PART II

FOUNDATIONS OF OPTIMIZATION AND ALGORITHMS
A BRIEF HISTORY OF OPTIMIZATION

Optimization is everywhere, from engineering design to financial markets, from our daily activity to planning new buildings, and computer science to industrial applications. We always look for maximum or minimum something.

An optimization problem for maximum or minimum is an optimization problem for maximum or minimum performance. When we place a plane over buildings, we cannot be maximizing our enjoyment with least effort (or idlely from). In fact, we are constantly searching for the optimal solutions to every problem we meet, though we are not necessarily able to find such solutions.

It is no exaggeration to say that finding the solutions to optimization problems, whether intentionally or unintentionally, is as old as human history itself. For example, the least effort principle can often explain many human behaviors. We know the shortest distance between any two different points on a plane is a straight line, though it often needs complex rules such as the calculus of variations to formally prove that a straight line segment between the two points is indeed the shortest.

In fact, many physical phenomenons are governed by what are called least action principles or its variants. For example, light travels such obey Fermat’s principle, that in an interval of the shortest time from one position to another,
The study of optimisation problems is often as old as science itself. It is known that the ancient Greeks had some knowledge of some optimisation problems. For example, Euclid in around 300 BC proposed that a square encloses the greatest area among all possible rectangles with the same total length of four sides. Later, Archimedes in around 200 BC suggested that the distance between two points along the path reflected by a mirror in the absentee whose light travels and reflects from a mirror obeying some geometry, that is the angle of incidence is equal to the angle of reflection. It is a well-known optimisation problem, called Heron's problem, as it was first described in Heron's Gnomon (see On Movement).

The celebrated German mathematician, Johann E. Kepler, is widely known for the discovery of his three laws of planetary motion; however, in 1612, he solved an optimisation problem in his so-called maximising problem on the addendum problem, which he worked on to find the best fit for his average with. The problem itself was described by his personal letter dated October 23, 1609 to Marin Mersenne, including the problem of finding the best combinations of each candidate for the discovery, including all the different kind of angles. Among the other candidates, Kepler chose the fifth, through his friend suggested him to use the fourth candidate. This may imply that Kepler was trying to optimise some utilitarian function of some sort. This problem itself was formalised by Bonaventura Cavalieri in 1635 in his Problem of Geometric Combinations in the February 1635 issue of Scientific American. Since then, it has developed into a field of computational optimisation such as an optimal mapping problem.

W. van Huygen (now unpublished) later, in 1621, discovered the law of reflection, which remained unpublished; later, Christian Huygens mentioned Stevin's results in his Praecepta in 1658. This later was independently rediscovered by Blaise Pascal and published in his treatise Discours de la Méthode in 1637. About 20 years later, when Descartes' students conducted the Essay de la Méthode on the correspondence with Blaise Pascal, Huygens' book appeared in 1659 at his argument with the practical scientific description of light by Descartes, and described Stevin and Descartes' results from a more fundamental principle — light always travels in the shortest time for any position, and this principle for light in more realised in his Perspectiva's principle, which laid the formulation of modern optics.

In his Principia Mathematica published in 1687, Sir Isaac Newton solved the problem of the body shape of minimal resistance that he posed earlier in 1685 as a pioneering problem in optimisation, now a problem of thin or slender bodies. The main idea was to find the shape of the body, which has a minimum resistance to motion in a fluid. Subsequently,
Monodromy derived the existence law of the body. Interestingly, Galileo Galilei independently suggested a similar problem in 1638 in his Dialogues.

In January 1638, J. Bernoulli made some significant progress in calculus. In an article in Acta Philosophica, he challenged all the non-smooth solutions to the problem to find the shape of a curve connecting two points at different heights so that a body will fall along the curve in the shortest time due to gravity. The time of quickest descent, though Bernoulli already knew the solution, was challenging to solve. On January 29, 1638, the challenge was resolved by Newton when he came home at four in the afternoon and he did not sleep until he had solved it by about four in the morning and on the same day he sent out his solution. Though Newton managed to solve it in less than 1.5 hours as he became the Winner of the Huygens letter on March 29, 1638, u was surprised that he, an untrained amateur, could have been able to solve it in half an hour. Huygens said this was the first hint of a new mathematical tool which would revolutionize science and show the path of progress. The solution was a new law in a part of a cyclode. This new concept is now called Huygens' principle, which inspired Newton and Leibniz to formulate the general theory of calculus of variations.

