

Final Report
**Maximising *entertainment* value in the
vote-reveal problem**
Final Year Project (CM3203) - 40 Credits

Author: Iain Johnston (1312579)

Supervisor: Richard Booth

Moderator: Xianfang Sun



Abstract

As yet very little research has been undertaken into maximising the "entertainment" of voting competitions, however the appeal of having this ability is immediately obvious. This project attempts to find the most entertaining order for votes to be revealed in competitions; such as the Eurovision Song Contest. The problem was first formalised into a framework designed in this project for solving optimisation problems, then functions to describe the entertainment were designed and then a range of optimisation algorithms were used to try to maximise the entertainment of the voting orders. A full analysis of the *entertainment* of the solutions and the properties of the algorithms was undertaken. A near optimal, if not globally optimal solution was found for the 2014 Eurovision Song Contest; and a novel problem specific algorithm was designed. The project also introduces a framework and generalisable methodology for future work in this field.

Acknowledgements

I would like to thank my supervisor, Richard Booth, for support and insights throughout the project.

Thanks to the European Broadcasting Union (EBU) for releasing the datasets for the Eurovision Song Contest for use in this project.

Contents

0.1	Introduction	6
0.2	Background	7
0.2.1	Project Context	7
0.2.2	Background Theory	8
0.3	Algorithm Designs and Approach	10
0.3.1	ε Functions	10
0.3.2	Neighbourhoods	13
0.3.3	Optimisation Algorithms	15
0.4	Implementation of System	22
0.4.1	ε function wrappers	22
0.4.2	ε functions	24
0.4.3	Neighbourhood functions	25
0.4.4	Simulated Annealing and Greedy search	26
0.4.5	Brute force	27
0.4.6	Piecemeal	28
0.4.7	Re-Use of Code for future work	28
0.4.8	Visualisation	29
0.5	Results and Evaluation	31
0.5.1	Overview of goals	31
0.5.2	Solutions	31
0.5.3	Testing	36
0.5.4	Critical Evaluation	41
0.6	Future Work	42
0.7	Conclusions	42
0.8	Reflection on Learning	43
	Glossary	45
	Appendices	45
	References	55

List of Figures

1	2014 Scoreboard	9
2	Uphill moves in optimisation problems	18
3	Calculating the cumulative scores per participant	22
4	refinedMaxMin method	25
5	getAdjacentNeighbour method	26
6	getRandomNeighbour method	26
7	Brute Force algorithm implementation in Python	28
8	Visualisation of solution to 2014 competition	30
9	Running algorithms	32
10	Snippet of Greedy results	33
11	Extracting scores from results file	33
12	Best solution	34
13	Chart of ε values found by Greedy, Simulated Annealing and Piecemeal	35
14	How to time the algorithms using Python	36
15	Extracting times from a results file	36
16	Average time taken to perform full and partial ε calculation on the same solution	38
17	Difference between min value found for MaxMin method and RefinedMaxMin .	39
18	Finding the difference between two orders	40
19	offsetGetEntertainment method	48
20	getEntertainment method	49
21	Simulated Annealing code	50
22	Greedy Search code	51
23	Piecemeal algorithm	52
24	Final Gantt Chart	54

List of Tables

1	Scores for simple example game	10
2	Scores per round	10
3	Difference per round for first order	11
4	Difference per round with Germany and Austria swapped	11
5	Score calculations needed with adjacent neighbour	14
6	Approximate 2014 Eurovision specific calculations	14
7	Worst case Greedy complexity per line	16
8	Simulated Annealing complexity per line	19
9	ε values found for small example game	21
10	Scores for simple example game	21
11	ε values found for bigger example game	22
12	Big-O of ε methods	25
13	Worst Case Big-O complexities of algorithms used	45
14	Final Simulated Annealing cooling values	53
15	Best and average solutions found by algorithms	53
16	Timing results for main algorithms (all times in seconds to 4 s.f)	53

0.1

Introduction

In many elections and competitions, a set of voters will rank a set of participants from best to worst, or will give scores to some of the candidates, with the winner then being the candidate that gets the highest total number of points. When it comes to revealing the result after all the votes have been cast but not revealed, some competitions proceed by having a roll-call of all the voters in which each voter announces their own scores. This is often done for entertainment purposes such as in the Eurovision Song Contest.[1]

The concept of *entertainment*, especially with respect to competition, is a heavily subjective matter and as such is difficult to quantify in simple terms. There are intuitive constituent parts to an *entertaining* competition (like Eurovision) such as, if the winner is known early or late, and how many teams are in the running to win. This project will attempt to find ways to translate these slightly nebulous concepts into more concrete mathematical functions.

This leads to the two main questions that this project will aim to answer:

1. How can we define the concept of "entertainment" in the context of an optimisation problem, and hence try to maximise it.
2. In which order should the votes be revealed in order to maximise that entertainment value?

To try and answer these questions the Eurovision Song Contest will be used as an example as it has a good quality datasets and the scoring rules of Eurovision are well documented and relatively simple (see Section 0.2).

Changes to the voting rules over the years show that the motivation behind this project is a current problem and there have been attempts to try and solve this problem already, at least by the Eurovision organisers.

There is no specific intended audience for this project, Eurovision is a test competition as it fits well with the more general optimisation problem that this project is trying to solve. The success of the project could be of interest to the organisers of the Eurovision Song Contest as they generally want a exciting competition. In a more general sense though it could be said that the beneficiaries of the project would be those who would also attempt to solve this problem for other datasets or competitions. This project can help set out a framework for approaching these problems and many parts could even be re-used as long as the problem can be formulated in a specific way (see Section 0.4.7 for details)

The scope of this project is to try and solve the optimisation problem by finding a globally optimum solution, or by finding a near optimal solution and hence have a voting order that is the most entertaining. Also within scope is to produce a way of visualising the solutions found and from this gaining a better insight into why they are entertaining. This project does not try and influence other parts of the voting that could also lead to more entertainment (such as who delivers the votes or how the votes are delivered, only the order itself)

The approach used to tackle this problem is a scientifically minded one. There are three main sections to the project. Firstly, designing and implementing functions that quantify the entertainment of a given voting order. Secondly, solving the optimisation problem of maximising the entertainment value given by those functions; using optimisation algorithms. Finally, analysing the solutions given and comparing and contrasting why there are entertaining and how well

the algorithms performed. This approach involves iteration on theoretical ideas whilst feeding back in results as the project goes along.

The important outcomes of this project are, a solution that delivers an entertaining competition when the votes are revealed, a way of describing entertainment mathematically, and a way of visualising the competition as the votes are revealed. Moreover an interesting and novel algorithm is devised that can match the more general optimisation algorithms in terms of quality of the solutions produced and do so faster.

The attached files include the algorithms used in this project; Brute force: *bruteForce.py*, Greedy: *greedySearch.py*, Simulated Annealing: *simulatedAnnealing.py* and Piecemeal: *step.py*. The main file that controls the problem: *order.py*, the problem specific supporting methods: *support.py*, and the full set of results. This full project can also been found and cloned from Github here: <https://github.com/iainrj/FinalYearProject>.

This report will first give some background to the problem and the Eurovision Song Contest as a dataset. Then the algorithms used to solve the problem will be introduced and their theory and necessity explained, as well as explaining the problem specific methods designed and why they are needed. How those methods are implemented is then explained with reference to the code produced. The results collected by running the algorithms are then analysed and the method to collect the results is justified, and interesting comparisons are made between the results. Finally more general testing of the system as a whole is explained and a critical evaluation of the project is undertaken.

0.2

Background

The more general reason for this project is to try and see if there can easily be found an optimal solution to a problem when given a mathematical representation of something that humans experience.

Furthermore as competitions are usually televised and watched live, it is generally in the interest of those in charge of the competition to produce an entertaining show as this helps them with many facets of their business such as through advertising, but also in building a following for the competition.

0.2.1 Project Context

Eurovision has been a topic of research for many years now. The main focus of that research has been in relation to the voting patterns that can be found over the course of many runnings of the competition. This research such as "*Comparison of Eurovision Song Contest Simulation with Actual Results Reveals Shifting Patterns of Collusive Voting Alliances*"[2], "*Geography, culture, and religion: Explaining the bias in Eurovision song contest voting*"[3] and "*The Eurovision Song Contest. Is voting political or cultural?*" [4] all use the Eurovision Song Contest as a basis to investigate political and cultural phenomena.

There has also been some research in more computational areas such as "*Using the Clustering Coefficient to Guide a Genetic-Based Communities Finding Algorithm*"[5] which attempts to find communities within the voting patterns.

The particular problem this project is addressing has not yet been published in a scientific setting. It does seem to have been at least tackled by those that run Eurovision, however as they are a private company they have never published any methodology on how they pick voting orders. They usually state on their website[6] when revealing the voting order to media that:

"The voting order has been determined by the results from last night's Jury Final. An algorithm has been created to try and make the voting as exciting as possible."

There is however no mention of how this algorithm works or what they constitute entertaining or exciting.

The biggest drawback that can be said about the current Eurovision algorithm, without seeing it and understanding its methodology, is that it only takes into account Jury votes as they can be decided on during the dress rehearsals. This is problematic as there are two sets of votes, one from the Jury and another from tele-voting. These votes can diverge for many reasons and hence the voting order that purports to be entertaining, may only be such when the viewers at home agree exactly with the Jury.

This project's method attempts to take into account *both* the sets of votes and hence could give a more correct picture of what will happen, entertainment wise, for any given voting order. One main constraint on this method is exactly that fact, as it is necessary to wait for all the votes to be cast before the voting order can be found. In real world terms this may not be feasible for the Eurovision Song Contest as they may have commitments that require that the voting order is known in advance. The voting order is usually announced early on the day of the Grand Final, which the method set out in this project could not do.[6]

Another drawback for the current Eurovision "algorithm" as mentioned above, is the fact that there is no explanation as to why their algorithm believes a given order to be more exciting than any other. This project will attempt to standardise mathematically a function for entertainment, which could in future be compared through human tests (see Section 0.6). By describing a concept mathematically not only can it be used by algorithms but it could also form part of a proof.

This project only looks at one specific area of the competition to describe and codify entertainment. This can be seen as a drawback of the solutions found, however the way this project has been undertaken leaves the opportunity for further work in the area of defining entertainment functions as described in Section 0.6.

The techniques used and implemented in this project have applications outside the strict problem that is being solved here. For example even though this problem is specifically trying to produce entertaining orders for Eurovision, there is no reason why other competitions who use the same type of ordered voting system (sometimes referred to as roll-call voting) could use the algorithms and framework designed and implemented in this project to solve their problems. Furthermore, with a little extension and modification this framework could feasibly be used for any optimisation problems that can be set out in a certain way (see Section 0.4.7)

0.2.2 Background Theory

Understanding how the Eurovision Song Contest voting works is key to this project, as this affects many parts of the design and implementation of the system. Furthermore the system for voting has been changed over many years and so to try and standardise the project I will be

using the system that was in use during the 2014[7] running of the competition, and the main dataset also comes from that year.¹

In this system, each country awards two sets of 12, 10, 8 – 1 points to their 10 favourite songs: one from their professional jury and the other from tele-voting. Moreover both constituent parts count 50% towards the final score. The *scores* used during the project will be those final composite scores.

In the 2014 contest, there were 37 countries that participated and of those 26 reached the final. All 37 countries vote in the final. From this point on in this report, the countries that vote are referred to as *voters* and those 26 competing in the final are referred to as *participants*. There is overlap between the two groups i.e the *participants* set is a subset of the *voters*; but it should be obvious from the context to whom is being referred.

Participant	UK	DE	FR	IT	ES	RU	PL	GR	TR	UA	RO	HR	CZ	SK	SE	NO	IS	AT	CH	PT	MT	DK	NL	SM	GB	Points	Place													
Ukraine		5	10	8	4	1	8		2		6	5	2		5	10	7	5		10	7	5					113	6												
Belarus	8		7																1								43	16												
Azerbaijan				3					7																		33	22												
Iceland	2				5			7		2	5				7												58	15												
Norway	1		4		6	3		7	2		5	3		1	7		5	8	2			7	3	1		5	3	5	10		88	8								
Romania	8	6	1	5			4								2	8	5			8	12		4		1		8				72	12								
Armenia		12		10		2	5	8	4	12	12	6	7	7	2		6		10		6	3	10		1	4	7	8	6		174	4								
Montenegro	6	12					12																				37	19												
Poland				2	7				5		5		10	1	3	3			8						2	4	2			1	2		62	14						
Greece	2	7		4	6																										2	35	20							
Austria	5			1		12	8	4	3	12	10	10	7	12	10	10	12	12	6	10	10	7	2	10		12	8	5		12	12	12	12	8	12	290	1			
Germany	4	6											5												8		2			2		7	5	39	18					
Sweden	7		6			10	12	10		10	4	2		2	8	7	4	10		8	7	7	6	3	8	4	8	12	2	10	8	10	6	8	12	7	218	3		
France									1																						1				2	26				
Russia		10		12	12			1	6			8	10						3		2	6	5	8						2			4	89	7					
Italy		10							2	1																	12		6				2		33	21				
Slovenia									1																									9	25					
Finland				4			3	4	6						4		6	5			6	3						7	3			3		6	4	2	6	72	11	
Spain		12	2				2	2			6		1						6	4			4	4				5	2		5		4		8	2	5	74	10	
Switzerland		5	3									1	3	4	1		5		2		2	3		5		10	7	6			3			3	1	64	13			
Hungary	8		7	8	5	7	3	7	10	5				6		6	1	7		1			4	12			6	10	6	7		2	7	1	4	3	143	5		
Malta	1			5						3									3		1										4			5	10	32	23			
Denmark		1				6			6	3		8			8								1				1	6	5	4		1	6	3	8	3	1	3	74	9
The Netherlands		4	10		2	8	10	12	7	8	8		12	8	12	12	10		4	12	12						12	12	10	3	3	2	10	7	10	10		8	238	2
San Marino		3	3		3																				1												14	24		
United Kingdom						1	7					3							4	8						4					3			5	5			40	17	

Figure 1: 2014 Scoreboard

This voting system creates a matrix of scores as seen in Figure 1[8]. The *voters* are arranged along the top alphabetically, while the *participants* are along the left hand side. Points refers to the final total of all the scores received at the end of all voting, and rank is their final position in the competition.

This is important to look at as it can help clarify what this project's aim is. A solution to the problem is found by swapping a *voter's* column along the top into different positions and then analysing the scores after each country has voted, to reach an *entertaining* order for those countries to reveal their votes.

Hence the problem that this project is trying to solve is about ordering who votes when so as to maximise some mathematical concept of entertainment.

¹2014 voting system is the system that had been in use from 2009 until changed in 2016

0.3

Algorithm Designs and Approach

This section should give an explanation of the main parts that are needed to solve this problem. This will include some pseudocode of the algorithms used, explanations of the neighbourhoods tried and the entertainment functions designed. It is intended to give an introduction to the why and how these parts of the project exists, before Section 0.4 goes into implementation details with code and explanations using Python.

Firstly to help clarify the problem it is easier to view a much smaller competition and then run through the main points of theory, that are then used in the full Eurovision system.

Firstly we design a competition that involves 4 teams total, of which 2 are *participants*. The scoring system is even more simple than the one used by Eurovision. Here each voting team gives 1 point the team they prefer and 0 to the other. A team cannot give itself 1 point. The scores given in this game are shown in Table 1.

