**EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS**
2026, Vol. 19, Issue 1, Article Number 7355
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

# Semigroup Codes for Diagnostic Sequences

Marshal I. Sampson[1], R. B. Abubakar[2], Reny George[3,*], Julie S. George[4]

[1] *Department of Mathematics, Akwa Ibom State University, Ikot Akpaden, Nigeria*
[2] *Department of Mathematics and Statistics, Federal University Otuoke, Bayelsa, Nigeria*
[3] *Department of Mathematics, College of Science and Humanities in Alkharj, Prince Sattam bin Abdulaziz University, Alkharj 11942, Saudi Arabia*
[4] *Department of Health and Rehabilitation Sciences, College of Applied Medical Science, Prince Sattam bin Abdulaziz University, Alkharj 11942, Saudi Arabia*

---

**Abstract.** This paper develops a semigroup-theoretic framework for the algebraic modelling of diagnostic processes. Diagnostic sequences are represented as words over finite alphabets of test actions, and their structural properties are examined via kernel congruences and Krohn–Rhodes-type decompositions. Beyond establishing these foundational results, we illustrate how algebraic reduction can potentially simplify diagnostic structures by yielding canonically reduced diagnostic words that eliminate redundant clinical transitions, thereby offering a principled mechanism for streamlining diagnostic pathways. The decomposition of diagnostic semigroups into reversible and irreversible components further provides a hierarchical modelling strategy that mirrors the layered structure of clinical decision-making. Stability properties derived from the semigroup action furnish an algebraic criterion for robustness under repeated evaluations, offering conceptual insights into diagnostic reliability. We also show that the associated kernel congruence aligns with Myhill–Nerode equivalence, enabling a direct interface with automata-theoretic and symbolic computation methods. Taken together, the theory provides a rigorous algebraic foundation for the formal analysis of diagnostic sequences, with illustrative implications for workflow design, decision support systems, and the structural evaluation of diagnostic protocols.

**2020 Mathematics Subject Classifications**: 20M35, 68Q70, 92C50

**Key Words and Phrases**: Semigroup codes, diagnostic semigroups, kernel congruence, Krohn–Rhodes decomposition, diagnostic optimization, oncology, infectious diseases, clinical protocols

---

## 1. Introduction

Algebraic structures have long provided rigorous tools for modelling deterministic systems governed by compositional rules. Among them, semigroup theory—the study of

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v19i1.7355

*Email addresses:* marshalsampson@aksu.edu.ng (M. I. Sampson), abubakarrb@fuotuoke.edu.ng (R. B. Abubakar), renygeorge02@yahoo.com (R. George), sanjujdr@gmail.com (J. S. George)

 

associative binary operations on sets—has been central to the algebraic representation of transformations, automata, and process dynamics [1, 2]. Beyond pure algebra, semigroup concepts have informed system theory, computation, and biological modelling, revealing how composition encodes information flow and operational dependencies.

In diagnostic and therapeutic contexts, each diagnostic action—such as a test, imaging, or decision rule—acts as a transformation on the state space of patient information. Sequences of such actions form composite transformations, whose totality naturally defines a *transformation semigroup*. This viewpoint permits the analysis of diagnostic efficiency, redundancy, and structural stability within a unified algebraic setting.

Historically, the foundational results of Rhodes and Krohn–Rhodes established that every finite semigroup admits a hierarchical decomposition into wreath products of simpler components [3–6]. Later extensions by Eilenberg and Pin connected these algebraic constructions to automata and language theory [7, 8]. Yet, such algebraic frameworks have rarely been applied to the formal study of diagnostic or decision-making systems. Semigroup-based formulations have found relevance in system composition, symbolic dynamics, and automata control, with recent extensions to stochastic and biological process modeling via transformation semigroups [6].

This paper introduces a *semigroup–theoretic model of diagnostic processes*, where diagnostic sequences are treated as algebraic words over finite alphabets of test actions. Within this framework, we develop a formal notion of *kernel congruence* to classify diagnostically equivalent sequences and apply a Krohn–Rhodes–type decomposition to derive hierarchical structures of diagnostic operations. The resulting model not only captures the logical composition of diagnostic actions but also establishes a foundation for optimization and complexity analysis. In particular, the proposed *diagnostic complexity index* provides an algebraic measure of the minimal hierarchical depth required to represent a diagnostic semigroup, thereby connecting semigroup structure to diagnostic efficiency and stability. This work is theoretical in nature and focuses on the algebraic structure of diagnostic processes rather than empirical validation. No clinical datasets are analysed, and no experimental performance evaluation is claimed. The examples and applications discussed throughout the paper are intended to illustrate modelling principles and structural implications of the theory, while empirical validation and implementation are identified as directions for future research.

## 1.1. Aim and Objectives

This paper aims to develop a unified algebraic and computational framework for formalizing diagnostic processes through semigroup–theoretic structures, while providing practical tools for their simplification and optimization. The approach begins by representing diagnostic actions as generators of a transformation semigroup acting on the state space of accumulated diagnostic information. It proceeds to analyse equivalence and redundancy in diagnostic sequences through the word problem and kernel congruence, thereby identifying opportunities for simplification. To manage complex diagnostic semigroups, Krohn–Rhodes decomposition is employed, breaking them down into interpretable and structurally meaningful components. Finally, the framework is applied to the optimization of diagnostic protocols, demonstrating how semigroup codes can streamline decision processes and enhance the efficiency of medical diagnosis.

### 1.2. Organisation of the Paper

The paper is organised as follows. Section 2 presents the algebraic preliminaries, including semigroups, the full transformation semigroup, homomorphisms, congruences, the word problem, codes, and the Krohn–Rhodes decomposition, together with the notation employed throughout the paper. Section 3 contains the principal theoretical contributions: it develops the methodological construction of diagnostic transformation semigroups, formulates redundancy and diagnostic equivalence via the kernel congruence, and establishes decomposition and structural optimisation results; the methodological material is presented in Section 3.1 and includes the minimal–word algorithm, the diagnostic decomposition theorem, and stability and reachability analyses. Section 4 discusses algebraic and computational implications, provides theoretical generalizations, and sets out design implications for clinical protocols. Section 4.7 illustrates applications and case studies, including diagnostic sequencing in oncology, infectious disease testing protocols, automated decision-support systems, networked diagnostic systems, and quantitative measures of diagnostic efficiency. Section 5 summarises the contributions, states limitations, and outlines directions for future research.

**Positioning and Scope of Contributions.** The algebraic tools employed in this paper, including kernel congruences, minimal-word constructions, and Krohn–Rhodes decompositions, are classical results in finite semigroup and automata theory. The novelty of this work does not lie in the introduction of new semigroup-theoretic theorems, but rather in the systematic formulation and interpretation of these classical structures within the context of diagnostic sequences and decision workflows. By reframing diagnostic actions as generators of transformation semigroups and interpreting decomposition, equivalence, and stability in diagnostic terms, the paper provides a unified algebraic modelling framework that bridges abstract semigroup theory with diagnostic reasoning and protocol design.

**Notation.** Throughout this paper, $\mathcal{M}$ denotes a finite alphabet of diagnostic actions, and $\mathcal{M}^+$ the set of all finite non–empty words over $\mathcal{M}$. Function composition is written from left to right, that is, $(fg)(x) = g(f(x))$. The identity transformation, when it exists, is denoted by $I$.

## 2. Preliminaries and Definitions

This section recalls the algebraic notions on which the proposed framework is founded. The definitions are presented in a self–contained manner, following the classical treatments in [1, 2, 7, 8].

