Turkish Journal of Mathematics

http://journals.tubitak.gov.tr/math/

Research Article

Turk J Math
(2022) 46: 3359 – 3372

doi:10.55730/1300-0098.3337

Maximising the number of connected induced subgraphs of unicyclic graphs

Audace A. V. DOSSOU-OLORY*
Department of Hydrology and Water Resources Management,
African Center of Excellence for Water and Sanitation National Water Institute,
University of Abomey-Calavi, Abomey-Calavi, Benin

Received: 11.03.2022 • **Accepted/Published Online:** 07.10.2022 • **Final Version:** 09.11.2022

Abstract: Denote by $\mathcal{G}(n, c, g, k)$ the set of all connected graphs of order n, having c cycles, girth g, and k pendant vertices. In this paper, we give a partial characterisation of the structure of those graphs in $\mathcal{G}(n, c, g, k)$ maximising the number of connected induced subgraphs. For the special case where $c = 1$, we find a complete characterisation of all maximal unicyclic graphs. We also derive a precise formula for the corresponding maximum number given the following parameters: (1) order, girth, and number of pendant vertices; (2) order and girth; (3) order.

Key words: Induced subgraphs, connected graphs, unicyclic graphs, girth, pendant vertices

1. Introduction

All graphs considered in this paper are simple (i.e. finite, no parallel edges, no loops, and undirected). A simple graph is called unicyclic if it has only one cycle. Unicyclic graphs are among the most popular tree-like structures studied in chemical graph theory. Various graph parameters have been studied in the class of unicyclic graphs of a given order (number of vertices). For instance, Gao and Lu [3] gave sharp lower and upper bounds on the Randić index of unicyclic graphs; Ou [8] investigated the unicyclic graphs with given girth and minimal Hosoya index. Xia and Chen [12] determined the unicyclic graphs with the first five largest, as well as the first two smallest Zagreb indices; Du et al. [1] found the unicyclic graphs with a given maximum degree that have the maximal sum-connectivity index, just to mention a few. In this work, we study the problem of maximising the total number of connected induced subgraphs of a unicyclic graph under some type of restrictions.

An induced subgraph of a simple graph G is a graph that contains a nonempty subset of vertices of G together with all edges incident with them in G. A graph is said to be connected if there is a path from vertex v to vertex w for any $v, w \in V(G)$ (where by $V(G)$, we mean the vertex set of G). Connected induced subgraphs have been studied extensively for trees (connected acyclic graphs): Chung et al. [4] determined the smallest asymptotic order of a tree that contain all trees of order n as subtrees; Jamison [6, 7] studied the average order of a subtree of a tree; Székely and Wang [10, 11] investigated some extremal trees for the number of subtrees. Recently, the present author [2] studied the following parameters in general graphs as well as unicylic graphs with prescribed order: the total number of subgraphs, the total number of induced subgraphs, the total number of connected induced subgraphs. In particular, he found the unicyclic graphs that have the smallest and the largest number of connected induced subgraphs, respectively. Things change decisively if additional

*Correspondence: audace@aims.ac.za

2010 *AMS Mathematics Subject Classification:* 05C30, 05C35, 05C05 restrictions are taken into account. For example, the extremal unicylic graphs for the number of connected induced subgraphs are not known if girth is prescribed, the girth of a simple graph G being the minimum number of vertices among all the cycles in G. In this paper, we provide a complete solution to this problem. We further consider an additional restriction, namely the number of pendant vertices (number of vertices of degree 1) and characterise those maximising unicyclic graphs for the number of connected induced subgraphs with given order, girth, and number of pendant vertices.

Counting and understanding graph structures with particular properties has many applications, especially in network theory, computer science, biology, and chemistry. For instance, graphs can represent biological networks at the molecular or species level (protein interactions, gene regulation, etc.). The topological structure of an interconnection network is a connected graph where, for example, vertices are processors and edges represent links between them. In chemical networks, vertices are atoms and edges represent their bonds. The other initial motivation for studying the parameter number of connected induced subgraphs was twofold: firstly, our main purpose was to extend many of the extremal results on the number of subtrees of a tree to more general classes of graphs such as connected graphs or unicylic graphs. Extremal results on the number of connected subgraphs (not necessarily induced subgraphs) appeared recently in [9]. We remark that in general, there is no monotone relationship between the number of connected subgraphs and the number of connected induced subgraphs. In other words, if graph G has more connected subgraphs than graph H, it is not necessarily true that G also contains more connected induced subgraphs than H. The novelty of our work is that it focuses on the number of connected induced subgraphs rather than the number of connected subgraphs.

While in the past many results on the number of connected induced subgraphs were obtained for trees, this is not yet the case for general graphs. Can we find a constructive characterisation of the graphs extremising the number of connected induced subgraphs over all connected graphs with prescribed order n and number of cycles $c > 0$? In general, this problem can appear out of reach due to the various ways in which the cycles may intersect in the graph. However, special cases of unicyclic graphs ($c = 1$) and bicyclic graphs ($c = 2$ or 3) are still of interest. As was mentionned above, the case $c = 1$ was treated in a paper by the present author. In this paper, we shall study the same graph parameter with the additional restriction that the number of pendant vertices is prescribed, which, therefore, parallels those results obtained by Pandey and Patra [9] on the number of connected (not necessarily induced) subgraphs.

Our approach consists of the following steps: we first introduce a main graph transformation (Lemma 2.1) and present some of its properties. Then we discuss certain techniques to characterise the graphs for which the maximum number of connected induced subgraphs is obtained (Proposition 2). Thereafter, we focus on unicyclic graphs with the following prescribed parameters: order, girth, and number of pendant vertices. We give further intermediate results (Lemmas 2.2, 2.3, 2.4, 2.5) and their combination allows us to give a complete characterisation of the structure of those maximising unicyclic graphs with given order, girth, and number of pendant vertices (see Propositions 2, 2, 2 and Theorem 2.7 in Section 2). In Section 3, we compare unicyclic graphs with different number of pendant vertices. More precisely, we characterise the maximal (with respect to the number of connected induced subgraphs) unicyclic graphs with given order and girth (Theorem 3.1), and with given order only (see Theorem 3.2). Finally, we notice that all our maximal graphs were previously shown to also minimise the Wiener index (sum of distances between all unordered pairs of vertices).

