Approximations of independent sets in graphs

Magnús M. Halldórsson*
University of Iceland
Reykjavik, Iceland

1 Introduction

The *independent set problem* is that of finding a maximum size set of mutually non-adjacent vertices in a graph. The study of independent sets, and their *alter egos*, cliques, has had a central place in combinatorial theory.

Independent sets occur whenever we seek sets of items free of pairwise conflicts, e.g. when scheduling tasks. Aside from numerous applications (which might be more pronounced if the problems weren't so intractable), independent sets and cliques appear frequently in the theory of computing, e.g. in interactive proof systems [6] or monotone circuit complexity [2]. They form the representative problems for the class of subgraph or packing problems in graphs, are essential companions of graph colorings, and form the basis of clustering, whether in terms of nearness or dispersion.

As late as 1990, the literature on independent set approximations was extremely sparse. In the period since Johnson [31] started the study of algorithms with good *performance ratios* in 1974 – and in particular showed that a whole slew of independent set algorithms had only the trivial performance ratio of n on general graphs – only one paper had appeared containing positive results [29], aside from the special case of planar graphs [34, 8]. Lower bounds were effectively non-existent, as while it was known that the best possible performance ratio would not be some *fixed* constant, there might still be a polynomial-time approximation scheme lurking somewhere.

Success on proving lower bounds for Independent Set has been dramatic and received worldwide attention, including the New York Times. Progress on improved approximation algorithms has been less dramatic, but a notable body of results has been developed. The purpose of this talk is to bring some of these results together, consider the lessons learned, and hypothesize about possible future developments.

The current paper is not meant to be the ultimate summary of independent set approximation algorithms, but an introduction to the performance ratios known and the strategies that have been applied, and offer glimpses of some of the results that have been proven.

We prefer to study a range of algorithms, rather than seek only the best possible performance guarantee. The latter is fine as far as it goes, but is not the only thing that matters; only so much information is represented by a single number. Algorithmic strategies vary in their time requirements, temporal access to data, parallelizability, simplicity and numerous other factors that are far from irrelevant. Different algorithms may also be incomparable on different classes of graphs, e.g. depending on the size of the optimal solution. Finally, the proof techniques are perhaps the most valuable product of the analysis of heuristics.

We look at a slightly random selection of approximation results in the body of the paper. A complete survey is beyond the scope of this paper but is under preparation. The primary criteria

*Adjunct affiliation: University of Bergen

for selection was simplicity, of the algorithm and the proof. We state some observations that have not formally appeared before, give some recent results, and present simpler proofs of other results.

The paper is organized as follows. We define relevant problems and definitions in the following section. In the body of the paper we present a number of results illustrating particular algorithmic strategies: subgraph removal, semi-definite programming, partitioning, greedy algorithms and local search. We give a listing of known performance results and finish with a discussion of open issues.

2 Problems and definitions

INDEPENDENT SET: Given a graph $G = (V, E)$, find a maximum cardinality set $I \subseteq V$ such that for each $u, v \in I$, $(u, v) \notin E$. The *independence number* of G, denoted by $\alpha(G)$, is the size of the maximum independent set.

CLIQUE PARTITION: Given a graph $G = (V, E)$, find a minimum cardinality set of disjoint cliques from G that contains every vertex.

κ-SET PACKING: Given a collection $\mathcal{C}$ of sets of size at most κ drawn from a finite set S, find a maximum cardinality collection $\mathcal{C}'$ such that each element in S is contained in at most one set in $\mathcal{C}'$.

These problems may also be weighted, with weights on the vertices (or on the sets in SET PACKING).

A set packing instance is a case of an independent set problem. Given a set system $(\mathcal{C}, S)$, form a graph with a vertex for each set in $\mathcal{C}$ an edge between two vertices if the corresponding sets intersect. Observe that if the sets in $\mathcal{C}$ are of size at most κ, then the graph contains no induced $\kappa+1$-*claw*, which is a subgraph consisting of a center node adjacent to $\kappa+1$ mutually non-adjacent vertices. The independent set problem in $\kappa+1$-claw free graphs slightly generalizes κ-SET PACKING, which in turn slightly generalizes κ-DIMENSIONAL MATCHING.