## 2.1. Semigroups and Transformations

**Definition 1.** *A* semigroup *is an ordered pair* $(S, \cdot)$ *consisting of a non–empty set* $S$ *together with an associative binary operation* $\cdot : S \times S \to S$*. That is, for all* $a, b, c \in S$*, we have* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$*. If there exists* $e \in S$ *such that* $e \cdot a = a \cdot e = a$ *for all* $a \in S$*, then* $e$ *is called an* identity element*, and* $(S, \cdot)$ *is said to be a* monoid*.*

**Definition 2.** *Let* $X$ *be a non–empty set. The collection* $\mathcal{T}_X$ *of all mappings* $f : X \to X$ *under the operation of composition forms a semigroup, called the* full transformation semigroup *on* $X$*. Any subsemigroup of* $\mathcal{T}_X$ *is called a* transformation semigroup on $X$*.*

Transformation semigroups provide the natural algebraic setting for representing deterministic systems. Each element corresponds to a transformation of the system state, and composition reflects the sequential application of transformations.

**Example 1.** *Let* $X = \{x_1, x_2, x_3\}$ *denote three possible diagnostic states. Consider transformations* $a, b : X \to X$ *defined by*

$$a: \ x_1 \mapsto x_2, \ x_2 \mapsto x_3, \ x_3 \mapsto x_3, \qquad b: \ x_1 \mapsto x_1, \ x_2 \mapsto x_2, \ x_3 \mapsto x_1.$$

*Then* $\{a, b\}$ *generates a finite transformation semigroup acting on* $X$*, where each composition of* $a$ *and* $b$ *encodes a potential diagnostic sequence.*

## 2.2. Homomorphisms and Congruences

**Definition 3.** *A* homomorphism *between semigroups* $(S, \cdot)$ *and* $(T, \circ)$ *is a map* $\varphi : S \to T$ *such that* $\varphi(a \cdot b) = \varphi(a) \circ \varphi(b)$ *for all* $a, b \in S$*. If* $\varphi$ *is bijective and its inverse is also a homomorphism, then* $\varphi$ *is called an* isomorphism*.*

**Definition 4.** *An equivalence relation* $\rho$ *on a semigroup* $S$ *is called a* congruence *if, for all* $(a, b), (c, d) \in \rho$*, the pairs* $(a \cdot c, b \cdot d)$ *and* $(c \cdot a, d \cdot b)$ *also belong to* $\rho$*. The quotient set* $S/\rho$ *obtained from the equivalence classes then inherits a natural semigroup structure, called the* factor semigroup *of* $S$ *by* $\rho$*.*

**Definition 5.** *Given a homomorphism* $\varphi : S \to T$*, the* kernel congruence *of* $\varphi$ *is*

$$\ker(\varphi) = \{(u, v) \in S \times S : \varphi(u) = \varphi(v)\}.$$

*Two elements are congruent under* $\ker(\varphi)$ *if they induce the same image in* $T$*.*

## 2.3. The Word Problem and Codes

Let $\mathcal{M}$ be a finite alphabet. The free semigroup $\mathcal{M}^+$ consists of all non–empty words over $\mathcal{M}$ under concatenation. For a semigroup $S$ generated by $\mathcal{M}$, any homomorphism $\varphi : \mathcal{M}^+ \to S$ is determined by the images of the generators. The *word problem* for $S$ with respect to $\varphi$ asks whether two words $u, v \in \mathcal{M}^+$ satisfy $\varphi(u) = \varphi(v)$, that is, whether they represent the same element of $S$.

A subset $C \subseteq \mathcal{M}^+$ is called a *code* if every element of the subsemigroup generated by $C$ admits a unique factorization as a product of elements of $C$. Codes play an essential role in the representation of diagnostic sequences, since each diagnostic procedure should be expressible uniquely by its constituent diagnostic actions.

## 2.4. Krohn–Rhodes Decomposition

The Krohn–Rhodes theorem establishes that any finite semigroup can be decomposed into a cascade (or wreath product) of simpler components, specifically finite simple groups and reset semigroups [5]. This decomposition provides a canonical hierarchical structure, offering a powerful tool for analysing the internal organization of transformation semigroups.

**Definition 6.** *Let* $(S_i)_{i=1}^n$ *be a family of semigroups. The* wreath product $S_1 \wr S_2 \wr \cdots \wr S_n$ *is an iterated semidirect product representing a cascade of transformations in which the output of one level governs the action of the next. A semigroup* $S$ *is said to* divide *such a wreath product if* $S$ *is isomorphic to a subsemigroup of a homomorphic image of it.*

**Theorem 1** (Krohn–Rhodes)**.** *Every finite semigroup divides a finite wreath product of finite simple groups and finite reset semigroups.*

This theorem forms the algebraic foundation for decomposing diagnostic transformation systems into minimal interacting subsystems, a concept elaborated in Section 3.1; recent work has refined aspects of minimal faithful representations and decomposition techniques in related settings [5, 9].

## 2.5. Diagnostic Actions and States

Let $\mathcal{S}$ denote the state space of diagnostic information for a given medical condition. Each elementary diagnostic action $\mu_i$ acts as a transformation $\mu_i : \mathcal{S} \to \mathcal{S}$, mapping one informational state to another according to the outcome of a specific test or procedure. The collection $\mathcal{M} = \{\mu_1, \mu_2, \ldots, \mu_m\}$ of such actions generates a transformation semigroup

$$\mathcal{T} = \langle \mathcal{M} \rangle \subseteq \mathcal{T}_{\mathcal{S}},$$

whose elements represent all possible compositions of diagnostic operations. A *diagnostic sequence* is then a word $w = \mu_{i_1}\mu_{i_2}\cdots\mu_{i_k} \in \mathcal{M}^+$, and the mapping $\Phi : \mathcal{M}^+ \to \mathcal{T}$ defined by $\Phi(w) = \mu_{i_1} \circ \mu_{i_2} \circ \cdots \circ \mu_{i_k}$ is the canonical homomorphism encoding the diagnostic semigroup structure.

**Remark 1.** *The semigroup representation of diagnostic actions provides an algebraically closed framework for analysing reachability, redundancy, and simplification within diagnostic pathways. In subsequent sections this representation will be exploited to construct semigroup codes and to derive optimization algorithms.*

# 3. Main Results

## 3.1. Methodology

This section develops the mathematical methodology used to construct and analyse semigroup codes for diagnostic sequences. The approach begins with the formal representation of diagnostic actions as transformations, proceeds to the identification of redundancy through the word problem, and culminates in the structural reduction of diagnostic semigroups via Krohn–Rhodes decomposition.

### 3.1.1. Semigroup Modeling of Diagnostic Actions

Let $\mathcal{S}$ denote the finite set of diagnostic states representing the information accumulated about a patient during the course of medical examination. Each diagnostic action $\mu_i$ is a transformation $\mu_i : \mathcal{S} \to \mathcal{S}$, and the collection $\mathcal{M} = \{\mu_1, \mu_2, \ldots, \mu_m\}$ generates a transformation semigroup $\mathcal{T} = \langle \mathcal{M} \rangle$. The semigroup operation corresponds to sequential composition of actions, while the resulting element encodes the overall diagnostic effect.

**Definition 7.** *The* diagnostic transformation semigroup *associated with the system $\mathcal{S}$ is the pair $(\mathcal{T}, \circ)$, where $\mathcal{T} = \langle \mathcal{M} \rangle$ is the subsemigroup of $\mathcal{T}_{\mathcal{S}}$ generated by $\mathcal{M}$, and $\circ$ denotes composition of mappings.*

To capture the composition of diagnostic processes, define a homomorphism

$$\Phi : \mathcal{M}^+ \to \mathcal{T}, \qquad \Phi(\mu_{i_1}\mu_{i_2}\cdots\mu_{i_k}) = \mu_{i_1} \circ \mu_{i_2} \circ \cdots \circ \mu_{i_k}.$$

Each word $w \in \mathcal{M}^+$ corresponds to a concrete diagnostic sequence, and the image $\Phi(w)$ represents its induced transformation on the diagnostic state space.

