An Enhanced Semantic Framework for Time-Constrained Clinical Decision-Making in Emergency Settings

Sivan Albagli-Kim
Department of Mathematics and Computer Science
Seton Hall University
South Orange, NJ, USA
email: sivan.albaglikim@shu.edu

Dizza Beimel
Department of Computer and Information Sciences
Dror (Imri) Aloni Center for Health Informatics
Ruppin Academic Center
Emek Hefer 4025000, Israel
email: dizzab@ruppin.ac.il

***Abstract*— Rapid and accurate decision-making is essential for identifying and treating life-threatening conditions in emergency medicine. This paper presents an enhancement to an existing Knowledge Graph-based clinical decision-making framework by integrating an emergency strategy layer to prioritize critical diagnoses. By categorizing diseases as life-threatening or non-life-threatening, our approach emphasizes the immediate exclusion of high-risk conditions. The enhancement is manifested on two primary levels: (a) we augmented the KG by incorporating conditional edges that are dynamically activated based on patient-specific indicators, such as age, gender, and pre-existing conditions. These conditional edges allow the framework to adapt to individual patient profiles, supporting a more precise and personalized diagnostic process; and (b) we refined the framework's algorithms to prioritize excluding life-threatening diseases. Future work will evaluate the framework with real-world clinical data and expand the KG's logic to include continuous data, further enhancing inference accuracy. Our contribution provides a foundation for expanding clinical decision-making frameworks to address urgent clinical needs, potentially improving patient outcomes in critical medical scenarios.**

***Keywords- knowledge graph; semantic reasoning; decision support systems; semantic technology*.**

I. INTRODUCTION

Healthcare 4.0 addresses key challenges related to the expansion, virtualization, and innovation of modern healthcare practices, such as home-based care, precision medicine, and personalized or remote drug therapies [1]. It represents the shift towards leveraging advanced technologies to overcome barriers in healthcare delivery. In particular, we focus on utilizing semantic technologies powered by large datasets and complex algorithms.

Advancements in healthcare technology have increasingly leveraged Knowledge Graphs (KG) - a graph data model that has gained popularity for representing complex knowledge structures [2] - constructed from electronic medical records to enhance clinical decision support systems. Rotmensch et al. demonstrated the potential of such an approach by learning a health KG from electronic medical records, which improved the structuring of complex patient data and facilitated more accurate inferences in diagnostic processes [3].

This aligns with our framework's utilization of a KG to support medical experts in making timely and informed decisions, especially in environments where time constraints impact diagnostic accuracy. In our ongoing research [4][5] in the medical domain, we investigate clinical decision-making processes that facilitate interactions between healthcare experts and patients. The goal is to assist medical experts in helping patients resolve health issues. These interactions typically consist of multiple iterations, where the expert asks questions, and the patient responds. With each iteration, the medical expert moves closer to making a decision concerning the patient's condition, which usually culminates in a medical diagnosis. However, time constraints often limit these interactions, which can impact the accuracy of diagnoses.

To address the goal mentioned above, we developed a framework based on semantic technologies that support the decision-making process. Each iteration suggests a question relevant to the patient's symptoms. In the final iteration, the framework produces a ranked list of hypotheses consisting of disease-symptom pairs, ordered by the likelihood that the disease is the correct diagnosis. This framework is built on a KG, which effectively models interconnected data [6], with nodes representing symptoms and diseases and edges connecting symptoms to diseases when relevant. We have developed a set of interactive algorithms that utilize both the KG and the patient's initial input to suggest relevant questions during the interaction.

The basic KG was enriched with semantic knowledge, extracted from symptom ontology, expanded the knowledge base, and added hierarchic layers. The framework was fully implemented in Python and evaluated via a set of tests [5].