Table 1: Scores for simple example game

	Germany	France	Spain	Austria
Austria	1	1	1	0
France	0	0	0	1

From Table 1 we can see how the scores will be calculated in this type of competition. For simplicity, the order that the teams vote is the order from left to right in Table 1. This means that after Germany has voted the scores are *Austria* : 1, *France* : 0, the competition is shown in Table 2

Table 2: Scores per round

	After Germany's vote	After France's vote	After Spain's vote	After Austria's vote
Austria	1	2	3	3
France	0	0	0	1

We can see that Austria wins this competition, with 3 points. This cumulative scoring extends to the full Eurovision competition in the exact same way, except for in that case there are 30+ voters.

All that is needed to turn this simple example into Eurovision is to add all the *voters* and *participants* and change the scores given out to be 12, 10, 8 – 1 instead of just 1 and 0.

0.3.1 ε Functions

From this small example it can be seen that the ordering used would not be very entertaining as after the second round of voting, Austria could not lose, only draw.

To justify this mathematically we could look at the difference between the scores of the two teams at each round (Table 3). By summing that difference over the whole competition we get a difference of 8, which represents the order [*Germany*, *France*, *Spain*, *Austria*].

Table 3: Difference per round for first order

	After Germany's vote	After France's vote	After Spain's vote	After Austria's vote
Difference	1	2	3	2

However if we swap when Austria and Germany vote we would get a total difference of **4** over all the rounds (Table 4). This would intuitively seem like a more entertaining order as well as having a lower total difference.

Table 4: Difference per round with Germany and Austria swapped

	After Austria vote	After France's vote	After Spain's vote	After Germany's vote
Difference	1	0	1	2

Any entertainment functions (ε functions) talked about in this section share some characteristics that need to be explained. The first is that they take a **solution** (Φ) to the problem (see Glossary-3) as input and they return a single **entertainment value** (ε) (see Glossary-4). Moreover they calculate the **scores** (S) (see Glossary-6) for each *participant* country every **round** (R) (see Glossary-5).

MaxMin

The first entertainment function that was designed is the one described in the small example at the start of this section, named **MaxMin**. As it's name suggests it works by finding the difference between the highest score (max) and the lowest score (min) after each voting round. It then sums these values which is the ε value for that solution. This function follows quite naturally from watching and analysing the Eurovision competition as the way roll-call style voting works, everything is building towards the later end of the voting order. Hence it seems to be a innate part of the competition that you would want every country to be in with a chance of winning as often as possible. Moreover as there is only one prize, for first place, it is even less important to worry about positions other than the top.

So each round the distance (Λ) for a set of scores, S is calculated as such:

$$\Lambda = \max(S) - \min(S) \quad (1)$$

Then keeping the distance per round i as Λ_i , the entertainment value (ε) is found by:

$$\varepsilon = \sum_{i=0}^{n_R} \Lambda_i \quad (2)$$

(See Glossary-5 for n_R)

This is a relatively simple equation and hence transfers simply to code. However more importantly, it encodes an intuitive part of entertainment mathematically. This functions encodes the fact that an entertaining competition is one in which the distance between first and last place is as low as possible as often as possible. In this case instead of **maximising** ε , we want to **minimise** it.

Over the course of a whole competition i.e: n_R voting rounds, it is intuitive to want that distance to stay as low as possible, by finding the sum of the distance (Λ) in each round and then using optimisation algorithms to try and find a solution (Φ) that minimises this value.

There is however one major drawback of the basic **MaxMin** method, which is that in some rounds, especially later in the competition, the last place country is actually mathematically eliminated from the race for first place.

RefinedMaxMin

This leads to a second version of the entertainment function for this problem. As it is essentially a refinement of the first function and not a brand new method it is called **RefinedMaxMin**.

Equation 2 is the same across both functions, however where **RefinedMaxMin** and **MaxMin** differ is how the $\min(S)$ part of equation 1 is calculated. As some of the countries cannot win the competition after some number of rounds, they should not be taken into account when seeing if a solution is entertaining or not.

To find whether a team can still win, the upper bound of points left available to be given is found. This means that we assume for each team, and for the remaining rounds left to be revealed, that they gain the maximum (12) points and the hence we find the highest score they could ever attain. This method does not take into account whether the country has already given itself a vote (in which case they would not be allowed to get 12 points in that round). This was done to simplify the method and make testing it's improvements against MaxMin easier. Furthermore finding the upper bound is generally safe, especially when removing the low scoring teams in Eurovision as voting is generally quite uniform past a certain point in the voting; i.e countries that are given high points already generally get more, and countries that have received few points continue to get few. Further refinements to this calculation are discussed in Section 0.6.

The method for finding if a team can still win is to first sort the scores for every country in round i in ascending order. Then iterating through that list of scores from the start and for each score checking whether equation 3 is true.

$$countriesScore + (maxScorePerRound * roundsRemaining) \leq currentTopScore \quad (3)$$

If equation 3 holds true then that country cannot win even if it received the maximum number of points (12 in Eurovision's case) and the leaders received the worst score possible (0), for the remaining voting rounds. As the scores are in ascending order the last minScore found for which Equation 3 held true is the minimum score to be returned.

It is quite simple to justify this refinement when looking at the competition. Those teams who cannot win should not be making a mark on the entertainment of the voting order as most viewers will not be paying attention to their scores and rank. Moreover this refinement only works in competitions where there is only interest at the top of the table as opposed to competitions with relegations or play-offs, where more than just who is the overall winner is important.

It is important to start with the design of the entertainment functions as they are the main workhorse for finding solutions to the problem. The approach to describe them was to analyse the competition and try and identify things in a competition that lead to excitement. From those ideas, it was a case of trying to formalise that theory into a concrete piece of maths that could then be programmed. Furthermore the entertainment functions are quite problem specific and may not transfer well into other problems, whereas the algorithms are entirely agnostic to the problem at hand. This means the entertainment functions follow much more closely from the problem itself than any other part of the project.

0.3.2 Neighbourhoods

An important part of solving optimisation problems is designing and implementing neighbourhoods[9] for given solutions. As a solution (Φ) is a list of *voters* in a certain order, then we define a neighbourhood as any other solution that has two members swapped. This means that there are only two changes between a solution and any of its neighbours.

In this regard an order is a *permutation* of the *voters*. Using Cauchy's two line notation for permutations[10], we can show the basic theory for our neighbourhoods.

$$\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ \sigma(x_1) & \sigma(x_2) & \sigma(x_3) & \cdots & \sigma(x_n) \end{pmatrix} \quad (4)$$

Equation 4 shows the first solution on the top row and a neighbour of that solution on the bottom row. The functions that map the elements of a solution to a neighbour (σ) are explained below.

These methods are examples of cyclic permutations i.e a permutation of a set which maps a subset of elements to each other in a cyclic way, while mapping all other elements of the set to themselves. The cyclic parts of a permutation are called *cycles*. A cycle with only two elements is called a *transposition*[11]. As both of the following methods only contains two indexes, they are *transpositions*.

Random neighbour

The first idea and most simple for this problem, was to swap two random elements of the ordering. This means that the function σ swaps the two elements at those indexes. This method can be described as such in equation 5 for two random indexes i and j .

$$\begin{aligned} \sigma &= x_i \mapsto x_j, \\ &\quad x_j \mapsto x_i, \\ &\quad \forall x \neq i \vee \forall x \neq j : x \mapsto x \end{aligned} \quad (5)$$

This means that we swap the elements at indexes i and j and swap the rest with themselves. The code for this will be explored in Section 0.4.3. This method gives a large set of possible neighbours as theoretically for any single solution there are $\{\text{length of solution} \times \text{length of solution} - 1\}$ possible neighbours. In the 2014 Eurovision competition this would mean $\{37 \times 36 = 1332\}$ neighbours.

This method may be simple, however, when a closer look is taken at the actual ε values of solutions that are very similar to each other i.e they have 3 or 4 elements swapped and most elements the same, it is clear that their ε values only differ by a little. This calls into question if the method detailed above is actually going to help the optimisation algorithm reach an optimal solution.

Adjacent neighbour

The questions about the efficacy of the random method leads to another method for finding a neighbour of a solution. As solutions that are close together are usually quite similar, it is intuitive that swapping adjacent elements may improve the solutions found.

This method works by finding a random value between 0 and length of solution - 1. Then swapping the element at that index with its immediate neighbour to the right i.e. index + 1. It can be expressed in the same way as in equation 5, except that in this case index i is a random number and index $j = i + 1$.

Another difference between this method and the random method is how it can actually be calculated. After finding a neighbour of a given solution, the ε value of that solution must be found. By swapping 2 adjacent countries in the solution a *partial re-calculation* of the entertainment value can be done using the entertainment value of the previous solution. As the previous solution and the new solution only differ by one position it is possible to only partially re-calculate the ε value. This idea is implemented separately to the **MaxMin** and **refinedMaxMin** methods in the code, and explanation as to how can be found in Section 0.4.1.

Re-calculation involves removing the distances of the swapped countries in the old solution and adding them back in their new positions in the new solution. The equation for this is shown in 6.

$$\varepsilon_{new} = \varepsilon_{old} - \Lambda_{ij_{old}} + \Lambda_{ij_{new}} \quad (6)$$

The ε_{new} and $\Lambda_{ij_{old}}$ values can be found by keeping them from the old solution and passing them into the next ε calculation. The $\Lambda_{ij_{new}}$ must still be re-calculated however only up to the higher of the two indexes.

The efficacy of this neighbourhood in reducing the number of calculations needed can be seen when calculating the ε values. A full calculation, such as is used in the random swap method involves {number of voters \times number of participants} or $V \times P$.

However due to the adjacency of the swaps Table 5 shows what this method needs to calculate.

Table 5: Score calculations needed with adjacent neighbour

	Number of calculations
Max	$V \times P$
Min	$2 \times P$
Average	$\approx 18 \times P$

Table 6 show that what this would mean in the 2014 competition.

Table 6: Approximate 2014 Eurovision specific calculations

	Number of calculations
Max	962
Min	52
Average	468

A deeper comparison and evaluation between the two methods can be found in Section 0.5.3

0.3.3 Optimisation Algorithms

In the previous section the necessary first steps to solve the problem were outlined. Namely calculating an entertainment value (ε) which can be optimised. To do this optimisation it is necessary to use a set of optimisation algorithms which maximise or minimise a value in order to settle on an optimal solution.

Greedy Search

The first algorithm that was posited was a simple Greedy search algorithm. Greedy search works on the idea that by selecting the optimal choice at every possible stage, it will lead to a globally optimal solution.[12]

At first glance it is not obvious to see if the problem is simple or complex. This lead to Greedy search as a good first try as it is sufficiently simple so that the code is easy to write, and more importantly it will likely find the optimal solution if the solution space is simple. If the solution space *is* in fact too complex then Greedy may still find a good approximate solution.

An important point to re-iterate, as mentioned when talking about the ε functions, is that even though in the report it may be mentioned that the algorithms are **maximising** entertainment to do this in the code, we have to **minimise** the ε value produced by the ε functions. All algorithms described in this report follow this methodology.

The pseudocode for Greedy is shown in Algorithm 1 in the appendix.

As can be seen, it is a very simple algorithm that takes possible solutions and compares them to the already found best solution. It's simplicity is, unfortunately, also it's downfall in this problem. Although it can find a good approximate solution it is not able to find the higher entertainment solutions that other algorithms find.

The stopping criterion defines when the algorithms should stop. To maintain parity between the algorithms, the stopping criterion chosen was to stop when a certain number of iterations had passed without a new solution being found. The value of the stopping criterion is given by the variable *num_iterations* in all algorithms that use it. To achieve this idea it is necessary to reset the value of *i* when a new solution is found. This can be seen on line 10 of Algorithm 1

Looking at the pseudocode, and with some simplifying assumptions the worst case time complexity for this Greedy search algorithm can be found.

This is done by giving all the instructions that make a difference to complexity, a value T_i from their line number. This means line 5 is T_5 , then working out how many times that instruction will need to be run given an input. The input is the *voters* and *participants*, V and P . Some lines can be ignored as they do very little to complexity. Table 7 shows the complexity added by some important lines.

Table 7: Worst case Greedy complexity per line

Line	Effect on complexity
T_1	run once with V steps
T_2	run once
T_3	run once with $V \times P$ steps
T_5	m times with constant steps
T_6	m times with $V \times P$ steps; as worst case, assume that full re-calculation is needed
$T_7 - T_{11}$	m times as is worst case, so assume this block always run
T_{12}	m
T_{14}	1

N.B The main loop (lines 4 - 13) is constant, given it comes from *num_iterations*. It adds m times to the instructions inside it.

Factoring this down into equation 7 gives:

$$f(n) = (V \times T_1) + T_2 + (V \cdot P \times T_3) + (m \times T_5) + (m \times V \cdot P \times T_6) + (m \times T_{7-11}) + (m \times T_{12}) + T_{14} \quad (7)$$

From this we can safely ignore anything that is less than the highest order as it will dominate the complexity growth. Moreover we can see that the highest order is T_6 , so we are left with equation 8

$$f(n) = (m \times V \cdot P \times T_6) \approx m \times V \cdot P \quad (8)$$

where $P \leq V \ll m$. This can be further simplified in terms of *voters* (V), as P is always less than or equal to V and m is always larger or equal to V , they can be roughly equated to V . The coefficients of m and P compared to V do not matter to the complexity, just that they are near the same size. This leaves us with a fair approximation of the complexity in equation 9.

$$O(n) \approx m \times V \cdot P \approx V \times V \times V \approx O(V^3) \quad (9)$$

A discussion and comparison with the other algorithms can be found in Section 0.5.3.²

Brute force

After running the Greedy algorithm and collecting some initial results it became clear that it would be difficult to know by just looking at the problem, what a good ε value would be. This lead to the decision about using a Brute force algorithm to try and enumerate all the solutions.

²A full table of Big-O complexity for all algorithms can be found in the appendix in table 13

The idea is extremely simple, try all possible solutions and keep track of the best found. Even though the pseudocode and the real code will differ, the pseudocode help understand the algorithm so is shown in Algorithm 2. The obvious difference with Brute force is the fact that it does not have to find neighbours of current solutions as it will try every one possible.

When some simple calculations are made about the solution space it is clear as to why the Brute force method is not ever going to be feasible for this problem. Searching all solutions, is not a simple task. As it must look at all permutations of solutions which are of length V , the complexity of Brute force follows the number of permutations for n distinct objects.[13]. This is shown in equation 10.

$$O(n) \approx O(V!) \quad (10)$$

This means that in all cases Brute force will run in $\approx V!$ time. Comparison to the other algorithms used can be found in Section 0.5.3.³

For example in the 2014 running of Eurovision this would be $37!$ which, after approximating the time to look at one single solution, would take about 1.2×10^{40} seconds, which is quite obviously computationally infeasible with the time and resources for this project.

Simulated Annealing

The performance of Greedy and Brute force suggests the fact that this problem is not a simple one at all and hence another, more capable algorithm is needed to find more optimal solutions. To do this the Simulated Annealing algorithm was used. First introduced by Kirkpatrick et al[14] it is a metaheuristic that as well as accepting better solutions, also occasionally accepts worse ones.

The pseudocode for Simulated Annealing can be found in Algorithm 3 in the appendix. It is quite a lot more complex than either Brute force or Greedy search.

The skeleton of the algorithm is shared with both the previous algorithms; that is, lines 22 - 25 in Algorithm 21 are the same as lines 5 - 7 in Algorithm 2 (Brute force) and lines 9 - 12 in Algorithm 1 (Greedy).

The important part where they differ is how they pick a solution to try. Simulated Annealing finds a neighbour of the current solution, however where Greedy always takes that solution and compares it to the best, Simulated Annealing does two things before that comparison.