For an independent set algorithm A, $A(G)$ denotes the size of the solution it obtains on G, and its *performance ratio* ρ_A is given by

$$\rho_A = \rho_A(n) = \max_{G, |G|=n} \frac{\alpha(G)}{A(G)}.$$

Notation

n	the number of vertices	$d(v)$	the degree of vertex v
m	the number of edges	$N(v)$	set of neighbors of v
Δ	maximum degree	$\overline{N}(v)$	non-neighbors of v
$\overline{d}$	average degree	$A(G)$	the size of solution found by A
δ	minimum degree	ρ_A	performance ratio of A
α	independence number	$\overline{\chi}$	clique partition number
κ	maximum claw size		

3 Ramsey theory and subgraph removal

The first published algorithm with a non-trivial performance ratio on general graphs was introduced in 1990. In appreciation of the heritage that the late master Erdős left us, we give here a treatment

Ramsey (G)	**CliqueRemoval** (G)
if $G = \emptyset$ then return $(\emptyset, \emptyset)$	$i \leftarrow 1$
choose some $v \in G$	$(C_i, I_i) \leftarrow$ Ramsey (G)
$(C_1, I_1) \leftarrow$ Ramsey$(N(v))$	while $G \neq \emptyset$ do
$(C_2, I_2) \leftarrow$ Ramsey$(\overline{N}(v))$	$G \leftarrow G - C_i$
return (larger of $(C_1 \cup \{v\}, C_2)$,	$i \leftarrow i + 1$
larger of $(I_1, I_2 \cup \{v\})$)	$(C_i, I_i) \leftarrow$ Ramsey (G)
	od
	return $((\max_{j=1}^{i} I_j), \{C_1, C_2, \ldots, C_i\})$

Figure 1: Independent set algorithm based on Ramsey theory

different from Boppana and Halldórsson [12] that more closely resembles the original Ramsey theorem of Erdős and Szekeres [17].

Theorem 3.1 Ramsey *finds an independent set* I *and a clique* C *such that* $\binom{|I|+|C|}{|C|} - 1 \geq n$. *In particular,* $|I| \cdot |C| \geq \frac{1}{4}\log^2 n$.

Proof. The proof is by induction on both $|I|$ and $|C|$. It is easy to verify the claim when either $|I|$ or $|C|$ are at most 1. By the induction hypothesis,

$$n = |N(v)| + |\overline{N(v)}| + 1 \leq (\binom{|I_1| + |C_1|}{|C_1|} - 1) + (\binom{|I_2| + |C_2| - 1}{|C_2|} - 1) + 1.$$

Recall that $|C| = \max(|C_1| + 1, |C_2|)$ and $|I| = \max(|I_1|, |I_2| + 1)$. Thus,

$$n \leq \binom{|I| + |C| - 1}{|C| - 1} + \binom{|I| + |C| - 1}{|C|} - 1.$$

The claim now follows from the equality $\binom{s+t}{s} = \binom{s+t-1}{t-1} + \binom{s+t-1}{t}$.

It is easy to verify that the product of $|I|$ and $|C|$ is minimized when they are equal. That is, when $n = \binom{2|C|}{|C|} \leq 2^{2|C|}$, hence $|I| \cdot |C| \geq (\frac{1}{2}\log n)^2$. ∎

The following simplified proof of a $O(n/\log^2 n)$ performance ratio also borrows from another of Erdős's work [15].

Theorem 3.2 *The performance ratio of* CliqueRemoval *is* $O(n/\log^2 n)$.

Proof. Let CC denote the number of cliques returned by CliqueRemoval and let CC_0 denote the number of cliques removed before the size of the graph dropped below $n_0 = n/\log^2 n$. Let t be the size of the smallest of these latter cliques, which without loss of generality is at most $\log^2 n$. Then $CC_1 \leq n/t$, and $CC \leq n/t + n_0 \leq 2n/t$.

If I is the independent set returned, we have that $|I| \geq 4\log^2 n_0/t \geq 2\log^2 n/t$. Consider the product of the two performance ratio of CliqueRemoval, ρ_α for independent sets, and $\rho_{\overline{\chi}}$ for clique partition:

$$\rho_\alpha \cdot \rho_{\overline{\chi}} = \frac{CC}{\overline{\chi}} \cdot \frac{\alpha}{|I|} \leq \frac{n}{\log^2 n}\frac{\alpha}{\overline{\chi}} \leq \frac{n}{\log^2 n}.$$

Clearly, either performance ratio is also bounded by $O(n/\log^2 n)$. ∎

For graphs with high independence number, the ratios are better.