While the existing framework provides a solid foundation for supporting decision-making among medical experts, its scalability allows for easy extension through KG's capacity to grow in volume and knowledge layers; for instance, it can be enhanced by adding information to edges that can yield deeper insights and more accurate hypotheses. Based on this, our current research introduces an *emergency strategy knowledge layer* to enhance decision-making processes further. This layer simulates an emergency room environment, focusing on promptly identifying emergency conditions and appropriate treatment despite time-constrained communication between patients and medical experts. Interviews with two physicians revealed key insights: (1) physicians observe additional symptoms through physical examinations and abnormal vital signs beyond those reported by patients; (2) personal patient information (e.g., age, gender, pre-existing conditions, medications) is crucial in diagnostics; (3) diseases are classified into life-threatening and non-life-threatening categories based on symptoms and patient data; and (4) the proposed strategy prioritizes the immediate exclusion of life-threatening conditions.

Following the above, the basic architecture of our framework has been enhanced to provide medical experts with a list of hypotheses related to life-threatening diseases. This enhancement includes augmenting the KG with conditional edges based on patient-specific indicators (such as age, gender, and pre-existing conditions) and refining the framework algorithms to prioritize excluding life-threatening diseases. The list of hypotheses is generated through an inference process that searches for symptoms to either rule out or confirm life-threatening conditions. By simulating an emergency room environment, this enhancement enables the framework to prioritize the rapid identification of critical conditions in time-sensitive settings.

The paper is organized as follows: Section 2 discusses knowledge representation and reviews studies that utilize KGs for healthcare applications. Section 3 details our framework, and Section 4 describes the framework enhancements developed to support the emergency strategy. Finally, Section 5 concludes with a summary of contributions and suggestions for future work.

II. BACKGROUND

In this section, we discuss how knowledge can be represented and provide an overview of researches that use KG in healthcare-related applications.

A. Knowledge Representation

Knowledge Representation (KR) serves several essential roles, such as enabling entities to predict the outcomes of actions, establishing frameworks for perceiving the world, providing foundations for intelligent reasoning, facilitating efficient computation, and acting as a medium for human expression [7]. Key methods of KR include KGs**,** ontologies**,** and semantic technologies**.**

KGs**,** also known as semantic graphs, represent information by encoding relationships between entities into graph structures. They offer semantically structured data that supports innovative solutions in tasks like question answering, recommendation systems [8], and information retrieval [9]. KGs hold significant promises for developing more intelligent machines.

Ontologies are explicit, machine-interpretable specifications of conceptualizations, defining entities within a domain, their attributes, and their interrelationships [10]. They establish a common vocabulary for humans and machines to share information, facilitating shared understanding, reuse of domain knowledge, and systematic analysis [11].

Semantic technologies aim to derive meaning from information by managing knowledge and integrating diverse data streams for inference. By representing both data and domain knowledge using graph models—since ontologies are often graph-based—graph algorithms can be employed to infer new insights.

B. Literature Review

Recent advancements in clinical decision support systems have increasingly leveraged KGs to enhance diagnostic accuracy and personalized care. For example, the construction and evaluation of causal KGs for diabetic nephropathy have demonstrated improved support in clinical decisions by modeling complex causal relationships within patient data [12]. Similarly, integrating KGs with large language models has been explored to enhance emergency decision-making, providing real-time support in critical care scenarios [13]. Furthermore, incorporating proteomics data into clinical decision-making through clinical KGs has shown promise in personalized medicine, allowing for more precise diagnostics and tailored therapies [14]. Additionally, enriching KGs from clinical narratives using natural language processing (NLP), Named Entity Recognition (NER), and biomedical ontologies has advanced healthcare applications by improving the extraction and structuring of valuable clinical information [15]. These studies underscore the significant potential of KGs in augmenting clinical decision support systems, particularly when combined with semantic technologies and patient-specific data. This aligns with our approach of integrating an emergency strategy layer into a KG-based framework to prioritize life-threatening conditions.

Recent studies have applied machine learning to clinical decision support, relying on large datasets for diagnosis prediction. While effective, these methods often require extensive labeled data and lack interpretability. Our approach, based on semantic reasoning within a KG, enables transparent and adaptive decision-making, allowing experts to incorporate new insights in real-time without retraining, enhancing explainability in time-sensitive medical settings.

III. THE FRAMEWORK

This section summarizes the framework we developed in our previous study [4][5], detailing its key algorithms and how they interact.

