THE INDEPENDENCE NUMBER PROJECT: INTRODUCTION

C. E. LARSON
DEPARTMENT OF MATHEMATICS AND APPLIED MATHEMATICS
VIRGINIA COMMONWEALTH UNIVERSITY

1. Introduction

Our research objective is to find an efficient algorithm for the independence number of a graph. Graphs can be used to model molecular structure [14, 19], the World Wide Web [43], social networks [39], GPS satellite communications [9], and the possibility of unambiguous message communication in a channel [34]. Calculating the independence number of a graph can lead to predictions of molecular stability [46], optimal network configuration, and the probability of unambiguous message transmission. Stable fullerenes have been shown to minimize their independence number; this is one of the best predictors of fullerene stability [17]. The independence number of a graph is a central concept of two of the most studied and important problems in graph theory: the P vs. NP question [18], and Hadwiger's Conjecture [15, 37, 11]. Many families of combinatorial objects including error-correcting codes, set packings in Hamming spaces, and balanced incomplete block designs can be viewed as maximum independent sets [42].

No efficient algorithm is known for computing the independence number of a general graph; all known algorithms require an exponential number of steps (as a function of the graph's representation). Many graph invariants, most famously the matching number [16], can be computed efficiently. Computing the independence number is computationally equivalent to hundreds of other widely-studied problems, and of enormous practical importance; these include scheduling, routing, and data storage problems [18], where small theoretical advances can lead to large practical payoffs. While most researchers believe that no algorithm is possible for the efficient calculation of the independence number of a general graph (and, equivalently, that P$\neq$NP), some, including Bollobás, think an algorithm is possible, that P=NP [6]. In any case there is a class of graphs, perhaps not all, where the independence number can be computed efficiently. It is of great importance to identify this class.

Date: August 4, 2012.
Original version: August 4, 2012.

The independence number can be efficiently computed for all graphs with up to n vertices, for some n. There will necessarily be graphs with $n+1$ vertices which are "difficult" with respect to existing independence number theory—a well-developed theory consisting of α-bounds, α-properties, and α-reductions. We will identify these difficult graphs, find new α-bounds, α-properties, and α-reductions, partly by computer identification of stable blocks, and by the use of automated conjecture-making programs. After difficult graphs are identified, the goal is to extend the theory to cover these difficulties. This process is then iterated.

Existing theory of the independence number include exact and approximate algorithms for calculating the independence number [22, 20, 3, 21, 24, 25], bounds, structure (including, for instance, well-covered graphs [49, 48, 23]), and asymptotics on random graphs [5, 7]. A large number of independence number bounds have been reported in the mathematical literature over the last 40 years. Lovász has written, "Deriving sharper and sharper upper bounds, more and more insight could be gained into the nature of independence number (a procedure vaguely reminiscent of the expansion of a function into, say, a Fourier series" [35].

While there are many NP-hard invariants, and hundreds of papers developing theory and algorithms for these, this proposed project represents a novel approach. There has been a reasonable amount of research on algorithms for the independence number of a graph; the goal in this case has been to improve the speed of existing algorithms and, hence, the size of graphs where the independence number can be calculated. This project proposes instead to identify exactly where existing independence number theory breaks down in the *efficient* calculation of the independence number of a graph, and then to improve the theory.

2. Background, Motivation, and Problem Description

An *independent set* in a graph is set of vertices which are pairwise non-adjacent; the *independence number* (or *stability number*) of a graph is the cardinality of a maximum independent set. The following notation is used throughout: the vertex set of a graph G is $V(G)$, the order of G is $n = n(G) = |V(G)|$, the set of neighbors of a vertex v is $N(v)$, the set of neighbors of a set $S \subseteq V(G)$ in G is $N(S) = \cup_{u \in S} N(u)$, the graph induced on S is $G[S]$, and the independence number is $\alpha = \alpha(G)$. All graphs are assumed to be finite and simple.

