Literature Review

Raufs Dunamalijevs

Contents

[1 Introduction 3](#_Toc88135253)

[2 Annealing in Combinatorial Optimisation 4](#_Toc88135254)

[2.1 Annealing 4](#_Toc88135255)

[2.2 Simulated Annealing 4](#_Toc88135256)

[2.3 Quantum Annealing 5](#_Toc88135257)

[2.4 Digital Annealer 6](#_Toc88135258)

[3 QUBO 8](#_Toc88135259)

[3.1 QUBO Model 8](#_Toc88135260)

[3.2 Constraints and Penalties 9](#_Toc88135261)

[3.3 Natural QUBO Formulation 10](#_Toc88135262)

[3.3.1 Minimum Vertex Cover 10](#_Toc88135263)

[3.3.2 Example Solution 11](#_Toc88135264)

[3.4 Non-Natural QUBO Formulation and Slack Variables 13](#_Toc88135265)

[2.4.1 QUBO Formulation 13](#_Toc88135266)

[3.4.2 Slack Variables 13](#_Toc88135267)

[3.5 Algorithms for Solving QUBOs 14](#_Toc88135268)

[3.6.1 Simulated Annealing 14](#_Toc88135269)

[3.6.2 Tabu Search 14](#_Toc88135270)

[3.6.3 Decomposing Solver 15](#_Toc88135271)

[3.6.4 Steepest Descent Algorithm 15](#_Toc88135272)

[4 Penalty Coefficient Optimisation Techniques 16](#_Toc88135273)

[4.1 Analytical 16](#_Toc88135274)

[4.2 Numerical 1 16](#_Toc88135275)

[4.3 Numerical 2 17](#_Toc88135276)

[4.4 Machine Learning 18](#_Toc88135277)

[5 Summary 21](#_Toc88135278)

[References 22](#_Toc88135279)

List of Figures

[Figure 1: Comparison of node transitions in Simulated and Quantum Annealing (Johnson et al., 2011). 6](#_Toc88151167)

[Figure 2: Visual representation of normal subsets and a vertex cover. 11](#_Toc88151168)

[Figure 3: Example MVC graph. 12](#_Toc88151169)

[Figure 4: A solution to the example MVC. 13](#_Toc88151170)

[Figure 5: Solver Surrogate training (Huang et al., 2021). 19](#_Toc88151171)

[Figure 6: Estimating penalty coefficient with Solver Surrogate (Huang et al., 2021). 19](#_Toc88151172)

List of Tables

[Table 1: Few known constraint/penalty pairs (Glover, Kochenberger and Du, 2019). 10](#_Toc88151174)

# 1 Introduction

Introduction.

# 2 Annealing in Combinatorial Optimisation

In this section, we will discuss a class of metaheuristics inspired by the physical *annealing* process. In the beginning, we will explain what it is. Then we will talk about multiple annealing metaheuristics that are relevant to this work. Finally, we will look at hardware named Digital Annealer.

## 2.1 Annealing

Annealing is a heat treatment process of matter like glass or metal (Collins, 1921; Wu and Fan, 2020). It involves heating the material and then carefully cooling it down at a controlled rate. By following that process, it is possible to achieve a stable crystal structure.

If, however, this rate is exceeded, the resulting crystal will have defects as only locally optimal structures were allowed to form (Kirkpatrick, Gelatt and Vecchi, 1983). The lack of equilibrium will make such crystal metastable.

## 2.2 Simulated Annealing

Simulated annealing (SA) is a metaheuristic that uses the concept of temperature to approximate the global optimum of a combinatorial optimisation problem (Kirkpatrick, Gelatt and Vecchi, 1983; Bertsimas and Tsitsiklis, 1993) The temperature (*T*) is a variable that controls the likelihood of making locally non-optimal uphill moves. Just as in physical annealing, we start the optimisation with a high temperature and decrease it at a controlled rate. If that rate is low enough, we are guaranteed to find a global optimum (Liang, Cheng and Lin, 2014). However, it is possible to find good approximations of the optima significantly faster at higher cooling rates.

The temperature concept allows the algorithm to explore the search space and find the optimal solution (Poole and Mackworth, 2017). The SA starts with a high temperature. At that point, the algorithm acts like a random walk, choosing random moves with high probability, which prevents us from being stuck at local optima – this is when exploration happens. At the end of the SA, when the temperature is low, the algorithm prioritises advantageous moves. It acts like a greedy algorithm and tries to find the best solution. As we shift from higher to lower temperatures, we gradually change from exploring the search space to finding the optimum.