For a graph G and two vertices v, w of G, we shall write $G - \{v\}$ (resp. $G - \{v, w\}$) to mean the induced subgraph consisting of all vertices of G except v (resp. all vertices of G except v, w). We shall also refer to a graph having the maximum number of connected induced subgraphs as optimal. We shall denote by $N(G)_v$ (resp. $N(G)_{v,w}$) the number of connected induced subgraphs of G that contain v (resp. both v and w), and by $\deg_G(v)$ the degree of v in G (i.e. the number of edges incident with v in G). By C_n, we mean the cycle of order n (so $n > 2$).

2. Getting to the optimal graphs

From now on, whenever we write subgraph, we always mean induced subgraph. We begin with some auxiliary results from which the proofs of our main theorems will be derived. The underlying technique is to apply a series of graph transformations that affect the total number of connected subgraphs while preserving a number of other parameters of the graph such as order, number of pendant vertices, etc.

The operation described in the following lemma increases the number of connected subgraphs of G at least by 1, while preserving the order and some other parameters of G. Lemma 2.1 has a counterpart for the Wiener index, see [5].

Lemma 2.1 *Let L, M, R be three connected graphs whose vertex sets are pairwise disjoints. Let $l \in V(L)$, $r \in V(R)$, $u, v \in V(M)$ be fixed vertices such that $u \neq v$. Denote by G the graph obtained from L, M, R by identifying l with u, and r with v. Similarly, let G' be the graph obtained from L, M, R by identifying both l, r with u, and G'' the graph obtained from L, M, R by identifying both l, r with v. See Figure 1 for a picture of the graphs G, G', G''.*

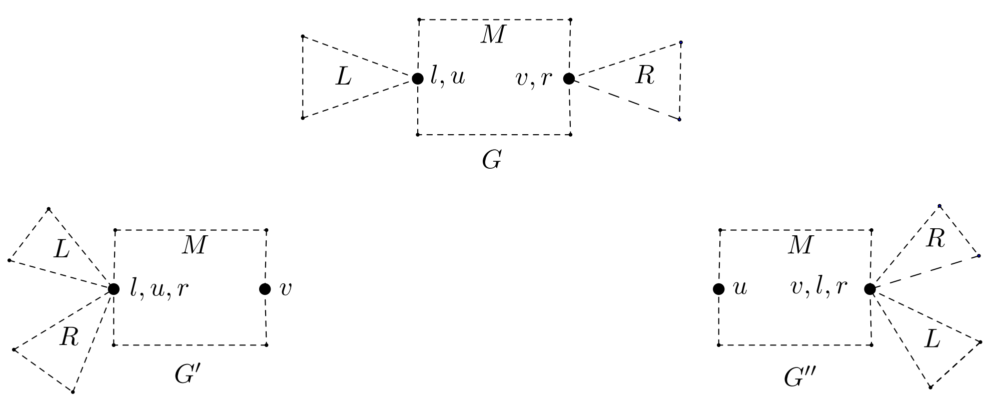

Figure 1. The three graphs G, G', G'' described in Lemma 2.1.

Assume that $|V(L)| > 1$ and $|V(R)| > 1$. Then we have

$$N(G') > N(G) \quad \text{or} \quad N(G'') > N(G)\,.$$

Proof Classify all connected subgraphs of each of the graphs G, G', G'' by the following cases:

1. those containing u and v;
2. those containing u but not v;
3. those containing v but not u;
4. those containing neither u nor v.

From this classification, we obtain

$$\begin{aligned} N(G) &= N(G)_{u,v} + N(G-\{v\})_u + N(G-\{u\})_v + N(G-\{u,v\}) \\ &= N(L)_l \cdot N(M)_{u,v} \cdot N(R)_r + N(L)_l \cdot N(M-\{v\})_u + N(M-\{u\})_v \cdot N(R)_r \\ &+ N(L-\{l\}) + N(M-\{u,v\}) + N(R-\{r\}) \end{aligned}$$

for the number of connected subgraphs of G. Likewise,

$$\begin{aligned} N(G') &= N(G')_{u,v} + N(G'-\{v\})_u + N(G'-\{u\})_v + N(G'-\{u,v\}) \\ &= N(L)_l \cdot N(M)_{u,v} \cdot N(R)_r + N(L)_l \cdot N(M-\{v\})_u \cdot N(R)_r \\ &+ N(M-\{u\})_v + N(L-\{l\}) + N(M-\{u,v\}) + N(R-\{r\}) \end{aligned}$$

for the number of connected subgraphs of G', and

$$\begin{aligned} N(G'') &= N(G'')_{u,v} + N(G''-\{v\})_u + N(G''-\{u\})_v + N(G''-\{u,v\}) \\ &= N(L)_l \cdot N(M)_{u,v} \cdot N(R)_r + N(M-\{v\})_u + N(L)_l \cdot N(M-\{u\})_v \cdot N(R)_r \\ &+ N(L-\{l\}) + N(M-\{u,v\}) + N(R-\{r\}) \end{aligned}$$

for the number of connected subgraphs of G''. It follows that

$$N(G') - N(G) = (N(R)_r - 1)(N(L)_l \cdot N(M-\{v\})_u - N(M-\{u\})_v)$$

and

$$N(G'') - N(G) = (N(L)_l - 1)(N(R)_r \cdot N(M-\{u\})_v - N(M-\{v\})_u) .$$

Since $N(R)_r - 1 > 0$ and $N(L)_l - 1 > 0$ by assumption, we deduce that

$$N(G') > N(G) \quad \text{if} \quad N(M-\{v\})_u \geq N(M-\{u\})_v$$

and

$$N(G'') > N(G) \quad \text{if} \quad N(M-\{v\})_u \leq N(M-\{u\})_v .$$

This completes the proof of the lemma. □

The setup presented in Lemma 2.1 also shows that except possibly vertices l, r, u, v, all parameters of G that solely depend of the feature of the single graphs L, M, R are preserved. For instance, the number of pendant vertices of G is preserved under this transformation provided that neither u, nor v has degree 1 in M (in which case the number of pendant vertices of both G' and G'' is at least one more that of G). Lemma 2.1 will be used repeatedly under specialisations.