Firstly as can be seen on lines 12 - 14 of Algorithm 3, it checks whether the new solution is better than the current one. If it is better, then it takes that solution for comparison to the best. However, if it is not better there is still a chance Simulated Annealing will take it for comparison.

In lines 16 - 19, a check is performed that allows worse solutions than the current to be checked against the best. This is usually called an "*Uphill move*" as, if ε values of solutions are plotted then better solutions are nearer to 0 and worse moves are further, meaning they will be moving upwards on the curve as can be seen in Figure 2.

³A full table of Big-O complexity for all algorithms can be found in the appendix in table 13

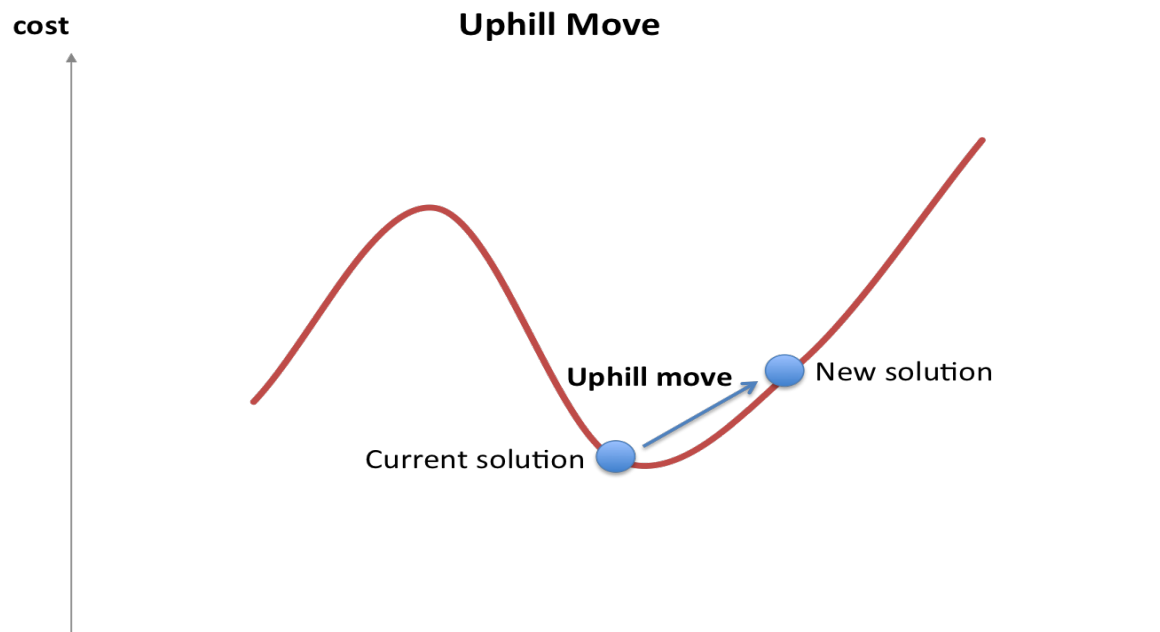


Figure 2: Uphill moves in optimisation problems

The reason for accepting a worse solution is because it allows for a more extensive search for the optimal solution. Line 17 is a feature that is designed to prevent the algorithm from becoming stuck at a local minima that is worse than the global one[15]. The local minima is represented by the bottom of the red curve in Figure 2. Moreover as the likelihood of accepting a worse solution varies with the value of t , it allows the algorithm to take worse solutions at the start, but not take the worse solution nearer the end, meaning the "*Uphill moves*" occur so that the algorithm does not get stuck in a local optima early.

Another part that is specific to Simulated Annealing is the concept of the cooling schedule. The values that make it up are ti , tl and $cr_coefficient$. ti is the initial temperature and should be high enough so that the final solution is independent of the initial solution, and so that it encourages a wide exploration of the feasible solutions at the start. tl is the number of iterations that occur at each temperature and the $cr_coefficient$ is the rate at which the temperature is reduced.[16]

An approximation of the worst case Big-O of Simulated Annealing is shown in Table 8.

Table 8: Simulated Annealing complexity per line

Line	Effect on complexity
T_1	run once with V steps
T_2	run once with $V \times P$ steps
T_3	run once with constant steps
T_4	run once with $V \times P$ steps
T_7	$m \times l$ times with constant steps
T_8	$m \times l$ times with $V \times P$ steps
T_9	$m \times l$ times with constant steps
$T_{10} - T_{19}^*$	$m \times l$ times with constant steps
$T_{20} - T_{23}$	$m \times l$ as it is the worst case, the assumption is that this block is always run
T_{26}	m times with constant steps
T_{27}	m times with constant steps
T_{29}	run once

* This if-else block can be treated as one because one of the two blocks will always be run

N.B The main loop (lines 5 - 28) is constant, given it comes from *num_iterations*. It adds m loops to the instructions inside it. The inside loop (lines 6 - 25) is also constant, given it comes from the temperature length, tl . It adds l times to the instructions inside it.

Factoring this down into equation 11 gives:

$$\begin{aligned}
 f(n) = & (V \times T_1) + (V \cdot P \times T_2) + T_3 + (m \times V \cdot P \times T_4) \\
 & + (m \times l \times T_7) + (m \times l \times V \cdot P \times T_8) + (m \times l \times T_9) \\
 & + (m \times l \times t_{10-19}) \\
 & + (m \times l \times T_{20-23}) \\
 & + (m \times T_{26}) + (m \times T_{27}) + T_{29}
 \end{aligned} \tag{11}$$

From this we can safely ignore anything that is less than the highest order as it will dominate the complexity growth. Moreover we can see that the highest order is T_8 , so we are left with equation 12

$$f(n) = (m \times l \times V \cdot P \times T_8) \approx m \times l \times V \cdot P \tag{12}$$

where $P \leq V \leq l \leq m$. As before this can be further simplified in terms of *voters* (V). In this case, there is an extra term with will be greater than or equal to V . This leaves us with a fair approximation of the complexity in equation 13.⁴

$$O(n) \approx m \times l \times V \cdot P \approx V \times V \times V \times V \approx O(V^4) \tag{13}$$

⁴A full table of Big-O complexity for all algorithms can be found in the appendix in table 13

A full comparison between the algorithms and their complexities is undertaken in Section 0.5.3.

Piecemeal

Part way through the project a slightly different approach to building a solution was suggested. The three algorithms discussed above all have a major thing in common, that they move from a **full** solution to another. It was suggested that another methodology for finding a solution could be used. The development of this algorithm was not directly due to the performance of the other algorithms. At this time in the project Simulated Annealing was performing very well; the decision to design this algorithm was more influenced by gaining a deeper understanding of the problem and the entertainment functions that had been designed. The Piecemeal method shares some of its theory with Constraint Satisfaction Problems (CSP's)[17] and more specifically backtracking algorithms used within them[18]. Backtracking algorithms contain the idea to incrementally build a full solution by trying different partial solutions.

This method begins with just one *voter* instead of a full solution. Then to pick the next *voter* to add to the solution, the algorithm looks at all the available *voters* it could add and it picks the one that gives the lowest overall Λ value if it were to be added. The Λ value is the distance between the best and worst who can still win. At the end this value is summed to get the solution's ε value.

This method hopes that by picking the best possible next choice this will lead to a good approximation or globally optimal solution. It acts very much like Greedy search that has been discussed, and the pseudocode supports that, however by working on partial solutions instead of full solutions, the idea is that it will at least approximate a good solution without having the need to check many full solutions, hence increasing performance, especially for large solution problems.

The pseudocode for it is shown in Algorithm 4 and the implementation is discussed in more detail in Section 0.4.6.

The main points of interest are that even though lines 7 - 9 look very similar to the three algorithms discussed before, there is a subtle difference. In this algorithm, the value that is being checked against the best is not entertainment (ε) but the distance (Λ).

Another point of interest is a more academic one. It is a question of choosing which *voter* to put at the start. The two most obvious choices are randomly choosing one, which complicates the code slightly, or picking the first from the given list of voters, making the code simpler. The choices of first *voter* and their effect on the ε returned is explored more in Section 0.5.2.

To come to a solution, the algorithm must look at a descending number of possible next solutions from the length of the solution down to 1. In the 2014 Eurovision, this would be 37 choices in the first round, 36 in the second, 35 in the third etc. until all *voters* are accounted for.

The formula in equation 14 describes the worst case that this algorithm will take. This can be represented as a triangular number[19] as it is the sum of the natural numbers from 1 to V .

$$O(n) = \sum_{k=1}^V k = V + (V - 1) + (V - 2) + \dots + 1 \approx \frac{V(V + 1)}{2} \quad (14)$$

which can then be simplified for Big-O notation into equation 15. This is because for looking at the complexity of algorithms it is most interesting to look at the term that does the most to the complexity, so in this algorithms case it is the V^2 term.⁵

$$O(n) = \frac{V(V+1)}{2} = \frac{V^2 + V}{2} \approx V^2 + V \approx O(V^2) \quad (15)$$

Looking at the Big-O complexity in equation 15 and comparing it to the others is done in Section 0.5.3.

One thing to note about this algorithm is that when the number of solutions to search is relatively low, this method is very good at finding the best solution. Moreover in smaller examples, this method tends to perform better than Simulated Annealing and Greedy. If the number of possible first choices is relatively low then this algorithm can try them all, hence it can always find the best solution that it could possibly find. This is shown by the fact that running against the small example introduced at the beginning of this section this algorithm will always find the best solution whereas the others find near best solutions. Table 9 shows the values found by the three main algorithms for the small example. As was discussed in Section 0.3.1 the optimal solution has a ε value of 4.

Table 9: ε values found for small example game

	ε value
Greedy	5
Simulated Annealing	5
Piecemeal	4, 4, 4, 4

Piecemeal produced 4 solutions (one with each *voter* as first choice) all with ε values of 4, whilst the other algorithms found a near optimal solution of 5 from a random initial solution.

However if the search space is increased to 5 *voters* and 7 *participants* such as in the game described by table 10, we find that Piecemeal begins to show signs of failures.

Table 10: Scores for simple example game

	Round 1	Round 2	Round 3	Round 4	Round 5
Federer	0	0	2	0	6
Murray	1	1	3	6	5
Djokovic	2	2	4	4	4
Nadal	3	3	5	5	3
Borg	4	4	6	2	2
Cash	5	5	1	3	1
Henman	6	6	0	1	0

The ε values found by the algorithms are shown in table 11, and we can begin to see that, even though Piecemeal is still beating Greedy and Simulated Annealing, it is also finding solutions similar to the other algorithms. It is not too difficult to see that for a competition as large as Eurovision, with 37 *voters* and 26 *participants*, Piecemeal could begin to fail to find the best solution.

⁵A full table of Big-O complexity for all algorithms can be found in the appendix in table 13

Table 11: ε values found for bigger example game

	ε value
Greedy	31
Simulated Annealing	31
Piecemeal	19, 19, 27, 27, 19

The hypothesis behind this theory is that as the number of choices this algorithm has to make increases, there is more opportunities for it to fail to choose the best next *voter*, hence making it fail to find the overall best solution.

0.4

Implementation of System

In this section, the main important parts of the project's code will be explored. This will take the theory and pseudocode described in Section 0.3 and discuss the actual code implemented as part of the project. It will both discuss some specific code as well as a discussion as to why parts of the system were implemented in certain ways. All code referenced in this section can be found in the attached source code files, submitted along with this report.

0.4.1 ε function wrappers

The ε functions form one of the most important parts of this project, however for them to be implemented they need some contextual data. The functions described below are larger functions that wrap around and allow the ε functions to do their work. The ε functions actually form just the $\min(S)$ part of equation 1.

These wrapper functions share an important part, which is shown in 3.⁶

Figure 3: Calculating the cumulative scores per participant

```

1 for j in range(len(solution)):
2     for i in range(len(countries)):
3         v = voters.index(solution[j])
4         scores[i] = scores[i] + score_board[i][v]
5     ...

```

The main point of interest is the two nested *for* loops and specifically lines 3 and 4.

As the distance (Λ) is found as the difference between the max and min scores in each round, it is necessary to go through and cumulatively add the scores given to each *participant* in every round. This is stored in the array *scores* at the index for that *participant*: *scores[i]*. The *score_board* is the matrix of all scores given, as discussed in Section 0.2.

⁶Section of code from **getEntertainment** in *support.py*. Full function can be found in attached source code

These lines are interesting to highlight as they are likely the most critical lines in all the system. Every algorithm must use them at some point and all ε functions will need them to calculate the ε value.

As the score lookup is done against a matrix of scores, it is necessary to find the correct column to look in. The row is the current *participant*: i , however finding the column is a little more complex.

The columns in the matrix are in a certain order, alphabetical in the case of the 2014 Eurovision dataset. This, along with the fact that each solution is a permutation of that order, means that the currently voting country in the solution, $solution[j]$, must be found in the original order to correctly retrieve the score it gives. By keeping track of the original order as it corresponds to the scoreboard, the system only ever has to keep one matrix in memory, along with 2 lists (of *voters* and *participants*). Any other method would include shuffling the scoreboard into the order of the current solution. As the scoreboard grows i.e in other, larger problems, this method would become untenable and possibly begin adversely affecting the algorithm's runtimes and complexities.

The method `array.index(element)` returns the index in the given *array* of the given *element*. In this case $solution[j]$, would be a country name such as "Germany", and the value of v will be an integer corresponding to the column.

Looking at the full code listings there are actually two methods that wrap around ε calculation functions to provide them with context and data. These are called **offsetGetEntertainment** and **getEntertainment**, and they both share the code shown in Figure 3. They are also fully independent of the **maxMin** and **refinedMaxMin** functions which are the actual ε functions. These methods differ because of when and why they can be used.

The main reason for these differences is shown in Section 0.3.2, as using different neighbourhood methods leads to different ways to calculate the entertainment.

offsetGetEntertainment

Firstly the more complex **offsetGetEntertainment**⁷ is used in conjunction with adjacent neighbour swaps as it relies on using the previous ε value. The theory behind why this works is explained in Section 0.3.2, specifically the Adjacent neighbours subsection and equations 5 and 6.

This function is shown in Figure 19 in the appendix.

The important parts to understand about this implementation are, that for it to work it must be given three things. Two keys that correspond to the indexes of the *voters* that were swapped to get this solution; the ε value of the old solution, and the array of the distances used to calculate that solution's ε value.

One obvious point of difference is between the *for* loops in Figure 19 and Figure 3. In the **offsetGetEntertainment** method the *for* loops only need to calculate the scores up to the higher of the two keys as the scores after the swaps are unchanged.

The actual recalculation described in Equation 6 is implemented in line 20. On lines 3 and 4, a copy of *countries* and *solution* are taken using Python's array copy syntax. This is most likely not really necessary but during implementation there was a couple of times where after

⁷Can be found in *support.py* in source code

amending one of the lists locally inside a method, the global version of that list that should have stayed the same, was also changed. This small problem is further discussed in 0.5.4.

Lines 22 and 23 are necessary so that the new distance values can be passed to this function again, with the distances in the correct positions. Essentially what this does is constantly pass around a list of distances, amending it by swapping values before passing it on.

offsetGetEntertainment is certainly faster and more performant for most solutions. This does depend on where in the solution the swaps have taken place (explored in Section 0.5.3.) The main drawback of this method is that it can only be used with adjacent neighbourhood and not the random neighbourhood method.

getEntertainment

The second entertainment function wrapper is simpler than **offsetGetEntertainment**. **getEntertainment**⁸ is the general purpose wrapper for calculating the entertainment of a given solution. It can work with either of the two neighbourhoods described. This makes it more general purpose, however it also makes it slower especially when the solution size is large.

