

How to construct the symmetric cycle of length 5 using Hajós construction with an adapted Rank Genetic Algorithm

Juan Carlos García-Altamirano* Mika Olsen† Jorge Cervantes-Ojeda

¹ UAM CUAJIMALPA, México

revisions 21st Oct. 2022, 31st Jan. 2023; accepted 3rd Feb. 2023.

In 2020 Bang-Jensen et al. generalized the Hajós join of two graphs to the class of digraphs and generalized several results for vertex colorings in digraphs. Although, as a consequence of these results, a digraph can be obtained by Hajós constructions (directed Hajós join and identifying non-adjacent vertices), determining the Hajós constructions to obtain the digraph is a complex problem. In particular, Bang-Jensen et al. posed the problem of determining the Hajós operations to construct the symmetric 5-cycle from the complete symmetric digraph of order 3 using only Hajós constructions. We successfully adapted a rank-based genetic algorithm to solve this problem by the introduction of innovative recombination and mutation operators from graph theory. The Hajós Join became the recombination operator and the identification of independent vertices became the mutation operator. In this way, we were able to obtain a sequence of only 16 Hajós operations to construct the symmetric cycle of order 5.

Keywords: Rank-based genetic operators, Dichromatic number, Directed Hajós join.

1 Introduction

A genetic algorithm can solve mathematical problems that are intractable by exhaustive search, as is the case of the problem described here. Evolutionary algorithms have been applied to a wide variety of engineering problems, for instance: Dasgupta and Michalewicz (2013), and they have also been applied to mathematics problems: Jong and Spears (1989) showed that Genetic Algorithms (GA) can be used to solve NP-complete problems, Jakobs (1996) applied a GA to a geometry problem, Pourrajabian et al. (2013) solved nonlinear algebraic equations by using a GA, and Cervantes-Ojeda et al. (2019) applied a Rank Genetic Algorithm (Rank GA) to the graph theory problem of finding the rainbow connection number of a graph. Bokal et al. (2004) proved that for a given digraph D , deciding whether the dichromatic number of a digraph is at most 2 is NP-Complete. Therefore, our motivation is to try using the Rank GA heuristic on this

*Supported by Conacyt Grant 963921.

†Supported by Conacyt Project 47510664

problem. To our best knowledge, this problem has not been approached in this way nor using any other heuristic algorithms.

In a genetic algorithm, we have an initial population of individuals. In our case, each individual represents a digraph obtained by Hajós constructions of the complete symmetric digraph on 3 vertices $D(K_3)$. Each individual in the population is evaluated according to a fitness function. This function measures the ability of the individual for reaching a predetermined objective. Once individuals are evaluated, the next generation is obtained by applying genetic operators to the population inspired by the evolution in nature, such as cross-over between individuals, selection of the fittest individuals, and mutations. This procedure is repeated until the genetic algorithm finds an individual that achieves the required objective or until a maximum number of generations is reached.

A vertex coloring of a digraph is acyclic if there are no monochromatic directed cycles. The dichromatic number of a digraph D was introduced in Neumann-Lara (1982) as the minimum number of colors of an acyclic coloring of D , denoted by $dc(D)$. The dichromatic number of a digraph is an extension of the chromatic number of a graph and several concepts and results for the chromatic number of a graph have been extended to digraphs using the dichromatic number. For instance, perfect digraphs by Andres and Hochstättler (2015), dichromatic polynomial by González-Moreno et al. (2022), Brooks theorem for digraphs by Harutyunyan and Mohar (2011), bounds in terms of the girth of a graph by Cordero-Michel and Galeana-Sánchez (2018), flow theory by Hochstättler (2017), and diachromatic number by Araujo-Pardo et al. (2017). A digraph D is **r-critical** if $dc(D) = r$ and $dc(H) < k$ for every proper subdigraph H of D . Bang-Jensen et al. (2019) extended the well-known Hajós construction for graphs to digraphs: any r -critical digraph can be obtained by Hajós constructions using complete symmetric digraphs on r vertices, $D(K_r)$. Although the result was proved, it is not a trivial task to obtain even simple digraphs such as symmetric cycles of odd length. In particular, the authors left as an open problem how to construct the symmetric cycle $D(C_5)$ using directed Hajós joins and identifying non-adjacent vertices. Another interesting problem is constructing digraphs of minimum order for a given dichromatic number r . The 3-dichromatic tournaments of minimum order were characterized in Neumann-Lara (1994).