**Example 2.** *Consider $\mathcal{S} = \{s_1, s_2, s_3, s_4\}$, representing successive refinement stages in an infectious disease diagnosis. Let*

$$\mu_1 : \ s_1 \mapsto s_2, \ s_2 \mapsto s_3, \ s_3 \mapsto s_3, \ s_4 \mapsto s_4; \qquad \mu_2 : \ s_1 \mapsto s_1, \ s_2 \mapsto s_4, \ s_3 \mapsto s_4, \ s_4 \mapsto s_4.$$

*Then $\mathcal{M} = \{\mu_1, \mu_2\}$ generates a transformation semigroup of eight distinct elements under composition. The word $w = \mu_1\mu_2\mu_1$ corresponds to the composite diagnostic process in which a preliminary test $\mu_1$ is followed by a confirmatory test $\mu_2$, and finally a refinement step $\mu_1$.*

**Remark 2.** *The above representation allows one to encode complex diagnostic processes algebraically, independent of probabilistic or causal assumptions. Each transformation captures the deterministic logical transition between diagnostic states, while the semigroup as a whole encapsulates the reachability structure of the entire diagnostic system.*

### 3.1.2. Redundancy and the Word Problem in Diagnostics

Redundancy arises when two different diagnostic sequences induce identical overall transformations. Formally, for $u, v \in \mathcal{M}^+$, if $\Phi(u) = \Phi(v)$ then both sequences lead to the same diagnostic outcome. The challenge is to determine minimal representatives among equivalent sequences.

**Definition 8.** *Two words $u, v \in \mathcal{M}^+$ are said to be* diagnostically equivalent *if $\Phi(u) = \Phi(v)$. The set*

$$\rho_\Phi = \{(u, v) \in \mathcal{M}^+ \times \mathcal{M}^+ : \Phi(u) = \Phi(v)\}$$

*is a congruence on $\mathcal{M}^+$, called the* diagnostic congruence *induced by $\Phi$.*

The corresponding factor semigroup $\mathcal{M}^+/\rho_\Phi$ partitions the space of diagnostic sequences into equivalence classes, each class representing a distinct diagnostic transformation. A canonical representative of each class can then be chosen to obtain a non–redundant encoding of diagnostic procedures.

**Definition 9.** *A* minimal diagnostic code *is a subset $C \subseteq \mathcal{M}^+$ containing exactly one representative from each equivalence class of $\rho_\Phi$. Every diagnostic process can thus be uniquely expressed as a composition of elements from $C$.*

The problem of constructing such a code can be viewed as an instance of the classical *word problem* for semigroups. Deciding whether two diagnostic sequences are equivalent amounts to testing whether their corresponding transformations in $\mathcal{T}$ coincide.

**Proposition 1.** *Let $\mathcal{M}$ be finite and $\mathcal{S}$ the corresponding finite diagnostic state space. Then the word problem for $\Phi : \mathcal{M}^+ \to \mathcal{T}_\mathcal{S}$ is decidable.*

*Proof.* Since $\mathcal{T}_\mathcal{S}$ is finite, each transformation can be represented as a finite table mapping $\mathcal{S}$ to itself. To determine whether $\Phi(u) = \Phi(v)$, compute the resulting transformations by successive composition of generators and compare their action on each element of $\mathcal{S}$. The algorithm terminates in finitely many steps, proving decidability; for related computational perspectives on practical transformation semigroup algorithms and applications to modelling, see [6, 10].

In practice, this decision process can be automated to identify redundant diagnostic paths. Let $\mathsf{Red}(u)$ denote the reduced form of a diagnostic sequence $u$ obtained by replacing any subword $v$ such that $(u, v) \in \rho_\Phi$ with its canonical representative. The resulting code ensures minimal length and avoids repetition of diagnostically equivalent subsequences.

### 3.1.3. Decomposition and Structural Optimization

The complexity of a diagnostic system grows rapidly with the number of possible tests and outcomes. Krohn–Rhodes decomposition provides a principled means of analysing this complexity by expressing the diagnostic semigroup as a composition of simpler components.

Let $\mathcal{T}$ be the diagnostic transformation semigroup. By the Krohn–Rhodes theorem, $\mathcal{T}$ divides a wreath product of simple groups and reset semigroups:

$$\mathcal{T} \text{ divides } G_1 \wr R_1 \wr G_2 \wr R_2 \wr \cdots \wr G_n \wr R_n.$$

Each group $G_i$ corresponds to a reversible component of the diagnostic process (representing reversible logical decisions), while each reset semigroup $R_i$ corresponds to irreversible steps such as definitive tests or terminal conclusions.

**Definition 10.** *The* diagnostic decomposition *of $\mathcal{T}$ is the minimal wreath product of simple and reset components through which $\mathcal{T}$ divides. The number of layers in this decomposition is called the* diagnostic complexity index*.*

**Proposition 2.** *If $\mathcal{T}$ is finite, then its diagnostic complexity index is finite. Moreover, each layer in the decomposition corresponds to a distinct level of decision granularity in the diagnostic process.*

*Proof.* Finiteness follows from the finiteness of $\mathcal{T}$ and the Krohn–Rhodes theorem. Each factor in the wreath product acts on a distinct projection of the state space, capturing either reversible permutations (group components) or absorbing reductions (reset components). Consequently, each level corresponds to a structural refinement of diagnostic reasoning.

The decomposition thus partitions the diagnostic system into hierarchically organized subsystems, facilitating targeted optimization. Redundant or repetitive subsystems correspond to repeated reset components, while essential decision nodes align with non–trivial group factors. This correspondence enables algebraic diagnosis of inefficiencies within the overall medical protocol.

**Remark 3.** *In computational implementation, the decomposition can be represented by a layered automaton whose transitions correspond to components of the wreath product. Simplifying the automaton by merging equivalent layers yields a reduced diagnostic semigroup that retains functional equivalence while minimizing redundant transformations.*

Next we formalize the principal theoretical outcomes derived from the methodology. We provide the minimal–word algorithm for constructing optimal diagnostic codes, state the decomposition theorem governing hierarchical structure, and analyse stability and reachability properties within the diagnostic semigroup.

## 3.2. Minimal Diagnostic Word Algorithm

We first introduce an explicit algorithmic framework for computing minimal representatives of diagnostic equivalence classes.

**Theorem 2** (Minimal Diagnostic Word Algorithm)**.** *Let $\Phi : \mathcal{M}^+ \to \mathcal{T}$ be the diagnostic homomorphism defined in Section 3.1. There exists an algorithm that produces, for every $u \in \mathcal{M}^+$, a canonical representative $w \in \mathcal{M}^+$ such that $\Phi(u) = \Phi(w)$ and $w$ is of minimal length among all such representatives.*

*Proof.* We give a constructive algorithm and prove its correctness and termination.

**Preparation: forming the (finite) transformation monoid.** Let $\mathcal{M} = \{\mu_1, \dots, \mu_m\}$ be the finite generating set and let $\mathcal{S}$ be the finite diagnostic state space. As in Section 2, each generator $\mu_i$ is a mapping $\mu_i : \mathcal{S} \to \mathcal{S}$. Consider the set

$$\mathcal{T}_0 = \{\mu_{i_1} \circ \mu_{i_2} \circ \cdots \circ \mu_{i_k} \ : \ k \geq 1, \ \mu_{i_j} \in \mathcal{M}\},$$

the semigroup generated by $\mathcal{M}$. Adjoin an identity transformation $I : \mathcal{S} \to \mathcal{S}$ (if one is not already present) to obtain the finite monoid $\mathcal{T} := \mathcal{T}_0 \cup \{I\}$. Since $\mathcal{S}$ is finite, the set of all mappings $\mathcal{S} \to \mathcal{S}$ is finite, hence $\mathcal{T}$ is finite.