The points of interest here are found in lines 13, 14 and 15 of Figure 20. These lines implement that which can be found in Equations 1 and 2 exactly as described. Line 13 finds the difference between the max of the scores and whatever min was returned by the current ε function. Line 14 adds the Λ value found to an array and then line 15 sums all the Λ values found to finally get a ε value. To allow it to interact with the **offsetGetEntertainment** method it also keeps track of the array of Λ values and returns them so they can be passed to **offsetGetEntertainment** in the next iteration.

Some test that compare and contrast these two methods can be found in Section 0.5.3.

0.4.2 ε functions

The **refinedMaxMin**⁹ method is the more interesting of the two ε functions to look at in depth. The original **maxMin**¹⁰ method is very simple and uses the Python in-built *min()* [20] method. Both methods use the *max()* [21] method from Python.

As can be seen in Figure 20 on line 12, the **refinedMaxMin** method is invoked, with the current list of scores for all *participants*, the current solution being tested and what round of voting the competition is at, *j*. The *maxScorePerRound* is used so that different competitions can define what it is instead of keeping it locally in the function allowing this method to be easily re-used.

The function is shown in Figure 4

⁸Can be found in *support.py* in source code

⁹Can be found in *support.py* in source code

¹⁰Can be found in *support.py* in source code

Figure 4: refinedMaxMin method

```

1 def refinedMaxMin(scores, solution, j, maxScorePerRound):
2     sorted_scores = sorted(scores[:])
3     currentTop = sorted_scores[-1]
4     minScore = sorted_scores[0]
5     del sorted_scores[-1] # remove highest score
6     roundsRemaining = (len(solution) - 1 - j)
7
8     for score in sorted_scores:
9         if score + (maxScorePerRound * roundsRemaining) <= currentTop:
10             minScore = score
11     return minScore

```

The theory behind this method was described in Section 0.3.1. Some of the implementation points to pick up on are, line 1 where the in-built Python function *sorted()*[22] is used to take the list of scores and sort them in ascending order. This does slow the method slightly, taking the method's complexity to $n \log n$ but it also simplifies the logic later on as the method does not need to keep track of anything other than the minimum found and also only needs to update it.

This implementation is likely not the most efficient way of finding the refined version of the minimum, however it is quite a simple method. Refinements to the performance are discussed in Section 0.6.

An important part of the two methods is their Big-O complexity. The complexities are shown in Table 12. The main point to recognise is that the methods are different in terms of complexity which can help in the evaluation such as in Section 0.5. The refined method is actually very similar to **maxMin** except for the sorting step which increases the complexity.

Table 12: Big-O of ε methods

	Big-O time complexity
maxMin*	$O(n)$
refinedMaxMin	$O(n \log n)$

* Complexity of Python's *min()* function from[23]

More in depth comparisons between the two methods can be found in Section 0.5.3, which take into account the complexity and the minimum scores that the methods produce.

0.4.3 Neighbourhood functions

The random and adjacent neighbourhood methods, **getNeighbour** and **getAdjacentNeighbour** are very similar to each other as can be understood from the theory in Section 0.3.2. Both can be found in *support.py* in the attached source code. This does mean they could be written as one single method which are supplied the keys needed to perform the swaps. This approach *was* given some thought during their implementation however in the end it was decided to favour duplication over complexity in the system. This means that there are two methods that are for all intents and purposes exactly the same, but are named differently.

The methods are shown in Figures 5 and 6.

The only difference occurs on lines 3 and 4 in the methods. Firstly in the adjacent neighbour method (figure 5) the first key is found by getting a random integer between 0 and 2 elements from the end of the list. It must be 2 elements from the end as the second key is taken as the first key + 1. In terms of a list this means the element directly to key1's right. The `len(neighbour)` part is using Python's in-built length function to calculate the length of the list and hence the last index it can return for correct swapping.[24].

The random neighbour method uses the Python *random* library in line 3 of Figure 6 to sample the given order and get two random elements back.[25]. As this method actually returns the elements and not indexes of those elements, like in the adjacent method, it is necessary to change them back into indexes (line 4) for simple swapping in line 5.

Figure 5: getAdjacentNeighbour method

```

1 def getAdjacentNeighbour(xNow):
2     neighbour = xNow[:]
3     key1 = random.randint(0, len(neighbour) - 2)
4     key2 = key1 + 1
5     neighbour[key2], neighbour[key1] = neighbour[key1], neighbour[key2]
6
7     return neighbour, key1

```

Figure 6: getRandomNeighbour method

```

1 def getNeighbour(xNow):
2     neighbour = xNow[:]
3     key1, key2 = random.sample(list(neighbour), 2)
4     index1, index2 = neighbour.index(key1), neighbour.index(key2)
5     neighbour[index2], neighbour[index1] = neighbour[index1], neighbour[index2]
6
7     return neighbour

```

0.4.4 Simulated Annealing and Greedy search

The implementation of the search algorithms is not the most interesting part of this project and hence only a little time will be spent explaining their implementation. Both Simulated Annealing and Greedy Search can be found in the appendices, and in the attached source code in the files *simulatedAnnealing.py* and *greedySearch.py*.

The main reason for not going into depth about their implementations is that both algorithms are essentially general search algorithms that receive the problem specific input from the other functions already described such as **getEntertainment** and **getAdjacentNeighbour**.

A point that is shared across both implementations is the increasing while loop that gets reset to 0 when a new solution has been found. Using Greedy Search as an example, in Figure 22, the while loop starts at line 10 with *i* equal to 0. It is incremented on line 22 after an iteration and comparison of a new solution. On line 18, the value of *i* is reset to 0. This allows

the *num_iterations* loop to be the number of iterations since a new best solution was found instead of a hard number of loops to go through. This minor implementation detail actually has a very measurable impact on the speed of the algorithms. Due to the nature of the problem space, it is likely that both Greedy search and Simulated Annealing reach local (possibly global) optima. When this happens, they are not going to choose a new solution. Furthermore having the *num_iterations* variable set as a hard number of iterations would require a very deep knowledge of the problem space and how the algorithms are traversing it; so that there is little overhead after they have found the best solution they are going to find. This is not really possible so this method also allows for slightly more flexibility.

Another implementation point that is shared is the fact that both algorithms must pass the *oldEntertainment* and *oldDistances* into the **offsetGetEntertainment** method. The reason for this is explained in Section 0.4.1. This also explains the need for the variable *key1* to be returned from the **getAdjacentNeighbour** method.

A minor point that is related to Python as a language is the ability to return multiple values from a function in a tuple by default. This hugely simplifies the code as methods can then for example return both a neighbour and the first key used to find it (**getAdjacentNeighbour**) or a possible solution and the individual distances that make up that solution (**offsetGetEntertainment**). In other languages there are ways to work around this problem, but this solution is one simple way in Python.

Specifically looking at the Simulated Annealing code in Figure 21, the main lines of interest in terms of implementation are line 32 and 34. Line 32 uses the Python *random*[26] library to get a random number between 0 and 1. It then uses the Python *math*[27] library to do the $e^{-\text{delta}C/t}$ part that is necessary for Simulated Annealing to work. These again, are two examples of the Python libraries making implementation very simple for this project.

The general cooling ratio values which were explained in Section 0.3.3 on Simulated Annealing, were experimented with throughout the project. The final values are shown in Table 14 in the appendix and how they are implemented is quite self explanatory.

0.4.5 Brute force

The implementation for the Brute force¹¹ algorithm was one of the most simple, and this has much to do with the libraries that Python provides. The library used for the Brute force implementation is *itertools*[28], which provides a set of common methods such as *permutations*[29]. This makes this implementation extremely simple as I did not have to write any code to correctly permute through the possible solutions without repetitions, as the library handled it for me. It can be seen on line 5 of Figure 7.

¹¹Can be found in *bruteForce.py* in the attached source code

Figure 7: Brute Force algorithm implementation in Python

```

1 def bruteForce(score_board, countries, voters):
2     best = voters[:]
3     bestCost = support.getEntertainment(voters, countries, score_board, voters)
4
5     for solution in itertools.permutations(voters):
6         currentCost, dist = support.getEntertainment(solution, countries, score_board,
7             voters)
8
9         if currentCost < bestCost:
10             best = solution[:]
11             bestCost = currentCost
12
13     return best, bestCost

```

The implementation shares many similarities with the other algorithms described so far. Namely, the if block on line 8 - 10 that checks if the current solution is better than the best found so far. Furthermore it uses the same **getEntertainment** method that is used by both the other search algorithms described so far. One small but important point is that, in contrast to Greedy and Simulated Annealing, during the main for loop (lines 5 - 10) in Figure 7, Brute force cannot use the **offsetGetEntertainment** method. This is the one drawback of using the *itertools* library. As it controls getting a new solution to try, the key needed for **offsetGetEntertainment** to work cannot be passed around. This further slows this Brute force implementation.

0.4.6 Piecemeal

The Piecemeal algorithm is the only break from the traditional style of search algorithms that have so far been discussed. Moreover it is the most novel approach to solving this problem in this project.

The theoretical differences were discussed in detail in Section 0.3.3. The code for this algorithm is shown in Figure 23 in the appendix and can be found in *step.py* in the attached source code.

The subtle difference in terms of comparing Λ values instead of ε values that was discussed in the design of this algorithm can be seen on lines 31 - 34.

Secondly, the concept of *ignoreIndexes* in lines 23, 24 and 35 is necessary so that the correct score is taken from the scoreboard. As the value of j is used as an index in the scoreboard matrix, it is necessary to ignore those indexes of countries that have already been added to the solution. As was discussed in Section 0.4.5, in the Brute force method Python took care of not having repeats, however in this method as the algorithm is slightly more complex and Python in-built methods cannot be used, this must be done manually.

0.4.7 Re-Use of Code for future work

One of the more general points that relates to how this project was implemented is the possibility of re-use of the work. Quite early in the project, when discussing other problems that are

similar in type to Eurovision, it became clear that the code that was being written would be quite general to all of these problems. Furthermore there was no established framework for this sort of problem which meant this implementation would have to do that job anyway. This led to the attempts to make all the implementations as re-usable as possible.

This can mainly be seen in that all loops have an upper bound that depends on the length of the lists given and do not use magic numbers that may change. For example Figure 20 on lines 8 and 9.

This is also seen when looking at the controlling file *order.py* in the attached source code. There are 4 inputs that all the algorithms need to run. These are: a list of *participants*, a list of *voters*, a scoreboard (which is a matrix of scores that *voters* give to *participants*) and finally the max score that can be given by each *voter* (12 in Eurovision). If a problem can be formalised into these 4 inputs then these methods and algorithms can quite easily be used for other problems, no matter their complexity or size. The supporting files that implement the entertainment and neighbourhood functions could also simply be swapped out for problem specific ones if needed.

The Eurovision Song Contest group released spreadsheets for the 2016, 2015 and 2014 Semi and Grand Finals. To allow the scoreboard to be taken directly from these results it was necessary to convert these spreadsheets to csv's first. This was done for the 2014 file and is attached. To read this into a matrix scoreboard for use by the algorithms a csv reader was written which can be found in the attached source code as *csvDataReader.py*.

As the spreadsheet and hence csv were not optimised for use by computers there were some modifications to the incoming data that had to be made. One such modification was to change new lines (`\n`) at the end of each row to a score of 0 points. This was due to the fact that the last element on each row was the composite score, and if a team was given 0 points, then the cell was left blank. When the spreadsheet was turned into a csv the blank cells became newlines. Another was to track who was currently voting and add an extra 0 points when they would have voted for themselves as the spreadsheet did not include the country that was currently voting in the list of countries that received votes. Once run it produces a matrix that can then be given to the *order.py* file and then used to solve the problem. Hence allowing the re-use of the algorithms and supporting functions on another Eurovision competition with relative ease.

0.4.8 Visualisation

A part of this project that was within scope, but not directly part of solving the problem, was creating a way to visualise the solutions that the algorithms produced. The main reason for doing this was to have a way to compare solutions that more closely correlates with how the real solutions would be viewed.

My idea for this was to produce a simple web application that could be given a competition and the solution found by one of the algorithms and then "play out" the competition for a viewer. A screenshot of the application is shown in Figure 8.

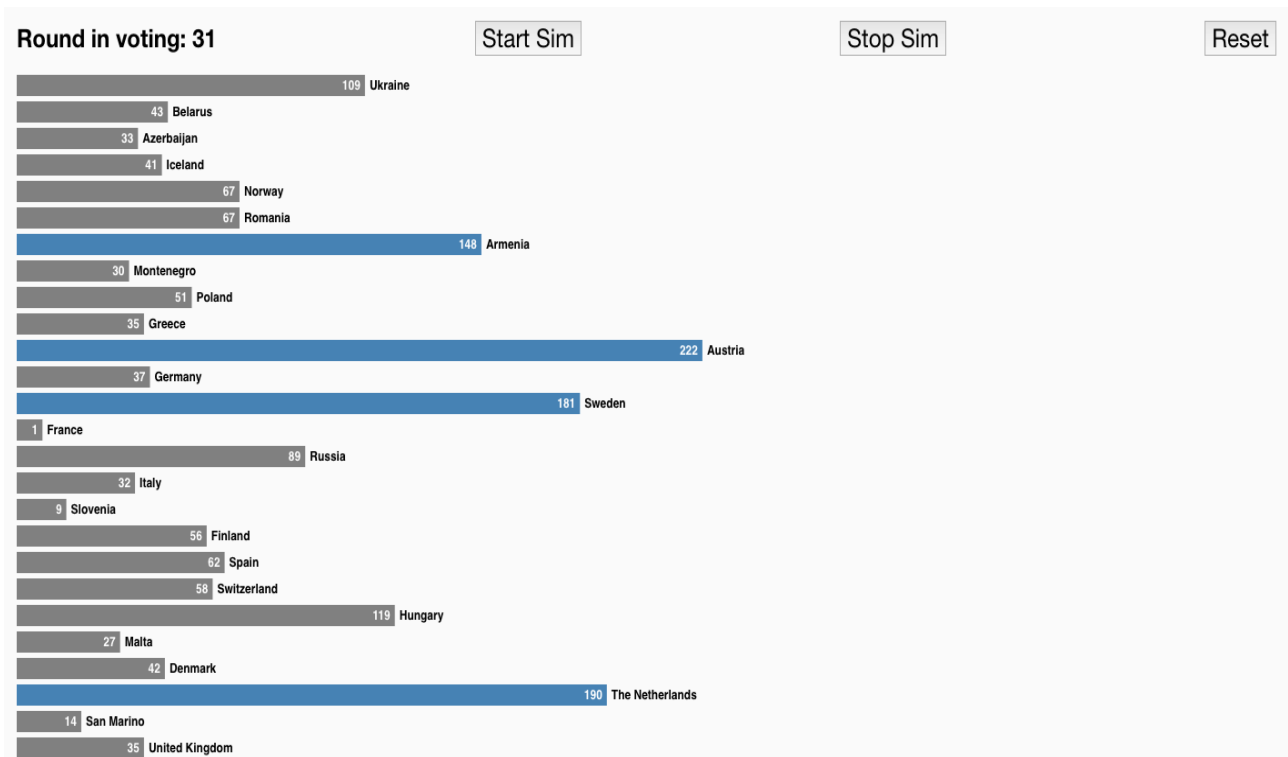


Figure 8: Visualisation of solution to 2014 competition

The application is a Node.js application that uses two main libraries for rendering the chart seen in Figure 8. The first is *d3.js* which is needed to scale the bars to the correct proportions as the scores are re-calculated. This uses the *d3-scale* library[30] instead of using the whole of *d3*. *d3* could have theoretically taken care of updating of scores each round, however I was not extremely familiar with the way in which it updated state. Moreover for this project I wanted to produce something rather than getting stuck in the implementation details of learning a new library, and the more important part of this project was to produce solutions.