Recall that our objective is to support collaborative decision-making through an efficient exchange between a domain expert and an end-user, where both parties share questions and answers. In the medical context, the domain expert is a medical expert, and the end-user is a patient. The questions and answers revolve around symptoms and potential diagnoses. The framework facilitates this interaction by suggesting relevant questions for the medical expert to ask the patient (e.g., “Does the patient exhibit a particular symptom?"), with the decision-making process advancing based on the patient's responses. The framework output is a ranked list of hypotheses, where each hypothesis links a specific disease to a related symptom. As a result, the key terms in this framework are *symptoms*, *diseases*, and *hypotheses*.

The framework utilizes a KG, a widely adopted approach for representing knowledge [5]. KGs have become increasingly popular due to their ability to represent interconnected data [16][17] naturally. In this context, the KG comprises `nodes` representing `symptoms` and `diseases`, with edges connecting a symptom to a disease when the symptom indicates that condition, named `symptomOf` within the KG. Building on the KG, we formulated an inference process comprised of a set of developed interactive algorithms that leverage both the KG and the patient's initial input to generate relevant questions for the medical expert.

The framework comprises two main stages: (1) an initial pre-processing phase upon framework initialization and 2) a subsequent processing phase triggered with each new patient interaction.

A. Pre-Processing Phase

In the pre-processing phase, a KG is constructed from raw data from Kaggle [18] using Neo4j Graph Database, Version 5 [19]. The dataset consists of patient records, each corresponding to a single patient. These records include each patient's diagnosed disease and the symptoms they reported. In total, the dataset covers 41 distinct diseases and 130 unique symptoms. Some symptoms appear only once, indicating they are linked to a single disease, while others are associated with multiple diseases.

The KG was enriched by semantic knowledge extracted from an ontology of symptoms (SYMP) [20] and their interrelationships. Key elements from this ontology, particularly its hierarchical structures, were incorporated into the KG as follows: the symptoms were defined in the KG as `ontology symptoms`, and the hierarchical relationships were defined as `ISA` edges. The enriched KG, with its expanded symptom representations and hierarchical organization, offers several advantages for the inference process. These enhancements include a wider range of recommended questions for the medical expert during each interaction with the patient and symptoms that can be represented by the patient (referred to as *evidence symptoms* or ES) [21]. Figure 1, a Neo4j screenshot, demonstrates a subgraph of the enriched KG, particularly the creation of the `cough` symptom node, which is linked by a `symptomOf` edge to the `GERD` disease node. Additionally, it shows the node for its descendant (e.g., `dry cough`), connected to the parent node via an `ISA` edge. Note that the `dry cough` node has its descendant, the `dry hacking cough` node.

Finally, we applied the Louvain hierarchical clustering algorithm [22] to the KG to identify clusters of diseases—called *communities*—that share similar symptoms. We named this step as Algorithm 1 [3].

B. Processing Phase

The processing phase begins whenever a new interaction between a medical expert and a patient starts, with the patient presenting evidence of symptoms (ES). During this interaction (named Algorithm 2 [3]), the framework executes inference algorithms that utilize the identified communities to determine which diseases are compatible with the patient's symptoms. Specifically, Algorithm 2 identifies the most probable diseases that align with the evidence symptoms. Next, Algorithm 3 [3] iteratively, as needed, suggests to the medical expert questions (i.e., symptoms) that point toward the community most likely to include the patient's disease. Finally, the processing phase concludes with Algorithm 4 [3], which infers and outputs a ranked list of hypotheses—ordered pairs of a disease and an indicative symptom—that the patient might have.

The entire framework was implemented in Python, and we conducted a series of tests to evaluate its output and effectiveness [4].