1.1 Definitions

We consider finite digraphs without loops and without multiple arcs. For all definitions not given here we refer the reader to the book of Chartrand, Lesniak and Zhang Chartrand et al. (2015). Let D be a digraph with vertex set $V(D)$ and arc set $A(D)$. The **in-neighborhood** of a vertex u is $N^-(u) = \{v \in V(D) \mid vu \in A(D)\}$ and the **out-neighborhood** of a vertex u is $N^+(u) = \{v \in V(D) \mid uv \in A(D)\}$. Two vertices in a digraph D are **independent** if there are no arcs between them in D , a set of vertices X is independent in a digraph D if any pair of vertices of X is independent in D . An arc $uv \in A(D)$ is **symmetric (asymmetric)** if $vu \in A(D)$ ($vu \notin A(D)$), and a digraph is symmetric (bidirected graph) if every arc of D is a symmetric arc. The **symmetric digraph** $D(G)$, of the graph G , is the digraph obtained by replacing each edge with a symmetric arc.

The Hajós construction was defined for digraphs in 2020 by Bang-Jensen et al. Bang-Jensen et al. (2019) as an extension of the well-known Hajós construction Hajós (1961); Jensen and Royle (1999); Urquhart (1997) for graphs. The class of **Hajós-k-constructible digraphs** is defined

as the smallest family of digraphs that contains all complete digraphs of order k and is closed under directed Hajós join and identifying independent vertices.

The digraph H obtained by **identifying** a non-empty set I of independent vertices is defined as the digraph $H = D - I$ adding a new vertex v and adding all arcs from v to $N_D^+(I) = \bigcup_{u \in I} N_D^+(u)$ and all arcs from $N_D^-(I) = \bigcup_{u \in I} N_D^-(u)$ to v . The new vertex v may preserve the label of one of the vertices of the independent set I .

We use Figure 1 to illustrate the definition of directed Hajós Join. Let D_1 and D_2 be two disjoint digraphs. Let $u_1v_1 \in A(D_1)$ and $v_2u_2 \in A(D_2)$. The **directed Hajós join** $D = (D_1, u_1, v_1) \nabla (D_2, v_2, u_2)$ or, briefly $D = D_1 \nabla D_2$ of D_1 and D_2 is defined as the disjoint union of D_1 and D_2 and deleting both arcs u_1v_1 and v_2u_2 , identifying the vertices v_1 and v_2 to a new vertex v and adding the arc u_1u_2 . The vertex v may be denoted by v_1, v_2 or v .

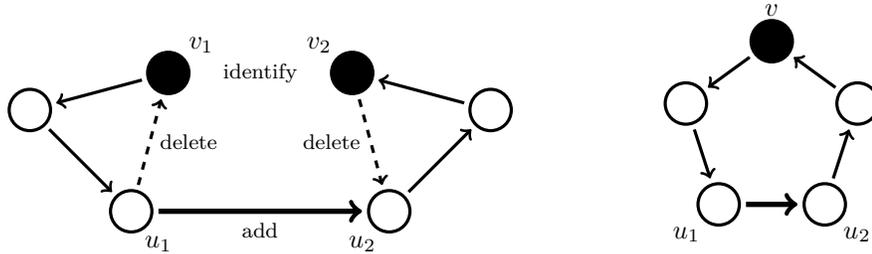


Fig. 1: The first digraph indicates the directed Hajós join of two directed triangles and the second is the resulting digraph.

2 The problem

Bang-Jensen et al. (2019) proved that any k -critical digraph is Hajós- k -constructible. In the same paper they posed the following question (Question 19):

How can a bidirected (symmetric) C_5 ($D(C_5)$) be constructed from copies of $D(K_3)$ by only using directed Hajós joins and identifying non-adjacent vertices?