**Cayley graph and shortest–word reduction.** Define the (right) labelled Cayley digraph $\Gamma = (V, E)$ of the monoid $\mathcal{T}$ with respect to the generating set $\mathcal{M}$ as follows:

- $V = \mathcal{T}$ (the vertices are the elements of the monoid);
- for each $t \in \mathcal{T}$ and each generator $\mu \in \mathcal{M}$ there is a directed edge

  $$t \xrightarrow{\mu} t \circ \mu,$$

  labelled by $\mu$.

By construction, a path in $\Gamma$ that starts at the identity $I$ and follows labels $\mu_{i_1}, \dots, \mu_{i_k}$ arrives at the vertex

$$I \circ \mu_{i_1} \circ \cdots \circ \mu_{i_k} = \mu_{i_1} \circ \cdots \circ \mu_{i_k} \in \mathcal{T},$$

so the label of the path is a word whose image under $\Phi$ equals the terminal vertex.

**Algorithm.** Given an input word $u \in \mathcal{M}^+$, let $t = \Phi(u) \in \mathcal{T}$ be the transformation produced by $u$. Execute the following steps:

(i) **Compute the monoid $\mathcal{T}$**: starting from the set $\mathcal{M} \cup \{I\}$, iteratively close under composition (compute all products $t \circ \mu$ for known $t$ and $\mu \in \mathcal{M}$) until no new map appears. Since $\mathcal{T}$ is a subset of the finite set of all functions $\mathcal{S} \to \mathcal{S}$, this process terminates and yields $\mathcal{T}$.

(ii) **Build the Cayley digraph** $\Gamma$ on vertex set $\mathcal{T}$ with labelled edges $t \xrightarrow{\mu} t \circ \mu$ for each $t \in \mathcal{T}$ and $\mu \in \mathcal{M}$.

(iii) **Breadth–first search (BFS)**: perform a BFS on $\Gamma$ starting from the source vertex $I$, recording for each visited vertex the predecessor vertex and the generator label used to reach it. Stop the BFS as soon as the target vertex $t$ is first discovered.

(iv) **Read off a shortest word:** follow the predecessor pointers from $t$ back to $I$ to obtain a shortest path; the concatenation of labels along this path (in order from $I$ to $t$) is a shortest word $w \in \mathcal{M}^+$ with $\Phi(w) = t$.

**Correctness.** Every word $v \in \mathcal{M}^+$ with $\Phi(v) = t$ corresponds to a path in $\Gamma$ from $I$ to $t$ whose label is $v$. BFS from $I$ finds shortest (in number of edges) paths to all reachable vertices; hence, the path returned by the algorithm is of minimal length among all words whose image equals $t$. Consequently, the associated word $w$ satisfies $\Phi(w) = t$ and has minimal length among all such words.

**Termination and complexity.** Step (i) terminates because there are only finitely many maps $\mathcal{S} \to \mathcal{S}$. Step (iii) (BFS) terminates because $\Gamma$ has finitely many vertices $|\mathcal{T}|$. Building $\mathcal{T}$ requires at most $O(|\mathcal{T}| \cdot |\mathcal{M}| \cdot |\mathcal{S}|)$ elementary operations if transformations are represented by their action on $\mathcal{S}$; BFS requires $O(|\mathcal{T}| \cdot |\mathcal{M}|)$ edge traversals. Thus the algorithm halts in finite time with polynomial dependence on $|\mathcal{T}|$, $|\mathcal{M}|$, and $|\mathcal{S}|$.

**Remark.** Since the congruence $\rho_\Phi$ partitions $\mathcal{M}^+$ into finitely many equivalence classes indexed by elements of the finite set $\mathcal{T}$, the above procedure produces for each class a canonical shortest representative. This completes the proof.

**Remark 4.** *The minimal–word algorithm described here corresponds algorithmically to a breadth-first search (BFS) approach for finding canonical representatives in each equivalence class of the kernel congruence. While the algorithmic procedure is classical in the context of finite automata and semigroup computation (see, e.g., [7, 8]), its application to the explicit representation of diagnostic sequences provides a concrete framework for formalizing redundancy and canonicalization in clinical or procedural workflows.*

*The algorithm ensures that each diagnostic sequence can be replaced by a shortest equivalent word, thereby eliminating redundancies in the operational workflow. The resulting set of canonical words constitutes the minimal diagnostic code defined earlier.*

**Proposition 3.** *Let $u, v \in \mathcal{M}^+$. Then $u$ and $v$ reduce to the same canonical form if and only if $\Phi(u) = \Phi(v)$. Hence the minimal–word algorithm defines a retraction from $\mathcal{M}^+$ onto a complete set of canonical representatives.*

*Proof.* If $\Phi(u) = \Phi(v)$, then $u, v$ belong to the same $\rho_\Phi$–class and hence share the same minimal representative. Conversely, if both reduce to the same canonical word $w$, then $\Phi(u) = \Phi(w) = \Phi(v)$, establishing the equivalence.

## 3.3. Decomposition Theorem for Diagnostic Processes

Having constructed minimal codes, we next characterise the internal algebraic structure of the diagnostic semigroup via decomposition.

**Theorem 3** (Diagnostic Decomposition Theorem)**.** *Let $\mathcal{T}$ be a finite diagnostic transformation semigroup. Then there exist simple groups $G_1, G_2, \ldots, G_r$ and reset semigroups $R_1, R_2, \ldots, R_s$ such that*

$$\mathcal{T} \text{ divides } G_1 \wr R_1 \wr G_2 \wr R_2 \wr \cdots \wr G_r \wr R_s.$$

*Moreover, the sequence of alternating components is unique up to divisibility equivalence.*

*Proof.* Let $\mathcal{T}$ be a finite diagnostic transformation semigroup acting on the finite state set $\mathcal{S}$.

**Existence.** The Krohn–Rhodes theorem (see [3, 11]) states that every finite semigroup divides a finite wreath product of finite simple groups and finite reset (aperiodic) semigroups. Because $\mathcal{T}$ is finite, the theorem guarantees the existence of a finite wreath product

$$W = G_1 \wr R_1 \wr G_2 \wr R_2 \wr \cdots \wr G_r \wr R_r$$

(consisting of alternating group factors $G_i$ and reset/aperiodic factors $R_i$) and a subsemigroup $U \leq W$ together with a surjective homomorphism $\psi : U \to \mathcal{T}$. In the usual terminology, $\mathcal{T}$ *divides* $W$. Thus a hierarchical (layered) wreath–product representation through which $\mathcal{T}$ divides always exists.

**Existence of a minimal layer count.** Let $\mathcal{D}$ denote the family of all such wreath products $W$ (of alternating group and reset factors) for which $\mathcal{T}$ divides $W$. For each $W \in \mathcal{D}$ define $L(W) \in \mathbb{N}$ to be the number of nontrivial factor layers appearing in the wreath product (for example, we count each $G_i$ and each $R_j$ that is not the trivial one–element semigroup). Since $\mathcal{T}$ is finite, $\mathcal{D}$ is nonempty by the existence paragraph. Hence the set of natural numbers $\{L(W) : W \in \mathcal{D}\}$ is nonempty and therefore admits a minimum $L_{\min} \in \mathbb{N}$. Choose $W_{\min} \in \mathcal{D}$ with $L(W_{\min}) = L_{\min}$. The wreath product $W_{\min}$ realises the smallest possible number of nontrivial layers through which $\mathcal{T}$ divides, so the layered representation induced by $W_{\min}$ is a minimal (canonical) hierarchical decomposition in the sense of having minimal depth. This proves existence of a canonical hierarchical decomposition and of the diagnostic complexity index $L_{\min}$.

