

Software Requirements Specification for ChemCode:

A program for solving chemistry problems

Samuel Crawford

February 3, 2023

Contents

1	Reference Material	iv
1.1	Table of Units	iv
1.2	Table of Symbols	iv
1.3	Abbreviations and Acronyms	v
1.4	Mathematical Notation	v
2	Introduction	1
2.1	Purpose of Document	1
2.2	Scope of Requirements	1
2.3	Characteristics of Intended Reader	1
2.4	Organization of Document	2
3	General System Description	2
3.1	System Context	2
3.2	User Characteristics	3
3.3	System Constraints	3
4	Specific System Description	3
4.1	Problem Description	4
4.1.1	Terminology and Definitions	4
4.1.2	Physical System Description	4
4.1.3	Goal Statements	5
4.2	Solution Characteristics Specification	5
4.2.1	Assumptions	5
4.2.2	Theoretical Models	6
4.2.3	General Definitions	7
4.2.4	Data Definitions	7
4.2.5	Instance Models	7
4.2.6	Input Data Constraints	9
4.2.7	Properties of a Correct Solution	9
5	Requirements	10
5.1	Functional Requirements	10
5.2	Nonfunctional Requirements	10
6	Likely Changes	11
7	Unlikely Changes	12
8	Traceability Matrices	12

Revision History

Nice!

Date	Version	Notes
Jan. 21, 2023	0.0	Start document, fill in title page, and add references
	0.1	Fill in Problem Description and Goal Statements sections, as well as all relevant definitions
Jan. 22, 2023	0.1.1	Replace the notion of “fractional oxidation state” with “non-stoichiometric compound”
	0.1.2	Fix capTemplate reference
	0.1.3	Update Abbreviations and Acronyms
	0.2	Add IM for balancing a chemical reaction (IM3), along with relevant reference information
Jan. 23, 2023	0.2.1	Add IM for determining if a chemical reaction is feasible (IM2), along with relevant reference information
Jan. 25, 2023	0.2.2	Add references between sections for notation and definitions
	0.2.3	Abstract IM2
	0.2.4	Add IM for converting chemical equation to a matrix (IM1), along with relevant reference information
	0.2.5	Rename and add input constraint to IM3
	0.2.6	Add note about IM3 returning the smallest solution
	0.2.7	Add TM for matrix equation (TM1)
	0.2.8	Improve referencing, mainly for TMs, and add note about potential TM for Conservation of Mass
	0.2.9	Add check for Conservation of Mass in IM2
	0.3	Add Likely Changes from Problem Statement, add assumption about difference between number of elements and compounds (A1), and add mole to Table of Units
Jan. 26, 2023	0.3.1	Update TM1 (and TM function) from feedback from presentation
	0.3.2	Fill in Scope of Requirements , fix A1 and change “infeasible” to “feasible”
	0.3.3	Improve section referencing in Revision History
Jan. 27, 2023	0.4	Fill in Functional Requirements
	0.4.1	Fill in main Nonfunctional Requirements , along with relevant reference information
Jan. 28, 2023	0.4.2	Add TM for the Law of Conservation of Mass (TM2)
	0.4.3	Represent input to IM1 as two sequences of chemical formulas, add assumptions about user input, and remove infeasible likely change
	0.4.4	Add NFRs for understandability (NFR2) and verifiability (NFR6) and improve quality of other NFRs

Date	Version	Notes
Jan. 29, 2023	0.5	Fill in Introduction based on Drasil examples
	0.5.1	Fill in Physical System Description and remove “product” and “reactant” from Terminology and Definitions
Jan. 30, 2023	0.5.2	Remove unnecessary template comments about SRS document and the ones from Specific System Description up to Instance Models
	0.5.3	Remove unnecessary content from Table of Symbols and add the notion of “gram” to the document
	0.5.4	Make NFR1 more accurate to current scope with potential future information for stretch goals added in comment
	0.6	Make System Context diagram (Figure 1) and start to fill in System Context
Jan. 31, 2023	0.6.1	Fill in rest of General System Description
Feb. 1, 2023	0.6.2	Change abbreviation of theoretical model from “T” to “TM” and make TM references consistent with other references
Feb. 2, 2023	0.6.3	Fill in Reference Material , renaming specific matrices/vectors in Instance Models
	0.6.4	Reformat and clarify Likely Changes
Feb. 3, 2023	0.6.5	Add type notation and information
	0.6.6	Fill in Unlikely Changes , add likely change for identifying more types of reactions (LC9), and reference A1 in IM1
	0.6.7	Fill in Traceability Matrices and remove more unnecessary content
	0.6.8	Fill in data constraints sections, update traceability matrices, and finalize formatting

