

An Application of Intelligent Algorithms in Solving Chemistry Problems

H. Jafari^{a,*}, S. Salehfar^b and M. Danesh Shakib^c

^aYoung Researchers and Elite Club, Arak Branch, Islamic Azad University, Arak, Iran,

^bDepartment of Computer Science, Arak Branch, Islamic Azad University, Arak, Iran,

^cDepartment of Industrial Management, College of Management and Accounting, Qazvin Branch, Islamic Azad University, Qazvin, Iran.

Abstract. Chemistry is undoubtedly the most practical science in our life, if we pay attention to the nature around us from the moment we wake up until the day ends, we will find that at the moment of the day and night, we are surrounded by different chemicals. And we have work. In fact, many of our daily activities are related to chemical processes. This has caused the expansion of chemistry and has caused the emergence of various problems. This study aims to analyze a classical chemistry problem known as the chemical reaction equilibrium, which has no uniform solution in different scenarios. In other words, different types of chemical reactions may require diverse methods for establishing equilibrium. This paper proposes a simple system for each chemical reaction through the law of conservation of mass (LCM) and basic mathematical concepts. Optimization algorithms (e.g., the genetic algorithm) are then employed to find a solution to this system.

Received: 20 July 2023, Revised: 31 August 2023, Accepted: 01 September 2023.

Keywords: Chemistry, Reaction equilibrium, Mathematics and computer, Genetic algorithm.

Index to information contained in this paper

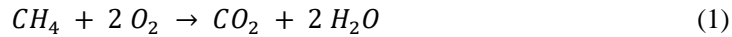
1. Introduction
2. Problem statement
3. Operations research
4. Phases of operations research
5. Genetic algorithm
6. Presentation the proposed approach
7. Conclusion

1. Introduction

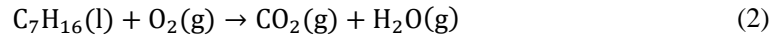
Chemistry is a branch of science that analyzes elements, compounds composed of atoms, molecules, and ions, chemical structures, properties and behaviors of substances, and their mutual effects as well as changes of a substance in reactions with other substances [20]. A chemical reaction is a process that results in the interaction between chemical substances (atoms). The substances participating in the chemical reaction are called reactants, whereas the substances produced during the reaction are referred to as products [3].

Since chemical reactions are typically characterized by a specific chemical change, the products are generally different from the reactants [17]. Both reactants and products are indicated by their chemical formulas. Reactants are usually placed on the left side, whereas products are on the right side. As an example, the combustion of methane in the presence of oxygen is expressed as below [16]:

*Corresponding author. Email: hossein_jafari_123@yahoo.com



This equation represents an irreversible reaction in which one methane molecule reacts with two oxygen molecules to generate one carbon dioxide molecule and two water molecules. In the first step of introducing a chemical equation, the equation is typically not balanced [5]. Consider the following chemical equation as an example:



Evidently, the above equation is not balanced. For instance, there is one carbon atom on the right side, whereas there are seven atoms on the left. In fact, the number of atoms in the reactants is not equal to the number of atoms in the products [15].

2. Problem statement

There is no uniform method for balancing chemical reactions. In simple terms, when it comes to balancing specific reactions, specific algorithms should be tailored to those reactions (similar reactions), something which is not particularly desirable [13]. In this regard, a powerful tool is the utilization of mathematical optimization techniques of operations research (OR). In fact, this study aims to propose a unified approach to balancing all chemical reactions through a mathematics-based system and the genetic algorithm (GA), which will be further discussed.

3. Literature review

Operations research (OR) is a field of applied mathematics and a key tool for engineering sciences, focusing on the study of problem analysis and solutions. The process of operations research can be divided generally into the following five stages [6]:

1. Problem definition
2. Mathematical modeling (expressing variables of the real problem in a mathematical language)
3. Model implementation
4. Testing solutions and assessing their success or failure
5. Real-world problem solving based on valid model outputs

Operations research shares significant similarities and overlaps with other decision-making methodologies such as statistical analysis, management science, game theory, optimization theory, artificial intelligence, and network analysis.

