > |I_c| - |Y_N|$. Note that $N(Y_I) \subseteq N(I_c) \setminus Y_N$. Then, $|Y_I| - |N(Y_I)| \geq |Y_I| - |N(I_c) \setminus Y_N| > (|I_c| - |Y_N|) - |N(I_c) \setminus Y_N| = |I_c| - (|Y_N| + |N(I_c) \setminus Y_N|) = |I_c| - |N(I_c)|$. Since Y_I is an independent set, I_c cannot be a critical independent set of G , contradicting the fact that it is. Thus, no independent set of $G[X]$ can have cardinality greater than I_c , and $\alpha(G[X]) = |I_c|$.

It follows from Butenko and Trukhanov's Theorem 3.9 that I_c is contained in a maximum independent set I of G . So $\alpha(G) = |I|$. $I \setminus I_c$ is an independent set in X^c . So $\alpha(G[X^c]) \geq |I \setminus I_c|$. Suppose there is an independent set $I' \subseteq X^c$ such that $|I'| > |I \setminus I_c|$. By construction, no vertex in I_c is adjacent in G to a vertex in X^c . Thus, no vertex in I_c is adjacent to a vertex in I' . Thus, $I_c \cup I'$ is an independent set in G , and $\alpha(G) \geq |I_c \cup I'| = |I_c| + |I'| > |I_c| + |I \setminus I_c| = |I| = \alpha(G)$, a contradiction. Thus, $I \setminus I_c$ is a maximum independent set in $G[X^c]$, and $\alpha(G[X]) + \alpha(G[X^c]) = |I_c| + |I \setminus I_c| = |I| = \alpha(G)$, proving (1).

Now suppose I_c is not a critical independent set in $G[X]$. Let Y be a minimum critical independent set of $G[X]$. So $|Y| - |N_{G[X]}(Y)| > |I_c| - |N(I_c)|$. Let $Y_I = Y \cap I_c$ and $Y_N = Y \cap N(I_c)$. (Note that $N(I_c)$ is unambiguous as $N_G(I_c) = N_{G[X]}(I_c)$.) Let $Y'_N \subseteq I_c$ be the set of neighbors of Y_N in I_c . It follows from the Matching Lemma 3.7 that there is a matching of the vertices in $N_{G[X]}(Y)$ to (a subset of) the vertices in Y . Since I_c is an independent set, and $Y'_N \subseteq I_c$, the vertices in Y'_N must be matched to vertices in Y_N . Thus, $|Y_N| \geq |Y'_N|$.

Suppose $|Y_N| = |Y'_N|$. Then $|Y_I| - |N(Y_I)| = (|Y_I| + |Y_N|) - (|N(Y_I)| + |Y_N|) =$

$|Y| - (|N(Y_I)| + |Y'_N|) \geq |Y| - |N_{G[X]}(Y)|$, implying that Y_I is a critical independent set of $G[X]$. Since $Y_I \subseteq Y$, and Y is a minimum critical independent set, it follows that $Y_I = Y$, and $Y_N = \emptyset$. Since $Y_I \subseteq I_c$, $N_G(Y_I) = N_{G[X]}(Y_I)$, and $|Y_I| - |N_G(Y_I)| \geq |Y| - |N_{G[X]}(Y)| > |I_c| - |N_G(I_c)|$, contradicting the fact that I_c is a critical independent set in G .

So $|Y_N| > |Y'_N|$. But then, for $I = I_c \setminus Y'_N$, $|I| - |N_G(I)| = |I_c \setminus Y'_N| - |N(I_c) \setminus Y_N| = |I_c| - |Y'_N| - (|N(I_c)| - |Y_N|) = (|I_c| - |N(I_c)|) + (|Y_N| - |Y'_N|)$. Since the last term is positive, it follows that $|I| - |N_G(I)| > |I_c| - |N(I_c)|$, again contradicting the fact that I_c is a critical independent set in G . Thus, I_c is a critical independent set in $G[X]$. Since $I_c \cup N(I_c) = X$, I_c is a maximum critical independent set in $G[X]$, and $\alpha'(G[X]) = |I_c|$. So $\alpha(G[X]) = \alpha'(G[X]) = |I_c|$ and $G[X]$ is totally independence reducible, proving (2).