1 Reference Material

This section records information for easy reference.

1.1 Table of Units

Throughout this document, SI (Système International d'Unités) is employed as the unit system. Several derived units are used as described below. For each unit, the symbol is given, followed by a description of the unit and the SI name.

symbol	unit	SI
g	mass	gram
mol	amount of substance	mole

\mathbb{N}^m (You don't need $\times 1$ for vectors)

1.2 Table of Symbols

The table that follows summarizes the symbols used in this document along with their units and types. The choice of symbols was made to be consistent with linear algebra conventions while distinguishing between specific matrices. The symbols are listed in alphabetical order.

symbol	unit	type	description
0	—	\mathbb{N}^m	zero vector
A	—	$\mathbb{R}^{m \times n}$	generic matrix
E	—	$\mathbb{R}^{m \times n}$	matrix representation of the inputted chemical equation
b	—	\mathbb{R}^m	generic vector
c	—	$(\mathbb{Z}^+)^n$	vector of the smallest coefficients that balance a chemical equation
<i>m</i>	—	\mathbb{Z}^+	generic whole number
<i>n</i>	—	\mathbb{Z}^+	generic whole number
x	—	\mathbb{R}^n	generic vector

1.3 Abbreviations and Acronyms

symbol	description
A	Assumption
DD	Data Definition
GD	General Definition
GS	Goal Statement
IM	Instance Model
LC	Likely Change
PS	Physical System Description
R	Requirement
SRS	Software Requirements Specification
TM	Theoretical Model
VnV	Verification and Validation

I ~~in~~ added a link to Gubb & Van Loan on the review issue.

1.4 Mathematical Notation

Matrices are written in bold uppercase letters (for example **A**), while vectors (one-dimensional matrices) are written in bold lowercase letters, (for example **x**) [1]. The type of a matrix will be denoted by the type of each value in the matrix with the matrix's dimensions as a superscript; for example, a matrix of natural numbers with two rows and three columns is of type $\mathbb{N}^{2 \times 3}$. [I couldn't find a source for this, but this is what I'll use. —SC] The zero matrix is a special type of matrix where each element is zero and is denoted by **0** [2]. Two (or more) matrices with the same number of rows can be joined together to form an “augmented matrix”, written as **(A|B)** [3].

The set of natural numbers is denoted by \mathbb{N} and consists of the non-negative integers (including zero). The set of whole numbers is denoted by \mathbb{Z}^+ and consists of the positive integers (excluding zero).

For completeness you should define \mathbb{R}

2 Introduction

Chemical equations are common ways of representing chemical reactions (as described in [Physical System Description](#)), but must be balanced for use in other calculations [1]. Therefore, it is useful to have a tool to automatically balance these chemical reactions for efficiency and accuracy.

The following section provides an overview of the Software Requirements Specification (SRS) for ChemCode. This section explains the purpose of this document, the scope of the requirements, the characteristics of the intended reader, and the organization of the document.

2.1 Purpose of Document

The primary purpose of this document is to record the requirements of the program for solving chemistry problems [2]. Goals, assumptions, theoretical models, definitions, and other model derivation information are specified, allowing the reader to fully understand and verify the purpose and scientific basis of ChemCode. With the exception of system constraints, this SRS will remain abstract, describing what problem is being solved, but not how to solve it.

