"A genetic algorithm and a local search procedure for a Hamiltonian path problem"

Eunice Ponce de León
Roberto Santana
eunice@cidet.icmf.inf.cu, rsantana@cidet.icmf.inf.cu
Centro de Inteligencia Artificial
Instituto de Cibernética, Matemática y Física
Calle 15, #551, entre C y D, Vedado
CP 10 400, C. Habana, CUBA

Abstract

This paper deals with Hável's conjecture that asserts that all bipartite graph formed with the middle levels of n-dimensional cube is Hamiltonian, when n is odd. This means that at least one Hamiltonian cycle should be found in the graph.

We use a circular representation of the vertices of n-dimensional cube which make it possible to introduce two group actions in order to reduce the problem to find out a Hamiltonian path in multi - level quotient graphs.

We report on a hybrid genetic algorithm for the Hável's conjecture problem. We present interesting results which show that this GA approach gives optimal solutions in multi-level quotient graph. We use bitstring representation, an evolutionary fitness function, restricted edge crossover operator, intelligent mutation, seeding and adaptive mutation rate. Three Hamiltonian path are constructed with this method.

Key words: Genetic algorithms, NP-complete, Hamiltonian circuit, cycle or path, Hável's conjecture, graph theory, local search procedure, hybrid methods.

1. Introduction

This paper deals with Hável's Conjecture that asserts that all bipartite graph formed with the middle levels of n-dimensional cube is Hamiltonian, when n is odd. This means that at least one Hamiltonian cycle should be found in the graph. An approach to Hável's Conjecture was reported by Dejter [1]. In these papers, he introduced the group action approach.

We use a circular representation of the vertices of n-dimensional cube which make it possible to introduce two group actions in order to reduce the problem to find out a Hamiltonian path in multi - level quotient graphs.

The *Hamiltonian path problem* (HPP) is one of the classic problems of combinatorial optimization. It consists of finding a path through a directed graph that touches all nodes exactly once. Clearly, if a graph is fully connected this is an easy task. However, as edges are removed the problem becomes much more difficult, and the general problem is known as NP- complete.

The contribution of this paper lies in two significant aspects:

1) We use the circular representation of the vertices of n-dimensional cube which make it possible to introduce two group actions. This result is discussed in a paper related to our research [4].
2) We introduce a hybrid genetic algorithm to solve the problem. It is possible to use this algorithm in any graph of the conjecture. An earlier version of this approach is presented in [3].

In this paper we report on a hybrid genetic algorithm for the Hável's conjecture problem. We present interesting results which show that this GA approach gives optimal solutions in multi-level quotient graph. We use bitstring representation, an evolutionary fitness function, restricted edge crossover operator, intelligent mutation, seeding and adaptive mutation rate. Three Hamiltonian path are constructed with this method.

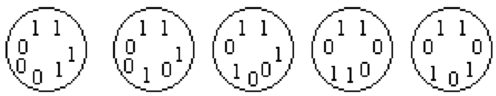
2. GA for Hamiltonian path problem in multi-level graphs

For this article a family of multi-level graph was used. Today, in terms of its complexity and structure, it constitutes a serious challenge in finding a Hamiltonian path. This graph family is the result of attempts to solve the conjecture of I. Havel, which asserts that the simple graph G(k+1) (k > 0) whose vertices are the subsets of cardinalities k and k+1 of the set {0,...,2k} and whose adjacency is given by subset inclusion is Hamiltonian [1]. As you can see the size of the search space grows exponentially when k is increased.

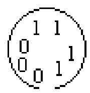
2.1 Multi - level graphs (H_{K+1}) [4]

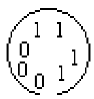
Let $H_{k+1} = (V_h E_h)$ be a graph whose vertex set is V_h and whose edge set is E_h. Each vertex $v \in V_h$ is representing a circular vector of Hamming weight k+1 that has no endpoints[4]. Let $|V_h|$ be a cardinal of set V_h, then $|V_h| = \binom{2k+1}{k} / (2k+1)$, which is a catalan number. This number represents all possible binary strings represented in circular form with exactly k coordinates in 0 and k+1 coordinates in 1.

For example: for k = 3, we have five vertices



The adjacency relation in H_{k+1} goes as follows:

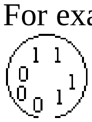
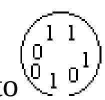
Consider for example the vertex  for H_4. Select a coordinate with value 1. There is an adjacent

vertex to  that can be seen as follows:

(1) write again 1 for the chosen coordinate

(2) write all the other values changed, for the remaining coordinates.