Denote by $\mathcal{G}(n,c,g,k)$ the set of all connected graphs with order n, having c cycles, girth g, and k pendant vertices. The following proposition gives a partial characterisation of the structure of every optimal graph in a special subset of $\mathcal{G}(n,c,g,k)$.

Let $H_{n,1,g,k}$ be a graph that maximises the number of connected subgraphs over all graphs belonging to $\mathcal{G}(n,1,g,k)$. Then $H_{n,1,g,k}$ has precisely the shape of the graph depicted in Figure 2, where H is a tree of order $n-g+1$, having k pendant vertices.

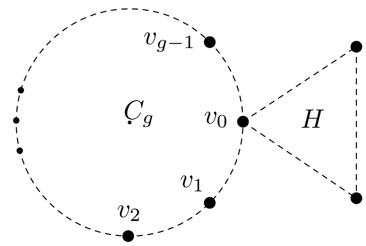

Figure 2. The shape of every optimal graph in $\mathcal{G}(n,1,g,k)$.

Proof Take M to be the cycle of order g. Then a repetitive application of Lemma 2.1 to every newly constructed graph (always choosing $M=C_g$) yields graphs all of which have precisely the shape of the graph shown in Figure 2. It is clear that the prescribed parameters: order, girth, and number of pendant vertices are all preserved at every step of the application of Lemma 2.1. This proves that H is indeed a tree of order $n-g+1$, having k pendant vertices. □

For the special case where H is a tree in Figure 2, one can even be more precise about the shape of every optimal graph. The same graph transformation presented in Lemma 2.1 can be used to increase the number of connected subgraphs further (for the special case $c=1$), while preserving all the parameters n,g,k. This is shown in the next lemma.

Lemma 2.2 *Let G be a connected graph and T be a subgraph of G. Assume that T is a rooted tree whose root is z and that $G-\{V(T)-\{z\}\}$ is connected. Further, assume that there are two distinct vertices $x\neq z$ and $y\neq z$ of T such that $\deg_T(x)\geq 3$ and $\deg_T(y)\geq 3$. Then G cannot be an optimal graph from $\mathcal{G}(n,1,g,k)$.*

Proof Denote by $x_1,x_2,\dots,x_p$ all neighbors of x in T and by $y_1,y_2,\dots,y_q$ all neighbors of y in T. Let x_1 (resp. y_1) be the unique neighbor of x (resp. y) that lies on the unique path from x to y in T (it is possible to have $x_1=y$ or $y_1=x$ or $x_1=y_1$). Furthermore, we let x_2 (resp. y_2) be the unique neighbor of x (resp. y) that lies on the unique path $P_{x,z}$ from x to z (resp. the unique path $P_{y,z}$ from y to z) in T if x_1 does not lie on $P_{x,z}$ (resp. y_1 does not lie on $P_{y,z}$). For every $i\in\{3,4,\dots,p\}$ (resp. $j\in\{3,4,\dots,q\}$), denote by L_{x_i} (resp. R_{y_j}) the subtree of T consisting of x_i (resp. y_j) and all its descendents in T. Let L (resp. R) be the rooted tree whose all branches are $L_{x_3},L_{x_4},\dots,L_{x_p}$ (resp. $R_{y_3},R_{y_4},\dots,R_{y_q}$). Thus, L (resp. R) is a rooted subtree of G rooted at vertex x (resp. y). Therefore, we can move R to L to produce a new graph G', and L to R to generate a new graph G'' through the operation given in Lemma 2.1 (see Figure 1) as $|V(L)|>1$ and $|V(R)|>1$. Hence, at least one of these moves strictly increases the number of connected subgraphs of G. This completes the proof of the lemma. □

It is important to note that the assumption that T is a tree (in Lemma 2.2) is essential to ensure that there is precisely one path between the two vertices of T that have degree greater than 2 in T. Of course, it may happen in the general case (where T is not necessarily a tree) that there are three distinct vertices x,y,z in T with precisely one path between any two of them: in this case, the same reasoning used in the proof of Lemma 2.2 can still be applied to them provided that $\deg_T(x)\geq 3$ and $\deg_T(y)\geq 3$.

An immediate consequence of Proposition 2 alongside Lemma 2.2 is that a graph $H_{n,1,g,k}$ maximising the number of connected subgraphs over all graphs from $\mathcal{G}(n,1,g,k)$ must have at most two vertices of degree greater than 2. On the other hand, the condition $n>g$ already guarantees at least one vertex of degree greater than 2 (which is vertex v_0 as shown in Figure 2) in $H_{n,1,g,k}$. This leaves us with only two main possibilities for the structure of every optimal graph in $\mathcal{G}(n,1,g,k)$, namely the graph H has precisely one vertex $w \neq v_0$ of degree greater than 2, or no vertex $w \neq v_0$ of degree greater than 2. In the next lemma, we find somewhat a condition that differentiates between these two main possibilities. We shall formulate it as part of a general result.

Let us first mention the following simple fact about the path P_n of order n, see for instance [11].

Lemma 2.3 *We have $N(P_n) = \binom{n+1}{2}$ and $N(P_n)_u = n$ if u is a pendant vertex of P_n.*

Lemma 2.4 *Let L, M, R be three vertex disjoint graphs such that $l \in V(L), m \in V(M)$ and $r \in V(R)$. From vertex m, draw a path of length $t+1 \geq 1$ and let w be the other pendant vertex of this path. Denote by H the resulting graph. Construct from H the two graphs G_1 and G_2 as follows:*

- *Identify both l, r with w to obtain the graph G_1; see Figure 3;*
- *Consider H: identify l with w, and r with m to obtain the graph G_2; see Figure 3.*