This document will be used as a starting point for subsequent development phases, including writing the design specification and the software Verification and Validation (VnV) Plan. The design document will show how the requirements are to be realized, including decisions on the numerical algorithms and programming environment. The VnV Plan will show the steps that will be used to increase confidence in the software documentation and the implementation. Although the SRS fits in a series of documents that follow the so-called “waterfall model”, the actual development process is not constrained in any way. Even when the waterfall model is not followed, as Parnas and Clements point out in [5], the most logical way to present the documentation is still to “fake” a rational design process.

2.2 Scope of Requirements

The scope of the requirements includes determining if a given chemical equation is feasible (see [Terminology and Definitions](#)) and, if it is, balancing it using the smallest whole number coefficients possible, regardless of whether or not the chemical formulas in the equation are correct or if the reaction would take place in a real-world setting.

2.3 Characteristics of Intended Reader

Reviewers of this documentation should have an understanding of high-school level chemistry, namely stoichiometry (see [Terminology and Definitions](#)), and undergraduate Level 1 linear algebra, namely solving systems of linear equations. The end user of ChemCode, as described in [User Characteristics](#), does not require knowledge of linear algebra, since relevant concepts will be implemented by the code and hidden from the user.

This implies there are calculations where the equations do not have to be balanced. I'm not sure this is the case, I think you can just say the equations need to be balanced.

productivity of scientist and engineers.

more the purpose here. Explain what balancing means.

related to balancing chemical equations.

chemistry problems seems to open ended, so I added some not words above

This is actually the purpose, not the scope also

this is the scope

2.4 Organization of Document

The organization of this document is based on a template from [6], which is based on the template for an SRS for scientific computing software proposed by [7], [8] and [9]. The presentation follows the standard pattern of presenting goals, theories, definitions, and assumptions. For readers that would like a more bottom-up approach, they can start reading the [Instance Models](#) and trace back to find any additional information they require.

The [Goal Statements](#) are refined to the [Theoretical Models](#), which are refined to the [Instance Models](#).

3 General System Description

This section provides general information about the system. It identifies the interfaces between the system and its environment, describes the user characteristics and lists the system constraints.

3.1 System Context

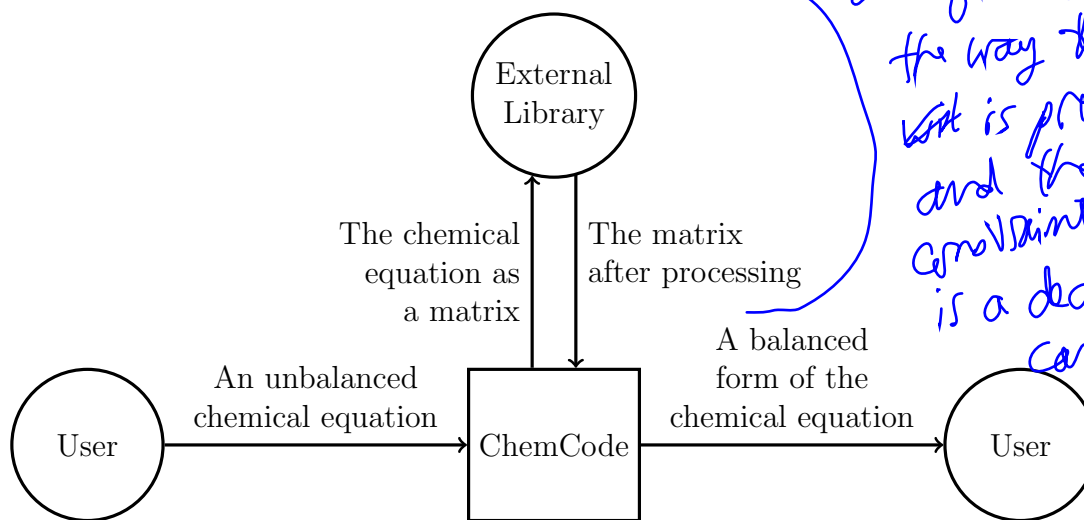


Figure 1: System Context

Figure 1 shows the system context. A circle represents an external entity outside the software. A rectangle represents the software system itself (ChemCode). Arrows are used to show the data flow between the system and its environment.