Suppose that $G[X^c]$ contains a non-empty critical independent set Z . So $|Z| \geq |N_{G[X^c]}(Z)|$. No vertex in I_c is adjacent to any vertex in Z as $N(I_c) \subseteq X$ and $Z \subseteq X^c$. So $I_c \cup Z$ is an independent set in G . Furthermore, $|N(I_c \cup Z)| = |N(I_c)| + |N_{G[X^c]}(Z)|$. So, $|I_c \cup Z| - |N_G(I_c \cup Z)| = (|I_c| + |Z|) - (|N(I_c)| + |N_{G[X^c]}(Z)|) = (|I_c| - |N(I_c)|) + (|Z| - |N_{G[X^c]}(Z)|) \geq |I_c| - |N(I_c)|$, contradicting the fact that I_c is a maximum critical independent set of G . Thus, $G[X^c]$ does not contain a non-empty critical independent set, $\alpha'(G[X^c]) = 0$, and $G[X^c]$ is irreducible, proving (3).

Now suppose that J_c is a maximum critical independent set of G . Thus, since J_c and I_c are both maximum critical independent sets, $|J_c| = |I_c|$. Since they are both critical, $|J_c| - |N(J_c)| = |I_c| - |N(I_c)|$. It then follows that $|N(J_c)| = |N(I_c)|$. Let $J = J_c \setminus (I_c \cup N(I_c))$. So $I_c \cup J$ is an independent set. Lemma 3.35 implies that $I_c \cup J$ is a critical independent set of G . But, since $I_c \subseteq I_c \cup J$ and I_c is a maximum critical independent set of G , $J = \emptyset$. A parallel argument yields that $I = \emptyset$.

The Matching Lemma 3.7 implies that there is a matching from the vertices in $N(J_c)$ into the vertices in J_c . Lemma 3.35 implies that $|I_c \cap N(J_c)| = |J_c \cap N(I_c)|$. So

if $v \in N(J_c) \setminus (N(J_c) \cap I_c)$, it must be matched to a vertex in $J_c \setminus (J_c \cap N(I_c) = I_c \cap J_c)$ and, thus, $v \in N(I_c \cap J_c) \subseteq N(I_c) \cap N(J_c)$. So every vertex in $N(J_c)$ is either in $N(J_c) \cap I_c$ or in $N(J_c) \cap N(I_c)$, which implies that $N(J_c) \subseteq I_c \cup N(I_c)$.

So both J_c and $N(J_c)$ are subsets of $I_c \cup N(I_c)$. Since $|J_c| + |N(J_c)| = |I_c| + |N(I_c)|$, it follows that $J_c \cup N(J_c) = I_c \cup N(I_c) = X$, proving (4).

The uniqueness of a set $X \subseteq V(G)$ satisfying the four conditions of the theorem follows immediately from (4).

□

3.4 An application: König-Egervary Graphs

The independence number of a graph is denoted α , the critical independence number is α' , the matching number is μ , and the vertex covering number is τ . One of the Gallai Identities is that, for any graph, $\alpha + \tau = n$ [70, p. 2]. For bipartite graphs, $\alpha + \mu = n$ (this is the König-Egervary theorem, [70]). A *König-Egervary graph* is a graph that satisfies this identity. There are non-bipartite König-Egervary graphs: the graph in Figure 3.4 is an example. König-Egervary graphs were first characterized by Deming in 1979 [18].

Ermelinda DeLaVina's program Graffiti.pc conjectured that, for any graph, $\alpha = \alpha'$ if, and only if, $\tau = \mu$. The proof of this conjecture yields a new characterization of König-Egervary graphs. The Graffiti.pc conjecture can be rewritten: for any graph, $\alpha = \alpha'$ if, and only if, $\alpha + \mu = n$; or, for a graph G , $\alpha(G) = \alpha'(G)$ if, and only if, G is a König-Egervary graph.