Several well-known graph invariants are closely related to the independence number. Some are computationally equivalent: an algorithm for any of these invariants can be efficiently converted into an algorithm for independence number. The clique number of a graph is the independence number of its complement. The covering number and the independence number sum to the number of vertices of the graph.

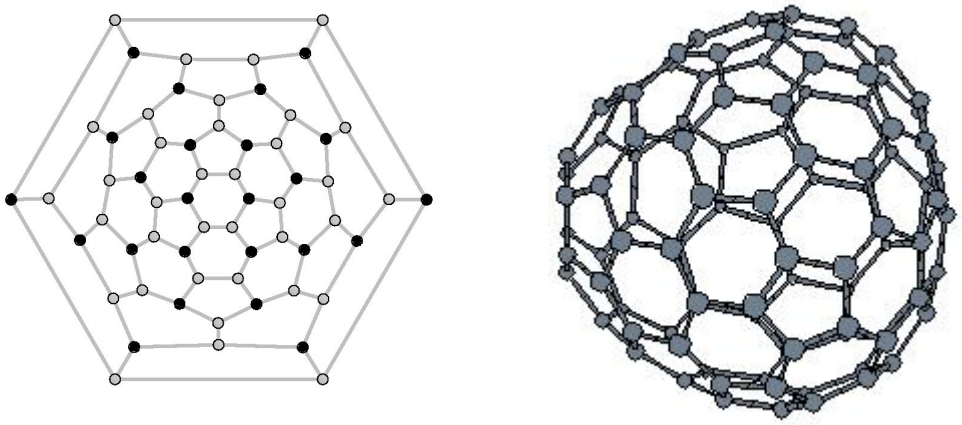

FIGURE 1. The graph on the left is Buckminsterfullerene. Of 1812 C_{60} isomers it is the only stable one, and the unique one which minimizes its independence number. The black vertices are a maximum independent set. The molecule on the right is tetrahedral C_{100}, the lone 100-atom IPR which minimizes its independence number (of 285, 914 isomers). Fajtlowicz and Larson predict this to be the stable 100-atom isomer [17].

Less directly, the domination number of a graph is never more than the independence number, and the product of the independence number and the chromatic number is at least the number of vertices of the graph.

A lower bound for the independence number of a graph is a graph invariant l such that, for every graph G, $l(G) \leq \alpha(G)$. Similarly, an upper bound for the independence number is a graph invariant u such that, for every graph G, $\alpha(G) \leq u(G)$. Many efficiently computable upper and lower bounds, called **α-bounds** here, have been published and these are surveyed in the following section. They can be used to predict the value of α. Suppose $l_1, l_2, \ldots, l_k$ are efficiently computable lower bounds for the independence number of a graph; then $l = \max\{l_1, l_2, \ldots, l_k\}$ is also an efficiently computable lower bound for the independence number. Similarly, if $u_1, u_2, \ldots, u_m$ are efficiently computable upper bounds for the independence number, then $u = \min\{u_1, u_2, \ldots, u_m\}$ is also an efficiently computable upper bound for the independence number. For some graphs G, $l(G) = u(G)$ and, in such cases, it follows that the independence number $\alpha(G) = l(G) = u(G)$ can be directly computed from its bounds. For instance, consider the graph consisting of the cycle C_4 with a diagonal. It is known that, for every graph $\alpha_c \leq \alpha$ and, for every graph, $\alpha \leq \alpha_f$, where α_c is the critical independence number and α_f is the fractional independence

number. These bounds are both efficiently computable and, for this graph, equal 2. Thus the theory implies that $\alpha = 2$.

New efficiently computable independence number bounds are also of practical interest: they can lead to faster independence number computations. New bounds can lead to new exact predictions of the independence number of a graph, without any need for computer search of subsets of vertices or calculating independence numbers of subgraphs of the given graph. If it is known that α must lie in the interval $[l, u]$ then only subsets of sizes in this range must be considered. In some instances theoretical upper and lower bounds for α can be used to predict the independence number with no further search (in this case the theory predicts that α lies in an interval where $l = u$). It is known that the independence of a graph is equal to the larger of the independence numbers of two proper subgraphs [51]. So computation of the independence number depends on the computation of the independence number of two proper subgraphs G_1 and G_2. If the theoretical bounds predict that either of these has independence number no more than l then no further computation for that graph is needed and, if the theoretical bounds imply that one of these graphs has independence number u then the original graph must have independence number u.