- User Responsibilities:

- Provide an unbalanced chemical equation, ensuring conformation to input data format required by ChemCode

If you are using numbers rather than the natbib style, you don't put the sentence like this. [7--9]

- Ensure required software assumptions from Subsection 4.2.1 are appropriate for the problem to which the user is applying the software
- ChemCode Responsibilities:
 - Detect data type mismatch, such as a string of characters instead of a rational number
 - Detect improperly formatted chemical equations
 - Format inputted chemical equation as a matrix
 - Determine if the inputted chemical equation is feasible
 - If the inputted chemical equation is feasible, find its balanced form with the smallest whole number coefficients
- External Library Responsibilities:
 - Solve the chemical equation matrix *system of equations* [Can this be more descriptive? —SC]

3.2 User Characteristics

The end user of ChemCode should have an understanding of high-school level chemistry, namely stoichiometry (see [Terminology and Definitions](#)).

3.3 System Constraints

ChemCode will be developed using Drasil [10], “a framework for generating high-quality documentation and code for Scientific Computing Software” [11, p. iii], with the goal of extending it by adding concepts relevant to the problem outlined in [Problem Description](#). One of the desired additions to Drasil is the notion of systems of linear equations, so this problem will be solved using matrices. Since Drasil is built on the idea of reusability, external libraries will be used to solve these systems of linear equations. This was previously done with ordinary differential equation (ODE) solvers, since “creating a complete ODE solver in Drasil would take considerable time, and there are already many reliable external libraries ... tested by long use” [12, p. 24]; these rationales also apply to solvers of systems of linear equations.

4 Specific System Description

This section first presents the problem description, which gives a high-level view of the problem to be solved. This is followed by the solution characteristics specification, which presents the assumptions, theories, definitions and finally the instance models.

4.1 Problem Description

ChemCode is intended to balance chemical equations (including ones with nonstoichiometric compounds; see [Terminology and Definitions](#)) so they can be useful for other computations [4]. Additionally, since molecules only exist in whole numbers (since dividing a molecule changes its composition into new types of molecules), the coefficients used to balance the equation must be whole numbers, and by convention should be as small as possible [4].

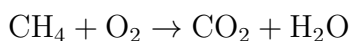
4.1.1 Terminology and Definitions

This subsection provides a list of terms that are used in the subsequent sections and their meaning, with the purpose of reducing ambiguity and making it easier to correctly understand the requirements:

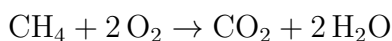
- **Feasible:** (Referring to a chemical equation) able to be balanced [13].
- **Hydrate:** “A compound formed by the chemical combination of water and some other substance in a definite molecular ratio” [14].
- **Nonstoichiometric Compound:** “Any solid chemical compound in which the numbers of atoms of the elements present cannot be expressed as a ratio of small whole numbers” [15].
- **Stoichiometry:** “The calculation of the quantities of reactants or products in a chemical reaction using the relationships found in a balanced chemical equation” [4, p. 337].

4.1.2 Physical System Description

Chemical reactions are interactions between different types of matter that result in new substances being formed [4, p. 286]. These are represented using chemical equations, with the reactant(s)—the substance(s) present at the beginning of the reaction—on the left-hand side and the product(s)—the substance(s) formed by the reaction—on the right. An example of a chemical equation representing the combustion of methane in the presence of oxygen to form carbon dioxide and water vapour is shown below. The subscripts indicate how many of each element is present in the given chemical compound.



[Is this paragraph helpful/needed? —SC] As mentioned in [Problem Description](#), chemical formulas must be balanced to be useful in other reactions [4]. This means that there must be the same number of each element before and after the reaction takes place to satisfy the Law of Conservation of Mass (TM2). Since changing the subscripts in a formula will change the type of molecule (for example, O_3 is ozone, not oxygen), equations are balanced by introducing coefficients before each compound. If no coefficient is present, there is an implicit coefficient of “1”. The above equation looks like this in its balanced form:



The physical system of ChemCode includes:

PS1: the products of a given chemical reaction;

PS2: the reactants of the same chemical reaction.

4.1.3 Goal Statements

Given a representation of a chemical equation, the goal statements of ChemCode are:

GS1: Determine if a given chemical reaction is feasible (see [Terminology and Definitions](#)).