A graph is *independence irreducible* if $\alpha' = 0$. For these graphs the number of neighbors of any independent set of vertices is greater than the number of vertices in the set. A graph is *independence reducible* if $\alpha' > 0$. A graph is *totally independence reducible* if $\alpha' = \alpha$. So Graffiti.pc's conjecture can be restated again: A graph G is

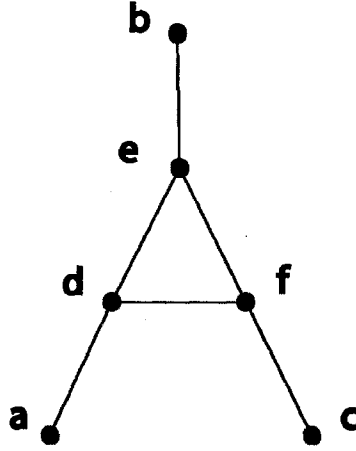


Figure 3.6: A non-bipartite König-Egervary graph. The graph has vertex set $V = \{a, b, c, d, e, f\}$ and edge set $E = \{ad, be, cf, de, ef, df\}$. So $n = |V| = 6$. The set $I = \{a, b, c\} \subseteq V$ is a maximum independent set. So $\alpha = |I| = 3$. The set $M = \{ad, be, cf\} \subseteq E$ is a maximum matching. So $\mu = |M| = 3$. Thus, $\alpha + \mu = n$ and the graph is König-Egervary.

totally independence reducible if, and only if, G is a König-Egervary graph.

Theorem 3.37. (*Graffiti.pc #329*) For any graph, $\alpha = \alpha'$ if, and only if, $\tau = \mu$.

Proof. Suppose that $\alpha(G) = \alpha'(G)$. It will be shown that $\tau(G) = \mu(G)$ or, equivalently, that $n - \alpha(G) = \mu(G)$.

Let I be a maximum critical independent set. So $\alpha(G) = \alpha'(G) = |I|$. Since $n - \alpha(G) = |N(I)|$, it remains to show that $\mu(G) = |N(I)|$. Since I is independent, $\mu(G) \leq |N(I)|$. It only remains to show that $\mu(G) \geq |N(I)|$. Since I is a critical independent set, the Matching Lemma 3.7 implies that there is a matching from $N(I)$ into I and, thus, that $\mu(G) \geq |N(I)|$.

Suppose now that $\tau(G) = \mu(G)$ or, equivalently, that $n - \alpha(G) = \mu(G)$. It will be shown that $\alpha(G) = \alpha'(G)$. $\alpha'(G) \leq \alpha(G)$. Suppose $\alpha'(G) < \alpha(G)$. Let I_c be a maximum critical independent set. Butenko and Trukhanov proved that every critical independent set is contained in a maximum independent set [10]. Let J be a maximum independent set such that $I_c \subseteq J$. Since $\mu(G) = n(G) - \alpha(G)$, J is

independent, and $|V \setminus J| = n(G) - \alpha(G)$, there is a matching from $V \setminus J$ into J . This implies that each vertex in $N(J) \setminus N(I_c)$ is matched to a vertex in $J \setminus I_c$. So $|J \setminus I_c| \geq |N(J) \setminus N(I_c)|$.

It will now be shown that $|J| - |N(J)| \geq |I_c| - |N(I_c)|$, implying that I_c is not a maximum critical independent set, as it was assumed to be. $|J| - |N(J)| = (|J \setminus I_c| + |I_c|) - (|N(J) \setminus N(I_c)| + |N(I_c)|) = (|I_c| - |N(I_c)|) + (|J \setminus I_c| - |N(J) \setminus N(I_c)|) \geq |I_c| - |N(I_c)|$. It follows that $I_c = J$, $|I_c| = |J|$, and $\alpha'(G) = \alpha(G)$, which was to be shown. \square

3.5 Open Problems on Critical Independent Sets

1. Critical Independent Sets

- (a) Is every critical weighted independent set contained in a maximum critical weighted independent set?
- (b) Is it possible to use critical independent sets to speed up algorithms for finding maximum independent sets and the independence number in independence irreducible graphs?
- (c) Do 2-connected independence irreducible graphs having an even number vertices have perfect matchings? Are 2-connected independence irreducible graphs having an odd number vertices pseudo-perfect?
- (d) Are power law graphs independence reducible?