Theorem 3.3 *If $\alpha(G) \geq n/k + p$, then* CliqueRemoval *finds an independent set of size $\Omega(p^{1/(k-1)})$ in G.*

This illustrates the strategy of *subgraph removal*, that is based around the concept that graphs without small dense subgraphs are easier to approximate.

4 Lovász theta function

A fascinating polynomial-time computable function $\vartheta(G)$, that was introduced by Lovász [37], has the remarkable *sandwiching* property that it always lies between two NP-hard functions, $\alpha(G) \leq \vartheta(G) \leq \overline{\chi}(G)$. This property suggests that it may be particularly suited for obtaining good approximations to either function. While some of those hopes have been dashed, a number of fruitful applications have been found and it remains the most promising candidate for obtaining improved approximations.

Karger, Motwani and Sudan [32] proved the following property in the context of coloring. The "soft-omega" notation $\tilde{\Omega}$ hides logarithmic factors.

Theorem 4.1 (Karger et al) *If $\vartheta(\overline{G}) \leq k$, then an independent set of size $\tilde{\Omega}(n^{3/(k+1)})$ in G can be constructed with high probability in polynomial time.*

Mahajan and Ramesh [38] showed how these and related algorithms can be derandomized. Alon and Kahale [4] applied the theta function further for independent sets.

Theorem 4.2 (Alon, Kahale) *If $\vartheta(G) \geq n/k + p$ (e.g. if $\alpha(G) \geq n/k + p$), then we can find an induced subgraph K in G on p vertices with $\vartheta(\overline{K}) \leq k$.*

Combining the two, they obtained a ratio for high-independence graphs that improves on Theorem 3.3.

Corollary 4.3 *For any fixed integer $k \geq 3$, if $\vartheta(G) \geq n/k + p$, then an independent set of size $\tilde{\Omega}(p^{3/(k+1)})$ in G can be found in polynomial time.*

Theta function on sparse graphs Karger et al. proved a core result in terms of maximum degree of the graph. In fact, their argument also holds in terms of average degree.

Theorem 4.4 (Karger et al) *If $\vartheta(\overline{G}) \leq k$, then an independent set of size $\tilde{\Omega}(n/\overline{d}^{1-2/k})$ in G can be constructed with high probability in polynomial time.*

Vishwanathan [40] observed that this, combined with Theorem 4.2, also yields an improved algorithm for bounded-degree graphs. This, however, has not been stated before in the literature, to the best of the author's knowledge.

Proposition 4.5 Independent Set *can be approximated within a factor of $O(\Delta \log\log\Delta/\log\Delta)$*[1].

Proof. Given G, if $\alpha(G) \geq n/k = n/2k + n/2k$, then we can find a subgraph K on $n/2k$ vertices with $\vartheta(\overline{K}) \leq k$ and maximum degree at most $\Delta(G)$, by Theorem 4.2. By Theorem 4.4, we can find an independent set in K (and G) of size $\Omega((n/2k)/\Delta^{1-2/k}) = \Omega(n/\Delta \cdot \Delta^{2/k}/k)$. If $k \leq \log\Delta/\log\log\Delta$, then the set found is of size $\Omega(n/\Delta \cdot \log\Delta)$, and the claim is satisfied since $\alpha \leq n$. Otherwise, $\alpha \leq n \log\log\Delta/\log\Delta$, and any maximal solution is of size $n/(\Delta+1)$, for a ratio satisfying the proposition. ∎

[1] We may need to assume that Δ be a large constant independent of n.

5 Partitioning and weighted independent sets

A simple strategy in the design of approximation algorithms is to break the problem into a collection of easier subproblems.

Observation 5.1 *Suppose we can partition G into t subgraphs and solve the weighted independent set problem on each subgraph optimally. Then, the largest of these solutions is a t-approximation of G.*

Proof. The size of the optimal solution for G is at most the sum of the sizes of the largest independent sets on each subgraph, which is at most t times the largest solution in some subgraph. ■

This gives us the first non-trivial ratio for weighted independent sets in general graphs [21].