GS2: Balance the ~~inputted~~ chemical equation with the smallest whole number coefficients possible.

4.2 Solution Characteristics Specification

The instance models that govern ChemCode are presented in Subsection 4.2.5. The information to understand the meaning of the instance models and their derivation is also presented, so that the instance models can be verified.

4.2.1 Assumptions

This section simplifies the original problem and helps in developing the theoretical model by filling in the missing information for the physical system. The numbers given in the square brackets refer to the theoretical model [TM], general definition [GD], data definition [DD], instance model [IM], likely change [LC], or unlikely change [UC] in which the respective assumption is used.

A1: For all inputted chemical equations, there is at most one more compound than element. [\[Should this be an assumption or a requirement? —SC\]](#) [IM1, UC2]

A2: All inputted chemical formulas describe real chemical formulas. [IM1]

A3: All inputted chemical formulas are formatted following some set of agreed-upon conventions. [IM1, LC6]

A4: All inputted chemical formulas are simple; i.e., only consist of atomic symbols and subscripts. [IM1, LC1]

Isn't this covered by A3?

You haven't defined these terms.

4.2.2 Theoretical Models

This section focuses on the general equations and laws that ChemCode is based on.

Number: TM1

Label: Matrix equation

Equation: $\mathbf{Ax} = \mathbf{b}$

Description: The above equation gives the general form of the matrix equation where \mathbf{A} is of type $\mathbb{R}^{m \times n}$, \mathbf{b} is of type $\mathbb{R}^{m \times 1}$, and \mathbf{x} is of type $\mathbb{R}^{n \times 1}$. [Is this type information redundant to have here? —SC] The values of \mathbf{x} are unknown and will be solved for [16].

Notes: For any \mathbf{A} and \mathbf{b} , there is either zero, one, or infinite values for \mathbf{x} [17].

Source: <https://textbooks.math.gatech.edu/ila/matrix-equations.html>

Ref. By: IM3

Derivation for TM1: Not Applicable

type information is helpful. I don't think it is redundant, just like units are not redundant.

} For an underdet. system you always have an infinite # of solutions (if there are any)

Don't you require an underdetermined system where $m < n$ (you have more unknown than equations.)

Your presentation of the TMs etc is too abstract. You have abstracted out all the information. I don't know what it means for an equation to be balanced. I don't even know what an equation is. I agree that we don't want to say the format of the input, but a mathematical abstraction of the input is fine. The input is a sequence of compounds.

we can write an equation

$$\forall e: E \mid \text{count}(R) = \text{count}(P)$$

You can define R2P
using the 2A4 rules
approach
or something
similar

Number: TM2

Label: Law of Conservation of Mass

all elements

products

↳ this can be defined in a data definition

Equation: Not Applicable

Description: This law states that “matter can neither be created nor destroyed in a chemical reaction ... but it may change forms to other substances” [4, p. 112].

Notes: None

Source: https://chem.libretexts.org/Courses/Anoka-Ramsey_Community_College/Introduction_to_Chemistry/03%3A_Matter_and_Energy/3.06%3A_Conservation_of_Mass

Ref. By: IM2

Derivation for TM2: Not Applicable

4.2.3 General Definitions

There are no general definitions.

4.2.4 Data Definitions

There are no data definitions.

4.2.5 Instance Models

This section transforms the problem defined in [Problem Description](#) into one which is expressed in mathematical terms. It uses concrete symbols defined in Section 4.2.4 to replace the abstract symbols in the models identified in Sections 4.2.2 and 4.2.3 [Should the reference to the general models be removed if there aren't any? —SC].

The goals GS1 and GS2 are solved by IM2 and IM3, respectively.

→ You can show how to get here if you formalize what an ~~eqn~~ chemical equation is. The rows come from TM2 for each element.

