Graphs whose vertices are graphs with bounded degree: Distance problems

Krystyna T. Balińska [a], Graham R. Brightwell [b] and Louis V. Quintas [c]

[a] *Computer Science Center, The Technical University of Poznań, Pl. M. Skłodowskiej-Curie 5, 60-965 Poznań, Poland*

[b] *Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK*

[c] *Mathematics Department, Pace University, New York, NY 10038, USA*

Received 4 April 1997; revised 5 May 1998

By an f-graph we mean a graph having no vertex of degree greater than f. Let $U(n, f)$ denote the graph whose vertex set is the set of unlabeled f-graphs of order n and such that the vertex corresponding to the graph G is adjacent to the vertex corresponding to the graph H if and only if H is obtainable from G by either the insertion or the deletion of a single edge. The *distance* between two graphs G and H of order n is defined as the least number of insertions and deletions of edges in G needed to obtain H. This is also the distance between two vertices in $U(n, f)$. For simplicity, we also refer to the vertices in $U(n, f)$ as the graphs in $U(n, f)$. The graphs in $U(n, f)$ are naturally grouped and ordered in levels by their number of edges. The distance $\lfloor nf/2 \rfloor$ from the empty graph to an f-graph having a maximum number of edges is called the *height* of $U(n, f)$. For $f = 2$ and for $f \geqslant (n-1)/2$, the diameter of $U(n, f)$ is equal to the height. However, there are values of the parameters where the diameter exceeds the height. We present what is known about the following two problems: (1) What is the diameter of $U(n, f)$ when $3 \leqslant f < (n-1)/2$? (2) For fixed f, what is the least value of n such that the diameter of $U(n, f)$ exceeds the height of $U(n, f)$?

1. Introduction

By an *f-graph* we mean a graph having no vertex of degree greater than f. Let $U(n, f)$ denote the graph whose vertex set is the set of unlabeled f-graphs of *order* n (number of vertices) and such that the vertex corresponding to the graph G is adjacent to the vertex corresponding to the graph H if and only if H is obtainable from G by either the insertion or the deletion of a single edge [1,3]. The *distance* between two graphs G and H of order n is defined as the least number of insertions and deletions of edges in G needed to obtain H and is denoted $d(G, H)$. This also denotes the distance between vertices in $U(n, f)$. The *diameter* of the graph $U(n, f)$ is defined as $\max\{d(G, H)\colon G, H \in V(U(n, f))\}$ and is denoted $\operatorname{diam} U(n, f)$. The graph $U(n, f)$ is the underlying graph of the transition digraph of the random f-graph process, a model of considerable interest in chemistry and physics [2]. Physically, the insertion or the deletion of an edge in this model can be interpreted as the creation or the breaking of

© J.C. Baltzer AG, Science Publishers

a bond between atoms or a link between molecules. For an application in medicinal chemistry and other distance functions between graphs, see [7,8,10].

Here we obtain results concerning distances in the graph $U(n,f)$.

The following theorem provides the starting point for our work in this paper. Parts (1), (3), and (4) appeared in [1]. Result (4) was noted earlier by Zelinka [12] as theorem 3.

Theorem 1.1. (1) $\operatorname{diam} U(n,2) = n$ for all $n \geqslant 3$;

(2) If $n = 20k$ for some integer k, then $\operatorname{diam} U(n,3) \geqslant 3n/2 + 2k = 8n/5$;

(3) If f is fixed, $f \geqslant 4$, and M is a positive number, then for any integer $x \geqslant 2M/f(f-3)$ and m the order of any f-regular graph with girth $\geqslant f+2$ it follows that, for $n = xm$, $\operatorname{diam} U(n,f) \geqslant nf/2 + M$; and

(4) $\operatorname{diam} U(n,n-1) = n(n-1)/2$ for all $n \geqslant 1$.

Before the proof of theorem 1.1 we first state a lemma, concerning the *size* (number of edges) of a graph, that we shall refer to repeatedly.

Lemma 1.2 (see [1, lemma 2.3]). Let G and H be two f-graphs of order n and I a maximum size unlabeled subgraph of order n common to both G and H. Then

$$d(G,H) = d(G,I) + d(I,H) = \left|E(G)\right| + \left|E(H)\right| - 2\left|E(I)\right|.$$

Remark. This is the distance function based on the notion of a maximal common subgraph introduced and studied in the 1980's for graphs with $f = n-1$ (see [7,8,10, 12]).

Proof of theorem 1.1. For parts (1), (3), and (4), see [1, lemmas 2.1, 2.4, and 2.2].

(2) Let G be the union of $2k$ disjoint copies of the Petersen graph and H the union of $5k$ disjoint copies of K_4 (where K_n is the *complete graph of order* n). For $k \geqslant 1$, G and H are 3-regular graphs of order $n = 20k$.