Theorem 5.2 *The weighted independent set problem can be approximated within $O(n(\log\log n/\log n)^2)$.*

Proof. The bound on Ramsey in Theorem 3.1 implies that it either outputs an independent set of size $\log^2 n$, or a clique of size $\log n/\log\log n$. We apply this algorithm repeatedly, like CliqueRemoval, but either removes a $\log^2 n$-independent set or a $\log n/\log\log n$-clique in each step.

We now form a partition where each class is either an independent set, or a (not necessarily disjoint) union of $\log n/\log\log n$ different cliques. This yields a partition into $O(n(\log\log n/\log n)^2)$ classes.

The weighted independent set problem on such classes can be solved by exhaustively checking all $(\log n/\log\log n)^{\log n/\log\log n} = O(n)$ possible combinations of selecting one vertex from each clique. Thus, by the above observation, the claimed ratio follows. ■

On bounded-degree graphs, we can apply a partitioning lemma of Lovász [35], which we specialize here to this application.

Lemma 5.3 *The vertices of a graph can be partitioned into $\lceil(\Delta+1)/3\rceil$ sets, where each induces a subgraph of maximum degree at most two.*

Proof. Start with an arbitrary partitioning into $\lceil(\Delta+1)/3\rceil$ sets, and repeat the following operation: If v is adjacent to three or more vertices in its set, move it to a set where it has at most two neighbors. Such a set must exist as otherwise v's degree would be at least $3\lceil(\Delta+1)/3\rceil \geq \Delta+1$. Observe that such a move increases the number of *cross edges*, or edges going between different sets, hence this process must terminate with a partition where every vertex has at most two neighbors in its set. ■

Dynamic programming easily solves the weighted maximum independent set problem on each such subgraph, and as shown in [23], the partitioning can also be performed in linear time by starting with a greedy partition.

Theorem 5.4 WEIGHTED INDEPENDENT SET *is approximable within $\lceil(\Delta+1)/3\rceil$ in linear time.*

Hochbaum [29] also used a form of a partition, a coloring, to approximate weighted independent set problems.

6 Greediness and Set packing

The general set packing problem can be shown to be equivalent to the independent set problem. Given a graph $G = (V, E)$, let the base set S contain one element for each edge in E, and for each vertex $v \in V$, form a set C_v containing the base sets corresponding to edges incident on v. It holds that the maximum number of sets in a set packing of $(S, \mathcal{C})$ is $\alpha(G)$.

There are four parameters of set systems that are of interest for Set Packing approximations: n, the number of sets, $|S|$, the number of base elements, κ, maximum cardinality of a set, and B, the maximum number of occurrences of a base element in $\mathcal{C}$.

In the reduction above, we find that $|\mathcal{C}| = n$, and therefore approximations of Independent Set as functions of n carry over to approximations of Set Packing in terms of $|\mathcal{C}|$. A reduction in the other direction also preserves this relationship.

As for B, observe that in the reduction above, $B = 2$, for arbitrary instances. Hence, we cannot expect any approximations as functions of B alone. It remains to consider approximability in terms of κ and $|S|$.

Local search

Just about any solution gives a modest approximation.

Theorem 6.1 *Any maximal solution is κ-approximate.*

Proof. We say that a vertex v *dominates* a vertex u if u and v are either adjacent or the same vertex. Any vertex can dominate at most κ vertices from an optimal solution. Yet, maximality requires that a maximal independent set dominates all vertices of the graph. Hence, an optimal solution is at most κ times bigger than any maximal solution. ∎

This can be strengthened using simple local search. Tight analysis was first given by Hurkens and Schrijver [30], whose article title seemed to obscure its contents since the results were reproduced in part or full by several groups of authors [33, 42, 20, 43, 7].

Local search is straightforward for problems whose solutions are collections of items: repeatedly try to extend the solution by eliminating t elements while adding $t + 1$ elements. A solution that cannot be further extended by such improvements is said to be *t-optimal.* It turns out that 2-optimal solutions, which are the most efficient and the easiest to analyze, already give considerably improved approximations.