Number	IM1
Label	Matrix representation of elements present in reaction
Input	Two sequences of representations of chemical formulas; one for the reactants and one for the products.
Output	E , a matrix such that each entry indicates the amount of a given element in a given compound; each row corresponds to a different element and each column corresponds to a different compound. Positive entries indicate the compound is a reactant, negative entries indicate the compound is a product, and zeroes indicate the corresponding element is not present in the corresponding compound.
Description	Each chemical formula present in the inputted sequences are assumed to be valid (A2), be formatted correctly (A3), and only consist of atomic symbols and subscripts (A4). The number of columns is assumed to be at most one more than the number of rows (A1).
Sources	[13]
Ref. By	IM2, IM3

Number	IM2
Label	Determination of feasibility
Input	E from IM1
Output	F if a row of E does not have a positive and a negative entry or if $(\mathbf{E} \mathbf{0})$ has no solutions and T otherwise
Description	E is a matrix representing the amount of each element in each compound in a reaction from IM1. [Is this necessary? —SC] If a row of E only has either a positive or negative entry, then there is an element that is either a reactant or a product but not both, which breaks the Law of Conservation of Mass (TM2). [Is this a good place for this? —SC] 0 is the zero matrix (see Mathematical Notation). Yes
Sources	[13]
Ref. By	IM3, UC1

Number	IM3
Label	Vector of compound coefficients in balanced equation
Input	E from IM1 The input is constrained so that E is feasible (IM2). [Is this correct? —SC]
Output	c , such that it is the smallest solution to $\mathbf{Ec} = \mathbf{0}$ from TM1 [Is this sufficient? —SC]
Description	E is a matrix representing the amount of each element in each compound in a reaction from IM1. [Is this necessary? —SC] c is a vector of whole number coefficients for each compound in the reaction (in the order inputted) so that the equation is balanced. 0 is the zero matrix (see Mathematical Notation).
Sources	[13]
Ref. By	UC1

4.2.6 Input Data Constraints

There are no input data constraints.

4.2.7 Properties of a Correct Solution

The equation from IM3 can be used as a “sanity” check, using the original input (formatted as a matrix according to IM1) as **E** and the outputted vector of coefficients as **c**. Performing this matrix multiplication should result in the zero matrix, **0**.

5 Requirements

This section provides the functional requirements, the business tasks that the software is expected to complete, and the nonfunctional requirements, the qualities that the software is expected to exhibit.

5.1 Functional Requirements

- R1: Input a representation of a chemical equation.
- R2: Convert the inputted equation to matrix form (from IM1).
- R3: Determine if the inputted equation is feasible (see [Terminology and Definitions](#); from IM2).
- R4: If the inputted equation is feasible, balance the chemical equation with the smallest whole number coefficients possible (from IM3).
- R5: If the inputted equation is infeasible, output a descriptive message. [\[Should this go before R4? —SC\]](#) *yes*
- R6: If the inputted equation is feasible, output a balanced form of the equation in the same format as the input, using the coefficients from IM3.

5.2 Nonfunctional Requirements

- NFR1: **Accuracy:** Chemical equations are only useful if they are balanced [4], so computed coefficients from IM3 should be exact.
- NFR2: **Understandability:** A new intended user (as described by [User Characteristics](#)) should be able to learn how to use ChemCode in an acceptable amount of time, as measured by the procedure in Section X of the Verification and Validation (VnV) Plan. ✓
- NFR3: **Usability:** An intended user (as described by [User Characteristics](#)) should find ChemCode easy to use, as measured by the procedure in Section X of the VnV Plan. ✓
- NFR4: **Maintainability:** The traceability between requirements, assumptions, theoretical models, general definitions, data definitions, instance models, likely changes, and unlikely changes is completely recorded in traceability matrices in the SRS.
- NFR5: **Portability:** ChemCode should be able to run on systems with the corresponding programming language installed, including systems running on Windows or macOS. The tests from the VnV Plan should pass in these environments.
- NFR6: **Verifiability:** ChemCode is tested following the VnV Plan.

This is for maintainability of the SRS. You want maintainability of the code.