2. The Binding and Isoperimetric Numbers

Zhang defined the notions of a critical set and a critical independent set in a graph partly because of their connection to the pre-existing interest of independent sets in graphs, and partly because of their connection to the *binding number* and *isoperimetric number* of a graph [84]. These concepts are defined below. The connection

between critical independent sets and the binding number is explained here. *What is the connection between critical independent sets and the isoperimetric number?*

- The **binding number** of a graph G is

$$b(G) = \min\left\{\frac{|N(U)|}{|U|} : U \subseteq V(G), U \neq \emptyset \text{ and } N(U) \neq V(G)\right\}.$$

- The **boundary** $\delta(U)$ of a set U is the set of edges incident to exactly one vertex in U .
- The **isoperimetric number** or **Cheeger constant** of a graph G is

$$i(G) = \min\left\{\frac{|\delta(U)|}{|U|} : U \subseteq V(G), U \neq \emptyset \text{ and } |U| \leq \frac{|V(G)|}{2}\right\}.$$

Let G be a graph and U be a set such that $b(G) = \frac{|N(U)|}{|U|}$. Then $|N(U)| = b(G)|U|$.

Theorem 3.38. *For a connected graph G with binding number $b(G)$:*

- $b \leq 1$ if, and only if, there is a non-empty critical independent set (that is, G is independence reducible).
- $b > 1$ if, and only if, the empty set is the only critical independent set (that is, G is independence irreducible).

Proof. (1) and (2) are obviously equivalent, so it is enough to prove (1). Suppose $b(G) = b \leq 1$. There is a non-empty set U , $N(U) \neq V$, such that $b = \frac{|N(U)|}{|U|}$. Then $|N(U)| = b|U| \leq |U|$ and $|U| - |N(U)| \geq 0$. Let $I = U \setminus N(U)$. Then $|I| - |N(I)| = |U| - |N(U)| \geq 0$. Suppose $I = \emptyset$. Then $U \subseteq N(U)$. So $|U| \leq |N(U)|$ and, thus, $|U| = |N(U)|$. Since G is connected, it follows that $N(U) = V(G)$, contradicting the definition of U .

Now, suppose G has a non-empty critical independent set I . That is, let I be a non-empty independent set that maximizes $|I| - |N(I)|$. So $|I| - |N(I)| \geq 0$, $|I| \geq |N(I)|$, and $\frac{|N(I)|}{|I|} \leq 1$. Since $b = b(G)$ is the minimum of $\frac{|N(U)|}{|U|}$ for all non-empty sets U , $U \neq V$. $b \leq \frac{|N(I)|}{|I|} \leq 1$. □

Bibliography

- [1] A. Ageev, On Finding Critical Independent and Vertex Sets, *SIAM J. Discrete Mathematics*, (1994) 7:293–295 .
- [2] H. Aldersey-Williams, *The Most Beautiful Molecule: The Discovery of the Buckyball*, Wiley, 1997.
- [3] D. Archdeacon, J. Ellis-Monaghan, D. Fisher, D. Froncek, P.C.B. Lam, S. Seager, B. Wei, and R. Yuster, Some Remarks on Domination, *Journal of Graph Theory*, (2004) 207–210.
- [4] A. Balaban and O. Ivanciuc, Historical Development of Topological Indices, in: *Topological Indices and Related Descriptors in QSAR and QSPR*, J. Devillers and A. Balaban (eds.), Gordon and Breach, The Netherlands, 1999, pp. 21–57.
- [5] N. Biggs, E. K. Lloyd, R. Wilson, *Graph Theory: 1736–1936*, Clarendon Press, Oxford, 1976.
- [6] B. Bollobas, *Random Graphs*, 2nd edition, Cambridge University Press, Cambridge, 2001.
- [7] B. Bollobas and P. Erdős, Graphs of Extremal Weights, *Ars Combinatoria* 50 (1998).
- [8] D. Bonchev and D. Rouvray (eds.), *Chemical Graph Theory: Introduction and Fundamentals*, Gordon and Breach, New York, 1991.