**Uniqueness up to divisibility (precise statement and justification).** The assertion that such a minimal layered decomposition is unique should be interpreted in the standard sense used in the Krohn–Rhodes literature: *any two wreath products of group and reset components through which $\mathcal{T}$ divides and which both realise the minimal layer count are equivalent in the divisibility preorder* (in other words, each divides the other, or equivalently they occupy the same equivalence class under mutual divisibility). Establishing this mutual divisibility is not an immediate combinatorial consequence of the existence argument above; it relies on deeper structural properties of finite semigroups and on refinement arguments for cascaded decompositions.

Concretely, suppose $W_1$ and $W_2$ are two wreath products ... appear in the foundational works of Krohn and Rhodes and in subsequent expositions; see in particular Krohn & Rhodes [11], Rhodes [3], Holcombe [4], Rhodes & Steinberg [12], and recent refinements and algorithmic discussions [5, 9, 13] for updated perspectives and technical details.

**In conclusion,** combining the existence of a minimal layer count with the mutual–divisibility equivalence of any two minimal realisations (as justified by the cited Krohn–Rhodes refinement theory) yields the claimed statement: every finite diagnostic semigroup admits a canonical hierarchical decomposition, and the minimal number of layers required equals its diagnostic complexity index.

**Remark 5.** *The structural decomposition stated above is an application of the classical Krohn–Rhodes theorem for finite semigroups [11, 12]. No new decomposition result is claimed. The contribution of this work lies in interpreting the Krohn–Rhodes structure in the context of diagnostic sequences, where group and reset components are viewed as reversible and irreversible stages in diagnostic workflows, respectively. This interpretation yields a hierarchical modelling framework that renders classical decomposition results meaningful for diagnostic analysis.*

*Krohn–Rhodes decompositions are not unique in general. Accordingly, any statement of uniqueness in the diagnostic setting should be understood only in the classical sense of Krohn–Rhodes theory, namely uniqueness up to mutual divisibility of the resulting wreath products. No claim of canonical or absolute uniqueness is made.*

*The final step of the proof, concerning mutual divisibility of minimal decompositions, relies on a classical but nontrivial refinement argument. For full technical details, we refer the reader to the foundational sources [3, 4, 11, 12].*

**Corollary 1.** *Every finite diagnostic semigroup admits a canonical hierarchical decomposition, and the minimal number of layers required equals its diagnostic complexity index.*

*Proof.* Existence of a hierarchical decomposition follows from the Diagnostic Decomposition Theorem (Theorem 3), which asserts that any finite diagnostic transformation semigroup $\mathcal{T}$ divides a finite wreath product of alternating simple group and reset components

$$G_1 \wr R_1 \wr G_2 \wr R_2 \wr \cdots \wr G_r \wr R_s.$$

By definition, any such wreath product provides a hierarchical (layered) representation of $\mathcal{T}$ in which the factors $G_i$ encode reversible decision layers and the $R_j$ encode irreversible (reset) layers. Hence every finite diagnostic semigroup admits at least one hierarchical decomposition.

To see that a canonical (minimal) hierarchy exists and that its number of layers equals the diagnostic complexity index, consider the class of all wreath products of simple groups and reset semigroups through which $\mathcal{T}$ divides. Select a wreath product $W_{\min}$ with minimal number of nontrivial layers. By construction, $W_{\min}$ realises the smallest number of layers among all decompositions through which $\mathcal{T}$ divides, which we call the *canonical hierarchical decomposition* of $\mathcal{T}$.

The diagnostic complexity index (Definition **??**) is defined as the minimal number of layers in a wreath product through which $\mathcal{T}$ divides. Hence the minimal number of layers in the canonical hierarchical decomposition equals the diagnostic complexity index.

**Uniqueness up to divisibility:** The invariance of the minimal layer count follows from the uniqueness up to divisibility of minimal wreath-product decompositions as established in classical semigroup theory [12, Ch. 9]. Any two decompositions realizing the minimal layer count are equivalent in the divisibility preorder; this justifies referring to $W_{\min}$ as canonical in the sense of minimal hierarchical depth.

**Remark 6.** *Corollary 1 illustrates a direct consequence of the Diagnostic Decomposition Theorem within the diagnostic context. While the corollary is algebraically classical, its interpretation emphasizes the canonical hierarchical representation of diagnostic sequences, linking the minimal layer count to the diagnostic complexity index. This framing is intended to provide a clear conceptual understanding for applications in diagnostic workflows, rather than to claim novelty in the underlying semigroup theory.*

*The corollary formalizes the observation that any finite diagnostic semigroup can be represented in a layered structure with reversible (group) and irreversible (reset) components. While the result follows directly from established theory, its presentation here illustrates how classical decomposition results can be interpreted and applied in the modeling of diagnostic sequences.*

**Remark 7.** *The decomposition theorem enables a modular understanding of the diagnostic system: reversible diagnostic checks correspond to group layers, and irreversible clinical conclusions correspond to reset layers. Optimizing diagnostic structure thus becomes equivalent to reducing the number of nontrivial layers in the wreath product representation.*

### 3.4. Stability and Reachability Analysis

To evaluate the robustness of diagnostic systems, we introduce an algebraic notion of stability and characterize reachability of states.

**Definition 11.** *Let $\mathcal{S}$ denote the diagnostic state space and $\mathcal{T}$ its transformation semigroup. A subset $X \subseteq \mathcal{S}$ is said to be* stable *under $\mathcal{T}$ if for all $t \in \mathcal{T}$, $t(X) \subseteq X$. An element $s \in \mathcal{S}$ is* absorbing *if $t(s) = s$ for all $t \in \mathcal{T}$.*

**Proposition 4.** *If $\mathcal{T}$ contains a zero element $0_{\mathcal{T}}$, then $\mathrm{im}(0_{\mathcal{T}})$ forms a unique minimal stable subset of $\mathcal{S}$.*

*Proof.* By definition of the zero element, $t \circ 0_{\mathcal{T}} = 0_{\mathcal{T}}$ for all $t \in \mathcal{T}$. Hence $0_{\mathcal{T}}$ maps every state to a fixed image, which is invariant under $\mathcal{T}$. Minimality follows since any smaller subset would not be invariant under all transformations.

**Definition 12.** *For $s_1, s_2 \in \mathcal{S}$, we write $s_1 \Rightarrow s_2$ if there exists $t \in \mathcal{T}$ such that $t(s_1) = s_2$. The* reachability graph *$\mathcal{G}_{\mathcal{T}}$ is the directed graph on $\mathcal{S}$ with edges $(s_1, s_2)$ whenever $s_1 \Rightarrow s_2$.*

**Theorem 4.** *Let $\mathcal{T}$ be finite. Then:*

*(i) Every strongly connected component of $\mathcal{G}_{\mathcal{T}}$ corresponds to a $\mathcal{T}$–invariant subset of $\mathcal{S}$.*

*(ii) The quotient graph obtained by collapsing each component is acyclic and reflects the causal hierarchy of diagnostic transitions.*

*Proof.* (i) Strong connectivity ensures mutual reachability; invariance follows because for any $t \in \mathcal{T}$, application of $t$ preserves reachability relations within the component. (ii) Collapsing components yields a directed acyclic graph because no cycle can exist between distinct components without contradicting maximality of strong connectivity.

**Corollary 2.** *The asymptotic behaviour of the diagnostic process is determined by the terminal strongly connected components of $\mathcal{G}_{\mathcal{T}}$, which correspond to stable diagnostic outcomes.*

**Remark 8.** *Stability analysis enables prediction of long–term diagnostic convergence. If all trajectories eventually enter a single absorbing component, the system exhibits deterministic diagnostic stability; otherwise, multiple terminal components indicate structural ambiguity in the diagnostic logic.*

## 4. Discussion, Implications and Applications

The discussion that follows focuses on structural and theoretical implications of the proposed framework rather than empirical performance evaluation.

The algebraic framework developed in this paper establishes a structural theory of diagnostic systems by representing diagnostic transitions as morphisms within finitely generated transformation semigroups. In this section, we analyse the formal implications of our results from both algebraic and computational standpoints, emphasizing how decomposition, congruence reduction, and stability contribute to diagnostic efficiency and reliability.