The current fastest algorithm for the independence number of a graph whose computational complexity is known is due to Robson; the complexity is $\mathcal{O}(2^{.276})$ [50]. Research on exact algorithms begins with the 1977 paper of Tarjan and Trojanowski [51]; this research has progressed steadily [50, 4, 45, 44, 1, 8], with the current fastest program being Östergård's CLIQUER program [41, 40].

A decision problem belongs to class P if there is a polynomial time algorithm for determining the answer to an instance the problem (as a function of the input size). A decision problem belongs to NP roughly if a solution to the problem can be checked in polynomial time. The INDEPENDENT SET decision problem is, given a graph G and an integer k, to determine whether the graph contains an independent set of size at least k. This problem belongs to NP. While no efficient algorithm is known for answering this question, it is possible, given an independent set, to determine in polynomial time if the given set has the required size. This problem is one of the Millennium Problems [13], with an accompanying million dollar prize for its resolution. If P=NP then there is an efficiently computable upper bound u and lower bound l so that $l = \alpha = u$. If P$\neq$NP then this is not possible but there will still be a large class of graphs for which such bounds exist—and it remains of mathematical importance to characterize this class.

A graph property P is an **α-property** if it can both be determined efficiently (1) whether or not an arbitrary graph has that property and, (2) in the case the graph does have property P, whether the independence number of the graph can be computed efficiently. An example of an α-property is the property of being bipartite. It

can be determined efficiently whether a graph is bipartite and also what its matching number μ is [36]; the König-Egerváry Theorem says that if a graph is bipartite then $\alpha = n - \mu$, so α can be computed efficiently.

An **α-reduction** is an efficient transformation of an independence number calculation on a graph G into one on a graph or graphs with fewer vertices. So, for instance, if G has a pendant vertex v, it can be included in some maximum independent set; so $\alpha(G) = 1 + \alpha(G - N[v])$; that is, the problem of finding the independence number for G can be reduced to that of finding the independence number for the smaller graph $G - N[v]$.

An important idea in calculating the independence number of a graph G is to try to partition the vertices into non-trivial sets A and B so that $\alpha(G) = \alpha(G[A]) + \alpha(G[B])$. What is always true, for any partition, is that $\alpha(G) \leq \alpha(G[A]) + \alpha(G[B])$. Let the *border* of a set of vertices S in a graph G, denoted $Bord(S)$, be the vertices which are adjacent to vertices in $V - S$. The *interior* of S in G, denoted $Int(S)$, is the set of vertices in S which are not adjacent to any vertices in $V - S$; so $Int(S) = S - Bord(S)$. So, for every set S, $\alpha(G) \geq \alpha(G[Int(S)]) + \alpha(G[V - S])$. Since $\alpha(G) \leq \alpha(G[S]) + \alpha(G[V - S])$, equality holds, and it is possible to reduce the original independence number problem to that on proper subgraphs—in the case that there is a non-trivial set of vertices S with the property that $\alpha(G[S]) = \alpha(G[Int(S)])$. Such a set S is called a *stable block*. So it is useful to have efficient algorithms for finding stable blocks; some are discussed below. If S is a stable block in a graph G then the problem of finding the independence number for G can be reduced to that of finding the independence numbers of the two smaller graphs $G[S]$ and $G[V - S]$.