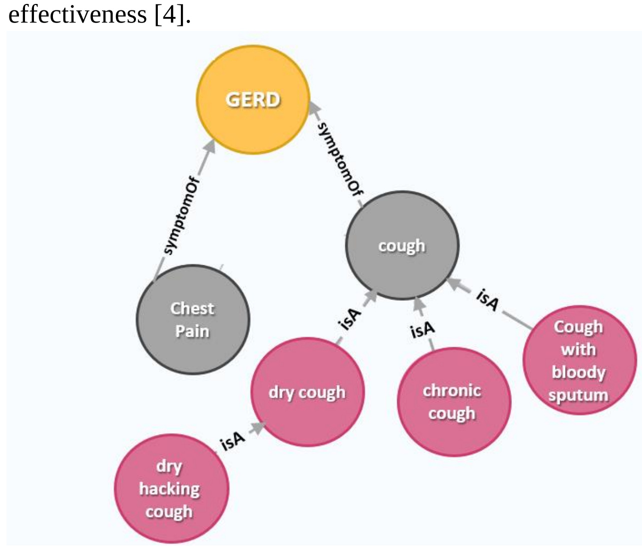

Figure 1. An example of integrating a hierarchical tree of symptoms into the KG. Disease nodes are represented in gray, symptom nodes in yellow, and ontology nodes in red.

IV. EXTENDING THE FRAMEWORK: INTEGRATING AN EMERGENCY STRATEGY

Within this section, we describe the framework enhancement we developed to support an emergency strategy along with its formal representation and its set of algorithms. In addition, we provide two simple examples that demonstrate how the enhanced framework can be utilized in emergency mode.

A. Motivation

The existing framework supports decision-making processes and is easily extendable through a scalable KG) that can incorporate additional insights. Our current research adds an emergency strategy knowledge layer to the KG, simulating an emergency room setting to prioritize identifying and treating life-threatening conditions under time constraints.

To develop this emergency strategy, we interviewed two physicians and gathered the following key insights:

1. In addition to the symptoms reported by the patient, there are other symptoms observed by the physician, which result from physical examination and abnormal vital signs (e.g., blood pressure outside the normal range).
2. Personal information about the patient (in particular, age, gender, pre-existing conditions, and medications) plays a crucial role in the diagnostic process.
3. given the symptoms and the above data, the possible diseases are classified into two categories: life-threatening and non-life-threatening.
4. The proposed strategy: first, rule out immediate life-threatening conditions.

In the following subsections, we describe the strategy and how we formulated it into a representation and a set of algorithms integrated into our framework.

B. Emergency Strategy Overview

To incorporate the aforementioned insights, we undertook two main actions: (a) enhancing the KG and (b) enhancing the processing phase to support the emergency strategy.

1) KG Enhancement

KG enhancement involves two steps, both conducted during the pre-processing phase:

A. Risk Attribute for Diseases: For all diseases in the graph, we add a Boolean attribute called *risk?*, which indicates whether a disease needs to be ruled out promptly or not.

B. Incorporating three indicators: age, gender, and pre-existing conditions into the KG. To incorporate the influence of these indicators on the presence of a life-threatening disease, we define a new type of `SymptomOf` edges: *conditional edges*. These edges are associated with an attribute formulated as a logical rule. The rule can involve one to three indicators connected with AND/OR operators.
For instance, to signal a life-threatening condition given a symptom *s1* indicating disease *d1*, if the patient is over 60 years old and male, the rule would be formulated as (*age > 60 AND gender = M*), and it assigns the conditional `SymptomOf` edge from *s1* to *d1*.

The second step (B) involves categorizing the `SymptomOf` edges in the KG into unconditional and conditional edges. Unconditional edges represent permanent relationships between symptoms and diseases that are universally applicable. In contrast, conditional edges are associated with logical rules involving patient-specific indicators. These conditional edges are incorporated into the patient's graph at runtime (processing phase) only when their associated logical rules are evaluated to be *true*. This mechanism allows the graph to adjust to individual patient profiles dynamically, enabling more precise and personalized inference during the diagnostic process. Figure 2 presents an example of a KG that was enhanced according to the described steps: It includes two diseases (*d2*, *d5*) that are characterized as high-risk, and conditional edges (e.g., the edge from *s5* to *d3*, marked with "*age<2*").