- [9] J.A. Bondy and U.S.R. Murty, Graph Theory with Applications (corrected first paperback edition), Macmillan Press, London, 1976.
- [10] S. Butenko and S. Trukhanov, Using Critical Sets to Solve the Maximum Independent Set Problem, *Operations Research Letters* 35(4) (2007) 519–524.
- [11] F. Chung, The Average Distance and the Independence Number, *Journal of Graph Theory* 12(2) (1988) 229–235.
- [12] V. Chvatal, and P. Erdős, A Note on Hamiltonian Circuits, *Discrete Mathematics*, 2 (1972) 111–113.
- [13] D. M. Cvetkovic, M. Doob, H. Sachs, Spectra of Graphs, 3rd revised ed., Johann Ambrosius Barth Verlag, Heidelberg, 1995.
- [14] E. DeLaVina, Graffiti.pc, *Graph Theory Notes of New York* XLII (2002) 26–30.
- [15] E. DeLaVina, Some History of the Development of Graffiti, S. Fajtlowicz, P. Fowler, P. Hansen, M. Janowitz, and F. Roberts (Editors), American Mathematical Society (2005), 81–118.
- [16] E. DeLaVina, S. Fajtlowicz, W. Waller, On Some Conjectures of Griggs and Graffiti, in: Graphs and Discovery, S. Fajtlowicz, P. Fowler, P. Hansen, M. Janowitz, and F. Roberts (Editors), American Mathematical Society (2005), 119–125.
- [17] E. DeLaVina, R. Pepper and B. Waller, Independence, Radius and Hamiltonian Paths, *MATCH Commun. Math. Comput. Chem: proceedings of the conference “Computers in Scientific Discovery III”*, Ghent, February 6-9, 2006 (eds. G. Brinkmann, P. W. Fowler) (2007) Vol. 58, 481–510.
- [18] R. W. Deming, Independence Numbers of Graphs—an Extension of the Koenig-Egervary Theorem, *Discrete Mathematics*, 27 (1979) 23–33.

- [19] P. Erdős, M. Saks, and V. Sos, Maximum Induced Trees in Graphs, *J. Combinat. Theory Ser. B*, 41 (1986) 61–69.
- [20] P. Erdős, J. Spencer and J. Pach, On the Mean Distance Between Points of a Graph, *Ars Combinatoria* (1986).
- [21] P. Erdős, S. Fajtlowicz and W. Staton, Degree Sequences in Triangle-free Graphs, *Discrete Mathematics* 92 (1991) 85–88.
- [22] S. Fajtlowicz, A Characterization of Radius-Critical Graphs, *Journal of Graph Theory* 12(4) (1988) 529–532.
- [23] S. Fajtlowicz, On Conjectures of Graffiti, *Discrete Mathematics* 72 (1988) 113–118.
- [24] S. Fajtlowicz, On Conjectures of Graffiti, II, *Congressus Numerantium* 60 (1987) 189–197.
- [25] S. Fajtlowicz, On Conjectures of Graffiti, III, *Congressus Numerantium* 66 (1988) 23–32.
- [26] S. Fajtlowicz, On Conjectures of Graffiti, IV, *Congressus Numerantium* 70 (1990) 231–240.
- [27] S. Fajtlowicz, On Conjectures of Graffiti V, in: *Graph Theory, Combinatorics, and algorithms: Proceedings of the Seventh Quadrennial Conference on the Theory and Application of Graphs*, Western Michigan University, vol. 1 (1995) 367–376.
- [28] S. Fajtlowicz, University of Houston, Personal Communication, 2006.
- [29] S. Fajtlowicz, Toward Fully Automated Fragments of Graph Theory, *Graph Theory Notes of New York* XLII (2002) 18–25.