Let *c* be the current node, *n* be the next node (neighbour of node *c*) and *evaluation(x)* be the function that evaluates the energy of the node *x*. Then the energy transition (*ΔE*) can be defined by the following formula:

*ΔE = evaluation(n) - evaluation(c)*

*ΔE* and *T* are the variables that will affect the probability of choosing node *n*. *ΔE* is an energy transition (negative when moving downhill) that the proposed node will introduce, and *T* is controlling how *ΔE* influences the probability of choosing this node (Johnson *et al.*, 1989).

The *acceptance probability function* is the function that maps the current node, proposed node, and temperature to a probability. Many functions suit this purpose, the original being exponential (Kirkpatrick, Gelatt and Vecchi, 1983).

If that acceptance probability is higher than a uniformly generated random number, node *c* is rejected, favouring node *n*. However, when (downhill movement), we will always choose the proposed node (Kirkpatrick, Gelatt and Vecchi, 1983).

Pseudocode of SA (Johnson *et al.*, 1989) is given by:

c ← initial node

**while** T > threshold

n ← generate next node(c)

ΔE ← evaluate(n) – evaluate(c)

**if** ΔE ≤ 0 t

c ← n

**else if** P(c, n, T) ≥ random(0, 1)

c ← n

**end**

T ← cooling(T)

**end**

## 2.3 Quantum Annealing

Quantum annealing (QA) is a metaheuristic that uses quantum-mechanical fluctuations to find a global minimum of a function (Ohzeki and Nishimori, 2010; Tanaka and Tamura, 2012). In QA, we start with a strong quantum field and gradually decrease its strength (quantum fluctuation) to find an optimal ground state. It is analogous to SA, where we reduced the temperature (thermal fluctuation) to achieve the same result.

QA begins in a uniform superposition of all possible states, but as we decrease the strength of the quantum field, we approach a single solution state (Farhi *et al.*, 2001).

In SA, neighbouring nodes lie one move away from the current node. In QA, the neighbouring nodes are in a particular range. This range is defined by the strength of the quantum field, slowly ceasing in the process of annealing. The transition from one state to another can happen within that range. If there is a short wall between the current state, which happens to be a local minimum, and the global minimum, the SA algorithm will need to make a series of disadvantageous moves to reach a better solution or, in other words, *“climb”* over the wall. This is avoided in QA as there is a capacity to tunnel through short walls if the quantum field is strong enough, making QA faster than SA when dealing with search spaces with many thin barriers (Crosson and Harrow, 2016). This effect is called *quantum tunnelling*. A comparison of thermal hopping and quantum tunnelling can be seen in Figure 1.

![Shape, arrow

Description automatically generated]()

Figure 1: Comparison of node transitions in Simulated and Quantum Annealing (Johnson et al., 2011).

The fundamental feature and the main benefit of QA are that it rapidly samples a wide range of configurations when exploring the energy landscape (Higham and Bedford, 2021). Quantum Annealer has been developed and commercialised by D-wave (Johnson *et al.*, 2011) and is currently used to solve real-world problems (D-Wave Systems, 2021).

## 2.4 Digital Annealer

Digital Annealer (DA) is quantum-inspired custom CMOS hardware developed by Fujitsu Laboratories (Aramon *et al.*, 2019). Its algorithm is based on SA but has three significant differences:

1. Instead of considering a single neighbouring node and choosing it if the acceptance probability is higher than a uniform random number, DA looks at all neighbouring nodes in parallel. If more than a single node were accepted, one of them would be chosen uniformly at random. This is advantageous as in SA, if the node was rejected, we need to generate and then consider a new one from the same position, which takes time. In DA, however, all the nodes are considered in parallel, and the acceptance probability per “turn” is higher.
2. DA uses *dynamic offset*, which means that if no nodes have been accepted, the following acceptance probability will be artificially increased to help the algorithm overcome narrow barriers.
3. All runs begin with the same random state to save time (in parallel SA the initial states are generated individually for each run). Otherwise, it would need to evaluate the new initial node and the neighbouring nodes for each run.