2) Processing Phase Enhancement

Enhancing the processing phase builds upon the original process by introducing new algorithms that support the emergency strategy. The evidence input process has expanded to include, beyond the symptoms reported by the patient, vital signs (such as blood pressure), and additional symptoms discovered by the medical expert during the patient's examination (e.g., a rigid abdomen). Despite the broader range of evidence entering the framework, all inputs are still characterized as *evidence symptoms*. Additionally, the patient's data, specifically the three noted indicators (age, pre-existing conditions, gender), are inputted. At this point, a new algorithm is introduced, which performs logical inference on the dependent edges. If the logical rule evaluates to true for each such edge, the associated edge is added to the patient's graph.

With the patient's graph now prepared for further analysis, identifying possible diseases and inferring potential communities proceeds with a slight modification to Algorithm 2: it now sorts the possible diseases as follows: a primary sorting of diseases with the attributed *risk? = true*, followed by a secondary sorting of all other diseases. Subsequently, the communities are ranked based on their disease scores. The disease score is decided, as before, by the degree of supporting evidence, that is, the number of evidence symptoms pointing to the disease, including the conditional edges becoming *true* (e.g., if three evidence symptoms indicate a disease, its score is 3).

Algorithms 3 and 4 are executed as described in [4][5]. Thus, for each community, we search for a symptom that can either rule out or confirm a life-threatening disease and the inference process concludes with a ranked list of hypotheses that the patient might have. Naturally, if the inference process identifies any life-threatening diseases, they will be prioritized first.

C. Formalizing the Emergency Strategy

In this section, we provide a formal description of how the strategy aligns with the KG, which includes the refined KG process and is supported by the algorithms.

1) KG and Pre-processing Formalizing

Refining the KG includes two main steps, as explained earlier. Both steps are implemented in the framework's pre-processing phase, as they do not involve the patient and remain consistent across patients.

A. Identify the diseases with high risk and add a Boolean attribute that recognizes them in the graph:
   a. Let D be the set of nodes representing the diseases in the KG. For every disease $d \in D$, add a Boolean attribute named risk? with the default value *false*.
   Let $D_{risk} \subseteq D$ Be the set of diseases with high risk. For each disease $d \in D_{risk}$, set risk? to *true*.

B. Incorporating the indicators age, gender, and pre-existing conditions into the KG: this step translates a set of rules R into conditional edges E_C in the processing step. Each rule $r \in R$ represented by a tuple $\langle s, d, f(i_{age}, i_g, i_{pre})\rangle$, where s is a symptom, d is a disease, and f is a boolean function that receives three personal indicators (i_{age}, i_g, i_{pre}) representing age, gender, and pre-existing conditions, respectively. The function f returns *true* if s indicates d according to the patient indicators. The set of conditional edges E_C are defined as follows: $E_C = \{(s, d) | f(i_{age}, i_g, i_{pre}) = true\}$. These edges will be evaluated during the processing step when a patient arrives.

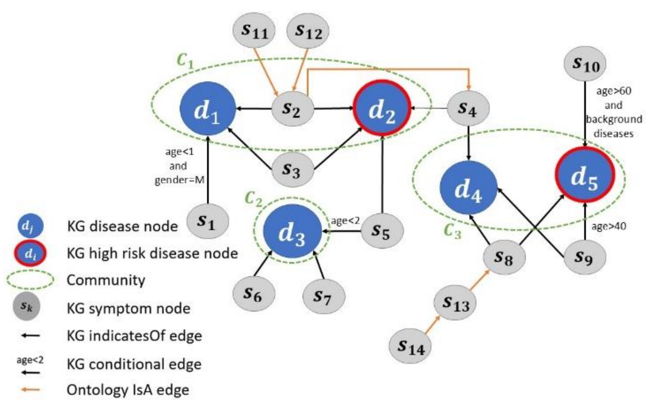

Figure 2. The Enhanced KG

2) Framework-specific Terminology

Table 1 (An extensive view is in Appendix 1) presents the terminology that we use to describe the algorithms. Additional terms supporting the emergency strategy are bold.

3) The Refined Framework Algorithms

We describe the refined algorithms developed in our framework to support the emergency strategy.