### 4.1. Algebraic Efficiency of Diagnostic Codes

Let $\mathcal{M} = \{\mu_1, \mu_2, \ldots, \mu_n\}$ denote the generating set of diagnostic actions acting on a finite state space $\mathcal{S}$, and let $\mathcal{T} = \langle \mathcal{M} \rangle \leq \mathcal{T}_{\mathcal{S}}$ be the induced transformation semigroup. For each diagnostic sequence $u \in \mathcal{M}^+$, define its canonical image under the diagnostic morphism $\Phi : \mathcal{M}^+ \to \mathcal{T}$, and let $\rho_\Phi$ be the kernel congruence given by

$$(u, v) \in \rho_\Phi \iff \Phi(u) = \Phi(v).$$

The minimal–word algorithm of Section 3.1 constructs a unique canonical representative in each equivalence class $[u]_{\rho_\Phi}$. By eliminating redundant compositions, this procedure realises a quotient semigroup $\mathcal{M}^+/\rho_\Phi$ that is isomorphic to $\mathrm{im}(\Phi)$ and therefore of strictly smaller cardinality whenever diagnostic redundancy exists.

Consequently, the algebraic complexity of the diagnostic system, measured by $|\mathrm{im}(\Phi)|$, is reduced without loss of semantic expressivity. From an algorithmic standpoint, construction of $\rho_\Phi$ and selection of minimal representatives can be implemented in time $\mathcal{O}(mn)$, where $m = |\mathcal{T}|$ and $n = |\mathcal{M}|$, ensuring polynomial-time realizability of canonical encoding. Hence, the algebraic minimalization procedure serves as an analogue of DFA minimization in automata theory and establishes a measure of *diagnostic compression* intrinsic to the semigroup structure.

**Remark 9.** *Although the present work develops an algebraic framework for diagnostic sequence optimization, it does not implement direct comparisons with existing clinical or* *computational diagnostic optimization methods. Traditional approaches in medical informatics often rely on decision trees, rule-based systems, or heuristic scheduling of tests. Our semigroup-theoretic perspective complements these methods by providing a formal algebraic structure that can potentially inform workflow design, hierarchical decomposition, and symbolic minimization of diagnostic sequences. Future studies may include benchmarking against established optimization frameworks to quantify relative performance.*

## 4.2. Diagnostic Decomposition and Hierarchical Modelling

By Theorem 3, each finite diagnostic semigroup $\mathcal{T}$ divides a finite wreath product of the form

$$G_1 \wr R_1 \wr G_2 \wr R_2 \wr \cdots \wr G_r \wr R_r,$$

where each $G_i$ is a finite simple group and each $R_i$ is a finite reset (aperiodic) semigroup. This decomposition induces a hierarchical architecture in which group layers correspond to reversible diagnostic decisions and reset layers correspond to irreversible progressions in diagnostic certainty.

Formally, the mapping

$$\psi : \mathcal{T} \longrightarrow \prod_{i=1}^{r} (G_i \wr R_i)$$

preserves the order of diagnostic refinement and provides an algebraic mechanism for isolating independent diagnostic subroutines. The minimal number of nontrivial layers in such a decomposition, termed the *diagnostic complexity index*, thus quantifies the hierarchical depth required to simulate $\mathcal{T}$. Reducing this depth corresponds to eliminating algebraic redundancy in diagnostic reasoning, yielding structurally stable and computationally efficient designs.

**Remark 10.** *In terms of computational considerations, while detailed empirical performance metrics are not provided, the minimal–word algorithm is polynomial in the size of the generating set and the semigroup, ensuring tractable computation for finite diagnostic systems. Formal analysis of complexity in large-scale or continuous domains remains an avenue for future work, and the framework currently illustrates conceptual potential for efficiency improvements rather than providing quantified performance guarantees.*

## 4.3. Stability as Diagnostic Robustness

Let $\mathcal{T}$ act on $\mathcal{S}$ by transformations. We recall that $\mathcal{T}$ is *stable* if the induced directed graph on $\mathcal{S}$, whose edges correspond to actions of generators in $\mathcal{M}$, has a unique terminal strongly connected component (SCC). Denote this terminal component by $\Omega$. For any initial state $s \in \mathcal{S}$, repeated application of arbitrary words in $\mathcal{M}^+$ satisfies

$$\lim_{k\to\infty} s \cdot u_1 u_2 \cdots u_k \in \Omega.$$

Thus, independent diagnostic trajectories converge to a single asymptotic class of conclusions, providing a rigorous algebraic characterization of diagnostic robustness.

If $\mathcal{T}$ fails to be stable, multiple terminal SCCs exist, corresponding to distinct absorbing subsemigroups $\mathcal{T}_1, \dots, \mathcal{T}_p$. Such multiplicity encodes the possibility of non-deterministic or ambiguous diagnostic outcomes. Formally, adding new diagnostic actions (tests) refines $\mathcal{T}$ by extending its generating set, which in turn refines $\rho_\Phi$ until $\mathcal{T}$ becomes stable. Hence, algebraic stabilization corresponds to the convergence of diagnostic decision processes.

## 4.4. Computational and Logical Implications

Let $\mathcal{A}_\mathcal{T} = (\mathcal{S}, \mathcal{M}, \Phi)$ denote the deterministic automaton associated with $\mathcal{T}$. Two words $u, v \in \mathcal{M}^+$ are equivalent in $\mathcal{A}_\mathcal{T}$ if and only if they induce identical transformations on $\mathcal{S}$, i.e. $(u, v) \in \rho_\Phi$. This is precisely the Myhill–Nerode equivalence relation in formal language theory [7, 8], and hence the canonical quotient $\mathcal{M}^+/\rho_\Phi$ corresponds to the syntactic semigroup of the diagnostic language recognized by $\mathcal{T}$; recent structural analyses of finite regular semigroups offer complementary insights into inverse and permutation behaviours that relate to reachability and stability [13, 14].

In computational terms, the algebraic minimization of diagnostic codes therefore parallels the minimization of deterministic finite automata (DFA). Furthermore, the decomposition of $\mathcal{T}$ into group and aperiodic factors mirrors the division of regular languages into reversible and irreversible components under the Eilenberg correspondence between language varieties and pseudovarieties of semigroups. This duality suggests that diagnostic hierarchies can be studied through algebraic language theory, linking symbolic decision processes with the logical framework of clinical reasoning.

From a computational perspective, the complexity guarantees established in this work apply primarily to finite diagnostic semigroups with explicitly given generating sets. In such settings, construction of the kernel congruence and computation of minimal diagnostic words can be carried out in polynomial time in the size of the transformation semigroup. For large-scale diagnostic systems, however, the cardinality of the state space or the generating set may grow rapidly, making explicit enumeration impractical. In continuous or high-dimensional diagnostic domains, the transformation semigroup may no longer be finite, and direct computation must be replaced by symbolic, approximate, or abstraction-based methods. In these cases, the algebraic framework should be interpreted as providing structural guidance rather than explicit algorithms, identifying invariants, decompositions, and stability properties that can inform scalable implementations. A detailed complexity analysis for such large-scale or continuous systems lies beyond the scope of the present work and constitutes an important direction for future research.