All operations research studies have three common features [7]:

Optimization: The objective of operations research is to achieve the best possible performance under arbitrary conditions. An optimization process includes comparing and constraining the number of possible solutions.

Simulation: In order to analyze potential solutions, a model of the problem is constructed prior to execution on the real problem.

Statistics and Probability: Mathematical algorithms and data are analyzed to better perceive the problem and its probable risks. Statistical and probabilistic processes enhance the credibility of predictions and potential solutions.

The significance of operations research lies in its high efficacy in solving complex problems across various fields. Operations research equips managers and decision-makers in organizations with several powerful decision-making tools. Experts in this field can assist different companies in collecting more comprehensive data for problem solving, considering all available options, exploring all probable outcomes, and estimating risks

[12]. The capabilities of operations research surpass those of ordinary data analysis software and tools. These capabilities benefit from high flexibility in adapting to processes across different domains. Hence, the applications of operations research are not limited to a particular field and are very broad in scope [8].

4. Phases of operations research

In operations research, there are always three subsequent phases that must be undertaken:

Phase one: Problem definition

The first stage of operations research is dedicated to defining the objectives and scope of the target problem. This stage involves the participation of all members of the operations research team. The main objective is to identify three crucial elements in the decision-making problem [19]:

1. Describing decision-making methods and alternative options
2. Determining the objective function
3. Specifying the constraints

Phase two: Modeling

The modeling phase is aimed at transforming the real-world problem into mathematical relations. If the resultant model conforms to one of the standard models (*e.g.*, linear programming), the problem can be solved through existing algorithms. In cases where the mathematical relationships are complex, simplification techniques or simulation methods are employed to solve the problem. In some instances, a combination of mathematical models, heuristics (empirical), and simulation models is employed to reach a solution [19].

Phase three: Model solution

The simplest phase of operations research involves solving the model through optimization algorithms. In fact, a key aspect of this phase is sensitivity analysis conducted to gather additional information regarding how the optimal solution changes with variations in different parameters. If it is impossible to estimate the model parameters accurately, sensitivity analysis becomes crucial. Therefore, it is essential to study the behavior of an optimal solution in the vicinity of estimated parameters [19].

5. Genetic algorithm

The genetic algorithm is a search technique in computer science. It is employed to find approximate solutions to optimization and search problems [11]. The genetic algorithm is a special evolutionary method that utilizes the evolutionary biology principles such as inheritance and mutation. This algorithm was initially introduced by John Henry Holland. In fact, genetic algorithms use the principles of Darwinian natural selection to find an optimal formula for prediction or pattern matching [14]. In artificial intelligence, the genetic algorithm is a programming technique that utilizes genetic evolution as a problem-solving pattern. The problem to be solved has some inputs that are transformed into solutions through a process patterned from genetic evolution. The solutions known as candidates are then evaluated with an evaluation function. The algorithm ends once the termination conditions are met. Generally, a genetic algorithm is an iterative method with many of its components selected as random processes [9]. The operators of this algorithm are presented as below.

5.1 Evaluation mechanism

The fitness function is obtained by applying appropriate transformations to the objective function, which is the function to be optimized. This function evaluates each string and assigns it a numerical value that indicates its quality. As the quality of the solution string increases, both the fitness value and the probability of participating in generating the next generation will increase [18].

5.2 Crossover mechanism

The most important operator in the genetic algorithm is the crossover operator. Crossover is a process whereby the chromosomes of the old generation are mixed and combined to generate a new generation of chromosomes. The pairs selected as parents in the selection process exchange their genes in this step, generating new members [10].

5.3 Mutation mechanism

Mutation is another operator that generates other possible solutions. In the genetic algorithm, when a member is created in the new population, each of its genes undergoes mutation with a certain probability. In mutation, a gene may be removed from the set of population genes, or a gene that has not existed in the population yet may be added [1].