Algorithm 1 builds the personalized subgraph from KG by adding the patient's personal information (the indicators). Algorithm 2 incorporates the patient's symptoms into the personalized sub-graph and uses inference to generate a ranked list of potential diseases. This list then serves as the input for Algorithms 3 and 4 [4][5], which return a set of hypotheses prioritized by their urgency. Each hypothesis is a pair consisting of a disease and a symptom indicating it.

Algorithm 1: personalized sub-graph
Input: K$G = (D \cup S, E)$, PI, ES
Output: personalized sub-graph PKG
Algorithm:
0. Let PKG be KG
1. For every s $\in ES$:
   a. For every $r \in R$ that contain s, that is $r = \langle s, d, f \rangle$:
   b. If f(PI)=true, add the edge (s, d) to PKG.
2. Return PKG

Figure 3. Algorithm 1: personalized sub-graph

Algorithm 2: identify possible diseases
Input: PKG, ES, C
Output: possible diseases, sorted according to their risk
Algorithm:
1. Let $PD \longleftarrow \{\}$
2. Let $C' \subseteq C$ be the set of communities having positive LinD.
3. Sort C' in a non-decreasing order according to their Risk (primary), and then according to their LinD (secondary).
   3.1. Let c be the community in the order:
      3.1.1 Go over the diseases in c.
         First go over the diseases d with risk?==true. Sort them according to their R^d(d) (in a decreasing order) and add them in that order into PD.
         Then add the rest of the diseases in c, sorted (in a decreasing order) according to their R^d(d).
4. return PD

Figure 4. Algorithm 2: identify possible diseases (sorted according to risk)

D. Simplified Example

We illustrate two distinct scenarios involving different patients who exhibit the same symptoms: s1, s5, s9, and s10. However, despite sharing identical symptoms, the patients in each case have unique personal characteristics and health profiles.

The first scenario involves a 75-year-old man with no prior health conditions. The resulting graph (PKG1), after his personal indicators were entered and processed, is presented in Figure 5.

The second scenario involves a 9-month-old baby without any prior health conditions. The resulting graph (PKG2) after inputting and processing his personal indicators is shown in Figure 6.

It is important to note that these two scenarios produce different graphs, meaning the algorithms process different inputs and generate distinct hypotheses. In the first scenario, only communities *C1* and *C3* are examined, and since $Risk(C3) \geq Risk(C1)$, the first disease to be ruled out is *d5*. In the second scenario, all communities are considered, and since: $Risk(C1) \geq Risk(C2) = Risk(C3)$, the first disease to rule out is *d2*.