For this reason I turned to *Preact.js*[31], which is a smaller and faster version of the currently very popular library *React*[32] from Facebook. I chose *Preact*, as I have experience with *React* however I did not need the large number of functions that *React* provides and I needed the state changing to be fast, as there are calculations that need to be done from round to round. Moreover to speed up starting up the visualisation, the code is based on a *Preact* example application.[33].

The visualisation is very simple, the user clicks a button to begin the competition voting and from there the scores are updated and shown on screen. A timer is used so that only one button press is needed to start the voting and from then on around every 1.5 seconds the new scores are shown, until the end of the competition. Moreover when a team cannot win, the colour of their bar is changed to grey, to help show who can still win. This provides anyone viewing the voting with a simple way to see how close certain teams are to each other and the current leader, and hence why the algorithms may have settled on this order. A further goal of producing this visualisation, which will be explained in Section 0.6, is using this visualisation as a tool for discovering new entertainment functions for this and other problems.

As with the other code in this project I have attempted to make this as re-usable as possible. It takes the same types of inputs as the algorithms need, and any work to modify the code should only be purely cosmetic, such as changing names of *voters* or *participants*.

The visualisation code can be found in the *visualisation* folder in the source code, along with explanations on how to run it to view it locally.

0.5

Results and Evaluation

0.5.1 Overview of goals

The main barrier to saying that this project achieved all of its stated goals is the fact that the optimal solution is not known. The goals set out in Section 0.1 were:

1. a solution that delivers an entertaining competition when the votes are revealed
2. a way of describing entertainment mathematically
3. a way of visualising the competition as the votes are revealed.

As it is the most important of the three goals, the goal of finding entertaining solutions is explored below in Section 0.5.2 in more depth.

Describing the entertainment of a given voting order mathematically was explored theoretically in Section 0.3.1 and their implementations in Section 0.4.2. In terms of success in achieving the goal, I think it was very successful. The main entertainment function (**refinedMaxMin**) provides a simple and elegant way of enumerating a concept that can be quite difficult to define. The functions not only define an entertaining solution mathematically, but they also allow them to be maximised or minimised easily, so in this regard they are also very accomplished in helping to achieve the first goal. A comparison between the two functions defined during the project and an explanation as to how they were tested against each other can be found in Section 0.5.3. Moreover this will also provide more support as to why I believe this goal has been successfully reached.

It is more difficult to quantify if the goal of creating a way to visualise solutions has been met or not. In my opinion, the visualisation web application does a very good job of allowing another dimension of this project to be shown. The side of the project that the visualisation helps to shed light on, is essentially ignored by the algorithms and only generalised heavily in the entertainment functions. That is how real people would feel when viewing an "entertaining" solution. The visualisation provides a more visual style of feedback to the viewer. Furthermore it helps to justify the other goals without having to go into as much detail about the correctness of the algorithms or the entertainment functions, such as is done in this report.

All raw results can be found in the *results* folder in the attached files.

0.5.2 Solutions

This section will discuss both the general solutions found by all the algorithms and compare how they performed in finding solutions. Moreover the methodology for collecting and analysing these solutions will be explored. A table of the best and average solutions along with average times for each of the four algorithms can be found in table 15 and table 16 in the appendix respectfully.

Methodology

To find a set of solutions for each algorithm in order to analyse them, the first step was to do some repeated runs. To facilitate this some code was added to the main controller *order.py* so that an optional command line parameter could be given which would repeat the running of the given algorithm how ever many times the user input.

The general method was to run each of the algorithms lots of times and record the Φ and ε values for that solution. From this we can be more confident about the performance of each algorithm and make some mathematical comparisons between them. To try and keep the experiment fair all the algorithms that take an initial solution were given the same one, that is the one given by the method *getInitialSolution()* in *support.py*.

An important point to recognise, before explaining the basic methodology, is that only testing of Greedy search and Simulated Annealing can be done in the exact same way, due to their similarities. Piecemeal and Brute force cannot be tested in the same way however the same idea of running them many times to collect many solutions still stands. Their methods are explained later on.

The experiments were designed to give each of the algorithms as much of a chance as possible to perform at their best. This meant that I attempted to keep the surrounding computation at a minimum so as not to add to the load on the CPU.

The experimental method for both Greedy search and Simulated Annealing are the same, apart from changing filenames. It begins with running a command in the terminal similar to that in Figure 9.

Figure 9: Running algorithms

```
>> python order.py simAnnealing 140 >> results/simAnnealingResults.txt
```

This command needs to be run in the same directory as the controller *order.py* and the algorithm that is being run. In this example this command will run the Simulated Annealing algorithm 140 times and then send the output to a file called *simulatedAnnealingResults.txt* in a *results* subdirectory.

The `>>` part is a **bash** command that means append the output of the preceding command into the file after it. Append is used as it was necessary to run this code in smaller chunks to reach 1000 total runs. So as not to create a new file each time, we append from wherever we left off. Saving the output could have been achieved quite easily from within Python, however I felt that it was a sound idea to take the saving of the output away from Python as the terminal commands are usually faster and, in my opinion safer to use. They are also much more simple to write and have very little room for error. Moreover, when doing the timing experiments I wanted the least amount of overhead so that as little external time was spent on other computations that could accidentally leak into the times for the algorithms.

After doing this I was left with a set of files that contained both a Φ and an ε value per line. These can be seen in the attached results files such as *greedyResults.txt*. A snippet of this file is shown in Figure 10.

Figure 10: Snippet of Greedy results

```
...
(['Russia', 'Portugal', 'Ukraine', 'Poland', 'Sweden', 'Switzerland', 'Montenegro',
 'Romania', 'San Marino', 'Italy', 'Latvia', 'Malta', 'The Netherlands', 'Slovenia',
 'FYR Macedonia', 'Ireland', 'Hungary', 'Iceland', 'United Kingdom', 'Lithuania',
 'Georgia', 'Norway', 'Germany', 'Greece', 'Israel', 'Spain', 'Armenia', 'Estonia',
 'Austria', 'France', 'Finland', 'Belgium', 'Denmark', 'Albania', 'Belarus', 'Moldova',
 'Azerbaijan'], 3409)
(['Italy', 'Russia', 'Ukraine', 'San Marino', 'Latvia', 'Norway', 'Israel', 'Belarus',
 'Slovenia', 'United Kingdom', 'Poland', 'Portugal', 'Romania', 'Georgia', 'Albania',
 'Denmark', 'Montenegro', 'Lithuania', 'Switzerland', 'Austria', 'The Netherlands',
 'Germany', 'Azerbaijan', 'Armenia', 'Ireland', 'Spain', 'Greece', 'Moldova',
 'Hungary', 'Finland', 'Belgium', 'Iceland', 'FYR Macedonia', 'Sweden', 'France',
 'Malta', 'Estonia'], 3298)
...
```

Although the Φ are very interesting in their own right, for them to make more sense and the find the best, it was necessary to extract the ε values from this file so that some analysis could be done. To do this I used *ack*[34], which is a tool that allows for searching of plaintext files for lines that match a given regular expression. This allowed me to extract just the scores from the files like *greedyResults.txt* and then send them to a file such as *greedyScores.txt*¹². The command for this is shown in Figure 11.

Figure 11: Extracting scores from results file

```
>> ack -o "(?<=,\s)\d{4}(?=\))" algorithmResults.txt > algorithmScores.txt
```

The *ack* tool uses Perl regular expressions for matching the given expression. In this case the regular expression is very simple but looks complex. The `\d{4}` part matches 4 of any digit. The parts before in brackets is a *positive lookbehind* and after in brackets is a *positive lookahead*. These match something before or after the digits but do not keep them in the output; and are used to find a comma and a space before, and the final bracket after the score. The `-o` part is a flag that is used to only print whatever is matched, otherwise the whole line where a match was found would be printed and there would be no difference between the results and score files. Once again we use the **bash** commands to send the output to a file, however this time we do not append, but create the file and write to it immediately.

After doing this we have files that contain 1000 ε values, which can then easily be copied into Microsoft Excel for analysis.

When running the two search algorithms there are some setup values that are needed. The main one is *num_iterations*, which both Simulated Annealing and Greedy Search use. To try and keep the tests as close to each other the same value was used (*num_iterations* = 500).

Specific to Simulated Annealing's case there are some more input values that are needed, and they were discussed in Section 0.3.3 in the pseudocode. The final values are shown in Table 14 in the appendix. To reach these values I took into account more general wisdom for each one as well as trying to apply more problem specific knowledge to reach the values that give the best solutions. Furthermore, the time taken to find a solution could be heavily effected by the

¹²*greedyScores.txt* can be found in the attached code and results.

values so a balance was struck so that the algorithm did not take too long, but also produced consistently good ε values.

The methodology for the Piecemeal algorithm was slightly different. This is due to the fact that given the same first choice the ε value will always be the same. The problem of which *voter* should be first, was brought up in Section 0.3.3. It was decided to try and find the best solution by trying all the possible first choices and saving them all. This takes away the random element but still allows the algorithm to return it's best possible solution.

To do this it was necessary to wrap the Piecemeal code as shown in Figure 23 inside a *for* loop and instead of hard-coding the first element on lines 10, 11 and 12, using the index from the loop.

Once again to extract the scores I used the same *ack* command from Figure 11 to find the scores and send them to a file.

As the Brute force algorithm can not be run in the same way as the other three, a different method was devised to test it. The method was to leave it running for 1 hour. This is a fair time to run it for, as will be discussed in the timing part of this section. The other algorithms produce solutions consistently in under 20 seconds. As before the experimental conditions were as similar as possible to the other tests. In this case all the solutions found were saved and the best solution found was the last one found after 1 hour.

Solutions found

The best Φ for the 2014 Eurovision was found by Simulated Annealing with a ε value of **2554** (using **refinedMaxMin**). Although it is difficult to say with certainty if this ordering is the *globally* optimal solution for the 2014 Eurovision competition, I believe that we can say with a high degree of confidence that this solution is at least close to optimal. This suggests that in terms of achieving the goal of finding a solution that delivers an entertaining competition, this project has been successful. This is true because even if the best solution found is not the "*optimal*" one, it is still very much an "*entertaining*" one.

Figure 12: Best solution

```
['Belarus', 'Albania', 'Poland', 'United Kingdom', 'Montenegro', 'Armenia', 'Malta',
'Russia', 'Azerbaijan', 'Germany', 'San Marino', 'Italy', 'FYR Macedonia',
'Moldova', 'Estonia', 'Austria', 'Romania', 'Switzerland', 'Ukraine', 'Latvia',
'Denmark', 'Georgia', 'Hungary', 'Finland', 'Ireland', 'Norway', 'Greece',
'Spain', 'Israel', 'Portugal', 'Lithuania', 'France', 'Belgium', 'Iceland',
'Sweden', 'Slovenia', 'The Netherlands']
```

To try and verify if this was the most optimal solution that could be found, this Φ was given to both Greedy and Simulated Annealing as the initial solution, and they were run around 10 times. They both returned this same solution, which gives good confidence that this is the best possible Φ .

Table 15 in the appendix shows some of the other interesting ε values found for all the algorithms used. The Φ corresponding to those values is not shown as they are very long, however finding them is a case of searching the results file for the ε value as they are saved together.

Figure 13 shows Simulated Annealing, Greedy and Piecemeal with their minimum, average and maximum ε values so they can be compared.

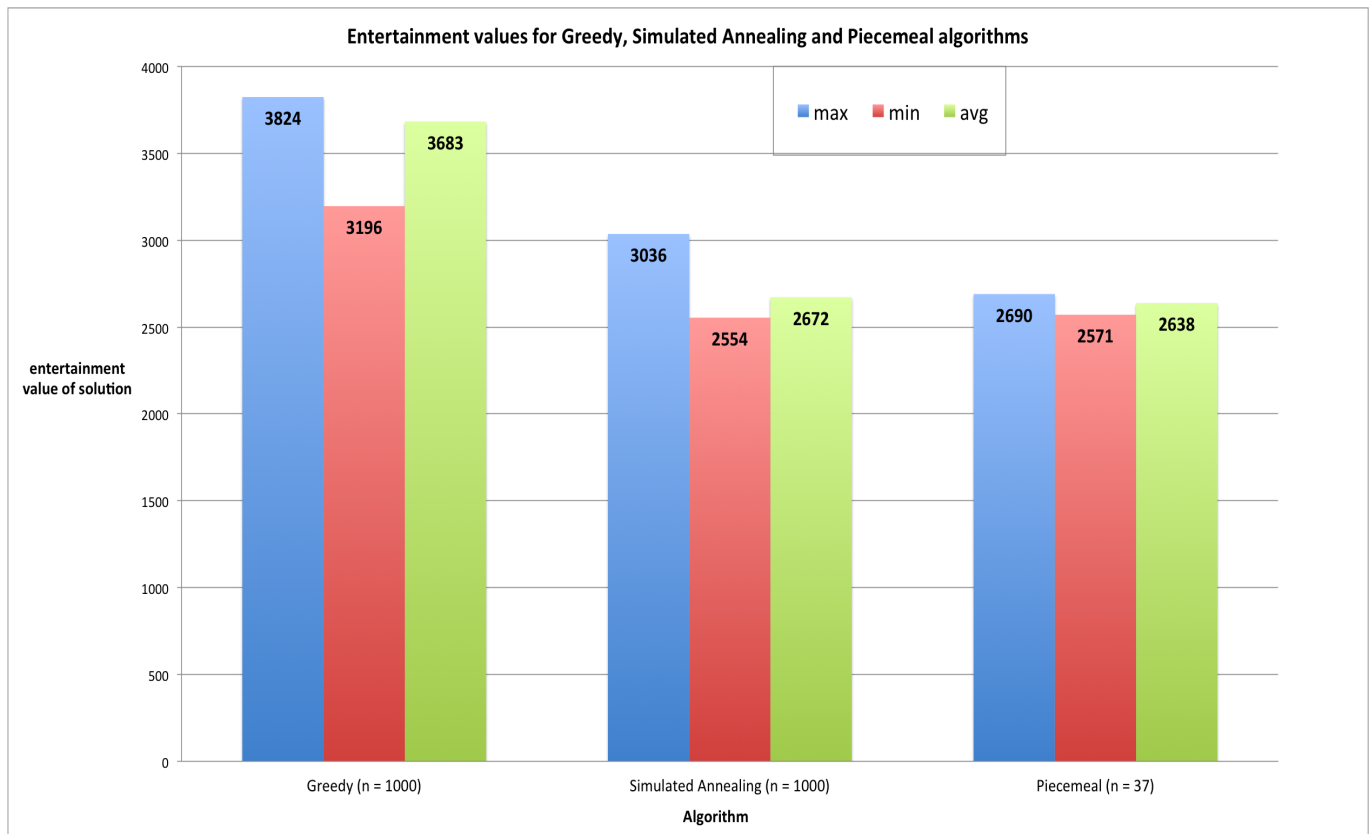


Figure 13: Chart of ε values found by Greedy, Simulated Annealing and Piecemeal

In terms of comparing the algorithms the first point is that Simulated Annealing is the best algorithm as it found the best ε value and hence solution as discussed above. This is in terms of ε value found only, and not with respect to any other measure of quality such as speed. How they compare in other regards is discussed in Section 0.5.3.