**Remark 11.** *While the framework outlines the algebraic structure and reduction mechanisms of diagnostic sequences, detailed computational complexity analyses for large-scale or continuous diagnostic systems are not provided. The primary purpose of the algebraic minimization is to illustrate conceptual potential for efficiency improvements, with polynomial-time feasibility guaranteed for finite semigroups. Future work may include empirical performance evaluation and implementation benchmarks.*

## 4.5. Theoretical Generalizations

Several extensions of the present framework are possible:

(i) *Stochastic and probabilistic diagnostics.* The present framework is formulated for deterministic diagnostic actions represented as transformations on a finite state space. However, many real diagnostic processes involve uncertainty, noise, or probabilistic outcomes. An extension of the theory may be obtained by associating each diagnostic action with a stochastic operator or Markov transition kernel acting on probability distributions over the diagnostic state space. In this setting, the diagnostic semigroup is replaced by a semigroup of stochastic matrices or Markov operators, and composition corresponds to sequential probabilistic testing. While the algebraic notions of composition and reachability remain meaningful, questions of stability and convergence are naturally expressed in terms of invariant or absorbing distributions. This perspective is not developed in detail here, but it illustrates how the deterministic semigroup model can serve as a structural backbone for probabilistic diagnostic reasoning.

(ii) *Probabilistic extensions.* Associating each generator $\mu_i$ with a stochastic kernel yields a Markov semigroup acting on probability measures, allowing analysis of uncertain or probabilistic diagnostic events.

(iii) *Fuzzy and multi-valued semantics.* Introducing a fuzzy semigroup structure on $\mathcal{S}$ enables representation of vague or partially defined clinical states, an important feature in early-stage disease modelling.

These generalizations preserve the algebraic essence of the theory while expanding its applicability to richer domains of computational medicine and artificial intelligence.

## 4.6. Design Implications for Clinical Protocols

The algebraic perspective also provides a structured approach for the design and evaluation of diagnostic protocols. A diagnostic process can be optimized by:

(i) minimizing the cardinality of the generating set $\mathcal{M}$ while preserving diagnostic completeness;

(ii) ensuring that the decomposition depth (number of group–reset layers) is minimal, thereby reducing complexity; and

(iii) verifying that the reachability graph $\mathcal{G}_{\mathcal{T}}$ has a unique terminal component, ensuring deterministic diagnostic convergence.

Each of these conditions can be translated into implementable checks within a computer-aided diagnostic system, confirming the direct operational relevance of the algebraic model.

## 4.7. Illustrative and Conceptual Applications

The applications presented in this section are illustrative and conceptual. They demonstrate how the proposed algebraic framework can be applied to diagnostic reasoning in oncology, infectious disease testing, automated decision support systems, and networked diagnostics. No real clinical datasets were analysed, and no empirical validation is claimed. These examples serve to clarify modelling principles, highlight structural patterns, and illustrate how algebraic reduction can potentially simplify diagnostic sequences and reduce redundant steps, providing a conceptual basis for efficiency improvements. Future work may extend these studies with empirical evaluation or simulation-based validation.

The algebraic theory of diagnostic semigroups acquires concrete meaning only when interpreted through real or simulated diagnostic procedures. We therefore demonstrate its relevance in several domains of medical and computational practice, emphasizing how the theoretical constructs introduced above improve interpretability, stability, and efficiency of diagnostic protocols.

### 4.7.1. Illustrative / Conceptual Case Study: Diagnostic Sequencing in Oncology

Consider a cancer–staging process where diagnostic actions correspond to clinical tests such as imaging, histopathology, and biomarker assays. Each action $a_i \in \mathcal{M}$ induces a transformation $\mu_i$ on the diagnostic state space $\mathcal{S}$, representing an update of clinical knowledge. The semigroup $\mathcal{T} = \langle \mu_1, \mu_2, \mu_3, \ldots \rangle$ encapsulates all possible compositions of these diagnostic actions.

In practice, redundant compositions often occur when two or more test sequences yield the same staging outcome. Application of the minimal–word algorithm (Theorem 4.1) identifies canonical sequences corresponding to distinct diagnostic outcomes. For instance, if magnetic resonance imaging (MRI) followed by a biopsy produces the same diagnostic state as biopsy followed by MRI, these two routes are algebraically equivalent under the kernel congruence $\rho_\Phi$. The algorithm retains only one representative, reducing unnecessary repetition in clinical testing. Thus, algebraic minimality translates directly into reduced procedural cost and time. Related applications of formal and computational methods to clinical diagnostic workflows can be found in the medical informatics literature, where rule-based and algorithmic decision models are employed for oncology and infectious disease diagnostics; see, for example, [15, 16].

### 4.7.2. Illustrative / Conceptual Case Study: Infectious Disease Testing Protocols

Infectious disease diagnostics often involve layered testing strategies: initial screening, confirmatory testing, and differentiation of pathogen types. Each of these steps can be modeled as elements of a transformation semigroup acting on the patient's infection–state space.

By applying the decomposition theorem (Theorem 4.2), the process decomposes naturally into reversible group layers (e.g., confirmatory re-tests) and irreversible reset layers (e.g., final classification as positive or negative). This structure clarifies which testing stages can be safely iterated without diagnostic distortion and which transitions mark permanent clinical commitments.

Furthermore, stability analysis (Section 4.3) identifies the absorbing states corresponding to final diagnoses. For example, repeated rapid–antigen and polymerase chain reaction (PCR) tests may eventually converge to a unique positive or negative state, showing algebraic stability of the testing protocol. Conversely, oscillation between inconclusive and negative results signals the presence of multiple terminal components in the reachability graph $\mathcal{G}_{\mathcal{T}}$, revealing procedural ambiguity that requires revision of testing thresholds.

#### 4.7.3. Illustrative / Conceptual Case Study: Automated Decision Support Systems

Modern decision support systems employ symbolic or rule–based modules integrated with probabilistic reasoning. A diagnostic semigroup can serve as the algebraic backbone of such systems. Each rule corresponds to a generator, and the semigroup of all rule compositions describes the system's inference capabilities.

The minimal–word algorithm becomes a method for logical simplification: redundant rule sequences are collapsed, yielding a smaller and faster decision module. The decomposition theorem then provides a hierarchy of subsystems: group components encode reversible logical modules (for example, consistency–checking routines), while reset components represent final decision points. This division parallels modular software design, where reversible computations are encapsulated in functions and irreversible state changes correspond to output commits.

From a computational standpoint, the algebraic reduction of diagnostic words ensures that every inference step corresponds to a transition in the canonical quotient semigroup $\mathcal{T}/\rho_\Phi$, thereby eliminating redundant rule applications. By reducing diagnostic sequences to their canonical forms, the minimal–word algorithm illustrates the potential to decrease redundant steps and streamline rule-based systems, providing a conceptual framework for improving computational efficiency.

#### 4.7.4. Illustrative / Conceptual Case Study: Algebraic Diagnostics in Networked Systems

Beyond clinical medicine, the diagnostic semigroup framework extends to cyber–physical and distributed computing systems, where diagnostic events correspond to fault–detection or status–update actions. Each node in a network can be modeled as possessing a local state space $\mathcal{S}_i$ and transformation semigroup $\mathcal{T}_i$. The global diagnostic semigroup $\mathcal{T}$ is then the direct product $\prod_i \mathcal{T}_i$, acting on the joint state space $\prod_i \mathcal{S}_i$.

Decomposition of $\mathcal{T}$ identifies subsystems that can diagnose independently (group layers) and those that must synchronize globally (reset layers). This insight assists in designing distributed monitoring protocols that minimize communication overhead while preserving collective reliability. Hence, the same algebraic principles that optimize medical diagnostics also enhance resilience and efficiency in large–scale engineered systems.

#### 4.7.5. Quantitative Evaluation of Diagnostic Efficiency

To provide a unified metric for diagnostic optimization, we define the *diagnostic efficiency index* $\eta(\mathcal{T})$ by

$$\eta(\mathcal{T}) = \frac{|\mathcal{M}|}{|\mathcal{T}|},$$

where $|\mathcal{M}|$ counts the number of primitive diagnostic actions and $|\mathcal{T}|$ denotes the cardinality of the resulting transformation semigroup. A smaller value of $\eta(\mathcal{T})$ indicates a more efficient diagnostic system since fewer generators suffice to produce the same overall transformation structure.

