On Restrictions of Balanced 2-Interval Graphs

Philippe Gambette¹ and Stéphane Vialette²

¹ LIAFA - Univ. Paris VII gambette@liafa.jussieu.fr ² LRI - Univ. Paris XI vialette@lri.fr

Abstract. The class of 2-interval graphs has been introduced for modelling scheduling and allocation problems, and more recently for specific bioinformatics problems. Some of those applications imply restrictions on the 2-interval graphs, and justify the introduction of a hierarchy of subclasses of 2-interval graphs that generalize line graphs: balanced 2 interval graphs, unit 2-interval graphs, and (x,x) -interval graphs. We provide instances that show that all inclusions are strict. We extend the NP-completeness proof of recognizing 2-interval graphs to the recognition of balanced 2-interval graphs. Finally we give hints on the complexity of unit 2-interval graphs recognition, by studying relationships with other graph classes: proper circular-arc, quasi-line graphs, $K_{1,5}$ -free graphs, ...

Keywords: 2-interval graphs, graph classes, line graphs, quasi-line graphs, claw-free graphs, circular interval graphs, bioinformatics, scheduling.

1 2-Interval Graphs and Restrictions

The interval number of a graph, and the classes of k-interval graphs have been introduced as a generalization of the class of interval graphs by McGuigan [\[McG77\]](#page-10-0) in the context of scheduling and allocation problems. Recently, bioinformatics problems have renewed interest in the class of 2-interval graphs (each vertex is associated to a pair of disjoint intervals and edges denote intersection between two such pairs). Indeed, a pair of intervals can model two associated tasks in scheduling $[BYHN⁺06]$ $[BYHN⁺06]$, but also two similar segments of DNA in the context of DNA comparison [\[JMT92\]](#page-10-1), or two complementary segments of RNA for RNA secondary structure prediction and comparison [\[Via04\]](#page-10-2).

RNA (ribonucleic acid) are polymers of nucleotides linked in a chain through phosphodiester bonds. Unlike DNA, RNAs are usually single stranded, but many RNA molecules have *secondary structure* in which intramolecular loops are formed by complementary base pairing. RNA secondary structure is generally divided into helices (contiguous base pairs), and various kinds of loops (unpaired nucleotides surrounded by *helices*). The structural stability and function of noncoding RNA (ncRNA) genes are largely determined by the formation of stable secondary structures through complementary bases, and hence ncRNA genes

A. Brandstädt, D. Kratsch, and H. Müller (Eds.): WG 2007, LNCS 4769, pp. 55[–65,](#page-10-3) 2007.

⁻c Springer-Verlag Berlin Heidelberg 2007

Fig. 1. Helices in a RNA secondary structure (a) can be modeled as a set of balanced 2-intervals among all 2-intervals corresponding to complementary and inverted pairs of letter sequences (b), or as an independent subset in the balanced associated 2-interval graph (c).

across different species are most similar in the pattern of nucleotide complementarity rather than in the genomic sequence. This motivates the use of 2-intervals for modelling RNA secondary structures: each helix of the structure is modeled by a 2-interval. Moreover, the fact that these 2-intervals are usually required to be disjoint in the structure naturally suggests the use of 2-interval graphs. Furthermore, aiming at better modelling RNA secondary structures, it was suggested in [\[CHLV05\]](#page-9-1) to focus on balanced 2-interval sets (each 2-interval is composed of two equal length intervals) and their associated intersection graphs referred as balanced 2-interval graphs. Indeed, helices in RNA secondary structures are most of the time composed of equal length contiguous base pairs parts. To the best of our knowledge, nothing is known on the class of balanced 2-interval graphs.

Sharper restrictions have also been introduced in scheduling, where it is possible to consider tasks which all have the same duration, that is 2-interval whose intervals have the same length $[BYHN⁺06, Kar05]$ $[BYHN⁺06, Kar05]$. This motivates the study of the classes of unit 2-interval graphs, and (x, x) -interval graphs. In this paper, we consider these subclasses of interval graphs, and in particular we address the problem of recognizing them.

