Unbalanced Chemical Equations Conversion to Mark-up Format and Representation to Vision Impaired Students

AZADEH NAZEMI,1 DAVID A. MCMEEKIN,2 IAIN MURRAY1

1Department of Electrical and Computer Engineering, Curtin University, Perth, Western Australia
2Department of Spatial Sciences, Curtin University, Perth, Western Australia

Received 24 December 2014; accepted 1 March 2015

ABSTRACT: This paper describes a method to represent unbalanced chemical equations to vision impaired students which allows them to navigate through classified data, such as species, elements, quantity numbers at the left and right hand sides of equations, reactants, and products. Then they can find appropriate coefficients and balance chemical equations without involving to mathematical aspects of balancing and remembering a lot of information. The goal of this research was the development of an application which assists vision impaired students enrolled in chemistry course to be able to read chemistry literature containing formulae, chemistry representations of elements, and other aspects of chemistry that has been difficult in the past to present in a way for vision impaired people to understand. Developed application by this research is an open source command line Bash Script application under Linux which accepts an unbalanced chemical equation as an input, processes, classifies information, and represents it as Mark-up format or Alternative Audio Description using Text to Speech. © 2015 Wiley Periodicals, Inc. Comput Appl Eng Educ 23:805–812, 2015; View this article online at wileyonlinelibrary.com/journal/cae; DOI 10.1002/cae.21651

Keywords: chemical equation balancing; vision impaired; navigation; mark-up format; bash script; assistive technology; algebraic balancing method; tactile representation; audio representation; Text to Speech (TTS)

INTRODUCTION

Chemical representations refer to various types of formulae, structures, and symbols used to represent chemical processes and conceptual entities (e.g., molecules and atoms) [1]. Chemistry at the symbolic level is represented by symbols, numbers, formulae, equations, and structures [2].

Ben-Zvi et al. studies have shown that understanding symbolic representations is especially difficult for students because these representations are invisible and abstract while students’ thinking relies heavily on sensory information. In addition, without substantial conceptual knowledge and visual-spatial ability, vision impaired students are unable to translate one given representation into another [3].

Mayo’s [4] research results indicate that visually impaired students had difficulty interpreting common representations of the structures encountered in introductory chemistry and organic chemistry courses. These students have problems with balancing equations and are unable to produce accurate pictorial representations. They were also unable to interpret subscripts, coefficients, and implicit meanings of the equations [5].
For studying science, especially chemistry, the role of visual and symbolic modes of representations is extremely important. However, students with visual impairments cannot utilize these forms of visual and symbolic representations because of their impairment. Hence, the study of chemistry then becomes an area that is virtually unattainable for them.

According to the results of the research undertaken by Bodner et al. [6], current pedagogical practices need to be revised to enhance the conceptual understanding that all students develop of the symbolic representations used to describe chemical reactions.

**Chemical Equations and Balancing Introduction**

Chemistry is the study of the properties of materials and the changes that those properties undergo. A chemical equation is the symbolic representation of a chemical reaction wherein the reactant entities are given on the left-hand side of the equation and the product entities are given on the right-hand side of the equation. According to the Law of Conservation of Mass and the Law of Conservation of Charge, the quantity of each element does not change in a chemical reaction. Thus, each side of the chemical equation must represent the same quantity of any particular element. Also, the same charge must be present on both sides of the balanced equation.

Chemical equations balancing plays a critical role in understanding the structure of chemical elements involved in chemical reactions. The ability to balance chemical equations in terms of mass and charge is a key skill that must be mastered by the serious chemistry student. Each reactant or product contains one, or more than one element and each element has a different quantity. In the result, the balancing task must be performed so that equal quantities of each element are on the left and right sides of the equation. Balancing a chemical equation involves multiple steps:

1. Recognition of elements in each reactant on left side of the equation;
2. Recognition of elements in each product on the right side of the equation;
3. Finding quantity of each element on left and right sides;
4. Comparing the quantity of each elements on left and right sides; and
5. Appropriate coefficient calculation to insert before each reactant and products to make entire equation balanced.

These coefficients are the absolute values of the stoichiometric numbers and lead to chemical equation balancing. There are two methods for balancing chemical equations or finding coefficients:

- **Inspection method** is the usual method of balancing chemical equations. It is fast but confusing for complicated equations. In addition, the final result must be double checked to make sure that the chemical equation is indeed balanced;
- **The algebraic method** uses algebraic equations and/or matrices. This method is systematic and can be applied to difficult reactions. It can be easily used with equation solvers but it is time consuming to define corrected algebraic equations [7].