- [30] S. Fajtlowicz, Written on the Wall: Conjectures of Graffiti, available on the WWW at: <http://www.math.uh.edu/~clarson/graffiti.html>
- [31] S. Fajtlowicz, P.W. Fowler, H. Hansen and K.M. Rogers, C60Br24 as a Chemical Illustration of Graph Theoretical Independence, *Journal of the Chemical Society, Perkin Transactions 2* (1998) 1531–1533.
- [32] S. Fajtlowicz, P. John and H. Sachs, On Maximum Matchings and Eigenvalues of Benzenoid Graphs, *Croatia Chemica Acta CCACAA* 78(2) (2005) 195–201.
- [33] S. Fajtlowicz, and C. E. Larson, Graph-theoretic Independence as a Predictor of Fullerene Stability, *Chemical Physics Letters* 377 (2003) 485–490.
- [34] S. Fajtlowicz, T. McColgan, T. J. Reid, and W. Staton, Ramsey Numbers of Induced Regular Subgraphs, *Ars Combinatoria* 39 (1995) 149–154.
- [35] S. Fajtlowicz and W. Waller, On Two Conjectures of Graffiti, *Congressus Numerantium* 55 (1986) 51–56.
- [36] O. Favaron, M. Mahéo, J.-F. Saclé, On the residue of a graph, *Journal of Graph Theory* 15 1 (1991) 39–64.
- [37] P. W. Fowler, Fullerene Graphs With More Negative Than Positive Eigenvalues; The Exceptions That Prove The Rule of Electron Deficiency, *Journal Chem. Soc. Faraday* 93 (1997) 1–3.
- [38] P. W. Fowler, P. Hansen, G. Caporossi and A. Soncini, Polyenes with Maximum HOMO-LUMO Gap, *Chemical Physics Letters* 342 (2001) 105–112.
- [39] P. Fowler and D. Manolopoulos, *An Atlas of Fullerenes*, Clarendon Press, Oxford, 1995.
- [40] P. Fowler, S. Daugherty, W. Myrvold, Independence Number and Fullerene Stability, *Chemical Physics Letters* 448(1-3) (2007) 75–82.

- [41] P. W. Fowler, K. Rogers, K. Somers, and A. Troisi, Independent Sets and the Prediction of Addition Patterns for Higher Fullerenes, *J. Chem. Soc., Perkin Trans. 2*, 1999, 2023–2027.
- [42] M. Garey and D. Johnson, *Computers and Intractability*, W. H. Freeman and Company, New York, 1979.
- [43] R. Gould, Updating the Hamilton Problem—A Survey, *J. of Graph Theory*, **15** (1991) 121–157.
- [44] J. Griggs and D. Kleitman, Independence and the Havel-Hakimi Residue, *Discrete Mathematics* 127 1-3 (1994) 209–212.
- [45] J. Gross and J. Yellen (eds.), *Handbook of Graph Theory*, CRC Press, Boca Raton, 2004.
- [46] G. Gutin, Independent Sets and Cliques, in: J. Gross and J. Yellen (eds.), *Handbook of Graph Theory*, CRC Press, Boca Raton, 2004, 389–402.
- [47] I. Gutman and O. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin, 1986.
- [48] J. Hadamard, *The Psychology of Invention in the Mathematical Field* Dover Publications, New York, 1954.
- [49] B. Hayes, Graph Theory in Practice, *American Scientist* 88 (2000) 9–13 (Part I), 104–109 (Part II).
- [50] G. H. Hardy *A Mathematician's Apology* Cambridge University Press, Cambridge, 1967.
- [51] G. Hopkins and W. Staton, Graphs with Unique Maximum Independent Sets, *Discrete Mathematics* 57 (1985) 245–251.

- [52] A. Jackson, Million-dollar Mathematics Prizes Announced, *Notices of the American Mathematical Society* 47 (8), 2000, 877-879.
- [53] A. Kotlov and L. Lovasz, The Rank and Size of Graphs, *Journal of Graph Theory* 23(1) (1996) 185–189.
- [54] L. Kucera, Combinatorial Algorithms, Adam Hilger, Bristol, 1990.
- [55] C. E. Larson, Intelligent Machinery and Mathematical Discovery, *Graph Theory Notes of New York XLII* (2002) 8–17.
- [56] C. E. Larson, A Survey of Research on Automated Mathematical Conjecture-Making, Graphs and Discovery, S. Fajtlowicz, P. Fowler, P. Hansen, M. Janowitz, and F. Roberts (Editors), American Mathematical Society (2005) 297–318.
- [57] C. E. Larson, A Note on Critical Independent Sets, *Bulletin of the Institute of Combinatorics and its Applications* 51 (2007) 34–46.
- [58] D. Lenat, AM: Discovery in Mathematics as Heuristic Search, in: R. Davis and D. Lenat, eds., *Knowledge-Based Systems in Artificial Intelligence* McGraw-Hill, New York, 1982, 1–225.
- [59] D. Lenat, Artificial Intelligence, *Scientific American*, Sept. 1995, 80–82.
- [60] D. Lenat, Automated Theory Formation in Mathematics, in: W. Bledsoe and D. Loveland, eds., *Automated Theorem Proving: After 25 Years* American Mathematical Society, Providence, 1984, 287–314.
- [61] D. Lenat, Computer Software for Intelligent Systems, *Scientific American*, Sept. 1984, 204–213.
- [62] D. Lenat, Eurisko: A Program that Learns New Heuristics and Domain Concepts, *Artificial Intelligence* 21 (1983) 61–98.