A graph $G = (V, E)$ of order *n* is a 2-interval graph if it is the intersection graph of a set of n unions of two disjoint intervals on the real line, that is each vertex corresponds to a union of two disjoint intervals $I^k = I^k_l \cup I^k_r$, $k \in [1, n]$ (l for "left" and r for "right"), and there is an edge between I^j and I^k iff $I^j \cap I^k \neq \emptyset$. Note that for the sake of simplicity we use the same letter to denote a vertex and its corresponding 2-interval. A set of 2-intervals corresponding to a graph G is called a realization of G. The set of all intervals, $\bigcup_{k=1}^n \{I_k^k, I_r^k\}$, is called the ground set of G (or the ground set of $\{I^1,\ldots,I^n\}$).

The class of 2-interval graphs is a generalization of interval graphs, and also contains all circular-arc graphs (intersection graphs of arcs of a circle), outerplanar graphs (have a planar embedding with all vertices around one of the faces [\[KW99\]](#page-10-5)), cubic graphs (maximum degree 3 [\[GW80\]](#page-10-6)), and line graphs (intersection graphs of edges of a graph).

Unfortunately, most classical graph combinatorial problems turn out to be NP-complete for 2-interval graphs: recognition [\[WS84\]](#page-10-7), maximum independent

Fig. 2. Graph classes related to 2-interval graphs and its restrictions. A class pointing towards another strictly contains it, and the dashed lines mean that there is no inclusion relationship between the two. Dark classes correspond to classes not yet present in the ISGCI Database [\[BLS](#page-9-2)⁺].

set [\[BNR96](#page-9-3)[,Via01\]](#page-10-8), coloration [\[Via01\]](#page-10-8), . . . Surprisingly enough, the complexity of the maximum clique problem for 2-interval graphs is still open (although it has been recently proven to be NP-complete for 3-interval graphs [\[BHLR07\]](#page-9-4)).

For practical application, restricted 2-interval graphs are needed. A 2-interval graph is balanced if it has a 2-interval realization in which each 2-interval is composed of two intervals of the same length [\[CHLV05\]](#page-9-1), unit if it has a 2 interval realization in which all intervals of the ground set have length 1 [\[BFV04\]](#page-9-5), and is called a (x, x) -interval graph if it has a 2-interval realization in which all intervals of the ground set are open, have integer endpoints, and length x [\[BYHN](#page-9-0)⁺06[,Kar05\]](#page-10-4). In the following sections, we will study those restrictions of 2-interval graphs, and their position in the hierarchy of graph classes illustrated in Figure [2.](#page-2-0)

Note that all (x, x) -interval graphs are unit 2-interval graphs, and that all unit 2-interval graphs are balanced 2-interval graphs. We can also notice that $(1, 1)$ -interval graphs are exactly line graphs: each interval of length 1 of the ground set can be considered as the vertex of a root graph and each 2-interval as an edge in the root graph. This implies for example that the coloration problem is also NP-complete for $(2, 2)$ -interval graphs and wider classes of graphs. It is also known that the complexity of the maximum independent set problem is NPcomplete on $(2, 2)$ -interval graphs [\[BNR96\]](#page-9-3). Recognition of $(1, 2)$ -union graphs, a related class (restriction of multitrack interval graphs), was also recently proven NP-complete [\[HK06\]](#page-10-9).

Fig. 3. The complete bipartite graph $K_{5,3}$ (a,b) has a balanced 2-interval realization (c): vertices of S_5 are associated to balanced 2-intervals of length 7, those of S_3 to balanced 2-intervals of length 11. Any realization of this graph is contiguous, *i.e.*, the union of all 2-intervals is an interval.

2 Useful Gadgets for 2-Interval Graphs and Restrictions

For proving hardness of recognizing 2-interval graphs, West and Shmoys consid-ered in [\[WS84\]](#page-10-7) the complete bipartite graph $K_{5,3}$ as a useful 2-interval gadget. Indeed, all its realizations are contiguous: for any realization, the union of all intervals in its ground set is an interval. Thus, by putting edges between some vertices of a $K_{5,3}$ and another vertex v, we can force one interval of the 2-interval v (or just one extremity of this interval) to be blocked inside the realization of $K_{5,3}$. It is easy to see that $K_{5,3}$ has a balanced 2-interval realization, for example the one in Figure [3.](#page-3-0)

However, $K_{5,3}$ is not a unit 2-interval graph. Indeed, each 2-interval $I = I_l \cup I_r$ corresponding to a degree 5 vertex intersect 5 disjoint 2-intervals, and hence one of I_l or I_r intersect at least 3 intervals, which is impossible for unit intervals. Therefore, we introduce the new gadget $K_{4,4} - e$ which is a $(2, 2)$ -interval graph with only contiguous realizations (the proof is omitted).