Theorem 6.2 *Any 2-optimal solution is $(\kappa + 1)/2$-approximate.*

Proof. Let us argue in terms of independent sets in $\kappa + 1$-claw free graphs. Let I be a 2-optimal solution, and let O be any optimal independent set. Partition O into O_1, those vertices in O that are adjacent to only one vertex in I, and O_2, those vertices in O that are adjacent to two or more vertices in I. Note that each vertex in I is adjacent to at most κ vertices in O, due to the lack of a $\kappa + 1$-claw. Then, considering the edges between I and O we have that

$$|O_1| + 2|O_2| \leq \kappa|I|.$$

Also, since I is 2-optimal

$$|O_1| \leq |I|.$$

Adding the two inequalities gives that

$$2|O| = 2(|O_1| + |O_2|) \leq (\kappa + 1)|I|,$$

or that the performance ratio is at most $(\kappa+1)/2$. ∎

Using t-opt we can prove a bound of $\kappa/2+\epsilon$. On bounded-degree graphs, local search can be applied with a very large radius while remaining in polynomial time, and using some additional techniques, Berman, Fürer, and Fujito [11, 10] obtained the best performance ratios known for small values of Δ, or $(\Delta+3)/5$.

Greedy algorithms

This leaves $|S|$ as the only parameter left to be studied for SET PACKING.

A related topic is the *Strong Stable Set* problem, where we seek an independent set in which the vertices are of distance at least two apart. Such a strong stable set corresponds to a set packing of the set system formed by the closed vertex neighborhoods in the graph. In this case, $|\mathcal{C}| = |S| = n$. The question is then whether this is easier to approximate than the general independent set problem.

Halldórsson, Kratochvíl, and Telle [22] recently gave a simple answer to this question, using a greedy set packing algorithm that always picks the smallest set remaining.

Theorem 6.3 *Set Packing can be approximated within $\sqrt{|S|}$ in time linear in the input size.*

Proof. Consider the following greedy algorithm. In each step, it chooses a smallest set and removes from the collection all sets containing elements from the selected set.

```
GreedySP(S,C)
  t ← 0
  repeat
    t ← t + 1
    X_t ← C ∈ C of minimum cardinality
    Z_t ← {C ∈ C : X ∩ C ≠ ∅ }
    C ← C − Z_t
  until |C| = 0
  Output {X_1, X_2, ..., X_t}
```

Let $M = \lfloor\sqrt{|S|}\rfloor$. Observe that $\{Z_1, \ldots, Z_t\}$ forms a partition of $\mathcal{C}$. Let i be the index of some iteration of the algorithm, i.e. $1 \le i \le t$. All sets in Z_i contain at least one element of X_i, thus the maximum number of disjoint sets in Z_i is at most the cardinality of X_i. On the other hand, every set in Z_i is of size at least X_i, so the maximum number of disjoint sets in Z_i is also at most $\lfloor |S|/|X_i| \rfloor$. Thus, the optimal solution contains at most $\min(|X_i|, \lfloor |S|/|X_i| \rfloor) \le \max_x \min(x, \lfloor |S|/x \rfloor) = M$ sets from Z_i.

Thus, in total, the optimal solution contains at most tM sets, when the algorithm finds t sets, for a ratio of at most M. ∎

Observe that this approximation is near the best possible. Since a graph contains $O(n^2)$ edges, Håstad's result [27] yields an $\Omega(m^{1/2-\epsilon})$ lower bound, for any $\epsilon > 0$.

Other greedy algorithms have been studied, especially the one that repeatedly selects vertices of minimum degree in the graph. It remains, e.g., the driving force for the best ratio known for sparse graphs, $(2\overline{d}+3)/5$ [26].

7 Summary of results

Table 1 contains a listing of the various ratios that have been proved for heuristics for the independent set problem, along with known inapproximability results. It is divided according to graph classes / graph parameter. Results in terms of other measures of graphs or pairs of measures are not included. The results hold for unweighted graphs except for the last category. Each entry contains the ratio proved, the algorithmic strategy used, the complexity of the method, and a citation.