To answer this question, we use an evolutionary algorithm that exclusively applies directed Hajós joins and identifications of non-adjacent vertices. The algorithm starts with a population that contains only copies of $D(K_3)$ and is guided using a fitness function whose optimum is in the digraph $D(C_5)$, which is the symmetric cycle of length 5. It is important to stress that as a consequence of Theorem 4 in Bang-Jensen et al. (2019), we already know that the symmetric cycle of order 5 can be obtained from $D(K_3)$ using only Hajós operations, but we do not know how. From the algorithm, we save the operations that are being carried out by the evolutionary algorithm to obtain each individual, in order to reconstruct the operations needed to obtain the optimum, which in this case is the digraph $D(C_5)$.

Observe that the digraphs in the populations obtained after applying Hajós constructions, may have different order and size.

3 Using a Genetic Algorithm

In this section, we describe how we use a genetic algorithm to solve our problem.

3.1 The Rank GA

The Rank GA is an algorithm that was first presented in Cervantes and Stephens (2009) as a solution to the limitations of existing mutation rate heuristics. It has been successfully applied to solve several problems, for instance Cervantes-Ojeda et al. (2019); López-Jaimes et al. (2018); Cervantes et al. (2010); Flores and Cervantes (2011). The claim is that this algorithm has a very good balance between exploration and exploitation by applying all genetic operations to the population based (or depending) on the fitness rank of each individual.

In the Rank GA, the individuals of the population are ranked from best to worst in terms of their fitness before the genetic operators (selection, recombination, and mutation) are applied. The application of these genetic operators depends on the rank of each individual in the population. The top ranked individuals tend to vary less than the bottom ranked ones. This is to make the latter try to escape from the local optima of the fitness function. Also, top ranked individuals tend to be cloned more than others who tend to disappear.

3.2 Adaptation to our problem

We use an adapted version of the Rank GA by Cervantes and Stephens (2009) to solve the problem. The adapted algorithm (Hajós RankGA) pseudocode is given in Algorithm 1 which we describe below. In that algorithm, the population size is equal to 50 and the number of generations has no limit. Here we give an overview of the features taken from Cervantes and Stephens (2009) and the modifications and adaptations made to that. For details, please consult that original paper.

Algorithm 1 Hajós Rank GA

```

procedure RANKGA()
  for each individual  $i$  do
    individual $_i$   $\leftarrow D(K_3)$ 
  Evaluation and Sort
  while not end Criteria met do
    rankRecombination()
    Evaluation and Sort
    rankMutation()
    Evaluation and Sort
    rankSelection() and Sort

```

Algorithm 2 rank Recombination

```

procedure RANKRECOMBINATION()
  for  $i$  in  $[0..individuals.size()-1]$  step 2 do
    arc1 = random(arcs of individual $_i$ )
    arc2 = random(arcs of individual $_{i+1}$ )
    individuals.add( Hajós( $i, (i + 1) \bmod individuals.size(), arc1, arc2$ ))

```

Algorithm 3 Hajós operation

```

procedure HAJÓS(INT  $i$ , INT  $j$ , ARC1, ARC2)
  result  $\leftarrow$  union(individual $_i$ , individual $_j$ )  $\triangleright$  disjoint union
  result.deleteArcs(arc1, arc2)
  identify(result, arc2.v2, arc1.v1)
  result.addArc(arc1.u1, arc2.u2)

```

Algorithm 4 identification operation

```

procedure IDENTIFY(INDIVIDUAL,  $u$ ,  $v$ )
  individual.vertex( $u$ ).addArcs(arcs of  $v$ )
  individual.delete( $v$ )

```

Algorithm 5 rank Mutation

```

procedure RANKMUTATION()
  for  $i$  in  $[0..2individuals.size()-1]$  do
     $r \leftarrow i / (2individuals.size-1)$ 
    nonAdjacentPairs  $\leftarrow$  binomial numVertices $_i, 2) - (asim_i + digon_i)$ 
    numAttempts  $\leftarrow$  ( $r$ )nonAdjacentPairs
    for  $j$  in  $[1..numAttempts]$  do
      select 1 random vertex  $u$  from individual $_i$ 
      if  $u$  has non adjacent vertices then
         $v =$  random(non adjacent vertices of  $u$ )
        identify(individual $_i, v, u$ )

```

3.2.1 Representation

Each individual in the population represents a digraph D by its adjacency matrix A , defined as follows

Let D be a digraph of order n with $V(D) = \{v_0, v_1, \dots, v_{n-1}\}$ being its set of vertices. The **adjacency matrix** of D is the matrix $A_{n \times n}$, where $A[i, j] = 1$ if $v_i v_j$ is an arc of the digraph D and $A[i, j] = 0$ otherwise.