5.4 Different steps in the genetic algorithm

Figure 1 demonstrates the flowchart of the main steps in the genetic algorithm.

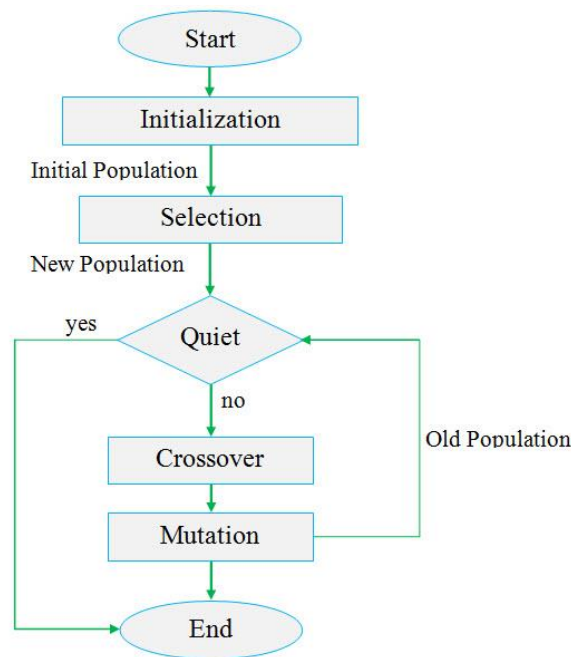


Figure 1. Genetic Algorithm Flowchart [1].

Figure 2 illustrates the structure of a chromosome. Figure 3 depicts an example of a crossover operator, and Figure 4 presents an example of a mutation operator.

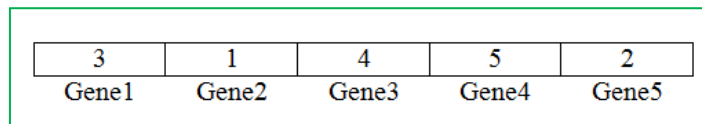


Figure 2. Structure of a chromosome [4].

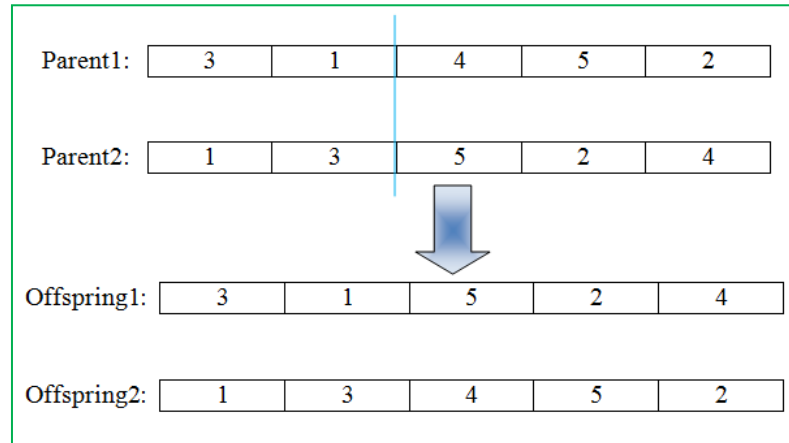


Figure 3. Example of a crossover operator [2].

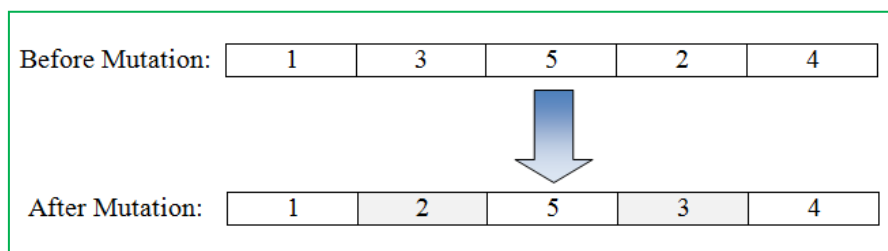


Figure 4. Example of a mutation operator [2].