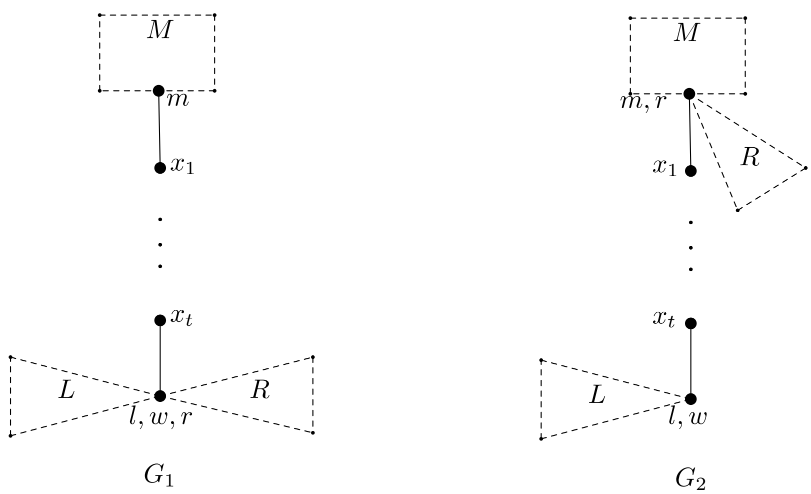

Figure 3. The graphs G_1 and G_2 described in Lemma 2.4.

Assume that $|V(R)| > 1$. Then $N(G_1) > N(G_2)$ if and only if $N(L)_l > N(M)_m$. Moreover, $N(G_1) = N(G_2)$ if and only if $N(L)_l = N(M)_m$.

Proof Categorise all subgraphs of G_1 and G_2 according to the subsets of $\{m, w\}$ that they contain as vertices. By grouping all 4 cases, i.e. getting expressions for $N(G_i - \{m\})_w, N(G_i - \{m, w\}), N(G_i - \{w\})_m, N(G_i)_{m,w}$, and using Lemma 2.3, we obtain

$$N(G_1) = N(L - \{l\}) + N(M - \{m\}) + N(R - \{r\}) + N(P_t) + (t+1)(N(M)_m + N(L)_l \cdot N(R)_r) + N(L)_l \cdot N(M)_m \cdot N(R)_r$$

for the number of connected subgraphs of G_1. Likewise, we have

$$N(G_2) = N(L - \{l\}) + N(M - \{m\}) + N(R - \{r\}) + N(P_t) + (t+1)(N(L)_l + N(M)_m \cdot N(R)_r) + N(L)_l \cdot N(M)_m \cdot N(R)_r$$

for the number of connected subgraphs of G_2. The difference $N(G_1) - N(G_2)$ is given by

$$N(G_1) - N(G_2) = (t+1)(N(R)_r - 1)(N(L)_l - N(M)_m)\,,$$

which proves the lemma, since $N(R)_r > 1$. □

The following lemma is also important for our analysis.

Lemma 2.5 *If v is a vertex of the cycle C_n then $N(C_n)_v = 1 + \binom{n}{2}$.*

Proof It was proved in [2] that the cycle C_n has $n^2 - n + 1$ connected subgraphs. Since deleting vertex v from C_n yields the path of order $n-1$, we get $N(C_n)_v = N(C_n) - N(P_{n-1})$. Thus, the proof of the lemma follows. □

For the rest of the paper, we always assume that $k \geq 2$ since there is only one unicyclic graph with at most one pendant vertex for every given order n and girth g: this graph corresponds to the specialisation $H = P_{n-g+1}$ (the path of order $n-g+1$) rooted at one of the pendant vertices of P_{n-g+1} in Figure 2.

An extended star is a tree in which all vertices have degree 1 or 2, except only one vertex w called central, which has degree greater than 2. Hence, if $H_{n,1,g,k}$ (see Figure 2) is an optimal graph from $\mathcal{G}(n,1,g,k)$, then H must be an extended star (provided that $k \geq 2$).

Let $H_{n,1,g,k}$ (see Figure 2) be an optimal graph in $\mathcal{G}(n,1,g,k)$. We have the following:

- If $\binom{g}{2} < \lfloor (n-g)/k \rfloor$, then H has precisely one vertex $w \neq v_0$ of degree greater than 2;
- If $\binom{g}{2} > \lfloor (n-g)/k \rfloor$, then H has no vertex $w \neq v_0$ of degree greater than 2;
- If $\binom{g}{2} = \lfloor (n-g)/k \rfloor$, then both possibilities are present for H.

By Proposition 2 alongside Lemma 2.2, for every optimal graph $H_{n,1,g,k}$ (see Figure 2) from $\mathcal{G}(n,1,g,k)$, the subgraph H of $H_{n,1,g,k}$ has either precisely one vertex $w \neq v_0$ of degree greater than 2, or no vertex $w \neq v_0$ of degree greater than 2.

- Assume that $\binom{g}{2} < \lfloor (n-g)/k \rfloor$. Suppose (for contradiction) that H has no vertex $w \neq v_0$ of degree greater than 2. Since H is a tree rooted at vertex v_0, let n_k be the maximum order among the k branches (actually paths) of H.

  We claim that $n_k \geq 2 + \binom{g}{2}$. To see this, simply note that if $n_k \leq 1 + \binom{g}{2}$, then

  $$n - g = |V(H - \{v_0\})| \leq k\Big(1 + \binom{g}{2}\Big)\,,$$

  which implies that $\binom{g}{2} \geq \lfloor (n-g)/k \rfloor$ (a contradiction). Therefore, $n_k \geq 2 + \binom{g}{2}$. Now take M to be the cycle of order g, L be the path of order n_k rooted at one of its pendant vertices l, and $t = 0$ as a specialisation in Lemma 2.4. Thus, $H_{n,1,g,k} = G_2$ in Figure 3 where R is the rest of H (note that $|V(R)| > 1$ as $k \geq 2$). Using Lemmas 2.3 and 2.5, we get

  $$n_k = N(L)_l > N(M)_m = 1 + \binom{g}{2}\,,$$

  which contradicts the optimality of $G_2 = H_{n,1,g,k}$ (see Lemma 2.4).