3 Balanced 2-Interval Graphs

We show in this section that the class of balanced 2-interval graphs is strictly included in the class of 2-interval graphs, and strictly contains circular-arc graphs. Moreover, we prove that recognizing balanced 2-interval graphs is as hard as recognizing (general) 2-interval graphs.

Property 1. The class of balanced 2-interval graphs is strictly included in the class of 2-interval graphs.

Fig. 4. The graph $K_{4,4} - e$ (a), a nicer representation (b), and a 2-interval realization with open intervals of length 2 (c)

Fig. 5. An example of unbalanced 2-interval graph (a) : any realization groups intervals of the seven $K_{5,3}$ in a block, and the chain of seven blocks creates six "holes" between them, which make it impossible to balance the lengths of the three 2-intervals I^1 , I^2 , and I^3

Proof. We build a 2-interval graph that has no balanced 2-interval realization. Let's consider a chain of gadgets $K_{5,3}$ (introduced in previous section) to which we add three vertices I^1 , I^2 , and I^3 as illustrated in Figure [5.](#page-4-0) In any realization, the presence of holes showed by crosses in the Figure gives the following inequalities for any realization: $l(I_l^2) < l(I_l^1)$, $l(I_l^3) < l(I_r^2)$, and $l(I_r^1) < l(I_r^3)$ (or if the realization of the chain of $K_{5,3}$ appears in the symmetrical order: $l(I_l^1) < l(I_l^3), l(I_r^3) < l(I_l^2),$ and $l(I_r^2) < l(I_r^1)$. If this realization was balanced, then we would have $l(I_l^1) = l(I_r^1) < l(I_r^3) = l(I_l^3) < l(I_r^2) = l(I_l^2)$ (or the symetrical equality): impossible! So this graph has no balanced 2-interval realization although it has a 2-interval generalization.

Theorem 1. Recognizing balanced 2-interval graphs is NP-complete.

Proof. We just adapt the proof of West and Shmoys [\[WS84](#page-10-7)[,GW95\]](#page-10-10): reduce the problem of Hamiltonian cycle in a 3-regular triangle-free graph to balanced 2 interval recognition.

Let $G = (V, E)$ be a 3-regular triangle-free graph. We build a graph G' which has a 2-interval realization (a special one, very specific, called H-representation and which we prove to be balanced) iff G has a Hamiltonian cycle. The construction of G , illustrated in Figure [6\(](#page-5-0)a) is almost identical to the one by West and Shmoys, so we just prove that G' has a balanced realization, shown in Figure [6](#page-5-0) (b), by computing lengths for each interval to ensure it. All $K_{5,3}$ have a balanced realization as shown in section [1](#page-0-0) of total length 79, in particular H_3 . We can thus affect length 83 to the intervals of v_0 . The intervals of the other v_i can have length 3, and their $M(v_i)$ length 79, so through the computation illustrated in

Fig. 6. There is a balanced 2-interval of G' (which has been dilated in the drawing to remain readable) iff there is an H-representation (that is a realization where the left intervals of all 2-intervals are contiguous) for its induced subgraph G iff there is a Hamiltonian cycle in G

Figure [6,](#page-5-0) intervals of z can have length $80 + 82 + 2(n-1) + 3$, that is $163 + 2n$. We dilate H_1 until a hole between two consecutive intervals of its S_3 can contain an interval of z, that is until the hole has length $165+2n$: so after this dilating, H_1 has length 79(165 + 2n). Finally if G has a Hamiltonian cycle, then we have found a balanced 2-interval realization of G of total length $13,273+241n$.

It is known that circular-arc graphs are 2-interval graphs, they are also balanced 2-interval.