- [63] D. Lenat, On Automated Scientific Theory Formation: A Case Study Using the AM Program, *Machine Intelligence* 9 (1979) 251–283.
- [64] D. Lenat, The Nature of Heuristics, *Artificial Intelligence* 19 (1982) 189–249.
- [65] D. Lenat, The Role of Heuristics in Learning by Discovery: Three Case Studies, in: R. Michalski, J. Carbonell, and T. Mitchell, eds., *Machine Learning* (Morgan Kaufmann, Los Altos, CA, 1983) 243–306.
- [66] D. Lenat, Theory Formation by Heuristic Search, *Artificial Intelligence* 21 (1983) 31–59.
- [67] D. Lenat, The Ubiquity of Discovery, *Artificial Intelligence* 9 (1978) 257–285.
- [68] X. Liu, T. G. Schmalz, D. J. Klein, Generation of Carbon-cage Polyhedra, *Journal of Computational Chemistry* 12 (10) 1265–1269.
- [69] X. Liu, T. G. Schmalz, D. J. Klein, Favorable Structures for Higher Fullerenes, *Chemical Physics Letters* 188 (1992) 550–554.
- [70] L. Lovasz, M. D. Plummer, *Matching Theory*, North Holland, Amsterdam, 1986.
- [71] R. Merrifield and H. Simmons, The Structure of Molecular Topological Spaces, *Theoretica Chimica Acta* 55 (1980) 55–75.
- [72] R. Pepper, On New Didactics of Mathematics: Learning Graph Theory via Graffiti, in: *Graphs and Discovery*, S. Fajtlowicz, P. Fowler, P. Hansen, M. Janowitz, and F. Roberts (Editors), American Mathematical Society, 2005, 341–350.
- [73] R. Pepper, Binding Independence, Ph.D. Dissertation, University of Houston, 2004.
- [74] K. Raghavachari, Ground State of C_{84} : Two Almost Isoenergetic Isomers *Chemical Physics Letters* 190 (1992) 397–400.

- [75] M. Ramras, Maximal Independent Sets in Graphs with Vertex m -covers, *Congressus Numerantium*, 39 (1983) 463–472.
- [76] G. Ritchie and F. Hanna, AM: A Case Study in AI Methodology, *Artificial Intelligence* 23(3) (1984) 249–268.
- [77] J. M. Robson, Algorithms for Maximum Independent Sets, *Journal of Algorithms* 7 (1986) 425–440.
- [78] D. Rouvray, The Origins of Chemical Graph Theory, in: D. Bonchev and D. Rouvray (eds.), *Chemical Graph Theory: Introduction and Fundamentals*, Gordon and Breach, New York, 1991, 1–39.
- [79] S. Smale, Mathematical Problems for the Next Century, *Mathematical Intelligencer* 20 (1998) No. 2, 7-15.
- [80] R. Tarjan and A. Trojanowski, Finding a Maximum Independent Set, *SIAM J. Comput.* 6, No.3, 1977, 537–546.
- [81] N. Triajstić, Graph Theory and Molecular Orbitals, in: D. Bonchev and D. Rouvray (eds.), *Chemical Graph Theory: Introduction and Fundamentals*, Gordon and Breach, New York, 1991, 235–279.
- [82] H. Wang, Computer Theorem Proving and Artificial Intelligence, in: W. Bledsoe and D. Loveland, eds., *Automated Theorem Proving: After 25 Years* American Mathematical Society, Providence, 1984, 49–70.
- [83] H. Wang, Toward Mechanical Mathematics, in: J. Siekmann and G. Wrightson, eds., *Automation of Reasoning: Classical Papers on Computational Logic 1957–1966* Springer-Verlag, Berlin, 1983, 244–264.
- [84] C.-Q. Zhang, Finding Critical Independent Sets and Critical Vertex Subsets are Polynomial Problems, *SIAM J. Discrete Mathematics*, 3:431–438, 1990.