**CURRENT METHODS TO REPRESENT CHEMICAL EQUATION TO VISION IMPAIRED STUDENTS**

Traditionally, there are two methods for representing symbolic level or chemical equations to vision impaired students. They are as follows:

1. **Tactile method**: The use of molecular models to teach balancing equations, which gives a concrete rather than abstract description of the problem, is an efficient approach [8]. Using Braille or embossed paper is another tactile representation of chemical equations [9] and students can actively interact with the implicit information. Generally, the molecular formula shows the composition of a compound and the number of each type of atom in a molecule. For example, molecular formula for benzene is: C6H6. In terms of representation molecular formula due to have subscript(s) are not entirely linear in nature. Since Braille is suitable for text representation, which is linear in nature, is not suitable for molecular chemical formulae representation.

2. **Audio method or using assistive technologies such as Text to Speech (TTS)**. In audio representation, students passively hear all information from the left to right hand side without opportunity to rehear a specific part and/or interact with it. As it is perceived from previous section(s), balancing can be challenging even for sighted students who are able to use visual aids to communicate with and understand the equation’s meaning.

For balancing, students need to have efficient knowledge of the chemical element structure as well as mathematical equation solving skills.

**PROPOSED SOLUTION**

Students with vision impairments will learn chemistry in the classroom and laboratory test and enjoy the most productive careers when they have access to the proper combinations of computer hardware and software and other assistive technology [10]. No vision impairment, including total blindness, should be a barrier to that process. Chemists who are blind use assorted assistive technology to work productively and safely in academia and industry [11]. Many adaptations are simple and readily available.

As stated previously, both current representation methods for chemical equation have disadvantages for vision impaired people; hence, finding a new method suitable for vision impaired people is an important priority.

Chemical Mark-up Language—CML XML is a mainstream approach providing semantics for most chemistry, especially:

- Molecules reactions [12]
- Spectra and analytical data [13]
- Computational chemistry
- Chemical crystallography and material

CML has been developed by Peter Murray-Rust and Henry Rzepa since 1995. It is the de facto XML for chemistry, accepted by publishers and with more than one million lines of Open Source code supporting it. This research used CML idea to convert unbalanced chemical equation to a Mark-up format by embedding essential tags. Unbalanced chemical equation representation in Mark-up format provides an opportunity to navigate through the
classified data. To achieve this purpose, an application was developed in two main modules:

1. Basic module is responsible for:
   - Classification of the implicit information in a chemical equation;
   - Tagging classified information; and
   - Representing Mark-up format.

2. Advanced module is responsible for generating algebraic equations related to chemical equation.

To develop the application to convert the chemical equations to a Mark-up format, a Linux Bash Script was used. The Bash scripting language has several powerful text manipulation tools in command line mode which are suitable for vision impaired users. Vision impaired users can run the application in text mode as a command line mode which are suitable for vision impaired users.

### METHODOLOGY

#### Species Classification and Reactants/Products Extraction

Species in a chemical reaction is a general term used to mean atoms, molecules, or ions. A species can contain more than one chemical element. The general format for species is:

\[
\text{Species} = \text{Symbol}^{\text{Charge}}_{\text{AtomicNumber}/\text{PhysicalStage}}
\]

The physical state of each reactant or product is represented by:

- (l)=liquid
- (g)=gas
- (s)=solid.

Coefficients are useful for keeping the same number of atoms. Chemical equation in most cases comes as a text line, therefore to convert it to a Mark-up format text, processing techniques are used. Suppose, a chemical equation reads as a parameter such as this:

\[
AgNO_3 + Cu \rightarrow Cu(NO_3)_2 + Ag
\]

Equation=$1

To classify and equation the script needs to:

- Extract reactants and product using “+” sign position as the separator:

  Reactant-{$(echo Reactants|sed 's/+ /g')}  
  Product-{$(echo Products|sed 's/+ /g')}  
  NoR-number of Reactants-$(#Reactant[@])  
  NoP-number of Products-$(#Product[@])

#### Chemical Elements Extraction and Symbol Replacement

This module is responsible to find the participating elements in a chemical reaction.