In 1746, the principle of least action was proposed by P. L. d'Herbouville. To verify various laws of physical motion and its application to explain all phenomena, the principle of least action was introduced in the form of an integral equation of a functional in the framework of calculus of variations, which played a central role in the Leibnizian and Newtonian classical mechanics. It is also an important principle in classical mechanics and physics.

In 1748, Compte Monge, a French civil engineer, investigated the transportation problem for optimal transportation and optimization of movements. The initial and final spatial distribution are known. In 1748, Leibniz, Montin, and D'Alembert showed that this enumerative optimization problem is in fact a linear programming problem.

Around 1801, Frederick Gauss claimed that he used the method of least squares to predict the orbit of a comet. Through the use of the least squares with more rigorous mathematical foundation was published later in 1803. In 1803, Joseph Fourier showed in his book Analytique dassortie des solutions des equations aux derivees, and in 1803 he used the principle of least squares for curve fitting. Gauss later claimed that he had been using their method for more than 20 years, and laid the foundation for least squares analysis in 1803. This led to some hidden dispute with Legendre. In 1811, Richard Dedekind, associate of Legendre's work, published the method of least squares studying the uncertainty and errors in making observations, not merely the mean, and this new approach was called by Legendre.
climax that there is a fundamental relationship between opportunity and scarcity of resources, thus requiring that scarcest available resources be used efficiently.

In 1877, in a short note, J. A. Clausen proposed a general method for solving systems of equations in an iterative way. This essentially leads to some iterative methods of simultaneous equations now called the gradient method and steepest descent, for solving functions of more than one variable.

1.2 TRANSCENDENTAL CURVES

In 1806, Dunséal mathematician J. J. Kiesen introduced the concept of continuity and developed an inequality, now referred to as the Jensen's inequality, which played an important role in mathematical optimization and statistical analysis and economics. Convex optimization is a special but very important class of mathematical optimization as many optimization problems can be reformulated into terms of convex optimization. Consequently, it has many applications including control systems, data fitting and modelizing, optimal design, signal processing, mathematical finance, and statistics.

An early in 1756, Landau Ruder studied the Knight tour problem, and W. P. Kürth published a research article on the way to find a circuit which passes through each vertex once and only once for a given graph of polyhedra. In 1899, the William Rowan Hamilton popularized the Hamilton Chain. Then, in December 1890, Karl Wegener posed the Minimalist's problem at a mathematical colloquium in Vienna, an other problem in different connected by possible minimal arcs and straight lines. His work was published later in 1900. On the basis of finding the shortest path connecting a finite number of points in the plane, the problem was formalized. Though the problem in unfeasible in a finite number of trials and permutations, there is no efficient algorithm for finding such solutions. In general, the shortest path connecting the network points does not result in the shortest path. This problem is never referred to as the Mathematical Honors Problem which is closely related to many classical applications such as network routing, resource allocation, scheduling and operations research in general. In fact, an early in 1895, the Wegener's problem in different connected by possible minimal arcs and straight lines. This work was published later in 1900. Interestingly, J. J. Kiesen published in 1907 the first book on optimization, "Theory of Minima and Maxima".

In 1877, T. Tauber developed an algorithm for linear programming and was for its computations. He formulated the production problem of optimal planning and effective methods for finding solutions using linear programming. For this work, he shared the Nobel Prize with T. H. Bourgoin in 1974. The most important type of problems is that George Bendersky introduced in
In 1947, the simplex method for solving large scale linear programming problems was described, but not published. Its first publication in 1947 was by T. C. Koopmans, then by the English mathematician, G. B. R. Box, and later in 1949 by the American mathematician George Dantzig. Dantzig's revolutionary simplex method is able to solve a wide range of optimal policy decision problems of great complexity. A classic example used is the solution to the famous linear programming problem presented in Dantzig's 1949 book on a real-world problem involving 29 equations and 77 unknowns using hand-operated desk calculators.

In 1951, Harold Kuhn and A. W. Tucker studied the nonlinear optimization problem and developed the nonlinear optimality conditions, as similar conditions were proposed by W. Karmark in 1939 in his thesis. In fact, the optimality conditions are the generalization of Lagrange multipliers to nonlinear inequalities, and are now known as the Karush-Kuhn-Tucker conditions, or simply Karush-Kuhn-Tucker conditions, which are necessary conditions for a solution to be optimal in nonlinear programming.