We have not described the Nemhauser-Trotter reduction [39] that was championed by Hochbaum [29], which allows one to assume in many cases without loss of generality that the maximum weight independent set is of weight at most half the total weight of the graph. The complexity of this procedure equals the complexity of finding a minimum cut in a network in the weighted case ($O(nm)$), and the complexity of bipartite matching in the unweighted case ($O(\sqrt{n}m)$).

Abbreviations: SR = subgraph removal, SDP = semi-definite programming, NT = Nemhauser-Trotter reduction, MIS = arbitrary maximal independent set.

Complexity: NT refers to the complexity of the Nemhauser-Trotter reduction. "Linear" means time linear in the size of the graph. $n^{O(1)}$ suggests time bounded by a polynomial of high degree; in the case of the $(\Delta + 3)/5$ ratio, the degree of the polynomial appears to be on the order of 2^{100} [24].

8 Discussion

A number of open issues remain.

General graphs There remains some gap between the best upper and lower bounds known for general graphs. Stated in terms of "distance from trivial", it is the difference between $\log^2 n$ and $n^{o(1)}$. It is not as presumptuous now to conjecture that the ultimate ratio is $n/polylog(n)$ as it was in 1991 [19]. It may be possible to extend the proof of [27] to argue a stronger lower bound than $n^{1-\epsilon}$ if given a stronger assumption, such as SAT not having $2^{o(n)}$ time algorithms. (Admittedly, such a task appears less than trivial [28]).

Performance of ϑ-function The theta function remains the most promising candidate for improved approximations. Some of the hopes attached with it have been dashed. Feige [18] showed that its performance ratio is at least $n/2^{O(\sqrt{\log n})}$. Can't we at least prove something better than the simple Ramsey-theoretic bound?

High-independence graphs Gaps in bounds on approximability are nowhere greater than in the case of independent sets in graphs with $\alpha(G) = n/k$, for some fixed $k > 2$. These problems are APX-hard, i.e. hard within some constant factor greater than one, but all the upper bounds known are some roots of n. These problems generalize the case of k-colorable graphs, for which a similar situation holds. Results of Alon and Kahale [4] indicate that some root of n is also the best that the theta function will yield in this case. The limited progress on the more studied k-coloring problem suggests that this is near best possible.

Vertex cover The preceding item has relevance to the approximability of VERTEX COVER, which is the problem of finding a minimum set of vertices S such that $V - S$ is an independent set. If VERTEX COVER can be approximated within less than 1.5, then INDEPENDENT SET in graphs with $\alpha = n/3$ is constant approximable and GRAPH 3-COLORING is $O(\log n)$ approximable, as first shown by Bar-Yehuda and Moran [9]. This gives support to the conjecture that factor 2 is optimal for VERTEX COVER, within lower order terms [29].

Bounded-degree graphs It is natural to extrapolate that the improved hardness ratio $n^{1-\epsilon}$ of [27] indicates that the hardness ratio $\Omega(\Delta^c)$ of [3] for bounded-degree graphs could be jacked up to $\Omega(\Delta^{1-o(1)})$.

