On the Independence Number of Sparse Graphs

James B. Shearer
Department of Mathematics, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598; email: jbs@watson.ibm.com

ABSTRACT

Let G be a regular graph of degree d on n points which contains no K_r $(r \geq 4)$. Let α be the independence number of G. Then we show for large d that $\alpha \geq c(r)n \frac{\ln d}{d \ln \ln d}$.

Let G be a regular graph of degree d on n points which contains no K_r $(r \geq 4)$. Let α [or $\alpha(G)$] be the independence number of G. We will show for large d that $\alpha \geq c(r)n \frac{\ln d}{d \ln \ln d}$. This improves the bound $\alpha \geq c'(r)n \frac{\ln \ln d}{d}$ proved in [1]; however, it does not settle the question (asked in [1]) of whether $\alpha \geq c'(r)n \frac{\ln d}{d}$ (which holds for triangle-free graphs).

We will actually establish a lower bound for $\bar{\alpha}$ the average size of an independent set in G. Clearly $\alpha \geq \bar{\alpha}$. The basic idea of the proof is to balance the probability that a vertex v of G is contained in a random independent set of G with the average number of neighbors of v which are contained in a random independent set of G. In what follows we will only concern ourselves with leading order terms for large d. We will need the following lemma.

Lemma 1. *Let S be a K_r-free graph $(r \geq 3)$. Let $I(S)$ be the number of independent sets in S. Let $\bar{\alpha}(S)$ be the average size of an independent set in S. Then, as $I(S) \rightarrow \infty$, $\bar{\alpha}(S) \geq c(r) \frac{\ln I(S)}{\ln \ln I(S)}$.*

Proof. Let m be the size of S. Let $k = \alpha(S)$. Let $\phi = \bar{\alpha}(S)/m$. Then we have the following inequalities:

Random Structures and Algorithms, Vol. 7, No. 3 (1995)
 CCC 1042-9832/95/030269-03

1. $I(S) \le 2^{mH(\phi)}$,
2. $\bar{\alpha}(S) = m\phi$,
3. $I(S) \ge 2^k$,

where H is the log 2 entropy function. 1 follows by considering the random probability distribution on the family of independent sets of S, which takes on each independent set with equal probability. The entropy of the distribution is $\log I(S)$, which must not exceed the sum of the entropies on each vertex which by the convexity of H is less than or equal to $mH(\phi)$. This is a standard method; see, for example, [2]. The inequality follows. 2 holds by definition and 3 by considering all the subsets of the maximum independent set.

Now by 2 and 1, $\bar{\alpha}(S) = m\phi = \frac{\phi}{H(\phi)} mH(\phi) \ge \frac{\phi}{H(\phi)} \frac{\ln I(S)}{\ln 2}$. $\frac{\phi}{H(\phi)}$ is an increasing function of ϕ so a lower bound on ϕ yields a lower bound on $\phi/H(\phi)$. By 1 and 3, $mH(\phi) \ge k$ or $H(\phi) \ge k/m$. Now it is easy to show by induction on r that $k \ge m^{1/(r-1)} - 1$. Since we are only considering leading order terms we have $H(\phi) \ge 1/m^{(\frac{r-2}{r-1})}$. Since $H(0) = 0$, this bounds ϕ away from 0. More precisely, to leading order $\phi \ge \frac{\ln 2}{(\frac{r-2}{r-1})(\ln m)m^{(\frac{r-2}{r-1})}}$ and $\phi/H(\phi) \ge \frac{\ln 2}{(\frac{r-2}{r-1}) \ln m}$. From 3, $I(S) \ge 2^k \ge 2^{m^{(\frac{1}{r-1})}-1}$. Therefore $\ln I(S) \ge (m^{\frac{1}{r-1}} - 1) \ln 2$ or to leading order $\ln I(S) \ge m^{\frac{1}{r-1}} \ln 2$ or $\ln \ln I(S) \ge \frac{1}{r-1} \ln m$ or $\ln m \le (r-1) \ln \ln I(S)$. Thus $\phi/H(\phi) \ge \frac{\ln 2}{(\frac{r-2}{r-1}) \ln m} \ge \frac{\ln 2}{(r-2) \ln \ln I(S)}$. Hence to leading order, $\bar{\alpha}(S) \ge \frac{\phi}{H(\phi)} \frac{\ln I(S)}{\ln 2} \ge \frac{1}{(r-2)} \frac{\ln I(S)}{\ln \ln I(S)}$.