Let I be a common subgraph of G and H. Then I must be a forest, since G contains no cycle of length less than 5, and H contains none greater than 4. Furthermore, there are at most $4k$ components of I of order 4, since each component of G hosts at most 2 of them. Therefore, all $5k$ components of H have 3 or fewer edges of I, and at least k of the components have 2 or fewer edges. Thus, the number of edges of I is at most $3(4k) + 2(k) = 14k$.

Thus, $d(G,H) \geqslant 30k + 30k - 2(14k) = 32k = 32n/20 = 8n/5 = 3n/2 + 2k$. (Indeed, one can readily verify that $d(G,H) = 3n/2 + 2k$.) □

Theorem 1.1 (2) solves a problem posed at the 1995 Prague Midsummer Combinatorial Workshop ([1, problem 1] and [3, problem 1]). Subsequent to the proof of theorem 1.1 (2), this result was also obtained by B. Guiduldi at the Prague Workshop.

Theorem 1.1 led us to consider the problem of determining the diameter of $U(n,f)$ for all n and f and to a sequence of results directed at determining the least integer n, for which there exist two f-graphs G and H of order n such that $d(G,H) > \lfloor nf/2 \rfloor$ or, equivalently, such that $\operatorname{diam} U(n,f) > \lfloor nf/2 \rfloor$. With respect to the diameter problem we have shown, if $f = 2$ or $f \geqslant (n-1)/2$, then $\operatorname{diam} U(n,f) = \lfloor nf/2 \rfloor$ (see theorem 1.1 (1) for $f = 2$ and theorem 2.9 for $f \geqslant (n-1)/2$). Study of the latter problem is motivated by the observation that the graphs (that is, the vertices) in $U(n,f)$ are naturally grouped and ordered by their number of edges into levels. Let H be an f-graph with a maximum number of edges, then $d(K_n^{\mathrm{c}}, H) = \lfloor nf/2 \rfloor$ (where K_n^{c} is the *complement* of K_n). This distance can be thought of as the "height" of $U(n,f)$. Thus, the problem is to determine the least n, for which the "width" of $U(n,f)$ exceeds its height.

Problem 1. For $3 \leqslant f < (n-1)/2$, determine $\operatorname{diam} U(n,f)$.

Problem 2. For fixed $f \geqslant 3$, determine $w(f)$, the least integer n, for which there exist two f-graphs G and H of order n such that $d(G,H) > \lfloor nf/2 \rfloor$ or, equivalently, such that $\operatorname{diam} U(n,f) > \lfloor nf/2 \rfloor$.

2. Results

Theorem 2.1. Given $f \geqslant 4$ and $n \geqslant f+1$, let $n = \alpha(f+1)+\beta$ so that $\alpha = \lfloor n/(f+1) \rfloor$ and $0 \leqslant \beta \leqslant f$, and let

$$H = \begin{cases} \alpha K_{f+1} \cup \beta K_1, & 0 \leqslant \beta \leqslant 4, \\ \alpha K_{f+1} \cup K_\beta, & 5 \leqslant \beta \leqslant f. \end{cases}$$

Then, for any f-regular graph G of order n and girth $\geqslant f+2$, we have

$$d(G,H) \geqslant \begin{cases} (1/2)nf + (1/2)\alpha f(f-3), & 0 \leqslant \beta \leqslant 4, \\ (1/2)nf + (1/2)\alpha f(f-3) + (1/2)(\beta-1)(\beta-4), & 5 \leqslant \beta \leqslant f. \end{cases}$$

Furthermore, if $f \geqslant 5$ and G has girth $\geqslant f$, we have

$$d(G,H) \geqslant \begin{cases} (1/2)nf + (1/2)\alpha(f+1)(f-4), & 0 \leqslant \beta \leqslant 4, \\ (1/2)nf + (1/2)\alpha(f+1)(f-4) + (1/2)(\beta-1)(\beta-4), & 5 \leqslant \beta < f, \\ (1/2)nf + (1/2)\alpha(f+1)(f-4) + (1/2)f(f-5), & \beta = f. \end{cases}$$

Proof. By lemma 1.2,

$$d(G,H) = \begin{cases} (1/2)nf + (1/2)\alpha(f+1)f - 2|E(I)|, & 0 \leqslant \beta \leqslant 4, \\ (1/2)nf + (1/2)\alpha(f+1)f + (1/2)\beta(\beta-1) - 2|E(I)|, & 5 \leqslant \beta \leqslant f. \end{cases}$$

Since G has girth at least $f+2$ and each of the α nontrivial components of H of order $f+1$ have circumference $f+1$, each such component can contribute at most a tree with f edges to I. Thus, $|E(I)| \leqslant \alpha f$ when $0 \leqslant \beta \leqslant 4$ and, when $5 \leqslant \beta \leqslant f$, the case where H has the additional nontrivial component K_β, we have $|E(I)| \leqslant \alpha f + \beta - 1$.