The illustrative examples suggest that applying the minimal–word algorithm can potentially reduce redundancy in diagnostic sequences, indicating conceptual improvements in diagnostic efficiency.

### 4.8. Limitations and Scope.

The present work is theoretical in nature and does not claim empirical validation or clinical evaluation of diagnostic protocols. All applied examples and case studies are illustrative and conceptual, serving to demonstrate how semigroup-theoretic tools can be used to model, analyze, and simplify diagnostic sequences. Statements concerning efficiency, reduction of redundancy, or workflow optimization are therefore intended to indicate potential structural benefits suggested by the algebraic framework, rather than measured performance gains. Empirical validation, simulation-based studies, and implementation on real diagnostic datasets constitute important directions for future work.

#### 4.8.1. Summary of Applied Benefits

The principal applied outcomes of the framework can be summarized as follows:

- (i) reduction of redundant diagnostic steps via canonical word representation;
- (ii) hierarchical decomposition of complex diagnostic workflows into reversible and irreversible modules;
- (iii) identification of stability and convergence properties ensuring diagnostic reliability; and
- (iv) generalization to computational and distributed contexts beyond medicine.

Collectively, these results confirm that algebraic semigroup theory provides both a rigorous and a practical foundation for the design of optimized diagnostic and decision systems.

## 5. Conclusion

The present study established a systematic algebraic framework for analysing diagnostic processes through the lens of transformation semigroups. By modelling diagnostic actions as semigroup generators, it formalised the composition, redundancy, and stability inherent in diagnostic reasoning. The resulting theory unifies key components of diagnostic logic, including rule–based decision support and hierarchical clinical protocols, within a common algebraic foundation.

The principal achievements of this work include the formal definition of diagnostic semigroups as transformation semigroups generated by diagnostic actions, together with the introduction of the kernel congruence for identifying redundant diagnostic sequences and constructing minimal diagnostic words. The study further develops a minimal–word algorithm that ensures a unique and irredundant representation of diagnostic processes. In addition, a decomposition theorem adapted from the Krohn–Rhodes framework reveals the internal hierarchical structure of diagnostic semigroups, while the analysis establishes a rigorous notion of stability and reachability that links algebraic invariance to diagnostic convergence and reliability. These theoretical contributions are complemented by demonstrations of how the framework applies across clinical, computational, and networked diagnostic systems. Taken together, these results show that algebraic semigroup methods provide a rigorous mathematical scaffold for designing and evaluating diagnostic systems, bridging the gap between abstract algebra and applied decision theory.

Beyond these specific results, the study offers conceptual insights by interpreting diagnostics as a compositional process in which each diagnostic action corresponds to a morphism within a category of diagnostic transformations, and the overall diagnostic pathway becomes a composition of such morphisms. In categorical terms, the kernel congruence induces a natural equivalence relation arising from a functor that maps diagnostic sequences to their outcomes, creating a deep link between algebraic semantics and practical reasoning systems. Moreover, the correspondence between diagnostic semigroups and deterministic automata reveals an algebra–automata duality in which diagnostic reasoning can be understood either as a semigroup of transformations or as an automaton recognising equivalent outcome classes. This dual viewpoint enables the transfer of results between algebraic and computational paradigms.

Despite the generality of the theoretical framework, certain limitations remain. The minimal–word algorithm, although polynomial for finite semigroups, may become computationally expensive for very large diagnostic alphabets or continuous domains, and the algebraic model currently assumes deterministic transitions, whereas real diagnostic processes often involve uncertainty and probabilistic variation. Addressing these challenges will require the incorporation of stochastic and fuzzy extensions, as discussed earlier.

Several avenues for future research emerge naturally from the present framework. One direction concerns the development of probabilistic diagnostic semigroups that embed the transformation structure within a Markovian or stochastic operator framework capable of capturing diagnostic uncertainty. Another involves category–theoretic generalisations that formalise diagnostic processes as objects in a category of semigroup actions, enabling functorial composition across different diagnostic domains. There is also scope for algorithmic implementation, including the creation of software libraries for diagnostic semigroup computation, decomposition, and visualisation of reachability graphs. A further direction focuses on empirical evaluation through the application of the model to real diagnostic datasets—such as oncology or infectious–disease registries—in order to validate its predictive power and computational advantages. Together, these directions aim to translate the algebraic theory presented in this work into an operational tool for data–driven diagnostic analysis.

In conclusion, semigroup theory, traditionally rooted in abstract algebra, finds in this study a novel and practical domain of application in the optimisation and formalisation of diagnostic reasoning. The demonstrated correspondence between algebraic structure and decision dynamics enhances theoretical understanding while directly supporting the engineering of efficient and reliable diagnostic systems. Future research will extend these results toward adaptive and learning–enabled diagnostic semigroups that integrate algebraic structure with data–driven parameter updates, thereby providing a mathematically grounded pathway toward intelligent diagnostic systems that remain both interpretable and operationally efficient.

## Funding

The authors extend their appreciation to Prince Sattam bin Abdulaziz University, Saudi Arabia for funding this research work through the project number (PSAU/2025/01/35396).

## Acknowledgements

The authors are thankful to the learned editor and reviewers for their valuable suggestions and comments which helped in bringing this paper to its present form.

**Conflict of Interest.** The authors declare that they have no conflict of interest. No data were collected or analysed in this study, and all examples are conceptual or simulated.

**Data Availability.** No datasets were generated or analysed during the current study.

## References

[1] J. M. Howie. *Fundamentals of Semigroup Theory*, volume 12 of *London Mathematical Society Monographs*. Clarendon Press, Oxford, 1995.

[2] A. H. Clifford and G. B. Preston. *The Algebraic Theory of Semigroups, Vol. I.* American Mathematical Society, Providence, RI, 1961.

[3] J. Rhodes. *Algebraic Theory of Machines, Languages, and Semigroups.* Academic Press, New York, 1969.

[4] M. Holcombe. *Algebraic Automata Theory.* Cambridge University Press, Cambridge, 1982.

[5] S. W. Margolis and B. Steinberg. On the minimal faithful degree of rhodes semisimple semigroups. *Journal of Algebra*, 633:788–813, 2023.

[6] K. Pichór and R. Rudnicki. Transformation semigroups and their applications. arXiv preprint arXiv:2402.01572, 2024.

[7] S. Eilenberg. *Automata, Languages and Machines, Volume B*, volume 59 of *Pure and Applied Mathematics*. Academic Press, New York and London, 1976.

[8] J.-É. Pin. Mathematical foundations of automata theory, 2010. Lecture notes, Université Paris Diderot, LIAFA.

[9] A. J. Cain, A. Malheiro, and T. Paulista. Commutative nilpotent transformation semigroups. *Semigroup Forum*, 109:60–75, 2024.

[10] R. I. Garba, S. A. Akinwunmi, A. O. Adeniji, and D. O. Oyewola. On the multiplicity order of spinnable star-like transformation semigroup $t\omega_n^*$. *Journal of the Nigerian Society of Physical Sciences*, 6(1):1962, 2024.

[11] K. Krohn and J. Rhodes. Algebraic theory of machines. i. prime decomposition theorem for finite semigroups and machines. *Transactions of the American Mathematical Society*, 116(2):450–464, 1965.

[12] J. Rhodes and B. Steinberg. *The q–Theory of Finite Semigroups*. Springer Monographs in Mathematics. Springer, 2010.

[13] M. Steindl. Not all nilpotent monoids are finitely related. *Algebra Universalis*, 85:13, 2024.

[14] P. M. Higgins. Finite regular semigroups with permutations that map elements to inverses. *Semigroup Forum*, 109:141–147, 2024.

[15] E. H. Shortliffe. *Computer-Based Medical Consultations: MYCIN*. Elsevier, New York, 1976.

[16] I. Kononenko. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in Medicine*, 23:89–109, 2001.