2.1. **König-Egerváry Theory and Fractional Independence.** One example of an α-reduction leads to the identification of a stable block and has a wide variety of applications. An independent set of vertices I_c is a *critical independent set* if $|I_c| - |N(I_c)| \geq |J| - |N(J)|$, for any independent set J. The definition of a critical independent set is due to Zhang [52], who showed that these could be found in polynomial time. The theory was then further developed by Ageev in [2]. 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. For some graphs the only critical independent set is the empty set; K_3 is an example. A *maximum critical independent set* is a critical independent set of maximum cardinality. It is easy to verify that, for any graph with at least three vertices, a maximum critical independent set must contain all pendant vertices; so a maximum critical independent set is a generalization of the set of pendants. The *critical independence number* of a graph G, denoted $\alpha_c = \alpha_c(G)$, is the cardinality of a maximum critical independent set. The author showed that maximum cardinality

critical independent sets, and thus α_c, can be found in polynomial-time [26]. Clearly α_c is a lower bound for α.

Butenko and Trukhanov showed that any critical independent set can be extended to a maximum independent set [10]. This result led to a number of recent papers [30, 32, 33, 31, 26, 27, 29, 12]. A maximum critical independent set in a graph is not unique but the union of a maximum independent set and its neighbors is unique, yielding the following Independence Decomposition Theorem (IDT) [27].

Theorem 2.1. *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 a König-Egerváry graph,
(3) for every non-empty independent set I in $G[X^c]$, $|N(I)| > |I|$, and
(4) for every maximum critical independent set J_c of G, $X = J_c \cup N(J_c)$.

A König-Egerváry (KE) graph is a graph where the independence number α and the matching number μ sum to the order n; they are generalizations of bipartite graphs. KE graphs have been widely studied, have a number of nice properties, can be identified in polynomial time and, significantly, their independence numbers can be computed in polynomial time [12]. Figure **??** provides an example of a decomposition according to the theorem. The vertices $I_c = \{a, b\}$ form a maximum cardinality critical independent set The sets $X = I_c \cup N(I_c) = \{a, b, c, d\}$ and $X^c = V \setminus X = \{e, f, g\}$ induce a decomposition of the graph which has the property that $G[X]$ is a KE graph and $G[X^c]$ has the property that every non-empty independent set has more neighbors than the cardinality of the set. The set X in this decomposition is necessarily a stable block: since $I_c \subseteq Int(X) \subseteq X$, it follows that $\alpha(G[X]) = \alpha(G[Int(X)])$. The graph in Figure **??** is an example of a graph where the problem of calculating its independence number can be efficiently α-reduced.

Jack Edmonds conjectured that the theory of critical independent sets and the IDT are equivalent to results on the linear programming relaxation of the integer programming formulation of the independence number problem due to Nemhauser, Trotter, Picard and Queyranne [38, 47]. This was proved by Larson [28]. The interaction between these theories may prove fruitful.

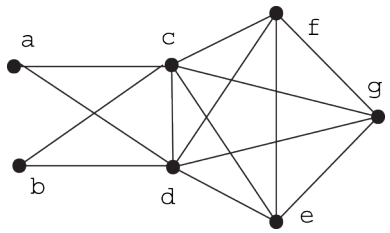

FIGURE 2. The sets $X = \{a, b, c, d\}$ and $X^c = \{e, f, g\}$ provide the unique decomposition guaranteed by the Independence Decomposition Theorem.

REFERENCES

[1] J. Abello, P. M. Pardalos, and M. G. C. Resende. On maximum clique problems in very large graphs. In *External memory algorithms (Piscataway, NJ, 1998)*, volume 50 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 119–130. Amer. Math. Soc., Providence, RI, 1999.

[2] Alexander A. Ageev. On finding critical independent and vertex sets. *SIAM J. Discrete Math.*, 7(2):293–295, 1994.

[3] Noga Alon and Nabil Kahale. Approximating the independence number via the θ-function. *Math. Programming*, 80(3, Ser. A):253–264, 1998.

[4] Egon Balas and Chang Sung Yu. Finding a maximum clique in an arbitrary graph. *SIAM J. Comput.*, 15(4):1054–1068, 1986.