Algorithm 6 rank Selection

```

procedure RANKSELECTION()
  clones ← null
  for  $i$  in  $[0..2\text{individuals.size}()-1]$  do
     $r \leftarrow i / (2\text{individuals.size}()-1)$ 
     $n \leftarrow \lfloor P(1-r)^{(2P-1)} \rfloor$ 
    for  $j$  in  $[0..n-1]$  do
      clones.add(individual $_i$ )           ▷ individual  $i$  is cloned  $n$  times
   $i \leftarrow 0$ 
  while clones.size() < individuals.size() do
     $r \leftarrow i / (2\text{individuals.size}()-1)$ 
     $n \leftarrow P(1-r)^{(2P-1)}$ 
     $f \leftarrow n - \lfloor n \rfloor$            ▷  $f$  is the fractional part of  $n$ 
    if random(0,1) <  $f$  then
      clones.add(individual $_i$ )         ▷ one extra clone of individual  $i$ 
     $i \leftarrow (i+1) \bmod 2\text{individuals.size}()$ 
  individuals ← clones                 ▷ replace population

```

3.2.2 Fitness Function

The fitness function to be minimized that we use is:

$$ft = |n - 5| + 2\frac{a}{n} + |n - s| + 15T_S + 5T, \quad (1)$$

where n is the order of the individual, a is the number of its asymmetric arcs, s is the number of its symmetric arcs, T_S is the number of its $D(K_3)$ subdigraphs and T is the number of its triangles with at least one non-symmetric arc.

The term $|n - 5|$ is introduced to favor those digraphs of order 5. The term $2\frac{a}{n}$ is introduced to favor digraphs with a low density of asymmetric arcs relative to their order. The term $|n - s|$ favors digraphs with the same number of vertices and symmetric arcs. The terms $15T_S$ and $5T$ are introduced to favor those digraphs with no triangles (symmetric and non-symmetric). Observe that if the fitness function is equal to 0, the digraph has 5 vertices, no asymmetric arcs, 5 symmetric arcs, and no triangles. The unique 3-dichromatic digraph that satisfies these conditions is the symmetric cycle of length 5, $D(C_5)$.

3.2.3 Recombination

Recombination is done based on the rank of the individuals (see Algorithm 2), so the population needs to be sorted by fitness value before any recombination. Mating is done in such a way that the best individual is recombined only with the second best. The third best is recombined only with the fourth best, etc.

We use the directed Hajós join between two digraphs as our Recombination operator.

Procedure *Hajós* in Algorithm 3 takes two individuals and one arc from each individual and returns the Hajós join between them. Procedure *identify* in Algorithm 4 takes an individual and

a pair of independent vertices and produces another individual that is the input individual with the identification operation done between the given vertices.

3.2.4 Mutation

The amount of mutation individuals get is a function of their rank (see Algorithm 5). The particular function to be used takes into account that the best individual should have 0 mutations and the worst individual should have the maximum number of mutations. First, we calculate the number of non-adjacent pairs of vertices and then we take a fraction of this to determine the number of mutation attempts to be done on the individual. This function is:

$$numAttempts = (r)numNonAdjacentPairs \quad (2)$$

where

$$r = \frac{i}{2popSize - 1} \quad (3)$$

where i is the index of the individual in the sorted population (starting at 0). Here we use $2popSize$ because after recombination we have doubled the population. A single mutation attempt selects one vertex randomly and, if it has non-adjacent vertices, one of those is selected randomly, or else the attempt is aborted. Once a pair of non-adjacent vertices was selected, the mutation is performed as the identification of these two (see Algorithm 4).