These differences can be observed in DA pseudocode (Aramon *et al.*, 2019):

initial state ← random state

**for** each run

current state ← initial state

Eoffset ← 0

**for** each step

update temperature if needed

**for** each neighbour *(in parallel)*

calculate ∆Eneighbour

consider neighbour using ∆Eneighbour - ∆Eoffset

if neighbour accepted, record

**end**

**if** accepted neighbours ≥ 1

choose one accepted neighbour uniformly at random

update current state and the energy landscape *(in parallel)*

Eoffset ← 0

**else**

increase Eoffset

**end**

**end**

**end**

Although it is hard to make comparisons as the technology is quickly developing, it has been shown that DA can be faster and produce more precise solutions compared to D-Wave 2000Q (penultimate generation of D-Wave’s QA devices) in certain scenarios (Ohzeki *et al.*, 2019). DA’s are used in industry and science (Fujitsu, no date; Snelling *et al.*, 2020)

Currently, QA devices are expensive and challenging to run as the technology is still under development (Şeker *et al.*, 2020). D-Wave 2000Q also suffers from qubit noise that results in lower precision (Katzgraber and Novotny, 2018) and sparse connectivity between spins that makes it less effective (Hamerly *et al.*, 2018). These problems are actively worked on, and some are addressed in more recent QA devices (Dattani, Szalay and Chancellor, 2019). As QA technology develops, it will likely outperform DA primarily due to its quantum parallelism (Boyd, 2018).

# 3 QUBO

Quadratic Binary Unconstrained Optimisation (QUBO) is a mathematical formulation that can be applied to many combinatorial optimisation (CO) problems (Kochenberger *et al.*, 2014). Problems formulated in QUBO can be subsequently solved using QUBO solvers. QA devices developed by D-Wave can solve CO problems formulated as QUBO or Ising models (Cruz-Santos, Venegas-Andraca and Lanzagorta, 2019). It has been shown that QUBO and Ising models are equivalent (Lucas, 2014); therefore, one can be derived from another. DA’s can only solve problems that are formulated in QUBO (Aramon *et al.*, 2019).

Many combinatorial optimisation problems from industry, government, and science can be reformulated in QUBO (Glover, Kochenberger and Du, 2019).

This section will discuss the structure of QUBO, the basics of reformulating constrained CO problems into equivalent unconstrained QUBO models, algorithms that can be used to solve these problems, and relevant programming libraries that can be useful when dealing with QUBO’s. We will introduce the concept of penalty, which will be better explained in the next section.

## 3.1 QUBO Model

QUBO models have been described in detail by Glover, Kochenberger and Du (2019). This subsection will summarise the essential information relevant to this project.

The QUBO model can be expressed in the following way:

Where *y* is the value to be optimised for, *x* is a vector of decision variables, and *Q* is a square matrix with coefficients. QUBO models are unconstrainted; the only restriction is that every variable in decision vector *x* should be 0 or 1. It is self-contained as all the information needed for the optimisation is stored in the matrix Q. QUBO problems are NP-hard.

A simple example of how to convert a Boolean function into QUBO can be demonstrated with the following minimisation problem:

Our function has two parts: linear and quadratic. As every *xj* belongs to {0, 1}, we can easily make the linear part quadratic (because 02=0 and 12=1), making the entire function quadratic.

We can then write it using matrixes in the form of , a QUBO model formula shown at the beginning.

## 3.2 Constraints and Penalties

In the previous subsection, it has been shown how an unconstrained minimisation problem can be reformulated as a QUBO. Constrained problems can be reformulated in this model too. The constraints will be expressed as quadratic penalties, which will be added to the original objective function (Glover, Kochenberger and Du, 2019). The more constraints are broken, the more significant is the overall penalty imposed. When none of the constraints is broken, the penalty imposed will be equal to 0.

Since we are trying to solve a minimisation problem, solutions that break constraints and impose a penalty on the objective function will be avoided.

*M* is a positive penalty scalar (Verma and Lewis, 2020). It controls the effect that the broken constraints will have on the objective function. If we have a soft constraint that is allowed to be broken, the penalty scalar can be decreased, which will reduce the effect that the broken constraint will have on our function. It is possible to use multiple scalars for various constraints if they are of different importance. But usually, only a single *M* is used (Glover, Kochenberger and Du, 2019).

If the penalty coefficient is too low, the broken constraints will be undervalued, and the solution produced by the optimiser will be infeasible. On the other hand, if the penalty coefficient is too large, the solution process will be negatively impacted as the penalties will overwhelm the objective function making it harder to differentiate between good and bad solutions (Glover, Kochenberger and Du, 2019).