Result	Method	Complexity	Reference
General graphs			
$O(n/\log n)$		$O(nm)$	[41]
$O(n/\log^2 n)$	SR	$O(nm)$	[12]
$\Omega(n^{1-\epsilon})$			[27]
High-independence graphs ($\alpha = n/k$)			
$O(n^{1-1/(k-1)})$	SR	$O(nm)$	[12]
$\tilde{O}(n^{1-3/(k+1)})$	SDP	SDP	[4]
$\Omega(1+c)$			[6]
Sparse graphs			
$\overline{d}+1$	Greedy	linear	[29], via [16]
$(\overline{d}+2)/2$	Greedy	linear	[26]
$(\overline{d}+1)/2$	Greedy + NT	NT	[29]
$(2\overline{d}+4.5)/5$	Greedy+SR	linear	[25]
$(2\overline{d}+3)/5$	Greedy + NT	NT	[26]
Bounded-degree graphs			
Δ	MIS	linear	
$\Delta/2$	Brooks+NT	NT	[29], via [36]
$(\Delta+2)/3$	Greedy	linear	[26]
$(\Delta+3)/5$	Local search +	$n^{O(1)}$	[11, 10]
$(\Delta+2)/4+\epsilon$	Local search	$\Delta^{O(\Delta)}n$	[24]
$\Delta/6+O(1)$	SR	$O(\Delta^{\Delta}n+n^2)$	[25]
$O(\Delta/\log\log\Delta)$	SR	$n^{O(1)}$	[25], via [1]
$O(\Delta\log\log\Delta/\log\Delta)$	SDP	SDP	[40], via [32, 4]
$\Omega(\Delta^c)$			[3]
$\kappa+1$-claw-free graphs and Set Packing			
κ	MIS	linear	
$(\kappa+1)/2$	Local search	$O(n^3)$	[33, 42]
$\kappa/2+\epsilon$	Local search	$O(n^{\log_\kappa 1/\epsilon})$	[30, 20, 43]
$\sqrt{\lvert S\rvert}$	GreedySP	linear	[22]
$\Omega(\kappa^c)$			[3]
$\Omega(\lvert S\rvert^{1-\epsilon})$			[27]
Weighted graphs			
$\Delta/2$	Brooks+NT	NT	[29]
$\lceil(\Delta+1)/3\rceil$	Partitioning	linear	[23], via [35]
$(\Delta+2)/3$	Partitioning+NT	NT	[23]
κ	Max-weight greedy	$O(n^2)$	[29]
$\kappa-1+\epsilon$	Local search	$n^{O(1/\epsilon)}$	[5, 7]
$(4\kappa+2)/5$	LS + greedy	$n^{O(\kappa)}$	[13]
$O(n(\log\log n/\log n)^2)$	SR+Partitioning	$O(n^2)$	[21]

Table 1: Results on approximating independent sets

From the upper bound side, it would be nice to extend the $o(\Delta)$ ratios of [25, 40] to hold for all values of Δ as a function of n. Demange and Paschos [14] have parametrized the strategy of [25] to give a ratio Δ/c for every c, that holds for every value of Δ in time $O(n^c)$.

$\kappa + 1$-claw-free graphs Claw-free graphs appear considerably harder than bounded-degree graphs. Any improvement to the $\kappa/2+\epsilon$ ratios would be most interesting. Observe that a maximum κ-set packing is within a factor κ from a *minimum hitting set* of a collection of sets of size κ, but we also do not have any better ratio than factor κ for the latter problem.

In the weighted case, we know that the greedy and the local search strategies do not improve on $\kappa - O(1)$. However, the combination of the two does attain asymptotically better ratios [13]. We conjecture that selective local search starting from a greedy solution does attain the unweighted bound of $\kappa/2+\epsilon$.

$o(\alpha)$-approximations While we do have $o(n)$-approximations of INDEPENDENT SET, these methods fail to give us anything beyond the trivial when, say, $\alpha = \sqrt{n}$. While it is probably too much to ask for a $\omega(1)$-size independent set in graphs with $\alpha \approx \log n$, it is not unfair to ask for, say, a $\alpha \log n / \log^2 \alpha$-approximation.

References

[1] M. Ajtai, P. Erdős, J. Komlós, and E. Szemerédi. On Turán's theorem for sparse graphs. *Combinatorica*, 1(4):313–317, 1981.

[2] N. Alon and R. B. Boppana. The monotone complexity of Boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[3] N. Alon, U. Feige, A. Wigderson, and D. Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60 – 75, 1995.

[4] N. Alon and N. Kahale. Approximating the independence number via the θ function. *Math. Programming*. To appear.

[5] E. M. Arkin and R. Hassin. On local search for weighted k-set packing. *ESA '97*, LNCS 1284.

[6] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. *FOCS '92*, 14–23.

[7] V. Bafna, B. O. Narayanan, and R. Ravi. Non-overlapping local alignments (weighted independent sets of axis parallel rectangles). *WADS '95*, LNCS 955, 506–517.

[8] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. ACM*, 41:153–180, Jan. 1994.

[9] R. Bar-Yehuda and S. Moran. On approximation problems related to the independent set and vertex cover problems. *Discrete Appl. Math.*, 9:1–10, 1984.

[10] P. Berman and T. Fujito. On the approximation properties of independent set problem in degree 3 graphs. *WADS '95*, LNCS 955, 449–460.

[11] P. Berman and M. Fürer. Approximating maximum independent set in bounded degree graphs. *SODA '94*, 365–371.

