

Abstract No: SO-08

QUBO formulation of the closest string problem

D. M. C. Dissanayake^{1*}

¹Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka
chandeepadissanayake@gmail.com*

The closest string problem is an NP-complete problem which appears more commonly in bioinformatics and coding theory. From a theorist's point of view, more interesting applications are found in automata theory, specifically in the context of Finite Automata and Push-Down Automata. The corresponding optimisation problem deals with determining the string corresponding to the minimal Hamming distance between the candidate string and all given strings. Less surprisingly, classical approaches have been pursued with two prominent algorithms being the genetic algorithm and simulated annealing. Latest improvements to quantum computing devices with a specialisation in optimisation tasks, such as D-Wave systems, suggest that an attempt to embed the problem in a model accepted by such systems is worthwhile. In this study, the interests lie in modelling the problem based on binary variables. More specifically, two slightly different Quadratic Unconstrained Binary Optimization (QUBO) formulations have been proposed, one using the Kronecker delta function and the other through the help of a bijective function. The interest in the use of a bijective function for a formulation was inspired by the fact that the Kronecker delta function can incur additional computational overhead and the potential to eliminate it by exploiting the inherent numerical representation of symbols. Subsequently, an evaluation based on a few simple test cases was carried out on both formulations. Here, the D-Wave Leap cloud solvers have been utilized and minor-embedding is implicitly handled. For evaluation purposes, a metric called the Occurrence Ratio (OR), which is based on the number of occurrences in D-Wave output was defined. Additionally, the Maximum Occurrence Ratio (MOR) was defined as the maximum value obtained for OR for any solution. With minimal hyperparameter tuning, the expected solutions were obtained for every test case, where MOR was the same as OR, except for one case ($MOR \neq OR$). Since only the basic constraints have been followed to set the values for the Lagrange parameters in the Hamiltonian (QUBO), hyperparameter tuning is essential for, cases such as $MOR \neq OR$. To address practical and implementation issues, an inherent decomposition strategy based on the possibility of having substrings has been elucidated to accommodate the restricted qubit count. It is deduced theoretically, that the required qubit count is mn , where m is the number of symbols in a string and n is the number of strings, whereas due to practical limitations, this number is usually higher. Conclusively, the need for further investigation on tuning the hyperparameters is emphasised.

Keywords: Quadratic binary optimisation, QUBO models, Quantum annealing, Quantum algorithms, Combinatorial optimization