We are now ready to prove our main theorem:

Theorem 1. *Let G be a regular graph of degree d on n points which contains no $K_r (r \ge 4)$. Let $\bar{\alpha}$ be the average size of an independent set of G. Then for large d, $\bar{\alpha} \ge c(r)n \frac{\ln d}{d \ln \ln d}$.*

Proof. Let F be the family of independent sets of G. For any vertex x of G, let p_x be the probability that x is contained in a random element of F. Let $d\bar{p}_x$ be the average number of neighbors of x that are contained in a random member of F. Let T be the set of neighbors of x. Let H be G with x and T deleted. Let $f(S)$, $S \subseteq T$, be the probability that a random independent set of H is not adjacent to any vertices in S but is adjacent to all vertices in $T - S$. Since the independent sets of G can be built up by adding vertices in $x \cup T$ to independent sets in H, we have the following expressions for p_x and $\bar{p}_x$:

$$p_x = \frac{1}{1 + \Sigma_{S \subseteq T} f(S)I(S)}, \tag{1}$$

$$\bar{p}_x = \frac{\Sigma_{S \subseteq T} f(S)I(S)\bar{\alpha}(S)}{d[1 + \Sigma_{S \subseteq T} f(S)I(S)]}, \tag{2}$$

where $I(S)$ is the number of independent sets in S and $\bar{\alpha}(S)$ is the average size of an independent set in S. Note since G is K_r-free T, S must be K_{r-1}-free. Therefore, Lemma 1 applies, and we have $\bar{\alpha}(S) \ge c(r-1) \frac{\ln I(S)}{\ln \ln I(S)}$. Now let λ be a parameter to be set later. Let $\Sigma_{S \subseteq T, I(S) \ge \lambda} f(S)I(S) = w$. Let $c(r-1) \frac{\ln \lambda}{\ln \ln \lambda} = y$. Then we may obtain the following inequalities from (1) and (2):

$$p_x \geq \frac{1}{1+\lambda+w}, \tag{3}$$

$$\bar{p}_x \geq \frac{wy}{d[1+\lambda+w]}. \tag{4}$$

Now (3) is a decreasing function of w and (4) is an increasing function of w. hence $p_x + \bar{p}_x \geq \max(p_x, \bar{p}_x) \geq \max(\frac{1}{1+\lambda+w}, \frac{wy}{d}\ \frac{1}{1+\lambda+w}) \geq \frac{1}{1+\lambda+d/y}$. Now choose $\lambda = d/\ln d$. Then, to leading order, we have $\lambda + \frac{d}{y} = \frac{d}{\ln d} + \frac{d \ln\ln d}{c(r-1)\ln d}$ and $\frac{1}{1+\lambda+d/y} = \frac{c(r-1)\ln d}{d \ln\ln d}$. Now $\bar{\alpha} = \Sigma_x p_x = \Sigma_x \bar{p}_x$. Hence, $2\bar{\alpha} = \Sigma_x p_x + \bar{p}_x \geq n \frac{c(r-1)\ln d}{d\ln\ln d}$ or $\bar{\alpha} \geq \frac{c(r-1)}{2}\ n\ \frac{\ln d}{d\ln\ln d}$, completing the proof of Theorem 1.

The regularity condition in Theorem 1 may be removed if we look at α instead of $\bar{\alpha}$ as shown by the following corollaries.

Corollary 1. *Let G be a graph on n points with maximum degree d and which contains no K_r $(r \geq 4)$. Let α be the maximum size of an independent set of G. Then, for large d, $\alpha \geq c(r)n \frac{\ln d}{d\ln\ln d}$.*

Proof. Suppose G is not regular. Consider the graph consisting of two copies of G with corresponding vertices connected iff they have degree less than d. Clearly, by repeating this construction as many times as necessary, we can produce a regular graph G^* of degree d consisting of many copies of G with some additional edges connecting corresponding vertices. If G is K_r-free, G^* will be also hence we may apply Theorem 1 to G^*. The corollary follows immediately since if an independent set in G^* contains at least the fraction $c(r)\frac{\ln d}{d\ln\ln d}$ of the points of G^*, it must also contain at least this fraction of the points of some copy of G within G^*.

Corollary 2. *Let G be a graph on n points with average degree d and which contains no $K_r (r \geq 4)$. Let α be the maximum size of an independent set of G. Then, for large d, $\alpha \geq c'(r)n \frac{\ln d}{d\ln\ln d}$.*

Proof. Let G be a graph on n points with average degree d and which contains no K_r $(r \geq 4)$. Note at most half the points of G have degree greater than $2d$. Let G' be G with these points deleted. G' has at least $n/2$ points and maximum degree $\leq 2d$. Hence, by Corollary 1, $\alpha(G) \geq \alpha(G') \geq c(r)\ \frac{n}{2}\ \frac{\ln 2d}{2d\ln\ln 2d} = c'(r)n\ \frac{\ln d}{d\ln\ln d}$, completing the proof.

REFERENCES

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

[2] D. J. Kleitman, J. Shearer, and D. Sturtevant, Intersections of k-element sets, *Combinatorica*, **1**, 381–384 (1981).

Received July 12, 1994
Accepted March 13, 1995