3.2.5 Selection

Rank selection defines that the number of clones of an individual is a function of its rank (see Algorithm 6). This function is monotonously decreasing and it is expected to produce a population of half the size of the previous population. The number of clones for $individual_i$ is:

$$clones = P(1 - r)^{2P-1} \quad (4)$$

where r is given by equation 3 and P is the selective pressure parameter that in our case is $P = 3$.

The integer part of $numClones$ determines a minimum number of clones for that individual, whereas the fractional part of it determines the probability that this individual gets one extra clone. The probabilistic part is performed after the integer part and only until the required population size is reached.

3.2.6 Stored information

Each population is stored and each individual is assigned its **origin** that describes how the individual is obtained in terms of the previous population. In case the individual was not modified, the origin indicates from which individual of the previous population it comes from, and in case the individual was modified, the origin indicates how it was obtained from one or two individuals from the previous population. Due to this information, we can reconstruct the operations that lead to the solution of the problem.

4 Results

The Hajós Rank GA used almost 5000 generations to obtain the symmetric 5-cycle, but we do not need 5000 Hajós operations. In each generation, the RankGA 1 stores the population and

the operations from which each individual was obtained, with this information we can recursively reconstruct the steps used by the algorithm to generate a particular individual. That is, we can reconstruct recursively each individual from $D(K_3)$ no matter the number of generations. In what follows, we describe the 16 Hajós operations we obtained from the algorithm to construct the symmetric 5-cycle from $D(K_3)$ using Hajós operations. Thus, we answer the question posed in Bang-Jensen et al. (2019).

The following sequence of directed Hajós joins and vertex identifications is the one obtained by the adapted version of the Hajós Rank GA:

The initial step is the directed Hajós join between two disjoint copies H and H' of $D(K_3)$, where $V(H) = \{v_0, v_1, v_2\}$ and $V(H') = \{v'_0, v'_1, v'_2\}$. In the left digraph in Figure 2, we indicate the arcs that must be deleted with a dotted line, and we indicate the vertices that we identify with a black background, and the right digraph in Figure 2 is the resulting digraph where v_0 is the vertex obtained by identifying the vertices v_0 and v'_0 , and with the vertices that correspond to the vertices of H' relabeled in cyclic order as v_3 and v_4 .

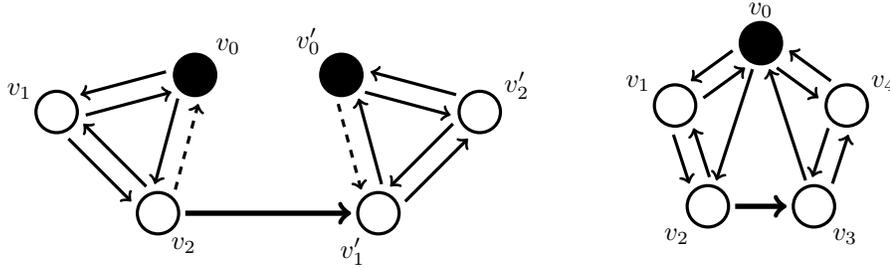


Fig. 2: The first digraph indicates the construction of the digraph $D_0 = (H, v_2, v_0) \nabla (H', v'_0, v'_1)$ using two copies of $D(K_3)$, and the second is the resulting digraph.

In each of the following steps, we consider 5 Hajós operations: a directed Hajós join and four vertex identifications. In Figures 2 to 5, the left digraph indicates the directed Hajós join, where the dotted arcs must be deleted, the thick arrow added and the black vertices identified, and the shades of gray indicate the four pairwise vertex identifications. The digraph on the right indicates the result of these operations.

The operation in Figure 3 uses as input the result in Figure 2, the operation in Figure 4 does so with the result in Figure 3 and the one in Figure 5 does so with the result in Figure 4.

Observe that the resulting digraph in Figures 2 and 3 is a symmetric 5-cycle with two diagonals, the resulting digraph in Figure 4 is a symmetric 5-cycle with only one diagonal and the resulting digraph in Figure 5 is a symmetric 5-cycle without diagonals.