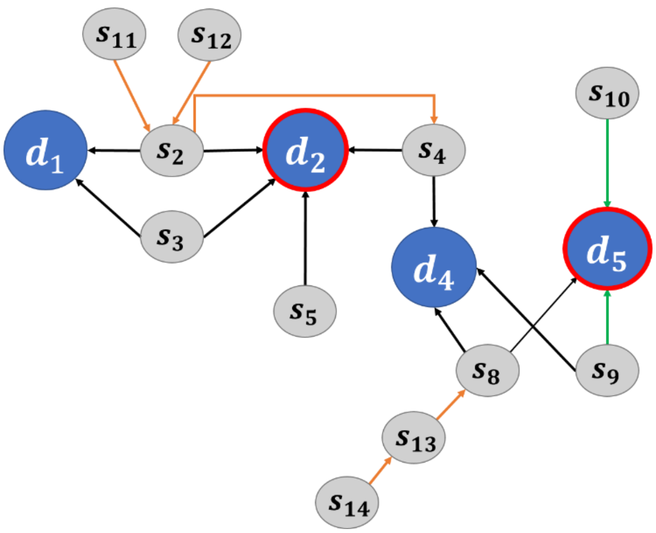

Figure 5. PKG1 – the graph for the 75-year-old man

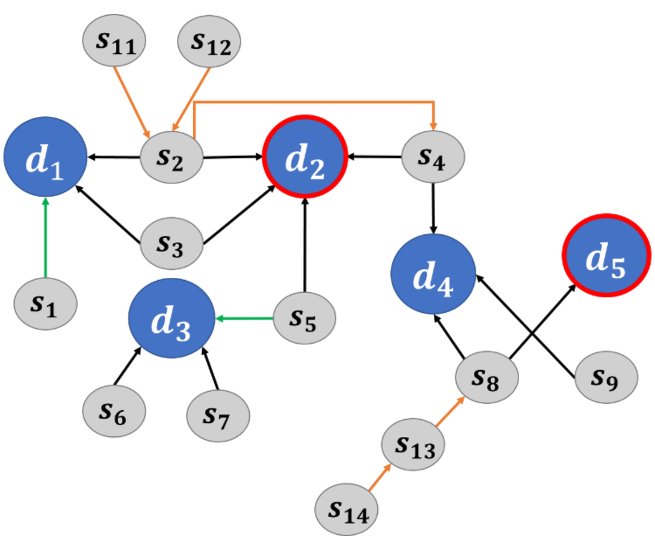

Figure 6. PKG2 – the graph for the 9-month-old baby

V. CONCLUSION AND DISCUSSION

In emergency medicine, the rapid identification and treatment of critical conditions are essential for optimizing patient outcomes. Horng et al. demonstrated the effectiveness of machine learning in developing an automated trigger for sepsis clinical decision support at emergency department triage, showcasing how advanced technologies can enhance patient care in high-stakes settings [23]. Our enhancement similarly integrates an emergency strategy layer into our KG-based decision-making framework to prioritize life-threatening conditions, thus improving decision-making efficiency and patient outcomes in urgent scenarios.

A. Summary

This study categorizes diseases based on symptoms and patient data into two main types: life-threatening and non-life-threatening. The proposed strategy focuses on the rapid exclusion of life-threatening diseases, which is crucial for optimizing patient outcomes in emergency care.

To improve the inference process for identifying life-threatening conditions, we augmented the KG by incorporating *conditional edges*. These edges, which rely on patient-specific indicators such as age, gender, and pre-existing conditions, are dynamically added to the patient's graph at runtime when specific conditions are met. This adaptive approach allows the decision support framework to tailor diagnostics to individual patient profiles, facilitating more precise and personalized recommendations.

B. Contribution

Our work advances clinical decision-making processes by formulating and integrating an emergency strategy prioritizing life-threatening conditions. We developed an enhanced KG with conditional edges informed by patient-specific data, allowing for real-time personalization. We also refined existing algorithms to incorporate this emergency strategy, enabling a diagnostic process that is more accurate and responsive to critical clinical needs. These contributions establish a more adaptable decision-making framework for emergency contexts, providing a robust foundation for further developments in emergency medical diagnostics.

C. Next Phase and Future Work

The next phase of this research will involve validating the emergency strategy using real-world clinical data to assess its effectiveness in supporting healthcare professionals in practice. Furthermore, we plan to refine the logic for conditional edges by incorporating continuous data, which will improve inference granularity and diagnostic accuracy within the KG. Expanding this work, we aim to integrate machine learning models that dynamically update the KG based on incoming data, thereby increasing the system's adaptability to evolving clinical practices and patient populations.

REFERENCES

[1] G. Aceto, V. Persico, and A. Pescapé, "Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0.," *Journal of Industrial Information Integration*, vol. 18, 100129, 2020.

[2] R. J Webber and E. Eifrem, *Graph Databases: New Opportunities for Connected Data*, O'Reilly Media, Inc.: Middlesex County, MA, USA, 2015.

[3] M. Rotmensch, "Learning a health knowledge graph from electronic medical records," *Scientific reports* vol. 7, no. 1, pp. 5994, 2017.

[4] S. Albagli-Kim and D. Beimel, "Knowledge graph-based framework for decision-making process with limited interaction," *Mathematics*, vol. 10, no. 21, pp. 3981, 2022.

[5] D. Beimel and S. Albagli-Kim, "Enhancing Medical Decision Making: A Semantic Technology-Based Framework for Efficient Diagnosis Inference," *Mathematics*, vol. 12. No. 4, pp. 502, 2024.

[6] E. Rajabi and S. Kafaie, "Knowledge graphs and explainable AI in healthcare," *Information*, vol. 13, no. 10, pp.459, 2022.