6. Presentation the proposed approach

The proposed approach in this research involves the use of a genetic algorithm and a new fitness function for balancing chemical reactions. The different stages of this approach are outlined below step by step:

Step (1): Start.

Step (2): Define a suitable fitness function for the chemical reaction.

Step (3): Create an initial population of chromosomes (at this stage, genes are initialized randomly with natural numbers).

Step (4): Generate a new population of chromosomes using crossover and mutation operators and the current generation.

Step (5): Select the top chromosomes and remove the rest from the new population.

Step (6): If the fitness function value is zero, proceed to Step (7), otherwise, go back to Step (4).

Step (7): Display the optimal solution.

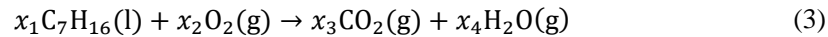
Step (8): End.

In the following discussion, the proposed approach will be presented and analyzed through several numerical examples. The solution process will then be described through the genetic algorithm (GA).

6.1 Numerical example 1

In this example, the goal is to balance the chemical reaction of heptane (C₇H₁₆) with oxygen through the genetic algorithm.

First, the unknown coefficients of the reaction are assigned as follows:



The goal is to find non-zero values for x_1, x_2, x_3, x_4 so that the chemical reaction under investigation can be balanced. The number of carbon atoms (C) must be equal on both sides of the reaction, which is expressed mathematically as below:

$$7x_1 = x_3 \quad (4)$$

The number of hydrogen atoms (H) must be equal on both sides of the reaction, which is expressed mathematically as below:

$$16x_1 = 2x_4 \quad (5)$$

The number of oxygen atoms (O) must be equal on both sides of the reaction, which is expressed mathematically as below:

$$2x_2 = 2x_3 + x_4 \quad (6)$$

Now, a descriptive system for the problem is presented as follows:

$$7x_1 = x_3 \Rightarrow 7x_1 - x_3 = 0 \quad (7.1)$$

$$16x_1 = 2x_4 \Rightarrow 16x_1 - 2x_4 = 0 \quad (7.2)$$

$$2x_2 = 2x_3 + x_4 \Rightarrow 2x_2 - 2x_3 - x_4 = 0 \quad (7.3)$$

$$x_1, x_2, x_3, x_4 \geq 1 \quad (7.4)$$

To solve the above-mentioned descriptive system, a fitness function can be expressed as:

$$f(x_1, x_2, x_3, x_4) = |7x_1 - x_3| + |16x_1 - 2x_4| + |2x_2 - 2x_3 - x_4| \quad (8)$$

The genetic algorithm is aimed at finding suitable values for the variables x_1, x_2, x_3, x_4 that minimize the value of f (or more precisely, equal to zero).

Adjusting the parameters of the genetic algorithm:

The initial population size for the genetic algorithm was set to 25.

The crossover percentage for the genetic algorithm was set to 50.

The mutation percentage for the genetic algorithm was set to 50.

The stopping criterion for the algorithm is the attainment of a zero value for f .

Execution of the genetic algorithm:

Figure 5 shows the performance of the genetic algorithm in balancing the chemical reaction under investigation. Accordingly, the GA managed to balance the mentioned reaction in the 10th generation (10th iteration).

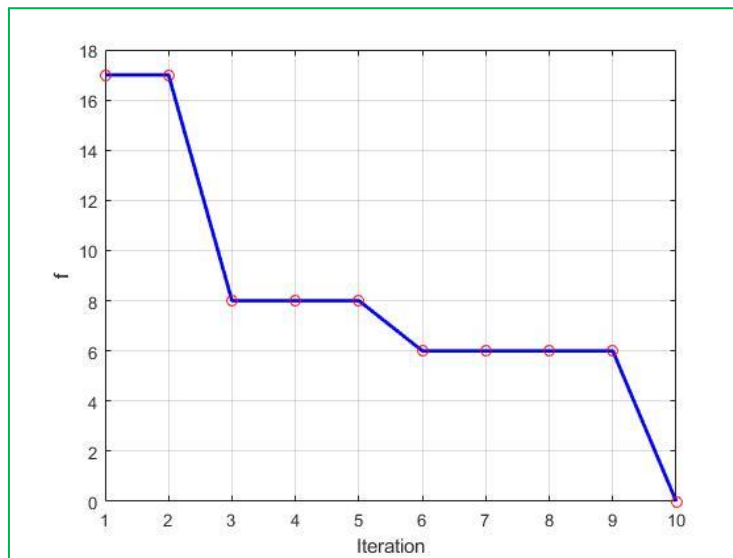
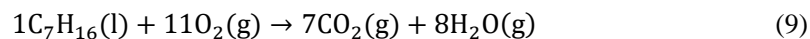