Thus,

$$d(G,H) \geqslant \begin{cases} (1/2)nf + (1/2)\alpha(f+1)f - 2\alpha f, & 0 \leqslant \beta \leqslant 4, \\ (1/2)nf + (1/2)\alpha(f+1)f + (1/2)\beta(\beta-1) & \\ \quad - 2(\alpha f + \beta - 1), & 5 \leqslant \beta \leqslant f, \end{cases}$$

and simplification yields the desired result.

For the case $f \geqslant 5$ and girth $\geqslant f$, we have the possibility that each of the K_{f+1} components of H can contribute up to $f+1$ edges to I in the form of an f-cycle with a pendant edge; thus $|E(I)| \leqslant \alpha(f+1)$ when $0 \leqslant \beta \leqslant 4$, $|E(I)| \leqslant \alpha(f+1) + \beta - 1$ if $5 \leqslant \beta < f$, and $|E(I)| \leqslant \alpha(f+1) + f$ when $\beta = f$.

Thus,

$$d(G,H) \geqslant \begin{cases} (1/2)nf + (1/2)\alpha(f+1)f - 2\alpha(f+1), & 0 \leqslant \beta \leqslant 4, \\ (1/2)nf + (1/2)\alpha(f+1)f + (1/2)\beta(\beta-1) & \\ \quad - 2(\alpha(f+1) + \beta - 1), & 5 \leqslant \beta < f, \\ (1/2)nf + (1/2)\alpha(f+1)f + (1/2)f(f-1) & \\ \quad - 2(\alpha(f+1) + f), & \beta = f, \end{cases}$$

and this simplifies to the second assertion of the theorem. □

Note that by replacing α and β in theorem 2.1 by expressions in n and f we can derive the following cruder, but more transparent result. Since $\operatorname{diam} U(n,2) = n$ for $n \geqslant 3$, most of what follows is written with a view to obtaining results for $f \geqslant 3$.

Theorem 2.2. Let f be a fixed integer at least equal to 3. Then, for sufficiently large n,

$$\operatorname{diam} U(n,f) \geqslant \frac{nf(f-1)}{f+1} - \frac{(f+1)^2}{8}.$$

Corollary 2.3. For fixed $f \geqslant 3$ and any two f-graphs G and H of order n, we have $d(G,H) \leqslant nf^2/(f+1)$ and, for sufficiently large n,

$$\frac{nf(f-1)}{f+1} - \frac{(f+1)^2}{8} \leqslant \operatorname{diam} U(n,f) \leqslant \frac{nf^2}{f+1}.$$

Proof. By Vizing's theorem, any f-graph can be edge-colored using at most $f+1$ colors. Thus, the f-graph G has a matching (the largest color class) of size at least $|E(G)|/(f+1)$. Now suppose that G and H are two f-graphs of order n with $|E(G)| \leqslant |E(H)|$. Then, G and H have a common subgraph I of size at least $|E(G)|/(f+1)$. Thus,

$$\begin{aligned} d(G,H) &= |E(G)| + |E(H)| - 2|E(I)| \\ &\leqslant |E(G)| + |E(H)| - 2|E(G)|/(f+1) \leqslant \frac{nf^2}{f+1}. \end{aligned}$$

Combining this with theorem 2.2 completes the proof. □

Corollary 2.4. If $\lim_{n\to\infty}(1/n)\operatorname{diam} U(n,f)$ exists for $f \geqslant 3$, then

$$1 \leqslant \frac{2(f-1)}{f+1} \leqslant \lim_{n\to\infty} \frac{\operatorname{diam} U(n,f)}{nf/2} \leqslant \frac{2f}{f+1} < 2.$$

Proof. Dividing the conclusion of corollary 2.3 by $nf/2$ yields

$$1 \leqslant \frac{2(f-1)}{f+1} - \frac{(f+1)^2}{4nf} \leqslant \frac{\operatorname{diam} U(n,f)}{nf/2} \leqslant \frac{2f}{f+1} < 2.$$

Applying the limit as n goes to infinity concludes the proof. □

Problem 3. Prove or disprove that $\lim_{n\to\infty}(1/n)\operatorname{diam} U(n,f)$ exists.

Remark. It should be possible to improve the upper bound in corollary 2.3 with a little more work. The result does, however, describe the general behavior of $\operatorname{diam} U(n,f)$ for n very large compared to f.

We now consider the case where f is fairly large compared to n, in particular if f is around cn (see theorem 2.9). We start with a simple probabilistic argument to show that any two graphs with many edges have a large common subgraph.

Lemma 2.5. Suppose G and H are graphs of order n with size R and S, respectively, and I is a maximum size unlabeled subgraph of order n common to both G and H. Then, $|E(I)| \geqslant RS/\binom{n}{2}$.