[5] B. Bollobás and P. Erdős. Cliques in random graphs. *Math. Proc. Cambridge Philos. Soc.*, 80(3):419–427, 1976.

[6] Béla Bollobás. The future of graph theory. In *Quo vadis, graph theory?*, volume 55 of *Ann. Discrete Math.*, pages 5–11. North-Holland, Amsterdam, 1993.

[7] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[8] Immanuel M. Bomze, Marco Budinich, Panos M. Pardalos, and Marcello Pelillo. The maximum clique problem. In *Handbook of combinatorial optimization, Supplement Vol. A*, pages 1–74. Kluwer Acad. Publ., Dordrecht, 1999.

[9] G. Brinkmann, S. Crevals, and J. Frye. An independent set approach for the communication network of the gps iii system. *Discrete Applied Mathematics*, 2011.

[10] Sergiy Butenko and Svyatoslav Trukhanov. Using critical sets to solve the maximum independent set problem. *Oper. Res. Lett.*, 35(4):519–524, 2007.

[11] Maria Chudnovsky. Hadwiger's conjecture and seagull packing. *Notices Amer. Math. Soc.*, 57(6):733–736, 2010.

[12] E. DeLaVina and C. E. Larson. A parallel algorithm for computing the critical independence number and related sets. to appear in Ars Mathematica Contemporanea.

[13] Keith Devlin. *The millennium problems*. Basic Books, New York, 2002. The seven greatest unsolved mathematical puzzles of our time.

[14] J.R. Dias. *Handbook of polycyclic hydrocarbons*, volume 1. Elsevier Science Ltd, 1987.

[15] P. Duchet and H. Meyniel. On Hadwiger's number and the stability number. In *Graph theory (Cambridge, 1981)*, volume 62 of *North-Holland Math. Stud.*, pages 71–73. North-Holland, Amsterdam, 1982.

[16] Jack Edmonds. Paths, trees, and flowers. *Canad. J. Math.*, 17:449–467, 1965.

[17] S. Fajtlowicz and C. E. Larson. Graph-theoretic independence as a predictor of fullerene stability. *Chemical physics letters*, 377(5-6):485–490, 2003.
[18] Michael R. Garey and David S. Johnson. *Computers and intractability*. W. H. Freeman and Co., San Francisco, Calif., 1979. A guide to the theory of NP-completeness, A Series of Books in the Mathematical Sciences.
[19] I. Gutman and S. J. Cyvin. *Introduction to the theory of benzenoid hydrocarbons*. Springer-Verlag Berlin, 1989.
[20] M. M. Halldórsson and J. Radhakrishnan. Greed is good: approximating independent sets in sparse and bounded-degree graphs. *Algorithmica*, 18(1):145–163, 1997.
[21] Magnús M. Halldórsson. Approximations of weighted independent set and hereditary subset problems. *J. Graph Algorithms Appl.*, 4:no. 1, 16 pp. (electronic), 2000.
[22] Magnús M. Halldórsson and Kiyohito Yoshihara. Greedy approximations of independent sets in low degree graphs. In *Algorithms and computations (Cairns, 1995)*, volume 1004 of *Lecture Notes in Comput. Sci.*, pages 152–161. Springer, Berlin, 1995.
[23] B. L. Hartnell. Well-covered graphs. *J. Combin. Math. Combin. Comput.*, 29:107–115, 1999.
[24] Akihisa Kako, Takao Ono, Tomio Hirata, and Magnús M. Halldórsson. Approximation algorithms for the weighted independent set problem. In *Graph-theoretic concepts in computer science*, volume 3787 of *Lecture Notes in Comput. Sci.*, pages 341–350. Springer, Berlin, 2005.
[25] Akihisa Kako, Takao Ono, Tomio Hirata, and Magnús M. Halldórsson. Approximation algorithms for the weighted independent set problem in sparse graphs. *Discrete Appl. Math.*, 157(4):617–626, 2009.
[26] C. E. Larson. A note on critical independence reductions. *Bull. Inst. Combin. Appl.*, 51:34–46, 2007.
[27] C. E. Larson. The critical independence number and an independence decomposition. *European J. Combin.*, 32(2):294–300, 2011.
[28] C. E. Larson. The fractional independence number and König-Egerváry graphs. *SIAM Discrete Mathematics conference talk, Halifax, Canada*, June 2012.
[29] C. E. Larson and R. Pepper. Graphs with equal independence and annihilation numbers. *Electronic Journal of Combinatorics*, 18(1), 2011.
[30] Vadim E. Levit and Eugen Mandrescu. Critical sets in bipartite graphs. 2011.
[31] Vadim E. Levit and Eugen Mandrescu. On the structure of the minimum critical independent set of a graph. 2011.
[32] Vadim E. Levit and Eugen Mandrescu. Vertices belonging to all critical independent sets of a graph. 2011.
[33] V.E. Levit and E. Mandrescu. Critical independent sets and könig–egerváry graphs. *Graphs and Combinatorics*, pages 1–8, 2011.
[34] L. Lovász. On the shannon capacity of a graph. *Information Theory, IEEE Transactions on*, 25(1):1–7, 1979.
[35] L. Lovász. Bounding the independence number of a graph. In *Bonn Workshop on Combinatorial Optimization (Bonn, 1980)*, volume 16 of *Ann. Discrete Math.*, pages 213–223. North-Holland, Amsterdam, 1982.
[36] L. Lovász and M. D. Plummer. *Matching theory*, volume 121 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29.
[37] F. Maffray and H. Meyniel. On a relationship between Hadwiger and stability numbers. *Discrete Math.*, 64(1):39–42, 1987.