It is possible to use domain expertise to select an acceptable *M* in some cases (Glover, Kochenberger and Du, 2019), but it has also been shown that some problems can have solutions of better quality if you choose penalty coefficient carefully (Şeker, Tanoumand and Bodur, 2020). Finding an optimal penalty coefficient value is not trivial and different techniques have been proposed to estimate it (Verma and Lewis, 2020; Huang et al., 2021).

Quadratic penalties for certain constraint types are already known. Some of them are shown in Table 1.

|  |  |
| --- | --- |
| Classical Constraint | Equivalent Penalty |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Table 1: Few known constraint/penalty pairs (Glover, Kochenberger and Du, 2019).

Consider a problem where x1 or x2 should be equal to 1, but they cannot be identical. This constraint can be described with the following formula:

Using the penalty equivalent to our constraint, we can express our objective function in the following way:

Where f(x) is the original objective function. We can see that the penalty is imposed when both decision variables are equal to 0 or 1 when the constraint is broken.

## 3.3 Natural QUBO Formulation

There exist combinatorial optimisation problems that can be naturally expressed as QUBO models. Some of them have been described by Glover, Kochenberger and Du (2019). One of such problems will be summarised in this subsection using their work to demonstrate the ‘*natural*’ QUBO formulation process.

### 3.3.1 Minimum Vertex Cover

Minimum Vertex Cover (MVC) is a CO problem, where we are given an undirected graph with vertexes and edges represented by sets *V* and *E,* respectively. A single vertex ‘covers’ the edges that it is incident to. *Vertex cover* is a subset of vertexes *V*, which covers all the edges *E*. Figure 2 compares vertex cover and standard subsets of *V*. In MVC, we aim to find a vertex cover with the minimum number of vertices.

Diagram, schematic

Description automatically generated

Figure 2: Visual representation of normal subsets and a vertex cover.

The decision vector *x* can be expressed in the following way. If vertex *j* is in the vertex cover, then is xj = 1. Otherwise, xj = 0. Then our minimisation function, before we consider any constraints, can be defined as the sum of all vertices in the cover.

To get a feasible solution, we need to cover all the edges. This can be expressed as the following constraint: every edge should have at least one vertex that belongs to the vertex cover. If it does not, the edge under consideration is not covered and, therefore, not all the edges are covered. This is infeasible. Thus, the constraints are:

Table 1 contains a penalty equivalent to this type of constraint. Using it, we can make the following minimisation function:

This formula can be expressed as , where *c* – is a constant that has no influence on the optimisation process and can be removed. This leaves us with an unconstrained QUBO model.

### 3.3.2 Example Solution

We are given the following graph, which we need to find an MVC for:

Schematic

Description automatically generated with low confidence

Figure 3: Example MVC graph.

Using the formula derived previously, we can express and then transform the problem in the following way:

If we remove the constant *6M*, which does not affect the optimisation (we can add it after the optimisation has finished if needed), from the end of the formula, we will get the following QUBO:

We can assign an arbitrary value of 8 to the *M* to get the following solution:

A close-up of a stethoscope

Description automatically generated with medium confidence

Figure 4: A solution to the example MVC.

## 3.4 Non-Natural QUBO Formulation and Slack Variables

Not all CO problems have a natural QUBO formulation. Also, quadratic penalties for some constraints might be unknown. Such problems can be formulated in QUBO too. In the first part of this subsection, we will show a general approach to QUBO formulation. The second part will show how any inequality constraint can be expressed as a quadratic penalty by introducing slack variables. Both processes have been accurately described by Glover, Kochenberger and Du (2019).

### 2.4.1 QUBO Formulation

Consider a constrained minimisation problem. Without quadratic penalties, it will have the following form:

Where *x* is a binary decision vector and *C* is a square matrix. If the constraints of the problem consist of integers, they can always be written as (this will be explained in more detail in the subsection devoted to slack variables):

Then the quadratic penalties that we need to add to the original objective function are:

And the resulting objective function with constraints will be:

Constant *c* can be removed as it does not affect the optimisation (it can be added after a solution is found if needed). That is how we get a standard QUBO model out of a general constrained problem that has no natural QUBO formulation.

### 3.4.2 Slack Variables

Consider the following inequality constraint:

It can be transferred into an equality constraint by adding a slack variable, *s.* The slack variable cannot be bigger than 7 because the lowest number that the left side of the inequality takes is -1 with (0,0,1) configuration. It also cannot be smaller than -3 as the maximum of the left side is 9 with (1,1,0) configuration.