Proof. Randomly label the vertices of G and H with elements from $\{1,2,\ldots,n\}$. Then, the probability that a (labeled) edge $\{i,j\}$ is in G is $R/\binom{n}{2}$, and the probability that $\{i,j\}$ is in H is $S/\binom{n}{2}$. Since these are independent events, the probability that $\{i,j\}$ is in the intersection of G and H is $RS/\binom{n}{2}^2$. Thus, the expected number of edges in the intersection of the labeled graphs G and H is equal to $RS/\binom{n}{2}$. Therefore, there is some labeling of G and H such that the size of their intersection is at least $RS/\binom{n}{2}$ and, consequently, $|E(I)| \geqslant RS/\binom{n}{2}$. □

Lemma 2.6. Let G and H be f-graphs of order n and size R and S, respectively. Then,

$$d(G,H) \leqslant R + S - 2RS\bigg/\binom{n}{2}.$$

Proof. By lemmas 1.2 and 2.5, $d(G,H) = R+S-2|E(I)| \leqslant R+S-2RS/\binom{n}{2}$. □

Remark. The bound $R+S-2RS/\binom{n}{2}$ in lemma 2.6 can be made more explicit for various values of R and S. For example, if either R or S is 0, then $d(G,H) \leqslant R+S$.

If $R \geqslant 1$ and $S \geqslant 1$, then $d(G,H) \leqslant R + S - 2$. For more results of this type, see [10].

Lemma 2.7. If G and H are graphs of order n and $|E(G)| \geqslant (1/2)\binom{n}{2}$, then

$$d(G,H) \leqslant d\bigl(G, K_n^{\mathrm{c}}\bigr) = \bigl|E(G)\bigr|.$$

Proof. Let G and H have size R and S, respectively. Then, by lemma 1.2,

$$d(G,H) = R + S - 2\bigl|E(I)\bigr|.$$

By lemma 2.5 we have

$$d(G,H) \leqslant R + S - 2RS\Big/\binom{n}{2} = R + S\left(1 - 2R\Big/\binom{n}{2}\right).$$

Since $R \geqslant (1/2)\binom{n}{2}$, we have $2R/\binom{n}{2} \geqslant 1$. Thus, $d(G,H) \leqslant R$, and this bound is realized. □

Lemma 2.8. If G and H are f-graphs of order n with $f \geqslant (n-1)/2$ and both G and H have size no greater than $(1/2)\binom{n}{2}$, then $d(G,H) \leqslant nf/2$.

Proof. Let G and H have size $R = (1/2)\binom{n}{2} - \alpha$ and $S = (1/2)\binom{n}{2} - \beta$, respectively, with $\alpha \geqslant 0$ and $\beta \geqslant 0$. By lemmas 1.2 and 2.5 we have

$$\begin{aligned} d(G,H) &= R + S - 2\bigl|E(I)\bigr| \\ &\leqslant R + S - 2RS\Big/\binom{n}{2} = R + S\left(1 - 2R\Big/\binom{n}{2}\right). \end{aligned}$$

Thus,

$$\begin{aligned} d(G,H) \leqslant \frac{1}{2}\binom{n}{2} - \alpha + \frac{\alpha}{(1/2)\binom{n}{2}} &= \frac{1}{2}\binom{n}{2} - \alpha + \left(\frac{1}{2}\binom{n}{2} - \beta\right)\frac{\alpha}{(1/2)\binom{n}{2}} \\ &= \frac{1}{2}\binom{n}{2} - \frac{\beta\alpha}{(1/2)\binom{n}{2}} \leqslant \frac{1}{2}\binom{n}{2}. \end{aligned}$$

Since $f \geqslant (n-1)/2$, we have $d(G,H) \leqslant nf/2$. □

The following theorem, announced in [3], follows from lemmas 2.5–2.8.

Theorem 2.9. If $f \geqslant (n-1)/2$, then $\operatorname{diam} U(n,f) = \lfloor nf/2 \rfloor$.

Proof. Let G and H be f-graphs of order n and size R and S, respectively.

(i) If $R \geqslant (1/2)\binom{n}{2}$ or $S \geqslant (1/2)\binom{n}{2}$, then, by lemma 2.7, $d(G,H) \leqslant \max(R,S) \leqslant nf/2$, and

(ii) if $R \leqslant (1/2)\binom{n}{2}$ and $S \leqslant (1/2)\binom{n}{2}$, then, by lemma 2.8, $d(G,H) \leqslant nf/2$.

By (i) and (ii) we have

$$\operatorname{diam} U(n,f) \leqslant nf/2.$$

For any n and f, there exists an f-graph G of order n and size $\lfloor nf/2 \rfloor$. Thus,

$$d\big(G, K_n^{\mathrm{c}}\big) = \lfloor nf/2 \rfloor \quad \text{and} \quad \operatorname{diam} U(n,f) \geqslant \lfloor nf/2 \rfloor.$$

Therefore, $\operatorname{diam} U(n,f) = \lfloor nf/2 \rfloor$. □

Lemma 2.10. If $f < (n-1)/2$, then $\operatorname{diam} U(n,f) \leqslant nf(1 - f/(n-1))$.