Property 2. The class of circular-arc graphs is strictly included in the class of balanced 2-interval graphs.

Proof. The transformation is simple: if we have a circular-arc representation of a graph $G = (V, E)$, then we choose some point P of the circle. We partition V in $V_1 \cup V_2$, where P intersects all the arcs corresponding to vertices of V_1 and none of the arcs of the vertices of V_2 . Then we cut the circle at point P to map it to a line segment: every arc of V_2 becomes an interval, and every arc of V_1 becomes a 2-interval. To obtain a balanced realization we just cut in half the intervals of V_2 to obtain two intervals of equal length for each. And for each 2-interval $[g(I_l), d(I_l)] \cup [g(I_r), d(I_r)]$ of V_1 , as both intervals are located on one of the extremities of the realization, we can increase the length of the shortest so that it reaches the length of the longest without changing intersections with the other intervals. The inclusion is strict because $K_{2,3}$ is a balanced 2-interval graph (as a subgraph of $K_{5,3}$ for example) but is not a circular-arc graph (we can find two C_4 in $K_{2,3}$, and only one can be realized with a circular-arc representation).

4 Unit 2-Interval and (x,x)-Interval Graphs

Property 3. Let $x \in \mathbb{N}, x \geq 2$. The class of (x, x) -interval graphs is strictly included in the class of $(x + 1, x + 1)$ -interval graphs.

Proof. We first prove that an interval graph with a representation where all intervals have length k (and integer open bounds) has a representation where all intervals have length $k + 1$.

We use the following algorithm. Let S be initialized as the set of all intervals of length k, and let T be initially the empty set. As long as S is not empty, let $I = [a, b]$ be the left-most interval of S, remove from S each interval $[\alpha, \beta]$ such that $\alpha < b$ (including I), add $[\alpha, \beta + 1]$ to T, and translate by $+1$ all the remaining intervals in S . When S is empty, the intersection graph of T , where all intervals have length $k + 1$ is the same as the intersection graph for the original S.

We also build for each $x \geq 2$ a $(x+1, x+1)$ -interval graph which is not a (x, x) interval graph. We consider the bipartite graph K_{2x} and a perfect matching $\{(v_i, v'_i), i \in [\![1, x]\!] \}$. We call K'_x the graph obtained from K_{2x} with the following transformations, illustrated in Figure [7\(](#page-7-0)a): remove edges (v_i, v'_i) of the perfect matching, add four graphs $K_{4,4} - e$ called X_1, X_2, X_3, X_4 (for each X_i , we call v_l^i and v_r^i the vertices of degree 3), link v_r^2 and v_l^3 , link all v_i to v_r^1 and v_l^4 , link all v_i' to v_i^2 and v_i^3 , and finally add a vertex a (resp. b) linked to all v_i , v_i' , and to two adjacent vertices of X_1 (resp. X_4) of degree 4. We illustrate in Figure [7\(](#page-7-0)b) that K'_x has a realization with intervals of length $x + 1$. We can prove by induction on x that K'_x has no realization with intervals of length x: it is rather technical, so we just give the idea. Any realization of K'_x forces the block X_2 to share an extremity with the block X_3 , so each 2-interval v'_i has one interval intersecting the other extremity of X_2 , and the other intersecting the other extremity of X_3 . Then constraints on the position of vertices v_i force their intervals to appear as two "stairways" as shown in Figure [7\(](#page-7-0)b). So v_r^1 must contain x extremities of intervals which have to be different, so it must have length $x + 1$.

The complexity of recognizing unit 2-interval graphs and (x, x) -interval graphs remains open, however the following shows a relationship between those complexities.

Lemma 1. {*unit 2-interval graphs*} =
$$
\bigcup_{x \in \mathbb{N}^*} \{(x, x) \text{-interval graphs}\}.
$$

Proof. The \supset part is trivial. To prove \subset , let $G = (V, E)$ be a unit 2-interval graph. Then it has a realization with $|V| = n$ 2-intervals, that is 2n intervals of the ground set. So we consider the interval graph of the ground set, which is a unit interval graph. There is a linear time algorithm based on breadth-first search to compute a realization of such a graph where interval endpoints are rational, with denominator $2n$ [\[CKN](#page-9-6)⁺95]. So by dilating by a factor $2n$ such a realization, we obtain a realization of G where intervals of the ground set have length $2n$.