If we express *s* using binary expansion, which will satisfy all configurations of decision variable *x*, we will get the following equality:

As both *xj* and *sj* are binary at this point, they can both be contained in a single vector: *x*. And all the coefficients can then be held in a single matrix *A*.

Then, as shown in the previous part of this subsection, this constraint can be expressed as a quadratic penalty and subsequently used to achieve a standard QUBO model.

## 3.5 Algorithms for Solving QUBOs

Many algorithms can effectively solve problems in QUBO formulation. A non-comprehensive list of them was drawn up by Kochenberger and Glover (2006). We have already described quantum annealing for quantum devices and the digital annealer algorithm for classical devices in section 1. In this subsection, we will briefly describe algorithms that are available in programming libraries dedicated to solving QUBO models.

### 3.6.1 Simulated Annealing

This algorithm was already described in section 1. When solving QUBO models, neighbouring nodes are located *n* bit-flips away from the current node, where *n* is a predetermined number, usually 1 (Alkhamis, Hasan and Ahmed, 1998). It is implemented in the ‘dwave-neal’ (2015) library with *n=1*.

### 3.6.2 Tabu Search

Tabu search solver was adapted to be used with QUBO models (Palubeckis, 2004). This algorithm can quickly find minimums in neighbourhoods but sometimes may struggle to escape them. It is implemented in the ‘dwave-tabu’ (2018) library.

### 3.6.3 Decomposing Solver

This algorithm splits large QUBO into smaller subQUBOs, solves them separately and combines the results (Booth and Reinhardt, 2017). The subQUBOs can be solved using a D-Wave quantum annealing device or tabu search with a classical machine. It is implemented in the ‘qbsolv’ (2017) library.

This is useful when using quantum devices because larger QUBOs cannot be mapped on them, but smaller subQUBOs can be, allowing us to solve larger problems. On classical devices, we achieve a speedup by decomposing large QUBO before applying tabu search.

### 3.6.4 Steepest Descent Algorithm

This greedy algorithm is like gradient descent, but instead of making moves based on the gradient, it uses local minimisation. At each step, we move to the state one bit-flip away that causes the highest energy drop.

It is implemented in the ‘dwave-greedy’ (2019) library.

# 4 Penalty Coefficient Optimisation Techniques

In this section, we will present multiple techniques of penalty coefficient optimisation. Two of them will be standard, and the other two will be state-of-art. This area of research is relatively new, and the number of unique approaches to this problem is limited.

As mentioned in previous sections, penalty coefficient optimisation is not trivial (Verma and Lewis, 2020). If the coefficient chosen is too low, the solution to the problem will be infeasible. If it is too large, penalties will dominate the search space, erasing the difference between good and bad states. This makes the optimisation problem numerically unstable, decreasing the accuracy of produced solutions (Vyskočil, Pakin and Djidjev, 2019).

## 4.1 Analytical

The first approach is analytical and involves using our domain knowledge to set a suitable *M* value (Glover, Kochenberger and Du, 2019). In the beginning, we need to make a rough estimate of the original objective function. Then, we set our penalty coefficient *M* to about 75% to 150% of this estimate. The resultant problem needs to be solved using a QUBO solver, and the solution should be checked for feasibility. If feasible, we can reduce the *M* and try again (or keep this coefficient). If it is not, we need to increase *M*. We then repeat the process of solving with the chosen *M*, checking for feasibility, and updating *M* until we settle with a penalty coefficient that satisfies us.

The benefit of this approach is its simplicity and that by the end of it, we get a good penalty coefficient. But at the same time, it is not always easy to estimate the original objective function “by eye”, especially when the problem is large.

Also, this trial-and-error procedure requires at least multiple calls to the solver. This is expensive and impractical for performance-critical applications (Huang *et al.*, 2021).

This method of penalty estimation is not automated. If we are trying to solve many problems, it is not efficient to estimate the penalty coefficient manually.

And even though this approach is somewhat systematic, it is not reproducible. Thus, we can get different penalty coefficients if we apply this method to the same problem multiple times.

## 4.2 Numerical 1

This method involves choosing an arbitrary value of matrix *Q* and making the penalty coefficient *M* larger than that number (Verma and Lewis, 2020).

Although this approach is fast and straightforward, it does not always work. We can try to improve our estimated *M* using the methodology used in the analytical method (Glover, Kochenberger and Du, 2019) described in the previous subsection, but that would impose the downsides of the analytical approach (inefficiency being one of them) on this method too.