Proof. Let G and H be any two f-graphs of order n and size R and S, respectively. Then, from lemmas 1.2 and 2.6,

$$d(G,H) = R + S - 2\big|E(I)\big| \leqslant R + S - 2RS \Big/ \binom{n}{2}.$$

Since G and H are arbitrary f-graphs, we can consider $B(R,S) = R + S - 2RS/\binom{n}{2}$ as a function of two independent variables R and S with domains $0 \leqslant R \leqslant nf/2$ and $0 \leqslant S \leqslant nf/2$. Since f is bounded by $(n-1)/2$, the maximum value of $B(R,S)$ is easily shown to occur at $R = S = nf/2$.

This yields

$$d(G,H) \leqslant nf/2 + nf/2 - 2(nf/2)^2 \Big/ \binom{n}{2} = nf\big(1 - 2nf2/4n(n-1)\big)$$
$$= nf\big(1 - f/(n-1)\big). \qquad \square$$

Our next goal is to find a lower bound for $\operatorname{diam} U(n,f)$, useful in the range $f = cn$ for large n. Our plan is, roughly speaking, to show that two random graphs almost surely have no very large common subgraph.

The first of the following results gives estimates for the tails of a binomial distribution with parameters n and p (see [6, p. 5] for definition and [6, pp. 5–14] for relevant discussion and other results of this type).

Lemma 2.11 (Chernoff's inequalities). Let X denote a binomial random variable with parameters n and p. Then

$$\Pr\big(X \leqslant pn(1-\varepsilon)\big) \leqslant \exp(-\varepsilon^2 pn/2)$$

and

$$\Pr\big(X \geqslant pn(1+\varepsilon)\big) \leqslant \exp(-\varepsilon^2 pn/3).$$

Let $G_{n,p}$ denote a random graph having n vertices and each edge present with probability $p = f/n$. Thus, $G_{n,p}$ is almost, but not quite, an f-graph. For the inequalities that we shall derive, we also assume that in lemmas 2.12–2.14 the parameter n is at least 10^6.

Lemma 2.12. With probability at least $1/2$, a random graph $G_{n,p}$, with $p = f/n$ and $n \geqslant 10^6$, can be made into an f-graph by the deletion of at most $4n\sqrt{pn}$ edges.

Proof. Since the degree $d(x)$ of any given vertex x in $G_{n,p}$ is a binomial random variable X with parameters $n-1$ and p, the probability that X is at least $p(n-1) + a\sqrt{p(n-1)}$ is at most $\exp(-a^2/3)$ for any positive real a. Thus, the probability that the degree is at least $pn + i$ is at most $\exp(-(i+p)^2/3p(n-1)) \leqslant \exp(-i^2/3pn)$.

Define the random variable $D = D_{n,p}$ to be the sum of $d(x) - f$ over all vertices x whose degree is greater than $f = pn$. Then the expectation of D is

$$n \sum_{i=1}^{n} \Pr\bigl(d(x) \geqslant f + i\bigr) \leqslant n \sum_{i=1}^{n} \exp(-i^2/3pn)$$
$$\leqslant n \int_{i=0}^{\infty} \exp(-i^2/3pn)\,\mathrm{d}i = n\sqrt{\frac{3\pi pn}{4}}.$$

Thus, the probability that D is greater than $n\sqrt{3\pi pn}$ is at most $1/2$, and this quantity is less than $4n\sqrt{pn}$, as required. □

Lemma 2.13. $G_{n,p}$ with $n \geqslant 10^6$ has at least $p\binom{n}{2} - 2n\sqrt{p}$ edges, with probability at least $9/10$.

Proof. The number of edges of $G_{n,p}$ is given by the binomial random variable X with parameters $\binom{n}{2}$ and p. Thus, by Chernoff's inequality (lemma 2.11) with $\varepsilon = 4/(n-1)\sqrt{p}$, the probability is at least $1 - \exp(-4n/(n-1)) > 9/10$ that $G_{n,p}$ will have the asserted number of edges. □

Let H be any fixed f-regular graph on vertex set $\{1, 2, \ldots, n\}$, or, if nf is odd, let H have $n-1$ vertices of degree f and one vertex of degree $f-1$. Thus, H has $nf/2$ or $(nf-1)/2$ edges, respectively.

Lemma 2.14. The probability that $G_{n,p}$ with $n \geqslant 10^6$ has more than $(1/2)pnf + (n/2)\sqrt{6pf \log n}$ edges in common with H is at most n^{-n}.