- It calculates the length of each species in terms of text processing, for example:
  \[
  H_2 + O = H_2O, \text{the length of } H_2O \text{ is } 3.
  \]

Species=$reactant[$i]  
SpeciesLength-$(echo $Species| wc -c)  
EquationLength-$(echo $Equation| wc -c)

- Allocate array[$SpeciesLength] to keep equation character by character;
- Investigate array members from the beginning of the array[$i]; and
- Parse equation from left to right to find elements.

In terms of text processing chemical element symbols, they are divided into two categories:

- Single uppercase character alphabet; and
- Double character elements containing one uppercase character followed by a lower case character (as shown in Table 1).

If the array member is inside this string: uppercase-="ABCDEFGHIJKLMNOPQRSTUVWXYZ" then read the next element. If the next element inside the string: lowercase-="abcdefghijklmnopqrstuvwxyz" consider these two adjacent members of the array as an element in category 2. Otherwise, the array member belongs to category 1. Figure 1 illustrates this procedure.

\[
AgNO_3 + Cu \rightarrow Cu(NO_3)_2 + Ag
\]

N.O: category 1  
Ag, Cu: category 2

For better understanding of the conceptual meaning of reaction and remembering the elements involved in the reaction which is shown in the equation, this module replaces the symbol of the extracted elements with a full name using the List of Periodic Table Elements. The following snippet shows the code for replacing the symbols by the full name from the elements.

```bash
for((k=0;k<$NumberElements;k++));do
  sym=$symbol[$k]
  ElementName[$k]=$(cat chemicalElements.lib | awk '$1
    />'$sym'/print $2'|head -1)
done
```
AgNO₃ + Cu²⁺ > Cu(NO₃)₂ + Ag

Ag with Silver, N with Nitrogen, O with Oxygen, and Cu with Copper are replaced.

**Calculations of Total Quantity of Element at Left and Right Sides and Comparison**

The quantity of each element appears in two forms:

- The first being: single quantity N: the number after an element indicates that there are \( N \) atoms of that particular element in each molecule. For example, the water molecule \( H₂O \) has two hydrogen atoms. The quantity comes immediately after the symbol name. This kind of quantity is considered as \( Q₁ \) in this research. \( Q₁ \) only belongs to the element came before it. It is worth noting that if there is only one atom of an element in a molecule it is not written. For example: \( H₂O \) the quantity of Oxygen is 1.

Figure 2 illustrates the procedure to find \( Q₁ \).

- The second being: parentheses or grouped quantity. Some formulae have parentheses and place a group of elements within parentheses \( (\)\). The number comes immediately after the \( (\)\) parenthesis is a grouped quantity. This quantity is called \( Q₂ \) by this research.

This equation includes grouped quantity. Number 3 (in the first reactant) which comes after the closed bracket, belongs to all the elements that are enclosed within the \( (\)\). It means the quantity of N is 3. \( Q₂ \) for all elements inside the parentheses.

The following example shows the quantity of each element inside the parentheses after finding \( Q₂ \).

\[
(E₁₁_{Q₁₁}E₂₃_{Q₁₃})Q₂
\]

\( E₁₁_{Q₁₁} \), \( E₂₃_{Q₁₃} \), \( Q₂ \)
Insert Unknown Coefficients Before Species Including Reactants and Products

The coefficients are the numbers in front of each species. They have very important meaning and indicate the relative amount of atoms in reaction. This module is responsible for considering a unique unknown coefficient for each reactant or product as shown below:

$$a(NH_4)3PO_4 + bPb(NO_3)4 = cPb_3(PO_4)4 + dNH_4NO_3$$

The following snippet shows the conversion process of species to species with coefficient:

```bash
alphabet=abcdefghijklmnopqrstuvwxyz
reactant=$(echo $1|sed 's/#//g'|sed 's/+/ /g')
product=$(echo $1|sed 's/#//g'|sed 's/+/ /g')
NoR=${#reactant[@]}
NoP=${#product[@]}
for ((r=0;r<$NoR;r++))
  do
    CoeffReactant[$r]=$alphabet[$r:0]$(reactant[$r])
    done
for ((p=0;p<$NoP;p++))
  do
    CoeffProduct[$p]=$alphabet[$p:$NoR:2]$(product=[$p])
    done
```

Figure 1  Flowchart for elements (symbols extraction).

