A Review of Literature for Solving the Maximum Independent Set Problem

Hayden D. Spinos

College of Graduate Studies, Georgia Southern University

CSCI 7501: Computational Intelligence

Dr. Weitian Tong

April 2, 2023

Literature Review

Introduction

Combinatorial optimization problems (COPs) deal with finding an optimal or near optimal solution given a finite set of feasible solutions. The goal in most COPs is to either maximize or minimize some metric that determines the overall quality of a solution. A well-known COP is the Maximum Independent Set Problem (MISP) that aims to find the largest subset of nodes from a given graph in which no two nodes in the solution are directly connected. While there are many methods for solving COPs, one popular method is using Evolutionary Algorithms (EAs) to produce quality solutions. This literature review will present a summary of various research studies done on COPs, the MISP, EAs, and the combinations of the three topics.

Combinatorial Optimization Problems:

As mentioned previously, the goal of COPs is to find the optimal solution from a finite set of feasible solutions. The difficulty however in COPs is due to many of them being classified as NP-hard which means that finding a definitive optimum in polynomial time has not yet been discovered. One of the most famous COPs is known as the Travelling Salesman Problem (TSP) which describes an environment consisting of n cities that a salesman must travel to and return home where the total distance traveled between cities along the tour is minimized. There have been many studies on the methodologies used to solve the TSP. One study done by Brezina and Čičková (2011) showed that by applying Ant Colony Optimization (ACO), acceptable solutions could be found in relatively low amounts of computation time which lended itself to be applicable in practical applications. Kumbharana and Pandey (2013) studied the effectiveness of the nature-inspired Firefly Algorithm in solving the TSP. They found that by discretizing the normally continuous-problem-oriented algorithm, they were able to achieve better results than ACO, EAs, and Simulated Annealing (SA). Since the patterns represented in CPOs can so often be mapped to real world applications, there is no shortage of literature on the many CPOs.

Maximum Independent Set Problem

The MISP is a well researched CPO and is the subject of this project's application. There are many applications of the MISP and its related problems. One example as described by Butenko (2003), involves solving the minimum connected dominating set problem in unit-disk graphs using approximation algorithms to construct a virtual backbone for wireless networks. A second example given by Butenko (2003) describes modeling massive data sets, such as financial market data, using graphs and applying principles from MISP to analyze various characteristics of the resulting graphs. As time passes, the applications for the MISP and other CPOs increase rapidly and continuously challenge researchers to propose new and faster methods of solving them. In 2017, Xiao and Nagamochi proposed a new algorithm using a measure-and-conquer method to achieve record computation times for exact solutions to the MISP. This is a notable feat as the difficulty of finding an exact solution for the MISP increases exponentially as the number of nodes increases.

Evolutionary Algorithms

Charles Darwin's theories on evolution provide immense inspiration in the world of optimization algorithms. These theories imply in a sense, that nature is the perfect optimizer with no limits to computational power. Populations of species evolve and their characteristics change to become more optimized for their environment. Bartz‑Beielstein et al. (2014) describe in their book the history of evolutionary algorithms and how researchers have attempted to model the natural phenomenon in computing since as early as the 1960s. A common form of EAs are Genetic Algorithms (GA). Forrest (1996) covers the principles of GAs in depth, however the premise is to model a population's genetics. Each individual, typically represented by a binary string, represents a single solution to the problem being solved. In each iteration (generation) of the algorithm, individuals are measured by a fitness function to determine their likelihood to reproduce and pass on their DNA. In their study published in 2014, Elsayed et al. proposed a new GA with a novel crossover and diversity operator that outperformed rival GAs and other "state-of-the-art algorithms" across multiple optimization problems. In recent years, the spotlight for solving complex computing problems has shifted more towards deep learning and neural networks, however EAs still provide more than acceptable solutions in reasonable time for many optimization problems.

Solving MISP with EAs

There have been many approaches to solving the MISP using various evolutionary methods and genetic algorithms. Back and Khuri (1994) applied a genetic algorithm to the MISP in which infeasible solutions were penalized in the fitness function. This represents a rather naive approach when compared to the study done by Javidi and Mehrabi in 2010 that guaranteed solutions to be feasible by implementing a simple greedy "repair" operator. If solutions became infeasible either during the crossover or mutation phase of the GA, they would be repaired greedily to overall improve the fitness of the population. As the industry's understanding of such problems increases, the technological barrier to entry drops as well. Fortin et al. (2012) highlight the DEAP (Distributed Evolutionary Algorithms in Python) framework that allows for rapid development and testing of evolutionary algorithms on all kinds of optimization problems using Python. The DEAP framework is still significant today as it will be used in this project's implementation.

Conclusion

There have been many studies done on the effectiveness of genetic algorithms in solving combinatorial optimization problems, the maximum independent set problem being a common topic. As the MISP still remains NP-hard, there still lies the challenge of finding an efficient method of solving the MISP especially for larger sized graphs. It has been shown that EAs are effective at quickly producing acceptable approximate solutions to the MISP under specific time constraints.

References

Back, T., & Khuri, S. (1994, June). An evolutionary heuristic for the maximum independent set problem. In Proceedings of the first IEEE conference on evolutionary computation. IEEE World Congress on Computational Intelligence (pp. 531-535). IEEE.

Bartz‐Beielstein, T., Branke, J., Mehnen, J., & Mersmann, O. (2014). Evolutionary algorithms. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(3), 178-195.

Brezina Jr, I., & Čičková, Z. (2011). Solving the travelling salesman problem using the ant colony optimization. Management Information Systems, 6(4), 10-14.

Butenko, S. (2003). Maximum independent set and related problems, with applications. University of Florida.

Elsayed, S. M., Sarker, R. A., & Essam, D. L. (2014). A new genetic algorithm for solving optimization problems. Engineering Applications of Artificial Intelligence, 27, 57-69.

Forrest, S. (1996). Genetic algorithms. ACM computing surveys (CSUR), 28(1), 77-80.

Fortin, F. A., De Rainville, F. M., Gardner, M. A. G., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. The Journal of Machine Learning Research, 13(1), 2171-2175.

Javidi, M. M., & Mehrabi, S. (2010). On evolutionary algorithms for maximum independent set problem. Journal of Artificial Intelligence: Theory and Application, 1(2), 54-59.

Kumbharana, S. N., & Pandey, G. M. (2013). Solving travelling salesman problem using firefly algorithm. International Journal for Research in science & advanced Technologies, 2(2), 53-57.

Xiao, M., & Nagamochi, H. (2017). Exact algorithms for maximum independent set. Information and Computation, 255, 126-146.