The Greedy algorithm does not perform particularly well as its average ε is much higher than the other two. Moreover the best Φ it found has a ε value of **3196**, which is about 150 worse than the worst ε value that Simulated Annealing found. The large range of values as seen from the large difference between the max and min means that Greedy is not a consistent algorithm when given consistent input, as in this problem. This backs up the idea that led to the development and use of Simulated Annealing and Piecemeal for this project as Greedy did not find very optimal solutions. If the results produced by Greedy were better, then there would have been little gained from the rest of the project.

One interesting point that is not shown in Figure 13 is that if we use a solution given by the Piecemeal algorithm as the initial solution, then Greedy *can* perform as well as Simulated Annealing (given the same initial solution as before). This increase in quality when given an already good Φ is shared by Simulated Annealing. This leads to the conclusion that Greedy is getting stuck in local minima and as it has no means of escaping them, returns the best Φ that it has.

The Piecemeal method was a major success for this problem, especially as it is not a general algorithm like the others, which have been proven and tested on many problems. Referring to Figure 13 the chart shows that the Piecemeal method held its own against a much more complex and established algorithm in Simulated Annealing. The lowest ε value that it found was **2690** which is only 17 worse than the globally best solution found. Moreover as can be seen from the range of values, this method is consistent in returning good solutions, when

varying the first *voter*. This may be because the only difference between the attempts is the first voter, however that is the only part that is changeable between runs. As the averages of both Simulated Annealing and Piecemeal are very similar we can say that Piecemeal is almost as good as Simulated Annealing at finding consistently good solutions.

The results in this section provide evidence for the theory that was put forward in Section 0.4.6, where it was suggested that the Piecemeal method would fail to find the best solution for larger problems. It was shown in that section that Piecemeal handily beats the other two search algorithms on small examples but begins to deteriorate as the problem space grows. The Eurovision dataset is a large problem, and Piecemeal did indeed fail to keep up it's high quality results.

0.5.3 Testing

The entertainment of a solution is not the only thing that gives it value. This is why some other testing was undertaken to try and give a more well rounded picture of how well the algorithms performed on this problem. Taking all the tests executed together gives a higher level of confidence in the results as there is more experimental data that can then be leaned on when evaluating the project as a whole and when making future decision about algorithm choice for other, similar, problems.

Timing and Big-O Complexity

As well as producing results that maximise entertainment, it is necessary for the algorithm to do so in a reasonable time. Table 16 contains some results for timing of the algorithms.

The methodology for analysing the timing was to run each of the algorithms 100 times each and have timing set to show how long it took for the function that runs the algorithm to finish. The timing is done using the *timeit*[35] library in Python, which purports to take care of many common pitfalls when timing code. It should be advised that these timing values are not necessarily the best, which is why they are averaged over 100 runs.

Figure 14: How to time the algorithms using Python

```
from timeit import default_timer as timer
...
start = timer()
print(gs.greedySearch(SCOREBOARD, PERFORMING_COUNTRIES, VOTING_COUNTRIES, 12))
end = timer()
print('algorithm took: ', end - start)
```

After sending those results into a file using the same method as for the solutions, I again used *ack* to extract the times and send them to a timing file. An example of the command to do this is shown in Figure 15.

Figure 15: Extracting times from a results file

```
ack -o "(?<=\\('algorithm took: ',\\s)\\d+\\.\\d+" greedyTiming.txt >> greedyTiming.txt
```

This varies slightly from the *ack* command for the solutions, as it must look for values with decimals, which it does by looking for any number of digits followed by a decimal place, followed by any number of digits.

From there I copied the values into the attached *Results* excel file. Direct comparison is quite difficult to visualise as there is a huge difference in the running times of the main algorithms. Due to Brute force's special situation no results are collected for it.

The first thing to notice is that Simulated Annealing is the slowest, taking an average of almost 13 seconds to return a solution. Greedy on the other hand is very quick taking on average half a second to return a solution. Finally by far the fastest is Piecemeal which on average takes around 8 milliseconds to return a solution.

So compared to Simulated Annealing, Greedy is ≈ 25 times faster (*4s.f*) and Piecemeal is ≈ 1606 times faster (*4s.f*). Moreover Piecemeal is ≈ 63 times faster than Greedy.

When these timing results are taken with the ε values that each algorithm found, we can make judgments as to why some algorithms are better than others. For example, even though Greedy returns solutions in less than a second the quality of those solutions are not very good. This is especially obvious when comparing Greedy to Piecemeal as the later is faster, and also returns solutions with a better ε value. This leads to the conclusion that Piecemeal should always be used over Greedy. When looking at Simulated Annealing, a large point in it's favour is the fact that it found the globally best solution. However a major point to it's detriment is the fact that it does so over 1600 times slower than Piecemeal, which finds a near to best solution. When deciding between Piecemeal and Simulated Annealing, a value judgment would have to be made in regards to what is more important speed or finding the best result. If speed is the main factor, and a close to best solution is good enough then Piecemeal is the obvious choice. If waiting ≈ 15 seconds for a solution but having a good chance of finding the best is the overriding goal then Simulated Annealing is the better choice. As was described previously Piecemeal performs better when the search space is smaller so that must also be taken into account when deciding on which algorithm to use.

Another inherent part of the algorithms is their Big-O complexity. This describes how the algorithms will scale to different size inputs. Explanation of each algorithms complexity in turn can be found in Section 0.3.3 and Table 13 summarises them. The complexity has a large effect of the runtime of the algorithms so it is useful to look at the theorised complexities and see if they compare to experimental runtimes. Although the algorithms have not been run against various sizes of input it can still be interesting to compare between the algorithms.

For example, following the complexities in Table 13, Piecemeal should be the fastest followed by Greedy and then Simulated Annealing, with Brute force taking a computationally infeasible time. This theory is backed up by the timing results that suggest the same. So in this way running the timing experiments is useful for proving the hypotheses explained in Section 0.3, as well as helping with decision making.

Partial and Full ε calculation

It was theorised that, as well as producing more realistic and exciting solutions, the partial calculation method described in Section 0.4.1 also reduces computation time for any algorithm that uses it.

To find out if that is true first it was necessary to run both of the functions **getEntertainment** (which uses full calculation each time) and **offsetGetEntertainment** (which uses partial cal-

ulation) side by side. When doing this the actual final solutions that were found were discarded as the extra overhead of doing 2 calculations of the same entertainment value could have adversely affected the results. The timing method was the same as used for the general algorithms (see Figure 14).

To get the results, the Greedy algorithm was run and its output saved to a file as before. To distinguish the times, the full calculation was output with "full" before it and the partial time with "reduced" before it. Once again an *ack* command similar to Figure 15 was used to extract the times and save them to another file specific for either reduced or full times. This produced 10,000 values for each of the methods which were then analysed in Excel.

Figure 16 shows the average time over those 10,000 iterations. The actual time taken is not the interesting part, more interesting is the fact that the reduced method takes around half the time. This result supports the hypothesis above in that it shows that the computation time for the reduced calculation is indeed lower. This also means that the computation time for any algorithm using it is also going to be lower. A further point that speaks to why the reduced method is better is that it is used many, many times during the course of finding a solution to this problem, so any saving in terms of computation time will be heavily felt by the overall time to reach a solution.

One drawback of this method is that it does not always guarantee the same reduction in time over the full method. As was described in Section 0.3.2, the amount of calculations needed is dependent on where in the order the swaps take place. The number of calculations range from being at worst the same as the full method to being only around 50 total calculations. If the reduced method must calculate the same or nearly the same number as the full method, it is likely that the time reductions will have been unfortunately lost due to setup overhead in the function.

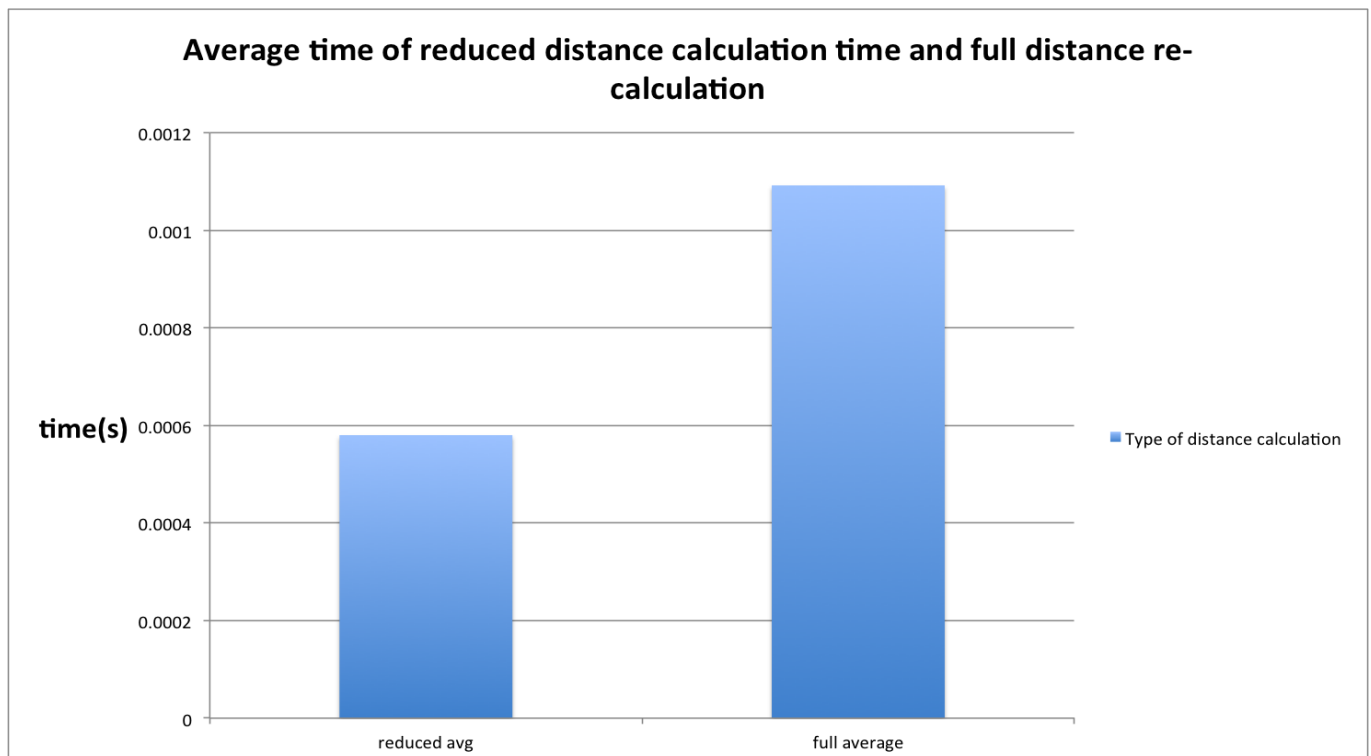


Figure 16: Average time taken to perform full and partial ε calculation on the same solution

To make sure that the reduced method was indeed correctly calculating the ε value it was

decided to run them side by side and compare their results. This could quite easily be done as the testing setup necessary to log the times required that they both be run every time. The test was essentially to equate the two outputs after they were run and to print "true" or "false" if they were the same or different. This adds to the confidence in the method as it was quite easy to return to test that at any time after any major changes to the algorithms or main support methods.

maxMin and refinedMaxMin

The theoretical reasons why the **refinedMaxMin** method is better, such as giving a more real picture of entertainment, have been discussed. To try and find some evidence for this, the minimum values that each method returned were found. To do this the same method as for the full and partial ε calculations test was followed. That was to run both methods side by side and return the minimum values that were found in two lists which could then be analysed.

The effect that the refined method has on the minimum value that is used in Equation 1, is shown in Figure 17. The graph shows the minimum scores that are returned by the two methods as green and blue points, as well as the difference between the two values (red line and red numbers) over a single typical competition. From this graph it can be seen that the **maxMin** method stays relatively low as in this competition the lowest score at the end was only 2 points. On the other hand, the **refinedMaxMin** method diverges at around round 22. This is when it can start ruling out *performers* from winning the overall crown. By the last few rounds the difference is quite large as there are only a few countries that could still win; so only their scores are being taken into account for entertainment.

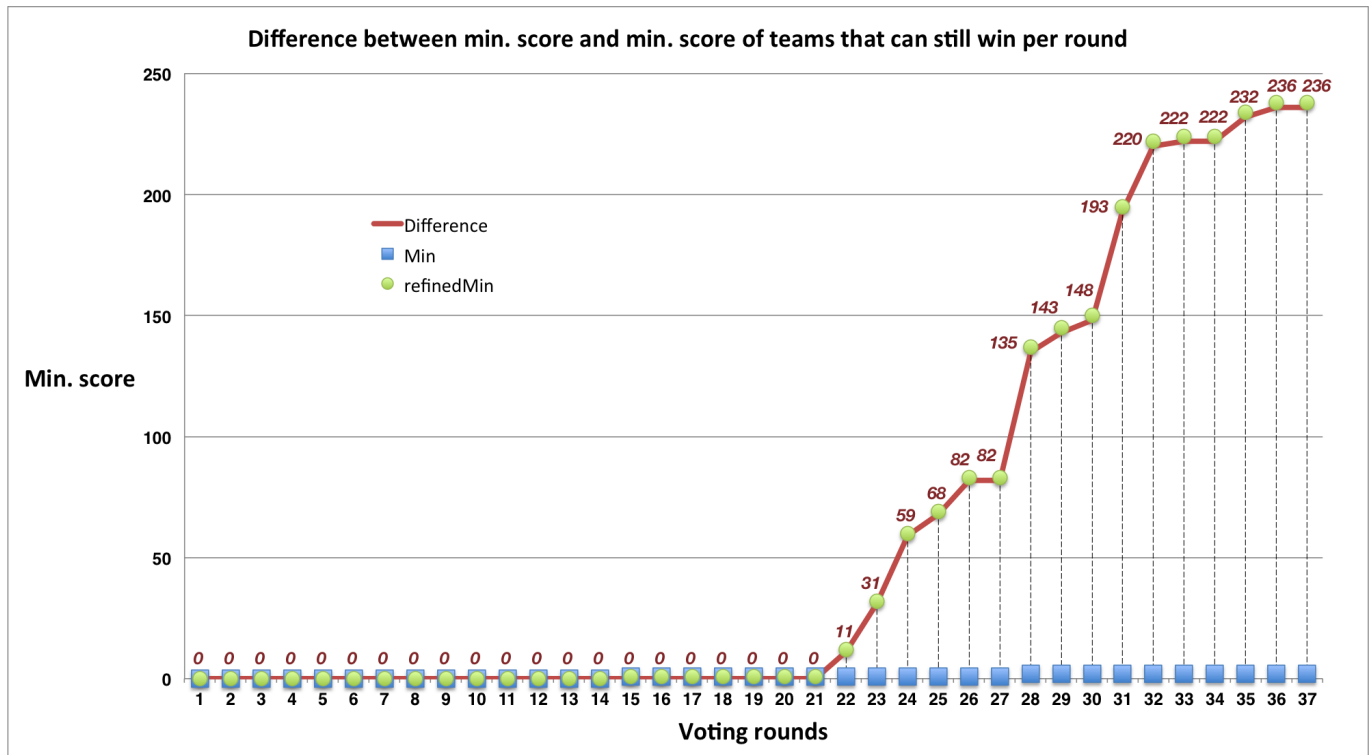


Figure 17: Difference between min value found for MaxMin method and RefinedMaxMin

This chart helps to explain why the idea of removing teams that can no longer win from the calculation can aid in achieving better solutions. The **maxMin** method only ever really returns

a value less than 5, which corresponds to the country currently in last place. It is easy to see that nearer the end of the competition those countries are not adding anything to the entertainment of who will win as their scores are near 0. On the other hand, the **refinedMaxMin** method begins to prune the losing countries, and by the end is only looking at a small group of countries that can still win.