Figure 2  Single quantity finding.
Defining Algebraic Equation

If the number of participating elements in the reaction is considered “n” that means there are “n” algebraic equations to be extracted by performing following steps for all participating elements in reaction one by one:

1. Consider an specific element;
2. Remove all species which do not contain the specific element at the left and right;
3. Keep all species containing the specific element at the left and right;
4. For remaining species obtain the mathematical production of the quantity of elements in species by the coefficient and replace species with obtained mathematical production.

Table 2 indicates the algebraic equation for the example from the previous section.

Tagging Classified Information and Mark-up Format Generation

Reviewing previous sections shows which classified data have been extracted. Mark-up Format is the collection of this information with meaningful tags (Table 3).

The following snippet indicates sample tags used by Mark-up Format:

Table 4 Results for Three Samples

| Equation with Coefficients | Reactants | Products | Elements | L | R | Algeaic Equation |
|----------------------------|-----------|----------|----------|---|---|-----------------|
| a Fe + b Cl2 = c FeCl3     | Fe        | FeCl3    | Cl= Chlorine | 2 | 3 | 2b=3c           |
|                            |           |          | Fe= Iron   | 1 | 1 | a=c             |
| a KMnO4 + b HCl = c KCl + d MnCl2 + e H2O + f Cl2 | KMnO4 | KCl | Cl=Chlorine | 1 | 5 | b=e+2f+2d       |
|                            |            |          | H=Hydrogen | 1 | 2 | b=2e            |
|                            |            |          | K=Potassium | 1 | 1 | a=c             |
|                            |            |          | Mn=Manganese | 1 | 1 | a=d             |
|                            |            |          | O=Oxygen   | 4 | 1 | 4a=e            |
| a K4Fe(CN)6 + b H2SO4 + c H2O = d K2SO4 + e FeSO4 + f (NH4)2SO4 + g CO | K4Fe(CN)6 | K2SO4 | C=Carbon | 6 | 1 | 6a=g           |
|                            |            |          | Fe=Iron    | 1 | 1 | a=e             |
|                            |            |          | H=Hydrogen | 4 | 8 | 2b+2c=8f       |
|                            |            |          | K=Potassium | 4 | 2 | 4a=2d           |
|                            |            |          | Mn=Manganese | 6 | 2 | 6a=2f           |
|                            |            |          | O=Oxygen   | 5 | 17| 4b+c+d=4e+4f+g |
|                            |            |          | S=Sulphur  | 1 | 4 | b=d+e+f         |

<body> Unbalanced chemical equation:
AgNO3 + Cu = Cu(NO3)2 + Ag
</body>
testing

Over 100 simple, intermediate, and complicated unbalanced chemical equation have been applied to the developed application in this research and equivalent description for these equations automatically provided as they are shown in Table 4.

CONCLUSION

The obtained results from testing module clearly demonstrated that this application is reliable for chemical equation balancing and, therefore, it can be trusted by students. Hence, the research undertaken here to develop an application which provides opportunity to both sighted and vision impaired students to balance chemical equations, regardless of the involved mathematical aspects, achieved its desired goal. The embedded navigation ability feature in this application allows students to skip through different parts of the equation and to simply replay ambiguous parts.

By the time visually impaired students enroll in chemistry, they should have acquired techniques for solving mathematical problems [14], thus the future work required, and already commenced, is the development of an evaluation module for the application. The evaluation module will assist users in inspecting the accuracies of the balancing of the equations which they have carried out.

BIOGRAPHIES

Azadeh Nazemi received her BS degree in Computer Hardware Engineering from Shiraz University, Iran. She started Master by Research and PhD in September 2010 and 2012 in Curtin University, WA, Australia respectively. She was awarded a Curtin University Postgraduate Scholarship (CUPS) and an Australian Postgraduate Award (APA) scholarship in 2012, respectively. She is presently working towards a PhD degree in Computer engineering at Curtin University, Perth, Australia. Her research area is Assistive Technology and specifically she is working to design Complete Reading System for vision impaired. On November 2012 her research paper has been awarded “Best Student Paper” in 3rd Annual International Conference on Computer Science Education Innovation & Technology (CSEIT 2012) in Singapore. She has nine years of industry experience as an engineer working in consultant companies.