Figure 5. Performance of the genetic algorithm in balancing the chemical reaction.

The chromosome associated with the optimal solution is $(x_1, x_2, x_3, x_4) = (1, 11, 7, 8)$. Hence, the balanced reaction can be written as follows:



To check the validity of the answer obtained from the genetic algorithm, the number of atoms on the left and right sides of the chemical reaction under investigation and comparison are placed in Table 1.

Table 1. Number of atoms on the left and right sides of the chemical reaction.

Atom	Number of atoms on the left side of the chemical reaction	Number of atoms on the right side of the chemical reaction	Is the balance in place?
C	7	7	Yes
H	16	16	Yes
O	22	22	Yes

According to Table 1, it can be said that the answer obtained is valid.

6.2 Numerical example 2

In this example, the aim is to balance the following chemical reaction through the genetic algorithm:



A descriptive system can be presented for the problem as follows:

$$Eq(Fe): x_1 = x_3 \quad (11.1)$$

$$Eq(Cl): 2x_2 = 3x_3 \quad (11.2)$$

To solve the above descriptive system, a fitness function can be expressed as below:

$$f(x_1, x_2, x_3) = |x_1 - x_3| + |2x_2 - 3x_3| \quad (12)$$

Figure 6 depicts the performance of the genetic algorithm in balancing the chemical

reaction under investigation. According to this figure, the GA managed to balance the mentioned reaction in the 6th generation (6th iteration).

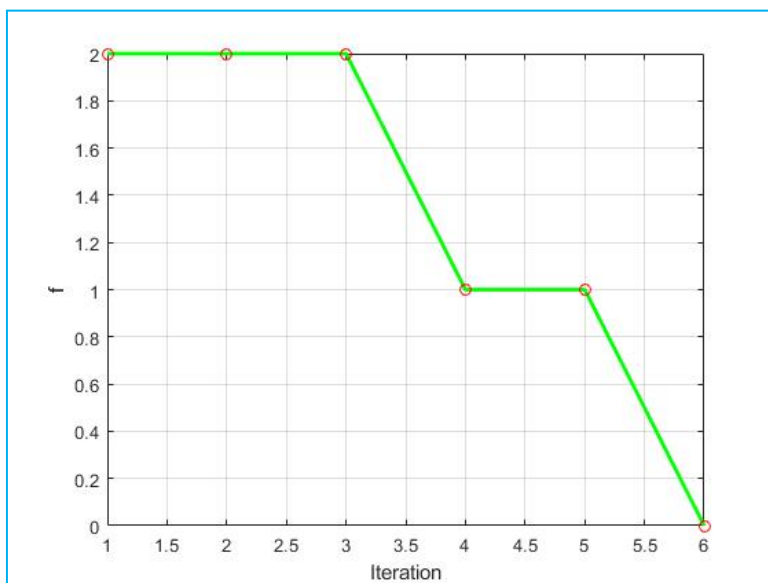


Figure 6. Performance of the genetic algorithm in balancing the chemical reaction.

The chromosome associated with the optimal solution is $(x_1, x_2, x_3) = (6, 9, 6)$. If these numbers are divided by three, the balanced reaction can be written as follows:



To check the validity of the answer obtained from the genetic algorithm, the number of atoms on the left and right sides of the chemical reaction under investigation and comparison are placed in Table 2.