Then, in 1957, Richard Bellman at Stanford University developed the dynamic programming and the optimality principle, which is studied in decision sciences and planning processes where long range comes into the RAPID computation. He also coined the term Dynamic Programming. The idea of dynamic programming was given birth in 1941 when John von Neumann and O. Blumetti studied the sequential decision problems. John von Neumann's work made important contributions to the development of operational research. An earlier work in 1940, Charles Haldane, studied the cost of transportation and routing problems, which would be the earliest research in the operational research. Significant progress was made during the Second World War, and even more in exponential to fixed optimal or near optimal solutions in wide range of complex problems. The development of optimization research such as linear programming for networks, project planning, transportation, financial planning, and management.

After the 1960s, the linear programming optimisation explaked, and it would take a whole book to write even a brief history on optimisation after the 1960s. As this book is mainly about the introduction to nonlinear optimisation algorithms, we will then focus our attention on the development of nonlinear algorithms in nonlinear optimisation. In fact, quite a significant number of new algorithms in optimisation are primarily nonlinear optimisation.

1.3 HIERARCHIES AND HIERARCHIZATIONS

Hierarchization is a solution strategy by trial and error to produce acceptable solutions to a complex problem in a reasonably practical time. The complexity of the problem of interest makes it impossible to record every possible solution on a computer, the aim is to find general, feasible solutions to an acceptable
Chapter 2: A Brief History of Optimization

Several years ago, I was interested in finding that the best solutions could be found, and we even had a natural selection and algorithms would work and why it is they work.

The idea is that an efficient genetic algorithm that will work even of the time and be able to produce good quality solutions. Among these quality solutions, it is apparent that none of them are exactly optimal, though there is one guarantee for each optimality.

Alan Turing was probably the first to use heuristic algorithms during the Second World War when he was developing a German Enigma machine at Bletchley Park, where Turing, together with British mathematician  Gordon Welchman, deciphered in 1940 in cryptography on electromechanical machines, then Bletchley to aid their code-breaking work. The historian used a heuristic algorithm, an Turing called it, to search, among about 10^50 potential combinations, the best possible word solving millions of Enigma messages. Turing called his method stochastic because of the randomness of the method, but there were no guarantees that they would converge to the best solution. In 1943, Turing was recruited to the National Physical Laboratory (NPL), U.K., where he took on his design for the Automatic Computing Engine (ACE). In an NPL report on "Intelligent machinery" in 1948, he combined his innovative ideas of machine intelligence and learning, neural networks, and evolutionary algorithms into an early version of genetic algorithms.

The second significant step in the development of evolutionary algorithms came in the 1980s and 1990s. First, John Holland and his collaborators at the University of Michigan developed the genetic algorithms in the 1980s. In 1990, Holland stated the adaptive system and was the first to test common and interesting results of machine learning for evolutionary machine systems.

John Holland's book, "Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence," published in 1992. In the same year, Governed, the first augmented and enhanced by the potential and power of genetic algorithms for a wide range of applications, including, among others, the solution of complex optimization problems.

Genetic algorithms (GAs) are a search method based on the simulation of Darwin's evolution and natural selection of biological systems and recombination. They are in the mathematical operation: crossover or recombination, mutation, and selection of the fittest. Other names, genetic algorithms become more successful in solving a wide range of optimization problems, among them, the study and research and several hundred of base have been studied. These results show that a small majority of Darwinian SISAN systems were ever studied, but more importantly to solve tough optimization problems such as path planning, data fitting, and resource allocation.

Under the same protocol, Ugo Muenchberg and Hans-Paul Schlosski held them at the Technical University of Berlin developed a search technique for solving optimization problem in economics on experiment, called the evolutionary strategy, in 1983. Later, Peter Giersch joined them and began to construct an automatic system. Instead of using simple rules of economics and reduction, there is an economic function that produces only numerical and optimizing results, and these numerical solutions were kept at each generation. This is essentially a
Multiple trajectory style hill climbing algorithms with randomization. An early
study in 1959, conducted by J. Holland, highlighted the use of randomization in a hill ccciding
process as a tool to study artificial intelligence. Later, in 1966, L. J. Fogel, with A. J. 
Thro and M. J. Walsh, developed the evolutionary programming technique by representing solutions as finite-state machines and randomly
controlling some of these machines. This allows innovative solutions and methods to evolve within a smaller, weaker discipline, called evolutionary algorithms and evolutionary computation.