Proof. The number of common edges between $G_{n,p}$ and H is a binomial random variable X with parameters $nf/2$ and p. Thus, by lemma 2.11 with $\varepsilon = \sqrt{6 \log n/pf}$, we have the desired result. □

Theorem 2.15. For every f-regular graph H on $n \geqslant 10^6$ vertices, there is an f-graph G such that

$$d(G, H) \geqslant f(n-f) - 5n\sqrt{f} - f\sqrt{6n \log n};$$

thus, for $n \geqslant 10^6$,

$$\operatorname{diam} U(n, f) \geqslant f(n-f) - 5n\sqrt{f} - f\sqrt{6n \log n}.$$

Proof. Let H be any f-regular graph of order $n \geqslant 10^6$ and set $p = f/n$. Let G' be a graph on n vertices such that:

(a) G' has at least $p\binom{n}{2} - 2n\sqrt{p}$ edges,

(b) G' can be made into an f-graph by the deletion of at most $4n\sqrt{pn}$ edges, and

(c) there is no relabeling of the vertices of G' so that G' has more than $(pnf + n\sqrt{6pf \log n})/2$ edges in common with H.

By lemmas 2.12–2.14, a random graph $G_{n,p}$ has all properties (a)–(c) with probability at least $1 - 1/2 - 1/10 - n!n^{-n} > 1/4$. Thus, such a graph G' exists. Now let G be any f-graph obtained from G' by the deletion of at most $4n\sqrt{pn}$ edges.

Since $d(G,H) = |E(G)| + |E(H)| - 2|E(I)|$ by lemma 1.2, we have

$$d(G,H) \geqslant p\binom{n}{2} - 2n\sqrt{p} - 4n\sqrt{pn} + nf/2 - 2\big(\big(pnf + n\sqrt{6pf \log n}\big)/2\big),$$

so that

$$\begin{aligned} d(G,H) &\geqslant f(n-1)/2 - 2n\sqrt{f/n} - 4n\sqrt{f} + nf/2 - f^2 - n\sqrt{6\big(f^2/2\big)\log n} \\ &= fn - f^2 - f/2 - 2\sqrt{nf} - 4n\sqrt{f} - f\sqrt{6n \log n} \\ &\geqslant fn - f^2 - 5n\sqrt{f} - f\sqrt{6n\log n}. \end{aligned}$$

□

Corollary 2.16. If $f < (n-1)/2$ and $n \geqslant 10^6$, then

$$f(n-f) - 5n\sqrt{f} - f\sqrt{6n \log n} \leqslant \operatorname{diam} U(n,f) \leqslant nf\left(1 - \frac{f}{n-1}\right).$$

Proof. This statement is the combination of theorem 2.15 and lemma 2.10. □

We can also assert the following corollary.

Corollary 2.17. For $n \geqslant 10^6$ and $100 < f \leqslant n/2 - 3\sqrt{n \log n}$, we have

$$\operatorname{diam} U(n,f) > nf/2.$$

Corollary 2.18. For any function $f = f(n)$ tending to infinity with n and such that $f < (n-1)/2$, we have

$$\operatorname{diam} U(n, f) = f(n - f) + \mathrm{o}(nf)$$

as n goes to infinity.

Proof. This follows from corollary 2.16. □

With respect to problem 2, we have the following result implied by corollary 2.17.

Theorem 2.19. If $f \geqslant 5 \times 10^5$, then $w(f) < 2f + 10\sqrt{f \log f}$.

Remark. It may be possible, say by adapting the methods used in the proof of theorem 2.15, to arrive at $w(f) < 2f + c\sqrt{f \log f}$ for all f and c some absolute constant. We also note that determining precise values of $w(f)$ remains an interesting open problem. Bounds for $w(f)$ for up to $f = 7$ are given in the following theorem.

Theorem 2.20.

(1) $8 \leqslant w(3) \leqslant 12$;

(2) $10 \leqslant w(4) \leqslant 20$;

(3) $12 \leqslant w(5) \leqslant 26$;

(4) $14 \leqslant w(6) \leqslant 49$; and

(5) $16 \leqslant w(7) \leqslant 52$.

Proof. By theorem 2.9, if $f \geqslant (n-1)/2$ or, equivalently, if $2f + 1 \geqslant n$, then $\operatorname{diam} U(n, f) = \lfloor nf/2 \rfloor$. Since $\operatorname{diam} U(n, f) \geqslant \lfloor nf/2 \rfloor$ is always the case, we have $\operatorname{diam} U(n, f) \neq \lfloor nf/2 \rfloor$ is equivalent to $\operatorname{diam} U(n, f) > \lfloor nf/2 \rfloor$. Thus, $\operatorname{diam} U(n, f) > \lfloor nf/2 \rfloor$ implies $n > 2f + 1$. Therefore, $w(f) \geqslant 2f + 2$. This is used as a (weak) lower bound in what follows.