## 4.3 Numerical 2

This method was proposed by Verma and Lewis (2020). They use the general QUBO model with penalties, , to estimate a lower bound of *M*.

First, they consider a problem, where the next node, *xto*, is better than the current node, *xfrom*, but the current node is feasible, and the next one is not:

, and .

In this scenario, penalty coefficient *M* should be large enough to prevent such transition, because otherwise, we would choose an infeasible solution over a feasible one:

The formula can be rearranged to get a lower bound of *M*:

To calculate the exact lower bound, we need to find a transition that will give us the maximum energy change with the minimum penalty imposed. Since there are no efficient methods for calculating the denominator, they assume that it equals 1, which is the lowest and best value it can take.

They propose an algorithm to find the maximum transition for a 1-flip solver in *O(2n2)* steps.

Both and flips can result in a positive change. For example, flipping from 0 to 1 will bring a positive transition to and flipping 1 to 0 will bring a positive transition to . That is why we must consider both flips.

We need to flip every *xi* of the binary decision vector *x* from 0 to 1 and vice-versa to find the maximum transition:

The maximum of (left expression in the brackets) part can be found by doing the following. For every *xi*, we need to take the coefficient of its linear part (*qii*) and add it to the sum of all the positive coefficients that the *xi* is associated with in its quadratic part (*qij > 0*). We only consider positive coefficients when they are negative. The second decision variable associated with (*xj*) will be 0 because we are looking for the maximal transition *xi* can bring.

The maximum of part can be found in the same way. Since such transition brings positive changes by nullifying the negative coefficients, we need to sum their negatives (will bring a positive change of 6, which is the negative of the coefficient *q1*). For the same reasons, instead of summing the positive coefficients of the quadratic part, we need to sum the negative ones (*qij < 0*).

All the steps described can be expressed mathematically to find the estimated penalty coefficient:

The estimates for *xto* nodes that are feasible (i.e., they have no constraints associated with them) will be discarded as their penalty () will be equal to 0, and the next largest *Mest will be used.*

This method is fast and produces good estimates of the lower bound of *M*, but it does not consider the quadratic penalties, , when estimating the penalty coefficient. That is why this approach can produce *M* values that are too large. One promising approach is to use Walsh analysis (Brownlee, 2009), which explicitly attaches energies to variable interactions in a natural way to the QUBO model. These energies may be used to estimate what values the quadratic penalties can take to decrease the lower bound estimate further.

## 4.4 Machine Learning

Huang *et al.* (2021) proposed using machine learning (ML) to optimise penalty coefficients. They have reviewed the analytical approach (described in section 3.1), where different penalty coefficients are used with the QUBO solver to find the optimal value. Calling a QUBO solver multiple times to solve a single problem is expensive and is not suitable for applications where performance is critical.

Alternatively, they propose using data obtained from the QUBO solver in the past to train an ML model called *Solver Surrogate* (SS) that will predict energy and feasibility using the problem features and a penalty coefficient. Then, instead of using a QUBO solver, we will use a faster SS, adjusting the coefficient repeatedly until a satisfactory result is achieved.

As conventional QUBO solvers usually return a batch of solutions with their energies, we can count how many of them are feasible and calculate the probability of feasibility (*Pf* ) with the penalty coefficient used. We can also calculate the mean energy of the solutions (*Eavg*) and the standard deviation (*Estd*). The SS will be trained to predict these three values.

During the training, features of the CO problem (*g*) and the penalty coefficient (*A*) from the dataset will be inputted into SS to make a prediction. This prediction and ground truth will be used to calculate the loss. The loss is then used to update the model with backpropagation. This will be repeated for the entire dataset multiple times until the accuracy of predictions reaches a plateau. The described process can be seen in Figure 5.

Diagram

Description automatically generated

Figure 5: Solver Surrogate training (Huang et al., 2021).

After training the surrogate, we can use it to find penalty coefficient estimates in the following way. We will use problem features and an arbitrarily chosen initial *A* to predict *Pf*, *Eavg* and *Estd*. It is also meaningful to estimate the initial *A* using Verma and Lewis (2020) method. Using the predicted values, we can adjust the penalty coefficient and rerun the SS. Huang *et al.* (2021) propose three strategies (called *Parameter Selection Strategies*) for changing coefficient *A*. This process is repeated until a suitable estimate of the penalty coefficient (*Ã*) is found. The described approach is demonstrated in Figure 6.

Diagram

Description automatically generated

Figure 6: Estimating penalty coefficient with Solver Surrogate (Huang et al., 2021).