For example:

 is adjacent to  .

2.2 Levels of H_{K+1} graph

In [4] we introduced a natural classification of the vertices of V_h, according to the number of unitary and zero blocks of the circular vectors that they represent.

Definition 2.- A block is the longest sequence of consecutive coordinates of a circular vector that has the same value. If all the coordinates are 1 the block is said to be a unitary block. The same happens with the zero blocks.

The circular vectors (respectively vertices) with the same number of unitary and zero blocks constitute one level of the graph H_{k+1}. There is an even number of blocks. We consider one level of graph H_{k+1} as a partial subgraph of graph H_{k+1}. We denote it $H_{k+1,b}$. where b is the level. When b=1, we have the first level and this represent that there is one unitary block and one zero block. When b=2 we are in the second level. There are two unitary blocks and two zero blocks, and so on.

The GA works in each graph $H_{k+1,b}$ at all levels, in a separate way [3]. An algorithm that generates all vertices is given in [5].

2.3 GA components

Solution representation.- Candidate solutions are represented or encoded, as binary string or bitstrings. Each bitstring comprises M bit that mask the M edge in graph $H_{k+1,b}$: bit value 1 stands for "select the corresponding edge", whereas bit value 0 stands for "do not select the corresponding edge". This correspondence is given by adjacency matrix of graph $H_{k+1,b}$. If the graph have N vertices, then the bitstring have N bit put in 1. This bitstring represents a subgraph of the graph $H_{k+1,b}$. The size of the search space is $\binom{M}{N-1}$.

Objective function and fitness of a solution.- For calculating the objective function, first we count the degree of each vertex of the subgraph generated by the chromosome. This information is entered in the vector $W = (w_1,...,w_k)$. Let $2 = (2,...,2)$ the vector whose coordinates are all 2, then $D = | 2 - W |$ is the vector of difference coordinate to coordinate between 2 and W. Thus, objective function F is defined as $F = 1000 / D$.

When two coordinates of vector W are 1 and the rest zeros, F will reach its maximum value, that is 500. This function determines a necessary condition for the existence of the path, but this is not sufficient. If two vertices have a degree of one and the rest a degree of two in the chromosome, it starts to take on the form of a path. In this moment, the necessary condition can be only satisfied for those chromosomes that have cycles and a subpath or have a Hamiltonian path. We will allow a period between the beginning of the genetic algorithm and the appearance of this type of chromosome as the first evolutionary stage.

Once this quality among individuals is reached in the population it is then convenient to verify by the sufficient condition. This condition is given by a measure of the length of the path and of the cycles that are codified in the chromosome.

We suppose that we have p cycles and each one connects $t_1,...t_p$ vertices respectively and the subpath has the length L, then

$F = 500 + t_1 (t_1 - 1) + ... + t_p (t_p - 1) + L (L - 1)$,

as $t_1 + ... + t_p + L = k$, it can be seen that the maximum is reached for $L = k$.

A second evolutionary stage starts when the first chromosome appears that fulfills the necessary condition. This stage is characterized by the both conditions checking simultaneously. The evolutionary objective function was introduced in natural way at the start of the GA, the graphs generated by the chromosomes are very complex. Many vertices could have a degree greater than two. This gave us the possibility of evaluating the first stage with the necessary condition and afterwards a second stage with the sufficient condition. In this way the computational cost of the objective function is reduced.

Initial population: We tested two different initialization schemes. One is random and the other is seeding subpaths randomly in the initial population.

Restricted edge crossover operator.- This crossover follows from the fact that offsprings should preserve the edges that are in both its parents and they should not have those that

are not present in either of its parents. On the sites where the values of both parents coincide the offsprings will keep this value. In the other case (When parents sites have different values) one of the offsprings is chosen randomly and it is assigned 1 or 0 values alternatively. It can be seen that this crossover guarantees that the restriction is adhered to, since both offspring have the same number of sites for each value.
Intelligent Mutation.- To speed up the convergence a local optimization step is used. We defined it as an Intelligent Mutation. Thus we have a hybrid genetic algorithm .