[38] G. L. Nemhauser and L. E. Trotter, Jr. Vertex packings: structural properties and algorithms. *Math. Programming*, 8:232–248, 1975.

[39] M.E.J. Newman, D.J. Watts, and S.H. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 1):2566, 2002.

[40] S. Niskanen and P.R.J. Östergård. *Cliquer User's Guide: Version 1.0*. Helsinki University of Technology, 2003.

[41] Patric R. J. Östergård. A fast algorithm for the maximum clique problem. *Discrete Appl. Math.*, 120(1-3):197–207, 2002. Sixth Twente Workshop on Graphs and Combinatorial Optimization (Enschede, 1999).

[42] Patric R. J. Östergård. Constructing combinatorial objects via cliques. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 57–82. Cambridge Univ. Press, Cambridge, 2005.

[43] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. 1999.

[44] Panos M. Pardalos, Jonas Rappe, and Mauricio G. C. Resende. An exact parallel algorithm for the maximum clique problem. In *High performance algorithms and software in nonlinear optimization (Ischia, 1997)*, volume 24 of *Appl. Optim.*, pages 279–300. Kluwer Acad. Publ., Dordrecht, 1998.

[45] Panos M. Pardalos and Jue Xue. The maximum clique problem. *J. Global Optim.*, 4(3):301–328, 1994.

[46] Ryan Pepper. An upper bound on the independence number of benzenoid systems. *Discrete Appl. Math.*, 156(5):607–619, 2008.

[47] Jean-Claude Picard and Maurice Queyranne. On the integer-valued variables in the linear vertex packing problem. *Math. Programming*, 12(1):97–101, 1977.

[48] M. D. Plummer. Well-covered graphs: a survey. *Quaestiones Math.*, 16(3):253–287, 1993.

[49] Michael D. Plummer. Some covering concepts in graphs. *J. Combinatorial Theory*, 8:91–98, 1970.

[50] J. M. Robson. Algorithms for maximum independent sets. *J. Algorithms*, 7(3):425–440, 1986.

[51] Robert Endre Tarjan and Anthony E. Trojanowski. Finding a maximum independent set. *SIAM J. Comput.*, 6(3):537–546, 1977.

[52] Cun Quan Zhang. Finding critical independent sets and critical vertex subsets are polynomial problems. *SIAM J. Discrete Math.*, 3(3):431–438, 1990.