David A. Mcmeekin received the PhD degree in Software Engineering from Curtin University. He is currently a senior research fellow in the Cooperative Research Centre for Spatial Information based in the Department of Spatial Sciences at Curtin University. His research interests are in the areas of the use of technology for people with differing physical and mental challenges as well as the real time delivery of spatially enabled data through semantic web technologies. He is a member of the IEEE Computer Society and Electronics Society.

REFERENCES

[1] W. Hsin-Kai, S. Krajcik Joseph, and S. Elliot, Promoting understanding of chemical representations: Students’ use of a visualization tool in the classroom, J Res Sci Teach 38 (2001), 821–842.

[2] D. Gabel, The complexity of chemistry and implications for teaching, In: International handbook of science education (pp. 233 + 248). B. J. Fraser and K. G. Tobin (Eds.), Kluver Academic Publishers, Boston, MA, 1998.

[3] R. Ben-Zvi, B. Eylon, and J. Silberstein, Students’ visualization of a chemical reaction, Educ Chem July, 117 ± 12 (1987).

[4] P. Mayo, Assessment of the impact chemistry text and figures have on visually impaired students’ learning, Doctoral Dissertation, Purdue University, West Lafayette, 2004.

[5] W. L. Yarroch, Student understanding of chemical equation balancing, J Res Sci Teach 22 (1985), 449–459.

[6] A. L. Bodner, M. Lewis, and M. George, Chemical reactions: What understanding do students with blindness develop? Chemistry education research and practice: 625–636. Mehmet Sahin, Nurettin Yorek 2009. Teaching Science to Visually Impaired Students: A Small-Scale Qualitative Study 6 pp (2013) 19–25.

[7] J. Santiago, Balancing chemical equations - Easy algebraic method, creative commons attribution 3.0 Unported (CC BY 3.0). Available at: http://jaminsantiago.files.wordpress.com/2013/04/balancing-chemical-equations-easy-algebraic-method.pdf 2013.

[8] D. L. Miner, R. Nieman, A. B. Swanson, and M. Woods, Teaching chemistry to students with disabilities: A manual for high schools, colleges, and graduate programs, 4th Ed., 2001, pp 10–96.

[9] P. Edman, Tactile graphics. American Foundation for the Blind, New York, 1992.

[10] L. A. Scadding, An overview of technology and visual impairment, Technol Disabil 1 (1991), 62.

[11] M. Woods, Working chemists with disabilities, expanding opportunities in science, T. A. Blumenkopf (Ed.), American Chemical Society, Washington, DC, 1996.

[12] G. L. Holliday, P. Murray-Rust, and H. S. Rzepa, Chemical markup, XML and the world wide web. Part 6. CMLReact; An XML vocabulary for chemical reactions, J Chem Inf Mod 46 (2006), 145–157. doi: 10.1021/ci0502698

[13] S. Kuhn, T. Helmus, R. J. Lancashire, P. Murray-Rust, H. S. Rzepa, C. Steinbeck, and E. L. Willighagen, Chemical markup, XML, and the world wide web. 7. CMLSpect, an XML vocabulary for spectral data, J Chem Inf Mod 47 (2007), 2015–2034. doi: 10.1021/ci600531a, PMID 17887743

[14] Tombaugh Dorothy, Chemistry and the visually impaired, J Chem Educ 58 (1981), 222–225.
Iain Murray received his BEng (Hons) in Computer Systems Engineering in 1998 and his PhD titled “Instructional eLearning technologies for the vision impaired” in 2008 both at Curtin University. He has worked in the field of assistive technology for more than 25 years both as a practitioner and researcher. Currently employed as a senior lecturer in the Department of Electrical and Computer Engineering, his research interest include learning environments for people with vision impairment, embedded sensors in health applications and assistive technology. He founded the “Cisco Academy for the Vision Impaired” in 2002 to deliver ICT training to vision impaired people globally. He is a Fellow of the Australian Computer Society and a member of the IEEE. Dr Iain Murray (B Eng(Hons) PhD (Curtin) MIEEE MACM FACS Senior Lecturer Course Co-ordinator - Bachelor of Technology (Computer Systems & Networking) & Double Degrees Rehabilitation Engineering & Assistive Technology Research Laboratory Department of Electrical & Computer Engineering Curtin University of Technology GPO Box U1987 Perth, WA 6845.