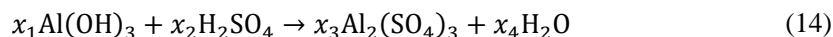
Table 2. Number of atoms on the left and right sides of the chemical reaction.

Atom	Number of atoms on the left side of the chemical reaction	Number of atoms on the right side of the chemical reaction	Is the balance in place?
Fe	2	2	Yes
Cl	6	6	Yes

According to Table 2, it can be said that the answer obtained is valid.

6.3 Numerical example 3

In this example, the aim is to balance the following chemical reaction through the genetic algorithm:



To solve the above descriptive system, a fitness function can be expressed as below:

$$f(x_1, x_2, x_3, x_4) = |x_1 - 2x_3| + |3x_1 + 4x_2 - 12x_3 - x_4| + |3x_1 + 2x_2 - 2x_4| + |x_2 - 3x_3| \quad (15)$$

Figure 7 shows the performance of the genetic algorithm in balancing the chemical

reaction under investigation. Accordingly, the GA managed to balance the mentioned reaction in the 11th generation (11th iteration). The GA parameters in this example are the same as in the Numerical Example (1).

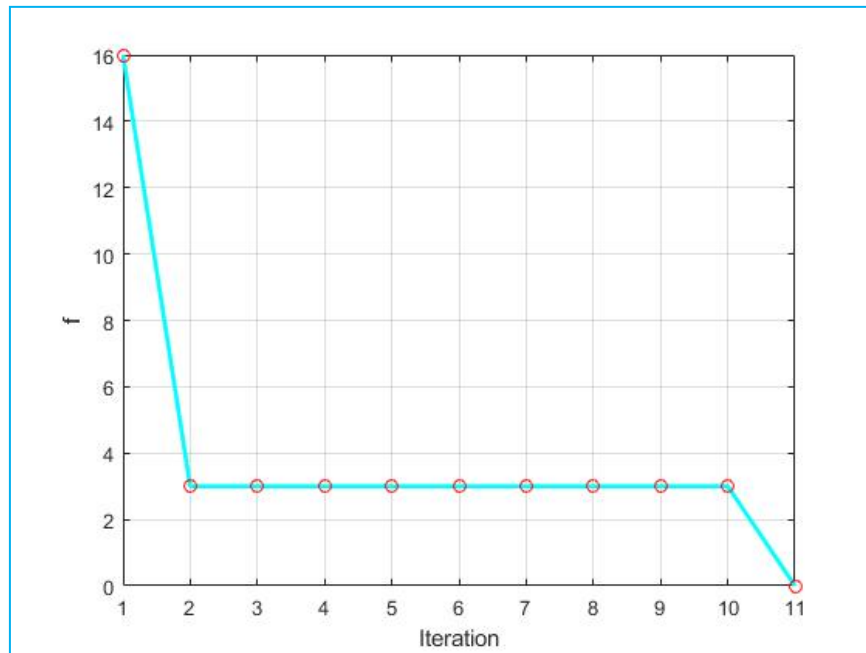
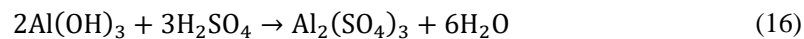


Figure 7. Performance of the genetic algorithm in balancing the chemical reaction.

The chromosome associated with the optimal solution is $(x_1, x_2, x_3, x_4) = (2, 3, 1, 6)$. Hence, the balanced reaction can be written as follows:



To check the validity of the answer obtained from the genetic algorithm, the number of atoms on the left and right sides of the chemical reaction under investigation and comparison are placed in Table 3.

Table 3. Number of atoms on the left and right sides of the chemical reaction.

Atom	Number of atoms on the left side of the chemical reaction	Number of atoms on the right side of the chemical reaction	Is the balance in place?
Al	2	2	Yes
O	18	18	Yes
H	12	12	Yes
S	3	3	Yes

According to Table 3, it can be said that the answer obtained is valid.