6 Likely Changes

make these assumptions of things that are excluded

- LC1: The system currently assumes that inputted chemical formulas are simple (A4). In the future, the user might be able to input more complex chemical formulas, such as hydrates (see [Terminology and Definitions](#)) or those with polymers or isotopes.
- LC2: In the future, ChemCode might be able to, given the amount of one substance (in moles) in a reaction, calculate the amount required/produced of every other substance (also in moles) in the reaction.¹ ✓
- LC3: In the future, ChemCode might be able to, given the amount of each reactant (in moles) in a reaction, determine the limiting reactant(s).¹ This is dependent on LC2.
- LC4: In the future, ChemCode might be able to, given the amount of more than one reactant (in moles) in a reaction, calculate the theoretical yield of each product (also in moles).¹ This is dependent on LC3.
- LC5: In the future, ChemCode might be able to, given the amount of more than one reactant (in moles) in a reaction, calculate the amount of excess reactant(s) (also in moles).¹ This is dependent on LC3.
- LC6: The system currently assumes that inputted chemical formulas are formatted according to some set of conventions (A3). In the future, ChemCode might be able to parse valid but incorrectly formatted chemical formulas inputted by the user and format them correctly when outputting them.
- LC7: In the future, the user might be able to enter the amounts required by LC2, LC3, LC4, and LC5 in terms of mass (e.g., in grams).¹
- LC8: In the future, ChemCode might be able to classify a chemical reaction as “combination (or synthesis), decomposition, combustion, single replacement, [or] double replacement” [4, p. 301].¹
- LC9: In the future, ChemCode might also be able to classify “oxidation-reduction reactions, ... acid-base reactions, and condensation reactions” [4, p. 301].¹ This should be done after LC8.
- LC10: In the future, ChemCode might allow the user to input phase labels.¹
- LC11: In the future, ChemCode might be able to, given the phase labels for the reactants of a reaction, identify the phase labels for the products, which would involve determining solubility.¹ This is dependent on LC10.
- LC12: In the future, ChemCode might be able to identify when a reaction will not take place.¹ This is dependent on LC11.

¹These examples of problems related to chemical equations were taken from [4].

7 Unlikely Changes

*nice to include
rationale*

- UC1: One of the system constraints in Subsection 3.3 is that ChemCode will be implemented using matrices and linear algebra. It is unlikely that a different method will be used to compute the values from IM2 and IM3, since “balancing chemical equations is not chemistry; it is just linear algebra” [18, p. 193].
- UC2: A1 assumes that for all inputted chemical equations, there is at most one more compound than element. Using WebQC’s chemical equation balancer [19], equations without this constraint, such as $O_3 + C \rightarrow O_2 + CO_2$, “can be balanced in an infinite number of ways” and are “combination[s] of two different reactions” [19]. Manual empirical analysis of these equations led to unexpected results. The constraint in A1 might be a law of nature and a byproduct of “balancing chemical equations [being] just linear algebra” [18, p. 193], but preliminary investigation did not find any proof of this. Therefore, this constraint will remain an assumption and is unlikely to be relaxed.

✓

8 Traceability Matrices

The purpose of the traceability matrices is to provide easy references on what has to be additionally modified if a certain component is changed. Every time a component is changed, the items in the column of that component that are marked with an “X” may have to be modified as well. Table 3 shows the dependencies of theoretical models, general definitions, data definitions, instance models, and data constraints on each other. Table 5 shows the dependencies of theoretical models, general definitions, data definitions, instance models, likely changes, and unlikely changes on the assumptions. Table 4 shows the dependencies of requirements on instance models.

	TM1	TM2	IM1	IM2	IM3	4.2.6	4.2.7
TM1							
TM2							
IM1							X
IM2		X	X				
IM3	X		X	X			X

Table 3: Traceability Matrix Showing the Connections Between Items of Different Sections


	IM1	IM2	IM3
R1			
R2	X		
R3		X	
R4			X
R5			
R6			X
NFR1			X
NFR2			
NFR3			
NFR4			
NFR5			
NFR6			

Table 4: Traceability Matrix Showing the Connections Between Requirements and Instance Models

	A1	A2	A3	A4
TM1				
TM2				
IM1	X	X	X	X
IM2				
IM3				
LC1				X
LC2				
LC3				
LC4				
LC5				
LC6			X	
LC7				
LC8				
LC9				
LC10				
LC11				
LC12				
UC1				
UC2	X			