(1) From theorem 1.1 (2), we have $\operatorname{diam} U(n, 3) \geqslant 3n/2 + n/10$ for $n = 20k \geqslant 20$. We further note that, if G denotes the Petersen graph union two isolated vertices and H consists of three disjoint copies of K_4, then G and H are 3-graphs of order 12 such that, for $n = 12k$, $d(kG, kH) = 15k + 18k - 2(7k) = 19k = 3n/2 + k$. Thus, $k = 1$ yields $w(3) \leqslant 12$ and

$$8 \leqslant w(3) \leqslant 12.$$

(2) Referring to works of Kárteszi, Singleton and Brown (see [5, p. 162]) it is known that the Moore graphs of degree δ and girth 6, are in one-to-one correspondence with the finite abstract projective planes with δ points on a line. The Moore graph corresponding to a projective plane P is the bipartite graph G with the set of points and the set of lines of P as the two color classes of G. A "point vertex" in G is adjacent to a "line vertex" in G if and only if the point in P is on the line in P. For girth 6, the bipartite graph G with degree $\delta = 4$ has order $2(3^2+3+1) = 26$. This graph is unique

and a drawing of it can be found in ([4, figure 3.5, p. 61]). Applying theorem 2.1 with $f = 4$ and $g = 6$ yields $d(G, H) \geqslant 26(4)/2 + \lfloor 26/5 \rfloor 2(4-3) = 52 + 5(2)$. This yields $10 \leqslant w(4) \leqslant 26$.

We now delete six vertices which lie on a 6-cycle of the 4-regular graph G of order 26 that we have just described. In this way we obtain a 4-graph G^* of order 20. Since the only adjacencies of the six vertices on the above mentioned 6-cycle are those determined by the 6-cycle itself, the deletion of these six vertices removes eighteen edges from G. Thus, G^* has size $52 - 18 = 34$. The girth of G^* is 6. Let $H^* = 4K_5$. Since $|E(I)| \leqslant 4(4)$, we obtain $d(G^*, H^*) \geqslant 34 + 40 - 2(4)4 = 42 = 20(4)/2 + 2$. This yields

$$10 \leqslant w(4) \leqslant 20.$$

(3) In [11], Wegner showed that the minimum order of a 5-regular graph with girth 5 is equal to 30. For a drawing of a 5-regular graph of order 30 with girth 5, see ([4, figure 3.6, p. 62]). Applying theorem 2.1 with $f = 5$ and $g = 5$ yields $12 \leqslant w(5) \leqslant 30$.

We can now define G^* as the minimum order 5-regular graph with girth 5 shown in ([4, p. 62]) with four independent vertices deleted. G^* has order $30 - 4 = 26$, size $75 - 20 = 55$ and girth at least 5. Let $H^* = 4K_6 \cup 2K_1$. Then, since $|E(I)| \leqslant 4(6)$, we obtain $d(G^*, H^*) \geqslant 55 + 60 - 2(4)6 = 67 = 26(5)/2 + 2 > 26(5)/2$. This yields

$$12 \leqslant w(5) \leqslant 26.$$

(4) In [9], Kárteszi showed that the minimum order of a 6-regular graph G with girth 6 is equal to 62. Apply theorem 2.1 with $f = 6$ and $g = 6$ to obtain $14 \leqslant w(6) \leqslant 62$.

We shall now modify G to obtain a graph G^* with order 49, size 117, and girth 6. Let S_1 denote a set of six vertices on a 6-cycle C of G. Let a and b be a pair of adjacent vertices on C, S_2 denote a set of four vertices, each of which is at distance two from a and not on C, and S_3 a set of three vertices at distance one from b and not on C. Since adjacency would imply the existence of either a triangle or a 5-cycle, the set $S_2 \cup S_3$ is a set of independent vertices. The deletion of the vertices in $S_1 \cup S_2 \cup S_3$ from G yields the graph G^*. Let $H^* = 7K_7$. Then, using $|E(I)| \leqslant 7(7)$, we obtain $d(G^*, H^*) \geqslant 117 + 147 - 98 = 166 > 49(6)/2$. Therefore,

$$14 \leqslant w(6) \leqslant 49.$$

(5) Let G be the 6-regular graph of order 62, size 186 and girth 6 and C the 6-cycle in G as given in (4). Let S_4 be the set of four vertices at distance one from b and not on C in G and, as before, let S_1 be the set of six vertices of C. Deletion of the vertices in $S_1 \cup S_4$ defines a graph G^* with order 52, size 136, and girth 6. Let $H^* = 7K_7 \cup 3K_1$ and, using $|E(I)| \leqslant 7(7)$, we have $d(G^*, H^*) \geqslant 136 + 147 - 98 = 185$. Since G^* has maximum degree 6, G^* can be considered a 7-graph. Noting that $185 > 52(7)/2 = 182$, we obtain

$$16 \leqslant w(7) \leqslant 52. \qquad \square$$

3. Summary

We have shown that

$$\operatorname{diam} U(n,f) = \lfloor nf/2 \rfloor, \quad \text{when } f = 2 \text{ or } f \geqslant (n-1)/2$$

(see theorems 1.1 (1) and 2.9).