Fig. 7. The graph K_4' (a) is $(5,5)$ -interval but not $(4,4)$ -interval

Theorem 2. If recognizing (x, x) -interval graphs is polynomial for any integer x then recognizing unit 2-interval graphs is polynomial.

5 Investigating the Complexity of Unit 2-Interval Graphs

In this section we show that all proper circular-arc graphs (circular-arc graphs such that no arc is included in another in the representation) are unit 2-interval graphs, and we study a class of graphs which generalizes quasi-line graphs and contains unit 2-interval graphs.

Property 4. The class of proper circular-arc graphs is strictly included in the class of unit 2-interval graphs.

Proof. As in the proof of Property [2,](#page-5-1) we cut the circle of the representation of a proper circular-arc graph G to get a proper interval realization, which we transform into a unit interval realization [\[Rob69\]](#page-10-11), which provides a unit 2-interval representation of G . To complete the proof, we notice that the domino (two cycles C_4 having an edge in common) is a unit 2-interval graph but not a circular-arc graph.

Quasi-line graphs are those graphs whose vertices are bisimplicial, *i.e.*, the closed neighborhood of each vertex can be partitioned into two cliques. They have been introduced as a generalization of line graphs and a useful subclass of claw-free graphs [\[Ben81,](#page-9-7)[FFR97,](#page-10-12)[CS05](#page-9-8)[,KR07\]](#page-10-13). Following the example of quasiline graphs that generalize line graphs, we introduce here a new class of graphs for generalizing unit 2-interval graphs. Let $k \in \mathbb{N}^*$. A graph $G = (V, E)$ is all-ksimplicial if the neighborhood of each vertex $v \in V$ can be partitioned into at most k cliques (note that quasi-line graphs are exactly all-2-simplicial graphs).

Property 5. The class of unit 2-interval graphs is strictly included in the class of all-4-simplicial graphs.

Proof. The inclusion is trivial. We show that it is strict. Consider the following graph which is all-4-simplicial but not unit 2-interval: start with the cycle C_4 , call its vertices $v_i, i \in [1, 4]$, add four $K_{4,4} - e$ gadgets called X_i , and for each i we connect the vertex v_i to two connected vertices of degree 4 in X_i . This graph is certainly all-4-simplicial. But if we try to build a 2-interval realization of this graph, then each of the 2-intervals v_k has an interval trapped into the block X_k . So each 2-interval v_k has only one interval to realize the intersections with the other v_i : this is impossible as we have to realize a C_4 which has no interval representation.

Property 6. The class of claw-free graphs is not included in the class of all-4 simplicial graphs.

Proof. The Kneser Graph $KG(7, 2)$ is triangle-free, not 4-colorable [\[Lov78\]](#page-10-14). The graph obtained by adding an isolated vertex v and then taking the complement graph, *i.e.*, $KG(7,2) \cup \{v\}$, is claw-free as $KG(7,2)$ is triangle-free. And if it was all-4-simplicial, then the neighborhood of v in $KG(7, 2) \oplus \{v\}$, that is $KG(7, 2)$, would be a union of at most four cliques, so $KG(7, 2)$ would be 4-colorable: impossible so this graph is claw-free but not all-4-simplicial.

Property 7. The class of all-k-simplicial graphs is strictly included in the class of $K_{1,k+1}$ -free graphs.

Proof. If a graph G contains $K_{1,k+1}$, then it has a vertex with $k+1$ independent neighbors, and hence G is not all-k-simplicial. The wheel W_{2k+1} is a simple example of $K_{1,k+1}$ -free graph in which the center can not have its neighborhood (a C_{2k+1}) partitioned into k cliques or less.

Unfortunately, all-k-simplicial graphs do not have a nice structure which could help unit 2-interval graph recognition.

Theorem 3. Recognizing all-k-simplicial graphs is NP-complete for $k \geq 3$.