This method benefits from good penalty coefficient estimates found quicker than if we made calls to QUBO solver. It can be used to solve many variations of the same CO problem, which is common in industry (Fujitsu, no date; D-Wave Systems, 2021).

However, if we do not have a dataset with the problems solved in past, solutions and penalty coefficients used, we will be unable to train the model. In this case, we will need another strategy to estimate penalty coefficients while collecting the data. Any other approach described in section 3 could be used for this purpose.

# 5 Summary

Summarise everything and form a research question (or repeat the one from the project proposal).

# References

Alkhamis, T. M., Hasan, M. and Ahmed, M. A. (1998) ‘Simulated annealing for the unconstrained quadratic pseudo-Boolean function’, *European Journal of Operational Research*, 108(3), pp. 641–652. doi: 10.1016/S0377-2217(97)00130-6.

Aramon, M. *et al.* (2019) ‘Physics-Inspired Optimization for Quadratic Unconstrained Problems Using a Digital Annealer’, *Frontiers in Physics*, 0(APR), p. 48. doi: 10.3389/FPHY.2019.00048.

Bertsimas, D. and Tsitsiklis, J. (1993) ‘Simulated Annealing’, *Statistical Science*, 8(1), pp. 10–15.

Booth, M. and Reinhardt, S. P. (2017) ‘Partitioning Optimization Problems for Hybrid Classical / Quantum Execution TECHNICAL REPORT’.

Boyd, J. (2018) ‘Fujitsu’s CMOS Digital Annealer Produces Quantum Computer Speeds’, *IEEE Spectrum*, 28 May. Available at: https://spectrum.ieee.org/ieee-courses-on-digital-transformation (Accessed: 12 November 2021).

Brownlee, A. E. I. (2009) ‘Multivariate Markov networks for fitness modelling in an estimation of distribution algorithm.’ Available at: https://rgu-repository.worktribe.com/output/247858/multivariate-markov-networks-for-fitness-modelling-in-an-estimation-of-distribution-algorithm (Accessed: 18 November 2021).

Collins, E. F. (1921) ‘ELECTRICALLY HEATED GLASS ANNEALING LEHR1’, *Journal of the American Ceramic Society*, 4(5), pp. 335–349. doi: 10.1111/J.1151-2916.1921.TB18664.X.

Crosson, E. and Harrow, A. W. (2016) ‘Simulated Quantum Annealing Can Be Exponentially Faster than Classical Simulated Annealing’.

Cruz-Santos, W., Venegas-Andraca, S. E. and Lanzagorta, M. (2019) ‘A QUBO Formulation of Minimum Multicut Problem Instances in Trees for D-Wave Quantum Annealers’, *Scientific Reports 2019 9:1*, 9(1), pp. 1–12. doi: 10.1038/s41598-019-53585-5.

D-Wave Systems (2021) *Volkswagen: Navigating Tough Automotive Tasks with Quantum Computing*. Available at: www.dwavesys.com/d-wave-launch (Accessed: 5 October 2021).

Dattani, N., Szalay, S. and Chancellor, N. (2019) ‘Pegasus: The second connectivity graph for large-scale quantum annealing hardware’.

‘dwave-greedy’ (2019). Available at: https://github.com/dwavesystems/dwave-greedy (Accessed: 17 November 2021).

‘dwave-neal’ (2015). Available at: https://github.com/dwavesystems/dwave-neal (Accessed: 17 November 2021).

‘dwave-tabu’ (2018). Available at: https://github.com/dwavesystems/dwave-tabu (Accessed: 17 November 2021).

Farhi, E. *et al.* (2001) ‘A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem’.

Fujitsu (no date) *Reducing traveling distance for parts picking operations by up to 45%*. Available at: https://www.fujitsu.com/global/services/business-services/digital-annealer/case-studies/201804-fjit.html (Accessed: 5 October 2021).

Glover, F., Kochenberger, G. and Du, Y. (2019) ‘Quantum Bridge Analytics I: a tutorial on formulating and using QUBO models’, *4OR*, 17(4), pp. 335–371. doi: 10.1007/S10288-019-00424-Y.

Hamerly, R. *et al.* (2018) ‘Scaling advantages of all-to-all connectivity in physical annealers: the Coherent Ising Machine vs. D-Wave 2000Q’.

Higham, C. F. and Bedford, A. (2021) ‘Quantum Deep Learning: Sampling Neural Nets with a Quantum Annealer’. Available at: https://arxiv.org/abs/2107.08710v1 (Accessed: 11 November 2021).