Another point to take into account is the complexities that these methods have and how they can affect the choice to use them. Table 12 in Section 0.4.2 shows the complexities of the two methods. Taking this with the results from Figure 17, even though the complexity of the refined method is slightly higher, the value that it adds means that the trade-off is worth it in my opinion.

Comparing solutions

Throughout the project the same problem kept appearing while trying to compare the intermediate or final solutions. For my piece of mind I wanted to know if similar Φ were giving similar ε values and have a quick way to compare two Φ without having to manually look at country names and compare in what position they were. Moreover having a way to almost immediately know if two orders were the same was very helpful as separating orders based solely on their ε values was not good enough as the values could be shared by vastly different orders. It was also helpful to find a way to see if there were common groupings or positioning of *voters* in different solutions.

To solve this problem a quick and simple Python script was written which can be seen in Figure 18. In the full source code that is attached, the code differs slightly as this code is not meant for use during the running of any algorithms so the file has manually hard-coded values for the orders, *order1* and *order2*.

The idea behind it is very simple, if there are two orders that are different, sum up by how many places the same *voters* are separated in each. This occurs on line 4, where the *dist* variable is the total sum of the difference between the indexes of the same *voter*. By using the *abs()* method from the Python standard library it was possible to not worry about which index was higher and hence keep the code simple.

Figure 18: Finding the difference between two orders

```

1 dist = 0
2 for country in (order1):
3     index1, index2 = order1.index(country), order2.index(country)
4     dist = dist + abs(index1 - index2)
5 print(dist)

```

Unit tests

As the main algorithms do not have repeatable behaviour, due to the fact that they rely on a random step to help pick solutions to look at, it is not possible to unit test them without defeating the purpose of unit testing. There are however some functions that can be unit tested. The file *orderTest.py* contains these tests. The methods that are tested are **getAdjacentNeighbour**, **getNeighbour**, and **refinedMaxMin**.

The tests themselves should be fairly self explanatory they call the function with set inputs and assert that the output from that function matches some value or set of values.

Although these tests are not strictly necessary for the completeness of the project, they help to provide another level of confidence along with the in-place testing described in this section. The tests are not the most comprehensive and only really test the simple working of the methods. The main reason for the lack of unit testing is that this project does not have an already known result, which is something that unit tests require.

0.5.4 Critical Evaluation

The main strength of the results lie in the analysis and the methodology for collecting and analysing them. These results, and the tests that were run to verify them, point to the fact that the project's results are correct and that they can hold up to scrutiny. The analysis has also shown that the results are significant in terms of reaching the goals set out at the start of the project.

On the other hand, a weakness of the results is the small amount of testing of the overall system as it runs. This could be seen as problematic as there are many moving parts that allow an algorithm to be run and produce solutions. Although some of these have been mitigated through other testing, it is always possible that there are still mistakes. Furthermore even though the manual testing attempted to reduce the likelihood of any problems showing up, as always with manual testing, human error can cause problems to be missed.

Another strength of this project was the scientifically-minded method in which each stage was undertaken. Each decision, such as writing a new algorithm or method, was based on intermediate experiments and experimental data. By doing this at every stage throughout the project, a higher level of confidence is gained in the final results, as there are explainable reasons for all the decisions made, and data to support them.

One important point to evaluate is the use of Python as the programming language for this project. Overall the biggest strength that Python brought to the project was it's ease of use. This manifested itself not only in the actual writing of code for the project, but also the supporting overheads that any programming language brings with it. For example, the good quality and reliable standard libraries helped to reduce workload for the more complex parts of the code. Moreover running the algorithms was simple, allowing manual testing to be done with little hassle, where using another language would have added extra setup and time.

A common weakness that is levelled at Python is it's lack of complex data structures and types. For this project, this was not a problem as although the algorithms may be complex, the data structures needed were not anything more complex than an array (list in Python). Having experience with Python also helped to mitigate any problems before they cropped up, such as stopping the global amendment of lists by copying them instead of passing the same list between functions. This can be seen in many places in the code as there is heavy use of code similar to `newList = oldList[:]`, which sets `newList` equal to a copy of `oldList` instead of setting it equal to `oldList` itself. In other languages such as Java this would not be problem, however it would likely bring its own quirks. Overall it is fair to say that Python performed extremely well for this project.

0.6

Future Work

The first and most obvious point for future work would be finding out if the best Φ found is globally optimal and if it is not, then finding the globally optimal solution. Although the Brute force method had to be abandoned due to its computational infeasibility, it may be possible that by using heuristics with the Brute force algorithm to cut down the search space, the globally optimal solution could be found.

Another thought for future work relates to testing the entertaining solutions that were produced, using people. This would be taking a more Psychological or Sociological approach to the project instead of just mathematical. This testing could take place as a sort of mock version of Eurovision where different groups of spectators are shown the voting in different orders and then must rank them in terms of excitement and entertainment. Moreover, by using questionnaires of those who undertake the study, it could be possible to uncover new entertainment functions that were as yet unrealised.

New functions could also be developed separately. In this project we have concentrated only on who can win the competition, however entertainment will contain many more parts. For example, maximising the entertainment for as many spectators as possible who are rooting for each country, possibly taking into account population and engagement of the viewers. Other functions could be ones that maximise those who can reach a certain position, in competitions that involve multiple prizes such as finishing above a certain place. Another would be for competitions with relegations, keeping as many participants in the fight to stay up for as long as possible.

Specifically related to the **refinedMaxMin** method, some future work would be to improve the upper bound score calculation, as it currently does not take into account the possibility that teams may have to vote for themselves and hence not be able to receive 12 points in *every* following round. This improvement should be as simple as tracking who has voted and who is yet to vote and remove 12 points if the current team is yet to vote.

A final piece of future work would be to try and optimise the code for performance. This would take a little more of a deeper understanding and investigation into Python as although I was experienced in using it, never for high performance code. Moreover there are likely places in the code that are not optimised as I prioritised getting correct results over the code being the most performant.

0.7

Conclusions

The main aims of the project were to find an ordering of voters that maximises entertainment, to develop some functions that could codify entertainment mathematically and to produce a way of visualising an ordering.

The order that was actually used by the Eurovision Song Contest in 2014 has a ε value of **3104**. This *is* a relatively entertaining value however this project has shown that there are significantly more entertaining orders, using the same definition of entertainment. Therefore it can be said that this project has at least reached the same level of quality as the current "algorithm" in use by the contest.

In terms of these goals, I believe all of them have been achieved to a high degree of quality. The algorithms produced found a solution to the problem that successfully maximised the entertainment. The functions that were produced to describe entertainment were very successful in taking the intuitive feelings of entertainment and turning them into mathematical functions. Finally the visualisation produced can clearly be shown to aid in the understanding of what makes these good solutions better than the bad ones. Moreover the experimental results outlined in Section 0.5 help to support the idea that these goals have been achieved.

From this report it should be possible to see that novel implementations for use in these types of problems have been produced. Moreover these novel methods have been properly backed up with experimental results in order that their correctness is not in doubt. Also the upmost has been done so that this project does not stand alone with no use for any future work in this area. I think it is fair to say that this project could lay some groundwork for solving more problems from this specific area of optimisation problems, albeit with some additions or modifications.

As can be seen from the context of this project in Section 0.2.1, there has not been a large amount of research into maximising entertainment, especially in the field of Computer Science. This means that although this project began with concrete goals, there was no real context to give an insight into how to best approach the problem, or how simple or complex finding a solution would be. In that respect, achieving the stated goals shows the overall success of the project.

Furthermore it can be said that due to the lack of current research, this project has added a good amount to the research area of optimising competitions for entertainment, by providing a good base for other similar projects, and some useful insights into how to approach these types of problems.

0.8

Reflection on Learning

I think that by assuming that the problem we were undertaking would be complex, we avoided a range of problems that could have arisen. This way of approaching problems (that have an unknown solution) is extremely transferable. From this project I have learnt that not underestimating the complexity of problems will also mean that you are more open to pivoting when problems do arise. Moreover the idea to continually look at data and use that to make decisions, could only be achieved if the problems complexity is thoroughly thought through before starting the project.

Giving myself the freedom to change tact based on changing results also relied on having a solid plan that was still flexible. For this plan to be useful whilst also following the learning from above, it needed to be sufficiently detailed so that I was never lost, but also allow me the freedom to swap tasks around depending on how things were going. In that respect I learnt a lot about planning complex projects in an Agile fashion. I think having the ability to react to changing contexts and information is a valuable skill that this project has given me.

On the point of planning in this project, another new skill gained is to be very industrious with my time. Even though throughout I had the aid of my supervisor, how I spent the majority of the time had to be decided by me. Especially for such a long project with many parts to it, this was complicated. I believe that I gained a deeper understanding of managing my time for a large and complex problem by splitting it up and attacking specific areas. For example by breaking

the project up into the three sections (Entertainment functions, optimisation algorithms and visualisation) in the Initial Plan and then looking at those as the major milestones, it took some pressure off me. I could happily change details of what tasks I needed to complete next and still reach that larger milestones. This again is an extremely useful skill to have gained by virtue of this project.

The work plan is attached in the appendix, and shows a fully updated version of the Initial Plan. The main overarching point is that for many of the tasks the *plan start* and *actual start* differ. As discussed this did not really adversely affect the project, as it was possible to bring forward the implementation of the visualisation and then return to implementing new algorithms easily. I learnt a lot about flexible planning by having this type of plan to refer to when changes were necessary.

Another major skill I have learnt from this project, is to always be sceptical of results until after a deeper analysis. On a couple of occasions during the project I was faced with intermediate results that did not match up with my hypothesis. When faced with this data, I did not immediately throw away the hypothesis and formulate a new one, but I used that initial data to start a deeper investigation. From this I learnt that in problems such as this, there can be hidden phenomena that do not necessarily match up with your first thoughts. These hidden problems can be causing side affects, that are observable, however the problems actually lie deeper.

For example after initially developing the Piecemeal algorithm it was producing solutions with very good ε values. This was not especially strange however as it was not using the same methods as the other algorithms, which had been more thoroughly tested, it seemed wise to investigate more. I compared the outputted solutions of Piecemeal and ran them through the full calculation method to see if the values matched up. When they did not I assumed the problem was with the Piecemeal code so I began investigating the code thoroughly. After finding nothing I decided to look at the supporting code that is shared between all the algorithms. As it turned out, the refinedMaxMin method was incorrectly calculating the min for the Piecemeal code but not for the others. This problem may have gone unnoticed had it not been because I was sceptical at it's initial results. I learnt that a deeper investigation is never a bad thing, and will always lead to a deeper understanding of the system.

Glossary of Terms

1. **Voters (V)**: All countries that are in the Eurovision song contest who vote in the final. The voters is a list of countries. It looks like ["Ukraine", "Austria", "France",....]
2. **Participants (P)**: A subset of the voters that perform songs in the final and receive points from the other voters.
3. **Solution (Φ)**: A solution to the optimisation problem this project is attempting to solve. A solution consists of two things:
 - An ordering of the voting countries as a list
 - An entertainment value found by an entertainment function
4. **Entertainment Value (ε)**: A value given to a solution that describes how entertaining it is. Calculated using an entertainment function, given a solution.
5. **Round (R)**: One round is when one *voter* has given all the *participants* a score. The competition is made up of n_R rounds where $n_R = \text{length of } V$.
6. **Scores (S)**: How many points each country has received per round. The scores are an array of the same length as the number of *participants*. It looks like [0,0,2,12,7,4,0,0,3,2,10,...]. When referring to it in the report along with **rounds** it is most likely cumulative so that after the final round the scores are the total scores for each country.

Appendices

Table 13: Worst Case Big-O complexities of algorithms used

	Big-O time complexity
Greedy Search	$O(V^3)$
Brute Force	$O(V!)$
Simulated Annealing	$O(V^4)$
Piecemeal	$O(V^2)$

All algorithms are expressed in terms of V : *voters* and P : *participants*. More explanation and calculations can be found in Section 0.3.

Algorithm 1 Greedy Search

Require: list of voters**Require:** scoreBoard**Require:** list of performers

```

1:  $xNow \leftarrow getInitialSolution()$ 
2:  $xBest \leftarrow getInitialSolution()$ 
3:  $entertainmentXBest \leftarrow getEntertainment(xBest)$ 
4: while  $i < num\_iterations$  do
5:    $xNow = getNeighbour(xNow)$ 
6:    $entertainmentXNow = getEntertainment(xNow)$ 
7:   if  $entertainmentXNow < entertainmentXBest$  then
8:      $xBest = xNow$ 
9:      $entertainmentXBest = entertainmentXNow$ 
10:     $i = 0$ 
11:   end if
12:    $i = i + 1$ 
13: end while
14: return  $XBest, EntertainmentXBest$ 

```

Algorithm 2 Brute Force

Require: list of voters**Require:** scoreBoard**Require:** list of performers

```

1:  $xBest \leftarrow getInitialSolution()$ 
2:  $entertainmentXBest \leftarrow getEntertainment(xBest)$ 
3: for all  $S \in Solutions$  do
4:    $entertainmentXNow = getEntertainment(S)$ 
5:   if  $entertainmentXNow < entertainmentXBest$  then
6:      $xBest = xNow$ 
7:      $entertainmentXBest = entertainmentXNow$ 
8:   end if
9: end for
10: return  $XBest, EntertainmentXBest$ 

```

Algorithm 3 Simulated Annealing

Require: list of voters
Require: scoreBoard
Require: list of performers
Require: ti: Initial Temperature
Require: tl: Temperature length
Require: cr_coefficient: cooling coefficient

```

1:  $xNow \leftarrow getInitialSolution()$ 
2:  $entertainmentXNow \leftarrow getEntertainment(xNow)$ 
3:  $xBest \leftarrow getInitialSolution()$ 
4:  $entertainmentXBest \leftarrow getEntertainment(xBest)$ 
5: while  $i < num\_iterations$  do
6:   for  $j$  to  $tl$  do
7:      $xPrime = getNeighbour(xNow)$ 
8:      $entertainmentXPrime = getEntertainment(xPrime)$ 
9:      $deltaC = entertainmentXPrime - entertainmentXNow$ 
10:    if  $deltaC \leq 0$  then
11:       $xNow = xPrime$ 
12:       $entertainmentXNow = entertainmentXPrime$ 
13:    else
14:       $q = \text{random int between 0 and 1}$ 
15:      if  $q < e^{-deltaC/t}$  then
16:         $xNow = xPrime$ 
17:         $entertainmentXNow = entertainmentXPrime$ 
18:      end if
19:    end if
20:    if  $entertainmentXNow < entertainmentXBest$  then
21:       $xBest = xNow$ 
22:       $entertainmentXBest = entertainmentXNow$ 
23:       $i = 0$ 
24:    end if
25:  end for
26:   $t = t \times cr\_coefficient$ 
27:   $i = i + 1$ 
28: end while
29: return  $XBest, EntertainmentXBest$ 

```

Algorithm 4 Piecemeal Search

Require: list of voters**Require:** scoreBoard**Require:** list of performers

```

1:  $[solution] \leftarrow chooseFirstVoter()$ 
2:  $best \leftarrow chooseFirstVoter()$ 
3:  $[distances] \leftarrow 12$ 
4: for  $k = 0$  to  $\text{length}(\text{voters}) - 1$  do
5:   for all  $Voter \in possibleNextVoters$  do
6:     calculate  $\Lambda$  with that voter as next in solution
7:     if  $\Lambda < bestDistance$  then
8:        $bestDistance = \Lambda$ 
9:        $best = Voter$ 
10:    end if
11:  end for
12:   $[solution] \leftarrow best$ 
13:   $[distances] \leftarrow bestDistance$ 
14: end for
15:  $entertainmentSolution \leftarrow \text{sum}(distances)$ 
16: return solution, entertainmentSolution