- Assume that $\binom{g}{2} > \lfloor (n-g)/k \rfloor$. Suppose (for contradiction) that H has one vertex $w \neq v_0$ of degree greater than 2. Denote by w_0 the unique neighbor of w that lies on the unique path from w to v_0 in H (it is possible to have $w_0 = v_0$), and $w_1, w_2, \ldots, w_k$ the other neighbors of w in H. Recall that H is a tree rooted at vertex v_0. For every $j \in \{1, 2, \ldots, k\}$, let L_j be the subtree of H consisting of w_j and all its descendents in H. Denote by $n_1 = |V(L_1)|$ the minimum order among the trees $L_1, L_2, \ldots, L_k$.
  We claim that $n_1 \leq \binom{g}{2} - 1$. To see this, simply note that if $n_1 \geq \binom{g}{2}$, then
  $$k\binom{g}{2} \leq |V(H - \{v_0\})| - 1 = n - g - 1\,,$$
  which implies that $\binom{g}{2} \leq \lfloor (n-g)/k \rfloor$ (a contradiction). Therefore, $n_1 \leq \binom{g}{2} - 1$. Now make the specialisation $M = C_g$ (the cycle of order g) and $L = P_{1+n_1}$ (the path of order $1 + n_1$) rooted at one of its pendant vertices l in Lemma 2.4. Thus, $H_{n,1,g,k} = G_1$ (for some $t \geq 0$) in Figure 3. Using Lemmas 2.3 and 2.5, we get
  $$1 + n_1 = N(L)_l < N(M)_m = 1 + \binom{g}{2}\,,$$
  which contradicts the optimality of $G_1 = H_{n,1,g,k}$ (see Lemma 2.4).

The proof that both possibilities are present for H in the case where $\binom{g}{2} = \lfloor (n-g)/k \rfloor$ makes use of Lemma 2.6 below. Indeed, let us consider the following scenario:

- Assume that $\binom{g}{2} = \lfloor (n-g)/k \rfloor$ and that k does not divide $n - g$.
  Suppose that H has no vertex $w \neq v_0$ of degree greater than 2, and let n_k be the maximum order among the k branches (actually paths) of H. Then $n_k = 1 + \lfloor (n-g)/k \rfloor = 1 + \binom{g}{2}$ using Lemma 2.6 below.

  Now take M to be the cycle of order g, L be the path of order n_k rooted at one of its pendant vertices l, and $t = 0$ as a specialisation in Lemma 2.4. Thus, $H_{n,1,g,k} = G_2$ in Figure 3 where R is the rest of H. Lemmas 2.3 and 2.5 give
  $$n_k = N(L)_l = N(M)_m = 1 + \binom{g}{2}\,,$$
  which shows that $N(G_2) = N(G_1)$, i.e. G_1 in Figure 3 is also an optimal graph.

We show in Figure 4 the two possibilities for the shape of every optimal graph from $\mathcal{G}(n, 1, g, k)$.

Proposition 2 raises the question for the number of vertices of each of the k paths $P_{n_1}, P_{n_2}, \ldots, P_{n_k}$, or $k+1$ paths $P_{n_0}, P_{n_1}, \ldots, P_{n_k}$ of H when H is an extended star (see Figure 4). Lemma 2.6 below determines, as a special case, the precise order of the paths $P_{n_1}, P_{n_2}, \ldots, P_{n_k}$.

Lemma 2.6 *Let $H_{l,r}$ be a graph obtained by identifying one pendant vertex of two vertex disjoint paths P_l and P_r with the same vertex z of another graph H. Assume that $|V(H)| > 1$ and $r \geq l \geq 1$. Then we have*
$$N(H_{l,r}) > N(H_{l-1,r+1})\,.$$

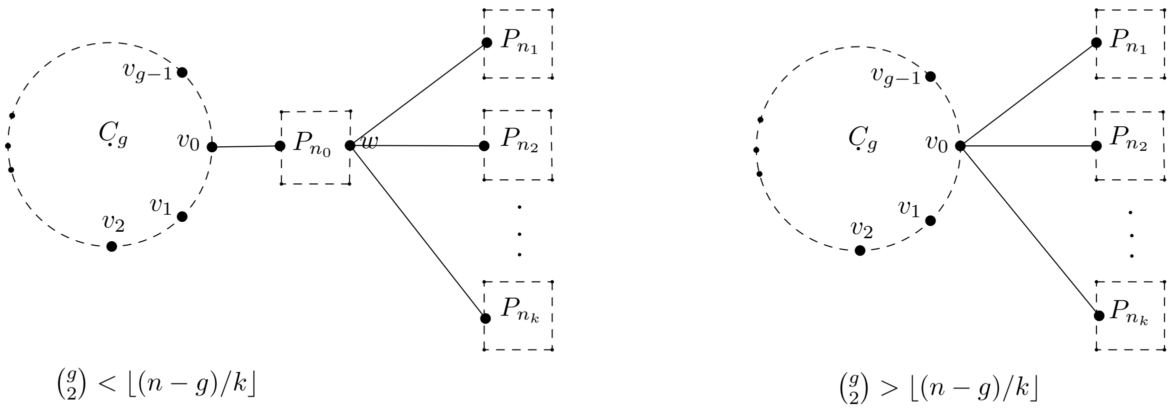

Figure 4. The two possibilities for the shape of every optimal graph in $\mathcal{G}(n,1,g,k)$ for $k \geq 2$ (see Proposition 2).

Proof Denote by u (resp. v) the fixed pendant vertex of P_l (resp. P_r) that is identified with vertex z of H. By distinguishing between subgraphs of $H_{l,r}$ that contain z and those that do not contain z, we obtain

$$N(H_{l,r}) = N(P_l)_u \cdot N(P_r)_v \cdot N(H)_z + N(P_l - \{u\}) + N(P_r - \{v\}) + N(H - \{z\})\,.$$

Using Lemma 2.3, we get

$$N(H_{l,r}) = l \cdot r \cdot N(H)_z + \binom{l}{2} + \binom{r}{2} + N(H - \{z\})$$

which implies that

$$N(H_{l,r}) - N(H_{l-1,r+1}) = (r - l + 1)(N(H)_z - 1) > 0\,.$$

□

Iterative application of Lemma 2.6 immediately shows that the order of all $k \geq 2$ paths $P_{n_1}, P_{n_2}, \ldots, P_{n_k}$ (Figure 4) must be as equal as possible. In particular, we know now the complete structure of every optimal graph for the case where H has no vertex $w \neq v_0$ of degree greater than 2; see Figure 4. It remains to determine the order of the path P_{n_0} for the case where H has precisely one vertex $w \neq v_0$ of degree greater than 2 (Figure 4). As it turns out, n_0 can only take on very few values.