5 Conclusions

Graph and digraph theory is a field of application of GAs that has been little explored, although there are examples where GAs have been applied successfully, for instance Bouazzi et al. (2021); Cervantes-Ojeda et al. (2019). The sequence of operations obtained by the GA inspired the

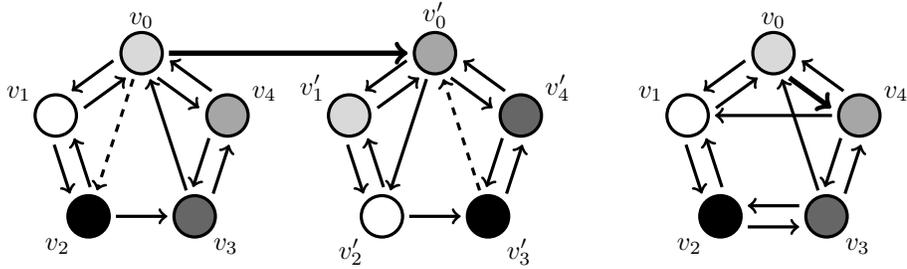


Fig. 3: The first digraph indicates the construction of the digraph $D_1 = (D_0, v_0, v_2) \nabla (D'_0, v'_3, v'_0)$ using two copies of D_0 , and the second is the resulting digraph.

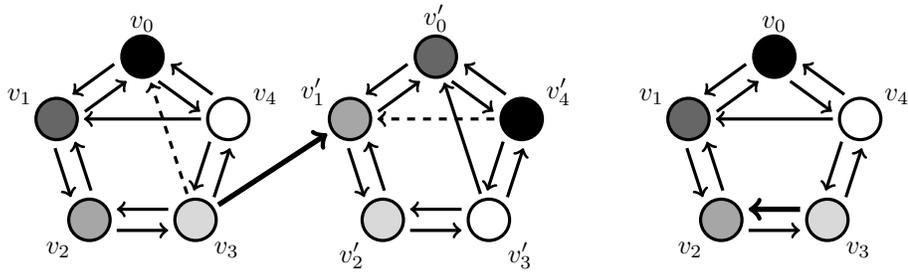


Fig. 4: The first digraph indicates the construction of the digraph $D_2 = (D_1, v_3, v_0) \nabla (D'_1, v'_4, v'_1)$ using two copies of D_1 , and the second is the resulting digraph.

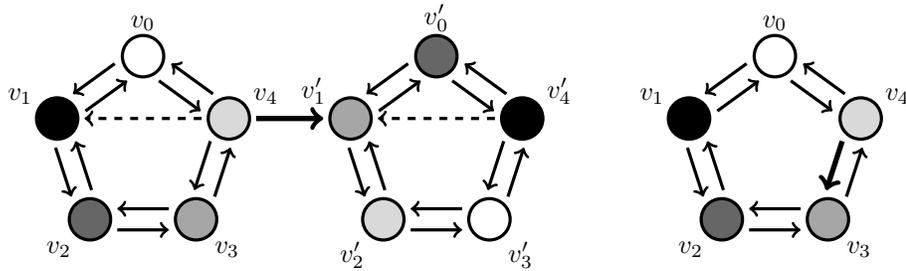


Fig. 5: The first digraph indicates the construction of the digraph $D_3 = (D_2, v_4, v_1) \nabla (D'_2, v'_4, v'_1)$ using two copies of D_2 , and the second is the resulting digraph.

authors to develop a method to obtain any symmetric odd cycle. The description of the operations of that method is out of the scope of this article, but it will be described in a forthcoming article by the authors. Thus, this article reiterates the viability of using genetic algorithms, with innovative ways for their adaptation, as an important heuristic in the development of graph theory.

Acknowledgment.

Research supported by CONACYT Project 47510664 and the Department of Applied Mathematics and Systems (DMAS) at Metropolitan Autonomous University (UAM) Cuaajimalpa, and Juan Carlos García-Altamirano was supported by Conacyt Grant 963921.