```

Figure 19: offsetGetEntertainment method

```

1  def offsetGetEntertainment(solution, countries, score_board, voters, key1,
2    oldEntertainment, oldDistances, maxScorePerRound):
3    entertainmentValue = 0
4    performing_countries = countries[:]
5    current_solution = solution[:]
6    distances = oldDistances[:]
7    key2 = key1 + 1
8
9
10   oldDistance1, oldDistance2 = oldDistances[key1], oldDistances[key2]
11
12   l_dist = []
13   scores = [0] * 26
14   for j in range(key2 + 1):
15     for i in range(len(countries)):
16       v = voters.index(solution[j])
17       scores[i] = scores[i] + score_board[i][v]
18       otherMin = refinedMaxMin(scores, solution, j, maxScorePerRound)
19       l_dist.append(max(scores) - otherMin)
20
21   newDistance1, newDistance2 = l_dist[key2], l_dist[key1]
22   entertainmentValue = oldEntertainment - (oldDistance1 + oldDistance2) +
23     (newDistance1 + newDistance2)
24
25   distances[key1] = newDistance2
26   distances[key2] = newDistance1
27
28   return entertainmentValue, distances

```

Figure 20: getEntertainment method

```
1 def getEntertainment(solution, countries, score_board, voters, maxScorePerRound):
2     entertainmentValue = 0
3     performing_countries = countries[:]
4     current_solution = solution[:]
5
6     distances = []
7     scores = [0] * 26
8     for j in range(len(solution)):
9         for i in range(len(countries)):
10             v = voters.index(solution[j])
11             scores[i] = scores[i] + score_board[i][v]
12             otherMin = refinedMaxMin(scores, solution, j, maxScorePerRound)
13             distance = max(scores) - otherMin
14             distances.append(distance)
15     entertainmentValue = sum(distances)
16
17     return entertainmentValue, distances
```

Figure 21: Simulated Annealing code

```

1  def simulatedAnnealing(score_board, performers, voters, maxScorePerRound):
2      num_iterations = 500
3      ti = 4000
4      tl = 40
5      cr_coefficient = 0.92
6      i = 0
7
8      t = ti
9
10     xNow = support.getInitialSolution(performers, score_board, voters,
11                                       maxScorePerRound)
12     entertainmentXNow, distxNow = support.getEntertainment(xNow, performers,
13                                                           score_board, voters, maxScorePerRound)
14
15     xBest = xNow[:]
16     entertainmentXBest = entertainmentXNow
17     oldEntertainment = entertainmentXNow
18     oldDistances = distxNow
19
20     while i < num_iterations:
21         for j in range(tl):
22             xPrime, key1 = support.getAdjacentNeighbour(xNow)
23
24             entertainmentXPrime, distXPrime = support.offsetGetEntertainment(xPrime,
25                                                                               performers, score_board, voters, key1, oldEntertainment,
26                                                                               oldDistances, maxScorePerRound)
27
28             deltaC = entertainmentXPrime - entertainmentXNow
29
30             if deltaC <= 0:
31                 xNow = xPrime[:]
32                 entertainmentXNow = entertainmentXPrime
33                 oldEntertainment = entertainmentXPrime
34                 oldDistances = distXPrime
35             else:
36                 q = random.randint(0, 1)
37
38                 if q < math.exp(-(deltaC)/t):
39                     xNow = xPrime[:]
40                     entertainmentXNow = entertainmentXPrime
41                     oldEntertainment = entertainmentXPrime
42                     oldDistances = distXPrime
43
44             if entertainmentXNow < entertainmentXBest:
45                 xBest = xNow[:]
46                 entertainmentXBest = entertainmentXNow
47                 i = 0
48
49             t = t * cr_coefficient
50             i = i + 1
51     return xBest, entertainmentXBest

```

Figure 22: Greedy Search code

```
1 def greedySearch(score_board, performers, voters, maxScorePerRound):
2     num_iterations = 500
3     xNow = support.getInitialSolution(performers, score_board, voters,
4         maxScorePerRound)
5     xBest = xNow[:]
6     entertainmentXBest, distances = support.getEntertainment(xNow, performers,
7         score_board, voters, maxScorePerRound)
8     oldEntertainment = entertainmentXBest
9     oldDistances = distances
10    i = 0
11
12    while i < num_iterations:
13        xNow, key1 = support.getAdjacentNeighbour(xNow)
14
15        entertainmentXNow, distancesXNow = support.offsetGetEntertainment(xNow,
16            performers, score_board, voters, key1, oldEntertainment, oldDistances,
17            maxScorePerRound)
18
19        if entertainmentXNow < entertainmentXBest:
20            xBest = xNow[:]
21            entertainmentXBest = entertainmentXNow
22            i = 0
23
24        oldEntertainment = entertainmentXNow
25        oldDistances = distancesXNow
26        i = i+1
27
28    return xBest, entertainmentXBest
```

Figure 23: Piecemeal algorithm

```

1  def stepByStepSolution(score_board, countries, voters, maxScorePerRound):
2      for p in range(37): # used for testing all solutions
3          entertainmentValue = 0
4          performing_countries = countries[:]
5          solution = []
6          voting_countries = voters[:]
7          distances = []
8          best = voting_countries[p]
9          ignoredIndexes = []
10         bestScores = [row[p] for row in score_board]
11         solution.append(voting_countries[p])
12         ignoredIndexes.append(p)
13         distances.append(12)
14
15         for k in range(1, len(voting_countries)): # go through and add to solution
16             scores = bestScores[:]
17             bestDistance = -1
18
19             for j in range(len(voting_countries)): # go through all voters and try
20                 their votes as next
21                 local_scores = scores[:] # get base scores
22                 current_country = voting_countries[j]
23
24                 if j in ignoredIndexes: # ignore those youve already tried
25                     continue
26
27                 for i in range(len(performing_countries)): # add on the votes from
28                     this country to all the participants
29                     local_scores[i] = score_board[i][j] + local_scores[i]
30                 minScore = support.refinedMaxMin(local_scores, voting_countries, k,
31                     maxScorePerRound) # find the min for that with that voters round
32                 distance = max(local_scores) - minScore
33
34                 if distance < bestDistance or bestDistance < 0:
35                     bestDistance = distance
36                     best = current_country
37                     bestScores = local_scores[:]
38                 ignoredIndexes.append(voting_countries.index(best))
39                 distances.append(bestDistance)
40                 solution.append(best)
41
42         entertainmentValue = sum(distances)
43
44     print(solution, entertainmentValue) # used for testing all solutions
45     # return (solution, entertainmentValue)

```

Table 14: Final Simulated Annealing cooling values

	Value
maxIterations	500
ti	4000
tl	40
cr_coefficient	0.92

Table 15: Best and average solutions found by algorithms

Algorithm	Type of solution	ε value
Brute force*	Best	3125
Greedy Search	Best	2908
	Average	3368
Simulated Annealing	Best	2426
	Average	2673
Piecemeal	Best	2519
	Average**	2586

Average taken over 1000 solutions unless otherwise stated

* Brute force was only run once so only has 1 solution, it's best.

** Piecemeal was run 37 different times, one for each different starting country.

Table 16: Timing results for main algorithms (all times in seconds to 4 s.f)










































	Greedy	Simulated Annealing	Piecemeal
Min	0.2377	11.8288	0.0063
Max	1.4933	17.4252	0.0130
Average	0.5030	12.8072	0.0080
Median	0.4074	12.5793	0.0075

Figure 24: Final Gantt Chart

Project Planner

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight 1  Plan Duration  Actual Start  % Complete

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE	PERIODS														
						1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
V/E/A - Write initial plan	1	1	1	1	100%															
E/A - Research	2	1	1	1	100%															
E - Define entertainment functions	3	2	2	1	100%															
First review meeting	4	1	4	1	100%															
A - Write skeleton algorithm	4	1	2	3	100%															
A - Gather and prepare data for use	5	1	3	3	100%															
A - Write problem specific algorithm	6	3	3	4	100%															
Second review meeting	8	1	8	1	100%															
V - Create visualisation of results	9	5	3	12	100%															
V/E/A - Run experiments on datasets	10	2	11	2	100%															
V/E/A - Analyse data from experiments	11	1	11	1	100%															
V/E/A - Write final report	12	3	11	4	100%															
Hand-in final report	15	1	15	1	100%															

References

- [1] Wikipedia, “Voting at the eurovision song contest — wikipedia, the free encyclopedia,” 2016. https://en.wikipedia.org/w/index.php?title=Voting_at_the_Eurovision_Song_Contest&oldid=756839176; [Online; accessed 28-January-2017].
- [2] D. Gatherer, “Comparison of eurovision song contest simulation with actual results reveals shifting patterns of collusive voting alliances,” *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 2, p. 1, 2006. <http://jasss.soc.surrey.ac.uk/9/2/1.html>.
- [3] L. Spierdijk and M. Vellekoop, “Geography, culture, and religion: Explaining the bias in eurovision song contest voting,” February 2006. <http://doc.utwente.nl/66198/>; [Online; accessed 28-January-2017].
- [4] V. Ginsburgh and A. G. Noury, “The eurovision song contest. is voting political or cultural?,” *European Journal of Political Economy*, vol. 24, no. 1, pp. 41 – 52, 2008. url-
<http://www.sciencedirect.com/science/article/pii/S0176268007000547>.
- [5] G. Bello, H. Menéndez, and D. Camacho, “Using the clustering coefficient to guide a genetic-based communities finding algorithm,” in *Intelligent Data Engineering and Automated Learning - IDEAL 2011: 12th International Conference, Norwich, UK, September 7-9, 2011. Proceedings* (H. Yin, W. Wang, and V. Rayward-Smith, eds.), pp. 160–169, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. http://dx.doi.org/10.1007/978-3-642-23878-9_20.
- [6] EBU, “‘good evening copenhagen’ - voting order revealed,” 10 May 2014. http://www.eurovision.tv/page/news?id=good_evening_copenhagen_-_voting_order_revealed; [Online; accessed 11-April-2017].
- [7] EBU, “Eurovision song contest 2014,” 2016. <https://eurovision.tv/event/copenhagen-2014/profile>; [Online; accessed 27-April-2017] - new website without scoreboard.
- [8] EBU, “Eurovision song contest 2014 - original,” 2016. <https://web-beta.archive.org/web/20160308084207/http://www.eurovision.tv/page/history/by-year/contest?event=1893#Scoreboard>; [Online; accessed 4-February-2017 - Original page with scoreboard].
- [9] Wikipedia, “Neighbourhood (mathematics) — wikipedia, the free encyclopedia,” 2016. [https://en.wikipedia.org/w/index.php?title=Neighbourhood_\(mathematics\)&oldid=735515564](https://en.wikipedia.org/w/index.php?title=Neighbourhood_(mathematics)&oldid=735515564); [Online; accessed 3-May-2017].
- [10] H. Wussing, *The Genesis of the Abstract Group Concept: A Contribution to the History of the Origin of Abstract Group Theory*, ch. 2, p. 94. Courier Dover Publications, 2007. “Cauchy used his permutation notation—in which the arrangements are written one below the other and both are enclosed in parentheses—for the first time in 1815.”.
- [11] Wikipedia, “Cyclic permutation — wikipedia, the free encyclopedia,” 2016. https://en.wikipedia.org/w/index.php?title=Cyclic_permutation&oldid=755419947; [Online; accessed 18-April-2017].
- [12] Wikipedia, “Greedy algorithm — wikipedia, the free encyclopedia,” 2017. [Online; accessed 15-February-2017].

- [13] Wikipedia, “Permutation — wikipedia, the free encyclopedia,” 2017. <https://en.wikipedia.org/w/index.php?title=Permutation&oldid=772638699>; [Online; accessed 19-April-2017]; ‘The number of permutations of n distinct objects is n factorial, usually written as $n!$,...’.
- [14] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [15] Wikipedia, “Simulated annealing — wikipedia, the free encyclopedia,” 2017. https://en.wikipedia.org/w/index.php?title=Simulated_annealing&oldid=775987067; [Online; accessed 19-April-2017].
- [16] R. Booth, “Sa parameters,” 2016. [Slides from Lecture 8(SA Parameters) - Combinatorial Optimisation(CM3109) - Cardiff University].
- [17] Wikipedia, “Constraint satisfaction problem — wikipedia, the free encyclopedia,” 2017. https://en.wikipedia.org/w/index.php?title=Constraint_satisfaction_problem&oldid=775578554; [Online; accessed 3-May-2017].
- [18] Wikipedia, “Backtracking — wikipedia, the free encyclopedia,” 2017. <https://en.wikipedia.org/w/index.php?title=Backtracking&oldid=775406206>; [Online; accessed 3-May-2017].
- [19] Wikipedia, “Triangular number — wikipedia, the free encyclopedia,” 2017. https://en.wikipedia.org/w/index.php?title=Triangular_number&oldid=777530447; [Online; accessed 3-May-2017].
- [20] Python-Software-Foundation, “Python built-in functions - min,” 2017. <https://docs.python.org/2/library/functions.html#min>; [Online; accessed 22-April-2017].
- [21] Python-Software-Foundation, “Python built-in functions - max,” 2017. <https://docs.python.org/2/library/functions.html#max>; [Online; accessed 22-April-2017].
- [22] Python-Software-Foundation, “Python built-in functions - sorted,” 2017. <https://docs.python.org/2/library/functions.html#sorted>; [Online; accessed 22-April-2017].
- [23] Python-Software-Foundation, “Common python method’s complexity,” 2017. <https://wiki.python.org/moin/TimeComplexity>; [Online; accessed 25-April-2017].
- [24] Python-Software-Foundation, “Python built-in functions - len,” 2017. <https://docs.python.org/2/library/functions.html#len>; [Online; accessed 22-April-2017].
- [25] Python-Software-Foundation, “Python random library - sample,” 2017. <https://docs.python.org/2/library/random.html#random.sample>; [Online; accessed 22-April-2017].
- [26] Python-Software-Foundation, “Python random library,” 2017. <https://docs.python.org/2/library/random.html>; [Online; accessed 22-April-2017].
- [27] Python-Software-Foundation, “Python math library,” 2017. <https://docs.python.org/2/library/math.html>; [Online; accessed 22-April-2017].
- [28] Python-Software-Foundation, “Python itertools library,” 2017. <https://docs.python.org/2/library/itertools.html>; [Online; accessed 22-April-2017].
- [29] Python-Software-Foundation, “Python itertools library - permutation,” 2017. <https://docs.python.org/2/library/itertools.html#itertools.permutations>; [Online; accessed 22-April-2017].

- [30] d3, “D3 - scalelinear,” 2016. <https://github.com/d3/d3-scale#scaleLinear>; [Online; accessed 17-February-2017].
- [31] Preact, “Preact.js,” 2016. <https://preactjs.com/>; [Online; accessed 11-February-2017].
- [32] Facebook, “React,” 2016. <https://facebook.github.io/react/>; [Online; accessed 11-February-2017].
- [33] J. M. (developit), “Preact-boilerplate,” 2017. <https://github.com/developit/preact-boilerplate>; [Online; accessed 11-February-2017].
- [34] “ack,” 2017. <https://beyondgrep.com/>; [Online; accessed 10-April-2017].
- [35] Python-Software-Foundation, “Python timeit library,” 2017. <https://docs.python.org/2/library/timeit.html>; [Online; accessed 23-April-2017].