Let $H_{n,1,g,k}$ (see Figure 2) be an optimal graph from $\mathcal{G}(n,1,g,k)$, where H is an extended star whose central vertex is w. Denote by $1 + n_0$ the order of the path joining v_0 to w in H (see Figure 4). Assume that $n_0 \notin \{1, n-g-k\}$. Then we have

$$\left\lceil \frac{n - n_0 - g}{k} \right\rceil - 1 - \binom{g}{2} \leq n_0 \leq 1 - \left\lfloor \frac{n - n_0 - g}{k} \right\rfloor - \binom{g}{2}.$$

In particular, the only possible values for n_0 are 1 and $n - g - k$.

Proof Let $H_{n,1,g,k}$ be as chosen in the statement of the proposition. Further, denote by $1+n_1, 1+n_2, \ldots, 1+n_k$ the order of the paths from w to each of the k leaves of H, respectively. Based on Figure 4, we first provide a formula for $N(H_{n,1,g,k})$. Using Lemma 2.3, we obtain

$$N(H_{n,1,g,k}) = N(C_g)_{v_0} \cdot \prod_{j=1}^{k}(1+n_j) + n_0 \cdot N(C_g)_{v_0} + n_0 \cdot \prod_{j=1}^{k}(1+n_j) + N(P_{g-1}) + N(P_{n_0-1}) + \sum_{j=1}^{k} N(P_{n_j}) \,.$$

Suppose (for contradiction) that

$$n_0 < \left\lceil \frac{n-n_0-g}{k} \right\rceil - 1 - \binom{g}{2} .$$

Construct from $H_{n,1,g,k}$ a new graph $H'_{n,1,g,k}$ obtained by replacing n_0 with $n'_0 = n_0+1$ and $n_k = \max_{1\leq j \leq k} n_j$ with $n'_k = n_k - 1$ (note that $n_k > 1$ as $k \neq n-g-n_0$). Thus, we have

$$N(H'_{n,1,g,k}) = N(C_g)_{v_0} \cdot (1+n'_k) \prod_{j=1}^{k-1}(1+n_j) + n'_0 \cdot N(C_g)_{v_0} + n'_0 \cdot (1+n'_k) \prod_{j=1}^{k-1}(1+n_j) + N(P_{g-1}) + N(P_{n'_0-1}) + N(P_{n'_k}) + \sum_{j=1}^{k-1} N(P_{n_j}) \,,$$

which implies that

$$N(H'_{n,1,g,k}) - N(H_{n,1,g,k}) = (n_k - n_0 - N(C_g)_{v_0}) \left(\prod_{j=1}^{k-1}(1+n_j) - 1 \right)$$

using Lemma 2.3 and after simplification. It follows from Lemma 2.5 that $N(H'_{n,1,g,k}) > N(H_{n,1,g,k})$, since

$$n_k = \max_{1\leq j\leq k} n_j = \left\lceil \frac{n-n_0-g}{k} \right\rceil > n_0 + 1 + \binom{g}{2} .$$

This contradicts the optimality of $H_{n,1,g,k}$. Likewise, suppose (for contradiction) that

$$n_0 > 1 - \left\lfloor \frac{n-n_0-g}{k} \right\rfloor - \binom{g}{2} .$$

Construct from $H_{n,1,g,k}$ a new graph $H''_{n,1,g,k}$ obtained by replacing n_0 with $n''_0 = n_0 - 1$ ($n_0 \geq 2$ by assumption) and $n_1 = \min_{1\leq j\leq k} n_j$ with $n''_1 = n_1 + 1$. Thus, we have

$$N(H''_{n,1,g,k}) = N(C_g)_{v_0} \cdot (1+n''_1) \prod_{j=2}^{k}(1+n_j) + n''_0 \cdot N(C_g)_{v_0} + n''_0 \cdot (1+n''_1) \prod_{j=2}^{k}(1+n_j) + N(P_{g-1}) + N(P_{n''_0-1}) + N(P_{n''_1}) + \sum_{j=2}^{k} N(P_{n_j}) \,,$$

which implies that

$$N(H''_{n,1,g,k}) - N(H_{n,1,g,k}) = (N(C_g)_{v_0} + n_0 + n_1 - 2)\left(\prod_{j=2}^{k}(1+n_j) - 1\right)$$

using Lemma 2.3 and after simplification. It follows from Lemma 2.5 that $N(H''_{n,1,g,k}) > N(H_{n,1,g,k})$ as

$$n_1 = \min_{1\le j\le k} n_j = \left\lfloor \frac{n-n_0-g}{k} \right\rfloor > 1 - n_0 - \binom{g}{2} = 2 - n_0 - N(C_g)_{v_0}\,.$$

We conclude that $n_0 \in \{1, n-g-k\}$. □

Next, we show that the situation $n_0 = n-g-k$ cannot occur; hence, $n_0 = 1$.

Let $H_{n,1,g,k}$ be an optimal graph (see Figure 2) from $\mathcal{G}(n,1,g,k)$ where H is an extended star whose central vertex is w (see Figure 4). Denote by $1+n_0$ the order of the path joining v_0 to w in H. Then $n_0 = 1$.