In a similar way to that of the fitness function, one vector $W=(w_1,...,w_n)$ is formed, containing each vertex's degree. For those vectors that don't reach the necessary condition defined in the fitness function, the next steps are followed:
1) One component (w_i) is randomly chosen among those w_x ($1 \leq x \leq k$) for which $w_x<2$.
2) Another component (w_j) is randomly chosen among those w_x for which $w_x>2$.
3) One edge adjacent to i is incorporated to the edges set represented in the chromosome.
4) One edge adjacent to j is extracted from the edges set represented in the chromosome.

If the chromosome already satisfies the necessary condition but not obey the sufficient condition, the intelligent mutation works the following way: the components wi and wj are chosen such that wi=1, wj=2, and the same previous steps 3 and 4 are executed.

Selection operator.- In our GA each bitstring is chosen with a probability proportional to its fitness.

Software.- All software needed for this study was programmed in the computer language C++. The GA was programmed using the software library GAL [2].

3. Experimental results

Figure 1 show different runs of the GA plotting the fitness of the best objects as a function of the generation count. The values shown are for the average of 40 runs each with a different seed for random subpaths.

Common parameters for all runs:
- Seeding of the initial population
- Number of generation
- Adaptative mutation rate
- Population size
- Restricted edge crossover operator

The GA with the above mentioned parameters and crossover rate of 0.7 achieved the best performance. This combination in GA was used to obtain the Hamiltonian paths of the graphs H_6, H_7 and H_8.

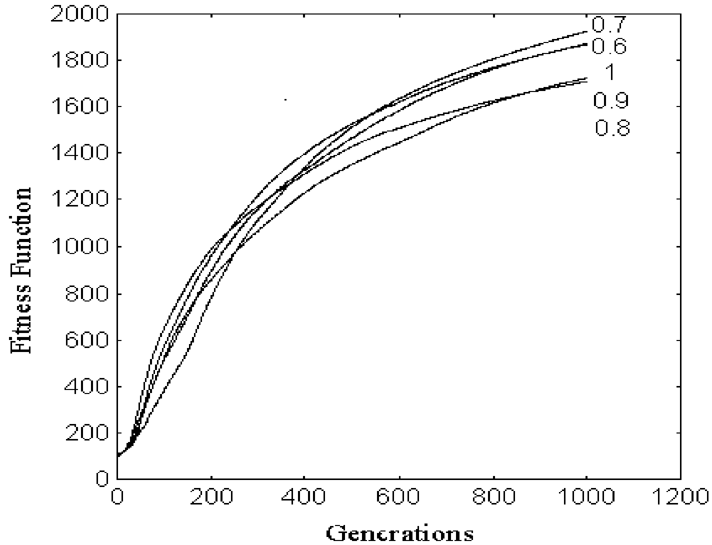


Figure 1.- The influence of the crossover rate in GHPA.

Conclusions

The formulation of the problem of finding a Hamiltonian path in multi-level graphs is presented. A new mutation operator that improve previous results is described. The GA was used to obtain the Hamiltonian paths of the graphs H_6, H_7 and H_8. The experimental results reveal the efficacy of the GA paradigm for the important problem of finding a Hamiltonian path in very complex graphs.

References

[1] Dejter, I.J. (1985): Hamilton cycles and quotients of bipartite graphs, in Y. Alavi et al. Eds., *Graph theory and its applications to algorithms and computer science* (J. Wiley, New York). 189-199.

[2] Ochoa, A.; Rossete, A. and Ponce de León, E. (1995): Programación de un algoritmo genético: un primer paso para la solución del problema. *Revista Ingeniería, Electrónica, Automática y Telecomunicaciones*. No. 2. Facultad Ingeniería Eléctrica. ISPJAE.

[3] Ponce de León, E.; Ochoa, A. and Santana, R. (1995): A genetic algorithm for a Hamiltonian path problem. *Reporte de Investigación ICIMAF*, 95 - 01, CENIA 95 - 1. ISSN 0138 - 8916.

[4] Ponce de León, E.(1997): La conjetura de Erdös y Havel: acciones de grupo, vectores circulares y grafos multiniveles. *Reporte de Investigación ICIMAF*, 97 - 39, CENIA 97 - 02. ISSN 0138 - 8916.

[5] Santana, R. and Ponce de León, E. (1996): Algoritmos genéticos para un problema de búsqueda de caminos Hamiltonianos. *Tesis de diploma*. Facultad de Matemática. Universidad de la Habana.