6.4 Numerical example 4

In this example, the aim is to balance the following chemical reaction through the genetic algorithm:



To solve the above descriptive system, a fitness function can be expressed as below:

$$\begin{aligned}
 f(x_1, x_2, x_3, x_4, x_5, x_6) & \\
 &= |3x_1 - x_4| + |2x_1 - x_6| + |8x_1 + 2x_2 - 3x_4 - x_5| \\
 &+ |x_2 - x_4|
 \end{aligned}
 \tag{18}$$

Adjusting the parameters of the genetic algorithm

The initial population size for the genetic algorithm was set to 90.

The crossover percentage for the genetic algorithm was set to 30.

The mutation percentage for the genetic algorithm was set to 70.

The stopping criterion for the algorithm is the attainment of a zero value for f .

Execution of the genetic algorithm

Figure 8 shows the performance of the genetic algorithm in balancing the chemical reaction under investigation. Accordingly, the GA managed to balance the mentioned reaction in the 10th generation (10th iteration).

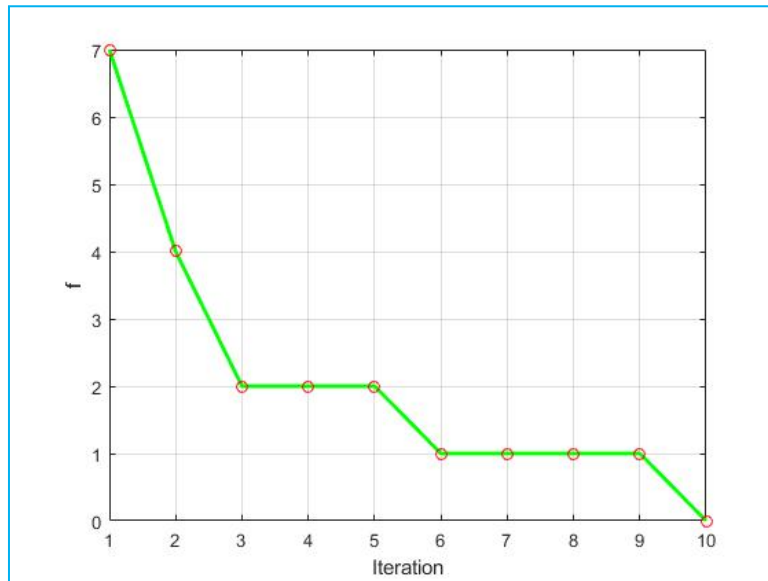
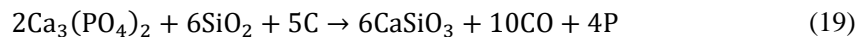


Figure 8. Performance of the genetic algorithm in balancing the chemical reaction.

The chromosome associated with the optimal solution is $(x_1, x_2, x_3, x_4, x_5, x_6) = (2, 6, 5, 6, 10, 4)$. Hence, the balanced reaction can be written as follows:



To check the validity of the answer obtained from the genetic algorithm, the number of atoms on the left and right sides of the chemical reaction under investigation and comparison are placed in Table 4.

Table 4. Number of atoms on the left and right sides of the chemical reaction.

Atom	Number of atoms on the left side of the chemical reaction	Number of atoms on the right side of the chemical reaction	Is the balance in place?
Ca	6	6	Yes
P	4	4	Yes
O	28	28	Yes
Si	6	6	Yes

According to Table 4, it can be said that the answer obtained is valid.

7. Conclusion

Balancing a chemical equation requires the selection of suitable coefficients for reactants and products in order to ensure an equal number of atoms for each element on both sides of the equation. However, it is not ideal for chemistry researchers to lack a standardized algorithm for balancing all reactions in the literature. This paper proposes a relatively straightforward approach by utilizing mathematical systems and genetic algorithms to establish a complete balance in chemical reactions.