However, the determination of $\operatorname{diam} U(n,f)$ for fixed $f \geqslant 3$ with $f < (n-1)/2$ is an open problem. Some bounds for the solution of the diameter problem are:

$$\operatorname{diam} U(n,3) \geqslant 3n/2 + k, \quad \text{when } n = 12k$$

(see proof of theorem 2.20 (1)),

$$\operatorname{diam} U(n,3) \geqslant 3n/2 + 2k, \quad \text{when } n = 20k$$

(see theorem 1.1 (2)), and, if there exists an f-regular graph with $4 \leqslant f < (n-1)/2$, order n and girth $f+2$, then for $n = \alpha(f+1) + \beta$ so that $\alpha = \lfloor n/(f+1) \rfloor$ and $0 \leqslant \beta \leqslant f$, we have

$$\left.\begin{array}{ll} (1/2)nf + (1/2)\alpha f(f-3) & (0 \leqslant \beta \leqslant 4) \\ (1/2)nf + (1/2)\alpha f(f-3) + (1/2)(\beta-1)(\beta-4) & (5 \leqslant \beta \leqslant f) \end{array}\right\} \leqslant \operatorname{diam} U(n,f) \leqslant nf\left(1 - \frac{f}{n-1}\right)$$

(see theorem 2.1 and lemma 2.10).

A cruder but more transparent result is given by corollary 2.3, namely, for sufficiently large n and $f \geqslant 3$, we have

$$\frac{nf(f-1)}{f+1} - \frac{(f+1)^2}{8} \leqslant \operatorname{diam} U(n,f) \leqslant \frac{nf^2}{f+1},$$

and from corollary 2.4, we have

$$1 \leqslant \frac{2(f-1)}{f+1} \leqslant \lim_{n\to\infty} \frac{\operatorname{diam} U(n,f)}{nf/2} \leqslant \frac{2f}{f+1} < 2,$$

provided that the limit exists.

If $f < (n-1)/2$ and $n \geqslant 10^6$, a sharper result is available; here we have

$$f(n-f) - 5n\sqrt{f} - f\sqrt{6n \log n} \leqslant \operatorname{diam} U(n,f) \leqslant nf\left(1 - \frac{f}{n-1}\right)$$

(see corollary 2.16).

For any function $f = f(n)$ tending to infinity with n and such that $f < (n-1)/2$, we have

$$\operatorname{diam} U(n,f) = f(n-f) + \mathrm{o}(nf)$$

as n goes to infinity (see corollary 2.18).

See theorems 2.19 and 2.20 for results concerning $w(f)$.

Acknowledgements

Some of this work was done while the authors were attending the 1995 Prague Midsummer Combinatorial Workshop at Charles University. We thank the organizers of this meeting for their hospitality. We also thank Fred Buckley (Baruch College, CUNY), Michael L. Gargano (Pace University) and John W. Kennedy (Pace University) for helpful discussions. G.R. Brightwell's visit to Prague was supported by the European Union HCM Project "DONET". Support of this work was provided in part for K.T. Balińska by grants from the Technical University of Poznań and for L.V. Quintas by grants from the Dyson College of Arts and Sciences, Pace University.

References

[1] K.T. Balińska, F. Buckley, M.L. Gargano and L.V. Quintas, The diameter of a graph whose vertices are graphs, CSIS Pace University Technical Report Series, Report No. 89 (1995).

[2] K.T. Balińska and L.V. Quintas, The random f-graph process, in: *Quo Vadis, Graph Theory*, Ann. Discrete Math. 55 (1992) 333–340.

[3] K.T. Balińska and L.V. Quintas, A diameter problem and some midsize graph problems, in: *1995 Prague Midsummer Combinatorial Workshop*, ed. M. Klazar, KAM Series 95-309, Department of Applied Mathematics, Charles University, Prague (1995) pp. 6–10.

[4] M. Behzad, G. Chartrand and L. Lesniak-Foster, *Graphs and Digraphs* (Wadsworth International Group, Belmont, CA, 1979).

[5] B. Bollobás, *Extremal Graph Theory*, London Math. Soc. Monographs, Vol. 11, eds. P.M. Cohn and G.E.H. Reuter (Academic Press, London, 1978).

[6] B. Bollobás, *Random Graphs* (Academic Press, New York, 1985).

[7] M. Johnson, Relating metrics, lines and variables defined on graphs to problems in medicinal chemistry, in: *Proceedings of the Fifth International Conference on Graph Theory*, eds. Y. Alavi, G. Chartrand, L. Lesniak and C. Wall (Wiley, New York, 1985) pp. 457–470.

[8] M. Johnson, An ordering of some metrics defined on the space of graphs, Czechoslovak Math. J. 37(112) (1987) 75–85.

[9] F. Kárteszi, Piani finiti ciclici come risoluzioni di un certo problema di minimo, Boll. Un. Mat. Ital. (3) 15 (1960) 522–528.

[10] M. Šabo, On a maximal distance between graphs, Czechoslovak Math. J. 41(116) (1991) 265–268.

[11] G. Wegner, A smallest graph of girth 5 and valency 5, J. Combin. Theory Ser. B 14 (1973) 205–208.

[12] B. Zelinka, Edge distance between isomorphism classes of graphs, Cas. Pest. Mat. 112 (1987) 233–237.