Table 5: Traceability Matrix Showing the Connections Between Assumptions and Other Items

References

- [1] D. O’Sullivan and D. J. Unwin, “Appendix A: Notation, Matrices, and Matrix Mathematics,” in *Geographic Information Analysis*, pp. 373–394, Hoboken, NJ, USA: Wiley, 2 ed., Mar. 2010.
- [2] E. W. Weisstein, “Zero Matrix,” 2023. Publisher: Wolfram Research, Inc.
- [3] M. Taboga, “Augmented matrix,” 2021.
- [4] L. Lund and Anoka-Ramsey Community College, *Introduction to Chemistry*. Cambridge and Coon Rapids, MN, USA: LibreTexts, Jan. 2023.
- [5] D. L. Parnas and P. Clements, “A rational design process: How and why to fake it,” *IEEE Transactions on Software Engineering*, vol. 12, no. 2, pp. 251–257, February 1986.
- [6] W. S. Smith, “capTemplate,” Sept. 2022. 
- [7] W. S. Smith and L. Lai, “A new requirements template for scientific computing,” in *Proceedings of the First International Workshop on Situational Requirements Engineering Processes – Methods, Techniques and Tools to Support Situation-Specific Requirements Engineering Processes, SREP’05* (J. Ralyté, P. Ågerfalk, and N. Kraiem, eds.), (Paris, France), pp. 107–121, In conjunction with 13th IEEE International Requirements Engineering Conference, 2005.
- [8] W. S. Smith, L. Lai, and R. Khedri, “Requirements analysis for engineering computation: A systematic approach for improving software reliability,” *Reliable Computing, Special Issue on Reliable Engineering Computation*, vol. 13, pp. 83–107, February 2007.
- [9] N. Koothoor, *A Document Driven Approach to Certifying Scientific Computing Software*. MASc thesis, McMaster University, Hamilton, ON, Canada, May 2013.
- [10] J. Carette, S. Smith, J. Balaci, T.-Y. Wu, S. Crawford, D. Chen, D. Szymczak, B. MacLachlan, D. Scime, and M. Niazi, “Drasil,” Feb. 2021.
- [11] B. MacLachlan, *A Design Language for Scientific Computing Software in Drasil*. MASc thesis, McMaster University, Hamilton, ON, Canada, June 2020.
- [12] D. Chen, *Solving Higher-Order ODEs in Drasil*. MEng thesis, McMaster University, Hamilton, ON, Canada, Sept. 2022.
- [13] I. Hamid, “Balancing Chemical Equations by Systems of Linear Equations,” *Applied Mathematics*, vol. 10, pp. 521–526, July 2019.
- [14] HarperCollins Publishers, “hydrate.”
- [15] The Editors of Encyclopaedia Britannica, “nonstoichiometric compound,” Oct. 2010.

[16] D. Margalit and J. Rabinoff, "Matrix Equations," in *Interactive Linear Algebra*, pp. 44–52, Atlanta, GA, USA: Georgia Institute of Technology, June 2019.

[17] D. Zwick, "Math 2270 - Lecture 16: The Complete Solution to $Ax = b$," 2012.

[18] B. Risteski, "A new complex vector method for balancing chemical equations," *Materiali in Tehnologije*, vol. 44, pp. 193–203, 07 2010.

[19] "Balance Chemical Equation - Online Balancer," 2023.

I don't think it would be that hard to formalize what we mean by an equation and the connection b/w the equation and the augmented matrix. You can also formalize what it means ~~to~~ to have a parametric solution to the underdetermined syst. of eq.

- type definition for R as a seq. of coefficients and compounds / a seq. P as a seq. of coefficients and compounds

(we keep coefficients separate because we'll use it for the matrix)

- type def for compounds
- type def for molecules

- each The coeffs of $AE = 0$ are the concatenation of the seq. from R & P
- each column of E can be built by taking each compound of R & P in the sequence. The R entries are positive, the P entries are negative. The entries rows in each column are the count of each element in the compound.
- we need a theory that extracts the subset of elements that are relevant for a specific chem. eq.