[12] R. B. Boppana and M. M. Halldórsson. Approximating maximum independent sets by excluding subgraphs. *BIT*, 32(2):180–196, June 1992.

[13] B. Chandra and M. M. Halldórsson. Approximating weighted k-set-packing. Manuscript, May 1998.

[14] M. Demange and V. T. Paschos. Improved approximations for maximum independent set via approximation chains. *Appl. Math. Lett*, 1996. To appear.

[15] P. Erdős. Some remarks on chromatic graphs. *Colloq. Math.*, 16:253–256, 1967.

[16] P. Erdős. On the graph theorem of Turán (in Hungarian). *Mat. Lapok*, 21:249–251, 1970.

[17] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[18] U. Feige. Randomized graph products, chromatic numbers, and the Lovász ϑ-function. *Combinatorica*, 17(1):79–90, 1997.

[19] M. M. Halldórsson. A still better performance guarantee for approximate graph coloring. *Inform. Process. Lett.*, 45:19–23, 25 January 1993.

[20] M. M. Halldórsson. Approximating discrete collections via local improvements. *SODA '95*, 160–169.

[21] M. M. Halldórsson. Approximation via partitioning. Res. Report IS-RR-95-0003F, Japan Adv. Inst. of Sci. and Tech., Mar. 1995.

[22] M. M. Halldórsson, J. Kratochvíl, and J. A. Telle. Independent sets with domination constraints. *ICALP '98*, LNCS.

[23] M. M. Halldórsson and H. C. Lau. Low-degree graph partitioning via local search with applications to constraint satisfaction, max cut, and 3-coloring. *J. Graph Algo. Applic.*, 1(3):1–13, 1997.

[24] M. M. Halldórsson and J. Radhakrishnan. Improved approximations of independent sets in bounded-degree graphs. *SWAT '94*, LNCS 824, 195–206.

[25] M. M. Halldórsson and J. Radhakrishnan. Improved approximations of independent sets in bounded-degree via subgraph removal. *Nordic J. Computing*, 1(4):475–492, 1994.

[26] M. M. Halldórsson and J. Radhakrishnan. Greed is good: Approximating independent sets in sparse and bounded-degree graphs. *Algorithmica*, 18:145–163, 1997.

[27] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *FOCS '96*, 627–636.

[28] J. Håstad. Private communication, 1997.

[29] D. S. Hochbaum. Efficient bounds for the stable set, vertex cover, and set packing problems. *Disc. Applied Math.*, 6:243–254, 1983.

[30] C. A. J. Hurkens and A. Schrijver. On the size of systems of sets every t of which have an SDR, with an application to the worst-case ratio of heuristics for packing problems. *SIAM J. Disc. Math.*, 2(1):68–72, Feb. 1989.

[31] D. S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. Syst. Sci.*, 9:256–278, 1974.

[32] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semi-definite programming. *FOCS '94*, 2–13.

[33] S. Khanna, R. Motwani, M. Sudan, and U. Vazirani. On syntactic versus computational views of approximability. *FOCS '94*, 819–830.

[34] R. J. Lipton and R. E. Tarjan. Applications of a planar separator theorem. *FOCS '77*, 162–170.

[35] L. Lovász. On decomposition of graphs. *Stud. Sci. Math. Hung.*, 1:237–238, 1966.

[36] L. Lovász. Three short proofs in graph theory. *J. Combin. Theory Ser. B*, 19:269–271, 1975.

[37] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, IT-25(1):1–7, Jan. 1979.

[38] S. Mahajan and H. Ramesh. Derandomizing semidefinite programming based approximation algorithms. *FOCS '95*, 162–169.

[39] G. L. Nemhauser and L. Trotter. Vertex packings: Structural properties and algorithms. *Math. Programming*, 8:232–248, 1975.

[40] S. Vishwanathan. Personal communication, 1996.

[41] A. Wigderson. Improving the performance guarantee for approximate graph coloring. *J. ACM*, 30(4):729–735, 1983.

[42] G. Yu and O. Goldschmidt. On locally optimal independent sets and vertex covers. Manuscript, 1993.

[43] G. Yu and O. Goldschmidt. Local optimality and its application on independent sets for k-claw free graphs. Manuscript, 1994.