References

- [1] M. Ahmadipناه, K. Chalaki and R. Shakeri, Designing cell production arrangement scenarios with the approach of artificial neural networks, *Journal of System Management*, **8 (4)** (2022) 49–64.
- [2] B. Ashrafijoo, N. Fegh-hi Farahmand, Y. Alavi Matin and K. Rahmani, Designing an optimal model using artificial neural networks to predict non-linear time series (case study: Tehran stock exchange index), *Journal of System Management*, **8 (4)** (2022) 65–80.
- [3] J. M. Campanario, Automatic ‘balancing’ of chemical equations, *Computers & Chemistry*, **19 (2)** (1995) 85–90.
- [4] P. C. Chang, W. H. Huang and C. J. Ting, Dynamic diversity control in genetic algorithm for mining unsearched solution space in TSP problems, *Expert System with Application*, **37** (2010) 1863–1878.
- [5] R. Hendi, P. Robbs, B. Sampson, R. Pearce and N. Rees, The electrochemical reduction kinetics of oxygen in dimethylsulfoxide, *Journal of Electroanalytical Chemistry*, **829 (15)** (2018) 16–19.
- [6] H. Jafari, M. Ehsanifar and A. Sheykhان, Finding optimum facility’s layout by developed simulated annealing algorithm, *International Journal of Research in Industrial Engineering*, **9 (2)** (2020) 172–182.
- [7] H. Jafari, M. Faraji and R. Farsi, Evaluation of performance of different units of water and wastewater company using DEA, *International Journal of Applied Operational Research*, **9 (3)** (2019) 21–27.
- [8] H. Jafari, M. Faraji and P. Khaleghi Ardehali, A heuristic method to calculate the real internal rate of return (RIRR), *International Journal of Applied Operational Research*, **10 (1)** (2020) 31–40.
- [9] H. Jafari, M. Faraji and M. Ehsanifar, Solve the facility’s layout problem by developed genetic algorithm, *International Journal of Applied Operational Research*, **8 (1)** (2018) 23–32.
- [10] H. Jafari and A. Sheykhان, Integrating developed evolutionary algorithm and Taguchi method for solving fuzzy facility’s layout problem, *Fuzzy Optimization and Modeling Journal*, **2 (3)** (2021) 24–35.
- [11] J. Jlassi, I. Rekik, S. Elloumi and H. Chabchoub, Genetic algorithm for patients scheduling in emergency department: A case study, *International Journal of Supply and Operations Management*, **10 (4)** (2023) 439–455.
- [12] S. Katoch, S. S. Chauhan and V. Kumar, A review on genetic algorithm: past, present, and future, *Multimedia Tools and Applications*, **80** (2021) 8091–8126.
- [13] J. C. Kotz, P. M. Trichel and G. C. Weaver, *Chemistry and Chemical Reactivity*, Thomson Publication, 6th edition, (2006) 11–13.
- [14] S. Nayeri, R. Tavakkoli-Moghaddam, Z. Sazvar and J. heydari, Solving an emergency resource planning problem with deprivation time by a hybrid metaheuristic algorithm, *Journal of Quality Engineering and Production Optimization*, **5 (1)** (2020) 65–86.
- [15] J. C. Phillips, Algebraic constructs for the graphical and computational solution to balancing chemical equations, *Computers & Chemistry*, **22 (4)** (1998) 295–308.
- [16] P. Ramasami, A concise description of an old problem: application of matrices to obtain the balancing coefficients of chemical equations, *Journal of mathematical chemistry*, **34 (2)** (2003) 21–26.
- [17] I. B. Risteski, A new algebra for balancing special chemical reactions. *Chemistry*, **21** (2012) 223–234.
- [18] L. M. Schmitt, C. N. Nehaniv and R. H. Fujii, Linear analysis of genetic algorithms, *Theoretical Computer Science*, **208** (1998) 111–148.
- [19] A. Slowik and H. Kwasnicka, Evolutionary algorithms and their applications to engineering problems, *Neural Computing and Applications*, **32** (2020) 12363–12379.
- [20] Z. Tóth, Balancing chemical equations by inspection, *Journal of Chemical Education*, **74 (11)** (1997)14–17.