Proof. We reduce from GRAPH k -COLORABILITY, which is known to be NPcomplete for $k \geq 3$ [\[Kar72\]](#page-10-15). Let $G = (V, E)$ be a graph, and let G' be the complement graph of G to which we add a universal vertex v. We claim that G is k colorable iff G' is all-k-simplicial. If G is k-colorable, then the non-neighborhood of any vertex in G is k-colorable, so the neighborhood of any vertex in G is a union of at most k cliques. And the neighborhood of v is also a union of at most k cliques, so G' is all-k-simplicial. Conversely, if G' is all-k-simplicial, then in particular the neighborhood of v is a union of at most k cliques. Let's partition it into k vertex-disjoint cliques X_1, \ldots, X_k . Then, coloring G such that two vertices have the same color iff they are in the same X_i leads to a valid k-coloring of G.

6 Conclusion

Motivated by practical applications in scheduling and computational biology, we focused in this paper on balanced 2-interval graphs and unit 2-intervals graphs. Also, we introduced two natural new classes: (x, x) -interval graphs and all-ksimplicial graphs.

We mention here some directions for future works. First, the complexity of recognizing unit 2-interval graphs and (x, x) -interval graphs remains open. Second, the relationships between quasi-line graphs and subclasses of balanced 2 intervals graphs still have to be investigated. Last, since most problems remain NP-hard for balanced 2-interval graphs, there is a natural interest in investigating the complexity and approximation of classical optimization problems on unit 2-interval graphs and (x, x) -interval graphs.

Acknowledgments

We are grateful to Vincent Limouzy in particular for bringing to our attention the class of quasi-line graphs, and Michael Rao and Michel Habib for useful discussions.

References

- FFR97. Faudree, R., Flandrin, E., Ryjáček, Z.: Claw-free graphs - a survey. Discrete Mathematics 164, 87–147 (1997)
- GW80. Griggs, J.R., West, D.B.: Extremal values of the interval number of a graph. SIAM Journal on Algebraic and Discrete Methods 1, 1–7 (1980)
- GW95. Gyárfás, A., West, D.B.: Multitrack interval graphs. Congress Numerantium 109, 109–116 (1995)
- HK06. Halldórsson, M.M., Karlsson, R.K.: Strip graphs: Recognition and scheduling. In: Fomin, F.V. (ed.) WG 2006. LNCS, vol. 4271, pp. 137–146. Springer, Heidelberg (2006)
- JMT92. Joseph, D., Meidanis, J., Tiwari, P.: Determining DNA sequence similarity using maximum independent set algorithms for interval graphs. In: Nurmi, O., Ukkonen, E. (eds.) SWAT 1992. LNCS, vol. 621, pp. 326–337. Springer, Heidelberg (1992)
- Kar72. Karp, R.M.: Reducibility among combinatorial problems, pp. 85–103. Plenum Press, New York (1972)
- Kar05. Karlsson, R.: A survey of split intervals and related graphs, Manuscript (2005)
- KR07. King, A., Reed, B.: Bounding χ in terms of ω ans δ for quasi-line graphs. Article in preparation (2007)
- KW99. Kostochka, A.V., West, D.B.: Every outerplanar graph is the union of two interval graphs. Congress Numerantium 139, 5–8 (1999)
- Lov78. Lovász, L.: Kneser's conjecture, chromatic number, and homotopy. Journal of Combinatorial Theory Series A 25, 319–324 (1978)
- McG77. McGuigan, R.: Presentation at NSF-CBMS Conference at Colby College (1977)
- Rob69. Roberts, F.S.: Indifference graphs. In: Proof Techniques in Graph Theory, Proceedings of the Second Ann Arbor Graph Theory Conference, pp. 139– 146 (1969)
- Via01. Vialette, S.: Aspects algorithmiques de la prédiction des structures secondaires d'ARN. PhD thesis, Universit´e Paris 7 (2001)
- Via04. Vialette, S.: On the computational complexity of 2-interval pattern matching. Theoretical Computer Science 312(2-3), 223–249 (2004)
- WS84. West, D.B., Shmoys, D.B.: Recognizing graphs with fixed interval number is NP-complete. Discrete Applied Math. 8, 295–305 (1984)