[7] R. Davis, H. Shrobe, and P. Szolovits, "What is a Knowledge Representation?," *AI Magazine*, vol. 14, no. 1, pp. 17-33, 1993

[8] Q. Guo, et al., "A survey on knowledge graph-based recommender systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3549–3568, 2020.

[9] L. Dietz, A. Kotov, and E. Meij, "Utilizing knowledge graphs for text-centric information retrieval." *The 41st International ACM SIGIR Conference on Research & development in Information Retrieval*, pp. 1387–1390, 2018.

[10] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?," *International journal of human-computer studies*, vol. 43, no. 5-6, pp. 907-928, 1995.

[11] N. F. Natalya and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," 2001.

[12] K. Lyu, et al., "Causal knowledge graph construction and evaluation for clinical decision support of diabetic nephropathy," *Journal of Biomedical Informatics*, 139: 104298, 2023.

[13] M. Chen, et al., "Enhancing emergency decision-making with knowledge graphs and large language models," *International Journal of Disaster Risk Reduction*, 113: 104804, 2024.

[14] A. Santos, et al., "Clinical knowledge graph integrates proteomics data into clinical decision-making," bioRxiv (2020), 2020.

[15] A. Thukral, S. Dhiman, R. Meher, and P. Bedi, "Knowledge graph enrichment from clinical narratives using NLP, NER, and biomedical ontologies for healthcare applications," *International Journal of Information Technology*, vol. 15, no. 1, pp. 53-65, 2023.

[16] M. Besta, et al., "Graph of thoughts: Solving elaborate problems with large language models," *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38, No. 16, pp. 17682-17690, 2024.

[17] Q. Wang et al., "Knowledge graph embedding: A survey of approaches and applications," *IEEE Transactions on Knowledge and data engineering*, vol.29, no. 12, pp. 2724-2743, 2017.

[18] Kaggle. Available online: https://www.kaggle.com [retrieved June, 2024].

[19] Neo4j. Available online: https://neo4j.com/ [retrieved: January 2025].

[20] Symptom Ontology (SYMP), Ontology Lookup Service (OLS). Available online: https://www.ebi.ac.uk/ols/ontologies/symp, [retrieved: January, 2025].

[21] S. Bonner, et al., "Understanding the performance of knowledge graph embeddings in drug discovery," *Artificial Intelligence in the Life Sciences*, vol.2: 100036, 2022.

[22] H. Lu, M. Halappanavar, and A. Kalyanaraman, "Parallel heuristics for scalable community detection," *Parallel Computing*, vol. 47, pp. 19-37, 2015.

[23] S. Horng, et al., "Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning," *PloS one*, vol.12, no. 4: e0174708., 2017.

APPENDIX 1

TABLE 1: THE EXTENDED ALGORITHMS TERMINOLOGY

Term	Definition
D	*The set of disease nodes*
$\boldsymbol{D_{risk}}$	***Let*** $\boldsymbol{D_{risk} \subseteq D}$ ***be the set of high-risk diseases.***
S	*The set of symptom nodes*
ES	*The set of evidence symptoms (i.e., the symptoms indicated by the patient)*
PI	***The patient's personalized indicators***
C	*The set of communities*
$\lvert c \rvert$	*The size of a single community* $c \in C$. *Defined by the number of diseases that belong to c*
Risk(c)	***Defined by diseases number of diseases in*** $\boldsymbol{D_{risk}}$ ***+ the number of evidence symptoms indicates a dieases in*** $\boldsymbol{D_{risk}}$
LinD(c)	*The Local-in-Degree of a given* $c \in C$. *Defined by the number of edges that point to diseases of c, by ES, hence, it is the sum of* $R^c(s,c)$*, for each* $s \in ES$ *and the given* c
PD's communities	*The set of communities* $c \in C$ *with a positive LinD(c), hence, a community in which at least one edge from* $s \in ES$ *points to c*
$R^d(d)$	*The Disease's Symptoms Rank.* *Defined by the number of symptoms the patient has that indicate D*