Huang, T. *et al.* (2021) ‘QROSS: QUBO Relaxation Parameter optimisation via Learning Solver Surrogates’, *2021 IEEE 41st International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pp. 35–40. doi: 10.1109/ICDCSW53096.2021.00013.

Johnson, D. S. *et al.* (1989) ‘Operations Research Society of America This content downloaded from 128’, 37(6).

Johnson, M. W. *et al.* (2011) ‘Quantum annealing with manufactured spins’, *Nature 2011 473:7346*, 473(7346), pp. 194–198. doi: 10.1038/nature10012.

Katzgraber, H. G. and Novotny, M. A. (2018) ‘How Small-World Interactions Can Lead to Improved Quantum Annealer Designs’, *Physical Review Applied*, 10(5), p. 054004. doi: 10.1103/PHYSREVAPPLIED.10.054004/FIGURES/13/MEDIUM.

Kirkpatrick, S., Gelatt, ; C D and Vecchi, ; M P (1983) ‘Optimization by Simulated Annealing’, *New Series*, 220(4598), pp. 671–680.

Kochenberger, G. *et al.* (2014) ‘The unconstrained binary quadratic programming problem: A survey’, *Journal of Combinatorial Optimization*, 28(1), pp. 58–81. doi: 10.1007/S10878-014-9734-0/TABLES/1.

Kochenberger, G. A. and Glover, F. (2006) ‘A Unified Framework for Modeling and Solving Combinatorial Optimization Problems: A Tutorial’, *Multiscale Optimization Methods and Applications*, pp. 101–124. doi: 10.1007/0-387-29550-X\_4.

Liang, F., Cheng, Y. and Lin, G. (2014) ‘Simulated Stochastic Approximation Annealing for Global Optimization With a Square-Root Cooling Schedule’, *https://doi.org/10.1080/01621459.2013.872993*, 109(506), pp. 847–863. doi: 10.1080/01621459.2013.872993.

Lucas, A. (2014) ‘Ising formulations of many NP problems’, *Frontiers in Physics*, 0, p. 5. doi: 10.3389/FPHY.2014.00005.

Ohzeki, M. *et al.* (2019) ‘Control of Automated Guided Vehicles Without Collision by Quantum Annealer and Digital Devices’, *Frontiers in Computer Science*, 1, p. 9. doi: 10.3389/FCOMP.2019.00009/BIBTEX.

Ohzeki, M. and Nishimori, H. (2010) ‘Quantum annealing: An introduction and new developments’.

Palubeckis, G. (2004) ‘Multistart Tabu Search Strategies for the Unconstrained Binary Quadratic Optimization Problem’, *Annals of Operations Research 2004 131:1*, 131(1), pp. 259–282. doi: 10.1023/B:ANOR.0000039522.58036.68.

Poole, D. L. and Mackworth, A. K. (2017) ‘Artificial Intelligence: Foundations of Computational Agents’, pp. 136–137. doi: 10.1017/9781108164085.

‘qbsolv’ (2017). Available at: https://github.com/dwavesystems/qbsolv (Accessed: 17 November 2021).

Şeker, O. *et al.* (2020) ‘Digital Annealer for quadratic unconstrained binary optimization: a comparative performance analysis’. Available at: http://arxiv.org/abs/2012.12264 (Accessed: 3 October 2021).

Snelling, D. *et al.* (2020) ‘A Quantum-Inspired Approach to De-Novo Drug Design’. doi: 10.26434/CHEMRXIV.12229232.V1.

Tanaka, S. and Tamura, R. (2012) ‘QUANTUM ANNEALING AND QUANTUM FLUCTUATION EFFECT IN FRUSTRATED ISING SYSTEMS’.

Verma, A. and Lewis, M. (2020) ‘Penalty and partitioning techniques to improve performance of QUBO solvers’, *Discrete Optimization*, p. 100594. doi: 10.1016/j.disopt.2020.100594.

Vyskočil, T., Pakin, S. and Djidjev, H. N. (2019) ‘Embedding Inequality Constraints for Quantum Annealing Optimization’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11413 LNCS, pp. 11–22. doi: 10.1007/978-3-030-14082-3\_2.

Wu, H. and Fan, G. (2020) ‘An overview of tailoring strain delocalization for strength-ductility synergy’, *Progress in Materials Science*, 113, p. 100675. doi: 10.1016/J.PMATSCI.2020.100675.