References

- S. D. Andres and W. Hochstättler. Perfect digraphs. *Journal of Graph Theory*, 79:21–29, 2015.
- G. Araujo-Pardo, J. J. Montellano-Ballesteros, M. Olsen, and C. Rubio-Montiel. The diachromatic number of digraphs. *Electron. J. Comb.*, 25:#P3.51, 2017.
- J. Bang-Jensen, T. Bellitto, T. Schweser, and M. Stiebitz. Hajós and ore constructions for digraphs. *Electron. J. Comb.*, 27:#P1.63, 2019.
- D. Bokal, G. Fijavz, M. Juvan, P. M. Kayll, and B. Mohar. The circular chromatic number of a digraph. *Journal of Graph Theory*, 46(3):227–240, 2004.
- K. Bouazzi, M. Hammami, and S. Bouamama. Application of an improved genetic algorithm to hamiltonian circuit problem. *Procedia Computer Science*, 192:4337–4347, 2021.
- J. Cervantes and C. R. Stephens. Limitations of existing mutation rate heuristics and how a rank ga overcomes them. *IEEE Transactions on Evolutionary Computation*, 13:369–397, 2009.
- J. Cervantes, M. Sánchez, and P. P. González. Emerging traits in the application of an evolutionary algorithm to a scalable bioinformatics problem. *IEEE Congress on Evolutionary Computation*, pages 1–8, 2010.
- J. Cervantes-Ojeda, M. del Carmen Gómez Fuentes, D. González-Moreno, and M. Olsen. Rainbow connectivity using a rank genetic algorithm: Moore cages with girth six. *J. Appl. Math.*, 2019:4073905:1–4073905:7, 2019.
- G. Chartrand, L. M. Lesniak, and P. Zhang. *Graphs & Digraphs, Sixth Edition*. Chapman and Hall/CRC., 2015.
- N. Cordero-Michel and H. Galeana-Sánchez. New bounds for the dichromatic number of a digraph. *Discret. Math. Theor. Comput. Sci.*, 21, 2018.
- D. Dasgupta and Z. Michalewicz. *Evolutionary Algorithms in Engineering Applications*. Springer, 2013.
- D. Flores and J. Cervantes. Rank based evolution of real parameters on noisy fitness functions: Evolving a robot neurocontroller. *2011 10th Mexican International Conference on Artificial Intelligence*, pages 72–76, 2011.
- D. González-Moreno, R. Hernández-Ortiz, B. Llano, and M. Olsen. The dichromatic polynomial of a digraph. *Graphs and Combinatorics*, 38(3):#85, 2022.

- G. Hajós. über eine konstruktion nicht n -färbbarer graphen. *Wiss. Z. Martin Luther Univ. Halle-Wittenberg, Math. Natur. Reihe*, 10:116–117, 1961.
- A. Harutyunyan and B. Mohar. Strengthened brooks' theorem for digraphs of girth at least three. *Electron. J. Comb.*, 18, 2011.
- W. Hochstättler. A flow theory for the dichromatic number. *European Journal of Combinatorics*, 66:160–167, 2017. Selected papers of EuroComb15.
- S. Jakobs. On genetic algorithms for the packing of polygons. *European Journal of Operational Research*, 88(1):165–181, 1996.
- T. R. Jensen and G. F. Royle. Hajós constructions of critical graphs. *Journal of Graph Theory*, 30(1):37–50, 1999.
- K. A. D. Jong and W. M. Spears. *Using Genetic Algorithms to Solve NP-Complete Problems*. IEEE, 1989.
- A. López-Jaimes, J. Cervantes-Ojeda, M. C. Gómez-Fuentes, and A. M. Alvarado-González. Simultaneous evolution of neuro-controller for multiple car-like robots. *Research in Computer Science*, 147(10):29–44, 2018.
- V. Neumann-Lara. The dichromatic number of a digraph. *Journal of Combinatorial Theory, Series B*, 33(3):265–270, 1982.
- V. Neumann-Lara. The 3 and 4-dichromatic tournaments of minimum order. *Discrete Mathematics*, 135(1):233–243, 1994.
- A. Pourrajabian, R. Ebrahimi, M. Mirzaei, and M. Shams. Applying genetic algorithms for solving nonlinear algebraic equations. *Applied Mathematics and Computation*, 219(24):11483–11494, 2013.
- A. Urquhart. The graph constructions of hajós and ore. *Journal of Graph Theory*, 26(4):211–215, 1997.