Proof Denote by $1+n_1, 1+n_2, \dots, 1+n_k$ the order of the paths from w to each of the k leaves of H, respectively. We know from Proposition 2 that $n_0 \in \{1, n-g-k\}$. Suppose (for contradiction) that $n_0 = n-g-k > 1$. Then we have $n_1 = n_2 = \cdots = n_k = 1$. Based on Figure 4 and using Lemma 2.3, we obtain

$$\begin{aligned} N(H_{n,1,g,k}) &= N(C_g)_{v_0}\cdot 2^k + (n-g-k)\cdot N(C_g)_{v_0} \\ &\quad + (n-g-k)\cdot 2^k + N(P_{g-1}) + N(P_{n-g-k-1}) + k\,. \end{aligned}$$

Replace $n_k = 1$ with $n'_k = 2$, and $n_0 = n-g-k$ with $n'_0 = n-g-k-1$ to obtain a new graph $H'_{n,1,g,k}$. It follows that

$$\begin{aligned} N(H'_{n,1,g,k}) &= N(C_g)_{v_0}\cdot 2^{k-1}(1+n'_k) + n'_0\cdot N(C_g)_{v_0} + n'_0\cdot 2^{k-1}(1+n'_k) \\ &\quad + N(P_{g-1}) + N(P_{n'_0-1}) + k - 1 + N(P_{n'_k}) \end{aligned}$$

and therefore, using Lemma 2.3, we get

$$N(H'_{n,1,g,k}) - N(H_{n,1,g,k}) = (2^{k-1}-1)(N(C_g)_{v_0} + n - g - k - 3)$$

after simplification. Hence, $N(H'_{n,1,g,k}) > N(H_{n,1,g,k})$ as $N(C_g)_{v_0} = 1 + \binom{g}{2}$ by Lemma 2.5. This contradicts the optimality of $H_{n,1,g,k}$. In view of Proposition 2, we conclude that $n_0 = 1$. □

Our main theorem can now be formulated as an immediate consequence of all that was discussed above:

Theorem 2.7 *Among all unicyclic graphs with order n, girth g, and $k \ge 2$ pendant vertices, the following hold:*

1. *If $\binom{g}{2} < \lfloor (n-g)/k \rfloor$, then the graph $O^1_{n,g,k}$ shown in Figure 5 uniquely realises the maximum number of connected subgraphs.*

2. *If $\binom{g}{2} > \lfloor (n-g)/k \rfloor$, then the graph $O^2_{n,g,k}$ shown in Figure 5 uniquely realises the maximum number of connected subgraphs.*

3. *If* $\binom{g}{2} = \lfloor (n-g)/k \rfloor$, *then the two graphs* $O^1_{n,g,k}$ *and* $O^2_{n,g,k}$ *shown in Figure 5 are the unique candidates that can realise the maximum number of connected subgraphs:*

$$N(O^2_{n,g,k}) = \big(1 + \binom{g}{2}\big)\Big(1 + \frac{n-g}{k}\Big)^k + \binom{g}{2} + k \cdot \binom{1 + \frac{n-g}{k}}{2}$$

for the case where k *divides* $n-g$.

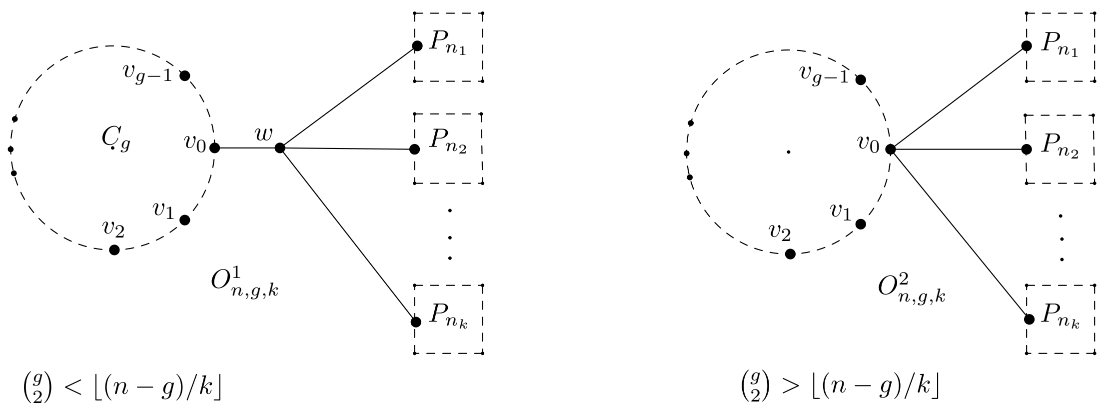

Figure 5. All optimal graphs in $\mathcal{G}(n,1,g,k)$: the values of $n_1, n_2, \ldots, n_k$ are all as equal as possible.

3. Ordering optimal graphs by number of pendant vertices

In this section, we find all unicyclic graphs having the maximum number of connected subgraphs given simultaneously order and girth. As a corollary to our results, we derive all unicyclic graphs having the maximum number of connected subgraphs given order only.

Theorem 3.1 *Let the order* n, *girth* g, *and number of pendant vertices* $k \geq 2$ *all be given. Then the following hold:*

1. *The graphs* $O^1_{n,g,k}$ *and* $O^1_{n,g,k+1}$ *(Figure 5) satisfy*

$$N(O^1_{n,g,k+1}) > N(O^1_{n,g,k})$$

provided that both graphs exist.

2. *The graphs* $O^2_{n,g,k}$ *and* $O^2_{n,g,k+1}$ *(Figure 5) satisfy*

$$N(O^2_{n,g,k+1}) > N(O^2_{n,g,k})$$

provided that both graphs exist.

In particular, the graph $O^2_{n,g,n-g}$ *uniquely realises the maximum number of connected subgraphs among all unicyclic graphs of order* n *and girth* g:

$$N(O^2_{n,g,n-g}) = \big(1 + \binom{g}{2}\big)2^{n-g} + \binom{g}{2} + n - g\,.$$

Proof We make use of the setup presented in Lemma 2.1.

1. By virtue of Theorem 2.7, both graphs $O^1_{n,g,k}$ and $O^1_{n,g,k+1}$ exist if $\binom{g}{2} < \lfloor (n-g)/(k+1) \rfloor$, or $\binom{g}{2} = \lfloor (n-g)/(k+1) \rfloor = \lfloor (n-g)/k \rfloor$. Also, we have $n-g-1 \geq k+2$ by assumption. Denote by $v_0, w_1, w_2, \ldots, w_k$ all $k+1$ neighbors of vertex w (see Figure 5) where v_0 is the vertex belonging to the cycle C_g. So without loss of generality, w_1 (resp. w_k) belongs to the path P_{n_1} (resp. P_{n_k}) and also $n_1, n_k > 1$. Let M be the graph obtained by deleting all vertices of the paths P_{n_1} and P_{n_k} except vertices w_1 and w_k. Further, let L and R be the rooted paths P_{n_1} and P_{n_k} rooted at vertices w_1 and w_k, respectively. Thus, by moving L to R, or R to L using the setup depicted in Lemma 2.1, we create precisely one more pendant vertex. Moreover, at least one of the moves increases the number of connected subgraphs of $O^1_{n,g,k}$. Hence, $N(O^1_{n,g,k+1}) > N(O^1_{n,g,k})$.