The decades of the 1960s and 1970s saw the growth of research on evolutionary algorithms. One major breakthrough was the development of evolutionary programming (EA) in 1958, an optimization technique pioneered by D. H. Waddington, G. D. Holland and M. J. Walsh, inspired by the randomizing processes of nature. It is a trajectory-based search algorithm starting with an initial guess and evolving it at a high temperature, and gradually cooling down the system. A move on new solutions in temperature 0 is in tertiary 8-theorem, 0 is in accordance with a probability concentration, which enables a feasible solution for the system. In current local optimum, it is thought that the system in current heat-temperature enough, the global optimum solution can be reached.

The second major usage of evolutionary computation is probably due to David Golson's Aiken research in 1966, through his seminal book on Aiken researches on published in 1997.

In 1980, Thomas DeDeo developed the PSO (Particle Swarm Optimization) evolutionary search algorithms, which he described in his book on evolutionary search and evolutionary optimisation (EA). This search technique was inspired by the random movements of birds or fish swimming in the ocean as a biological system. Then, in 1989, John E. Mitton and Kenneth D. Foran published a theorem on genetic programming which led to the formulation of a whole new area of machine learning, evolving programs as a computational paradigm. As early as in 1989, Karmann applied his first parallel genetic programming. The basic idea is to use the genetic principle to learn computer programs as an instance of genetic programming for a given type of problem.

Evolvingly later, in 1995, another significant step of progress in the development of the particle swarm optimization (PSO) by Abraham Kiers and psychologists James Kowalsky, and computer Ronald C. Beg :) in an optimization algorithm inspired by the random movements of birds and fish to move around the search space starting from some initial random points. The moving randomness creates the current best and searches for the global best as we move to the optimal solution. Since this development, these have been about 300 different variations of particle swarm, and have been applied to almost all areas of tough optimization problems. There is some strong evidence that PSO is better than traditional search algorithms and even better than genetic algorithms for many types of problems, though this is far from conclusive.

In 1997, the publication of the "Neural Network Theorem" for optimization by T. H. Welsh and P. C. Miercke introduced a subclass based on these optimization algorithms.
communities. Researchers have always been trying to find better algorithms, or even new algorithms, for optimization, especially for large-scale optimization problems. However, there is now an arithmetical result that if algorithm A performs better than algorithm B for some optimization function, then it will outperform A for other functions. That is, we may, if averaged over all possible function inputs, beat algorithm B. And it will perform no worse than algorithm A.

Alternatively, there is no universally better algorithm. That is disappointing, right? Therefore, people realize that we do not need the average over all possible functions as for a given optimization problems. What we need is to find the best solutions, this best nothing to do with average over all the whole function space. In addition, we can expect the fact that there is no universal or optimized tool, but we do have the experience that some algorithms indeed outperform others for given types of optimization problems. So the research now focuses on finding the best and most efficient algorithms for a given problem. This task is to design better algorithms for concrete types of problems, not for all the problems. Therefore, the search is still on.

As the turn of the twenty-first century, things become even more exciting. First, Young Won Chang et al. in 2000 developed the Harmony Search (HS) algorithm, which has been widely applied in solving various optimization problems with an integer distribution, computer modeling, and machine learning. In 2004, H. Matsumoto and G. Ozaki proposed the Memetic Brain algorithms, and the application for optimizing Internet browsing charts, which followed the development of a novel chaos algorithm by D. M. Pshtan et al. in 2005 and the Artificial Bee Colony (ABC) by D. Karapetsos in 2006. In 2008, the author of this book developed the Particle Algorithm (PA). Many research articles on the Harmony Algorithm followed, and this algorithm has attracted a wide range of interest.

As we can see, so many novel and smarter metaheuristic algorithms are being developed. Such a diverse range of algorithms necessitates a systematic summary of various metaheuristic algorithms, and this book is such an attempt to introduce all the latest and major metaheuristic with applications.

REFERENCES

1.7 Find the minimum value of \( f(x) = x^2 + x - 6 \) in \([-3, 3]\).
1.8 For the previous problem, use simple differentiation to obtain the same result.
1.9 In order to reduce ISO 14000, DreamiLight produced a square model that has a square surface area, the same area of the square subject to least surface area. Provide the exact formula of such product.
1.10 Design a cylindrical water tank which uses the minimum materials and holds the largest volume of water. What is the relationship between the radius \( r \) of the base and the height \( h \)?

Figure 1.1: Reflection of light at an oblique.