2. By virtue of Theorem 2.7, both graphs $O^2_{n,g,k}$ and $O^2_{n,g,k+1}$ exist $\binom{g}{2} > \lfloor (n-g)/k \rfloor$, or $\binom{g}{2} = \lfloor (n-g)/(k+1) \rfloor = \lfloor (n-g)/k \rfloor$. Also, we have $n-g \geq k+1$ by assumption. Denote by $v_{0,1}, v_{0,2}, \ldots, v_{0,k}$ all k neighbors of vertex v_0 (see Figure 5) that do not belong to the cycle C_g. So without loss of generality, $v_{0,1}$ (resp. $v_{0,k}$) belongs to the path P_{n_1} (resp. P_{n_k}).

   - Assume that $n-g \geq k+2$. Then without loss of generality, we have $n_1, n_k > 1$. Let M be the graph obtained by deleting all vertices of the paths P_{n_1} and P_{n_k} except vertices $v_{0,1}$ and $v_{0,k}$. Further, let L and R be the rooted paths P_{n_1} and P_{n_k} rooted at vertices $v_{0,1}$ and $v_{0,k}$, respectively. Thus, by moving L to R, or R to L using the setup depicted in Lemma 2.1, we create precisely one more pendant vertex. Moreover, at least one of the moves increases the number of connected subgraphs of $O^2_{n,g,k}$. Hence, $N(O^2_{n,g,k+1}) > N(O^2_{n,g,k})$.
   - Assume that $n-g = k+1$. We use direct calculations which yield

$$N(O^2_{n,g,k}) = N(C_g)_{v_0} \cdot 3 \cdot 2^{k-1} + N(P_{g-1}) + k + 2\,,$$
$$N(O^2_{n,g,k+1}) = N(C_g)_{v_0} \cdot 2^{k+1} + N(P_{g-1}) + k + 1\,.$$

   Hence, $N(O^2_{n,g,k+1}) > N(O^2_{n,g,k})$.

It follows that $O^2_{n,g,n-g}$ uniquely realises the maximum number of connected subgraphs among all unicyclic graphs with order n, girth g, and at least two pendant vertices, we have

$$N(O^2_{n,g,n-g}) = N(C_g)_{v_0} \cdot 2^{n-g} + N(P_{g-1}) + n - g\,.$$

However, it is easy to see that $N(O^2_{n,g,n-g})$ has superiority over the two unicyclic graphs with order n and girth g that have at most one pendant vertex. This completes the proof of the theorem. □

As a consequence of Theorem 3.1, we obtain:

Theorem 3.2 *Among all unicyclic graphs of order $n \geq 6$, precisely the graph $O^2_{n,3,n-3}$ maximises the number of connected subgraphs.*

Proof Assume $n \geq g+2$. Then we have

$$N(O^2_{n,g,n-g}) = \big(1 + \binom{g}{2}\big) \cdot 2^{n-g} + \binom{g}{2} + n - g\,,$$

$$N(O^2_{n,g+1,n-g-1}) = \big(1 + \binom{g+1}{2}\big) \cdot 2^{n-g-1} + \binom{g+1}{2} + n - g - 1\,.$$

Consequently, we get

$$N(O^2_{n,g,n-g}) - N(O^2_{n,g+1,n-g-1}) = (1 - g + \binom{g}{2})2^{n-g-1} - g + 1 > 0$$

provided that $g > 3$. On the other hand, if $n = g+1$ then $O^2_{n,g+1,n-g-1}$ is precisely the cycle C_{g+1} and so

$$N(O^2_{n,g,n-g}) - N(O^2_{n,g+1,n-g-1}) = \binom{g}{2} - 2g + 2 > 0$$

provided that $g > 3$. It is now easy to see that $O^2_{n,3,n-3}$ uniquely maximises the number of connected subgraphs if $n \geq 6$. □

An alternative proof of Theorem 3.2 was given in [2].

References

[1] Du Z, Zhou B, Trinajstić N. Sum-connectivity indices of trees and unicyclic graphs of fixed maximum degree. https://doi.org/10.48550/arXiv.1210.5043

[2] Dossou-Olory AAV. Graphs and unicyclic graphs with extremal connected subgraphs. Accepted for publication in Ars Combinatoria 2020. https://doi.org/10.48550/arXiv.1812.02422

[3] Gao J, Lu M. On the Randić index of unicyclic graphs. MATCH Communications in Mathematical and in Computer Chemistry 2005; 53 (2): 377-384.

[4] Graham FRK, Chung RL, Coppersmith D. On trees containing all small trees. The Theory and applications of graphs 1981; John Wiley & Sons, pp. 265-272.

[5] Hong Y, Liu H, Wu X. On the Wiener index of unicyclic graphs. Hacettepe Journal of Mathematics and Statistics 2011; 40 (1): 63-68.

[6] Jamison RE. Monotonicity of the mean order of subtrees. Journal of Combinatorial Theory, Series B 1984; 37 (1): 70-78.

[7] Jamison RE. On the average number of nodes in a subtree of a tree. Journal of Combinatorial Theory, Series B 1983; 35 (3): 207-223.

[8] Ou J. On extremal unicyclic molecular graphs with prescribed girth and minimal Hosoya index. Journal of Mathematical Chemistry 2007; 42 (3): 423-432.

[9] Pandey D, Patra KL. On the number of connected subgraphs of graphs. Indian Journal of Pure and Applied Mathematics. 2021; 52: 571-583.

[10] Székely LA, Wang H. Binary trees with the largest number of subtrees. Discrete Applied Mathematics 2007; 155 (3): 374-385.

[11] Székely LA, Wang H. On subtrees of trees. Advances in Applied Mathematics 2005; 34 (1): 138-155.

[12] Xia F, Chen S. Ordering unicyclic graphs with respect to Zagreb indices. MATCH Communications in Mathematical and in Computer Chemistry 2007; 58 (3): 663-673.