Integrating Declarative and Procedural Knowledge in Infectious Disease Scenario for Epidemiological Monitoring*

Evellin Cardoso[1]

[1]Federal University of Goias, Goias, Brazil

Abstract

Although promising, the synergies between declarative and procedural knowledge have been little explored in Medicine [1]. To tackle this problem, this paper proposes an approach that combines declarative and procedural knowledge. The approach uses ontologies as a knowledge artifact that expresses medical declarative knowledge. This ontology is used as a starting point for an approach that derives a BPMN procedural specification from the knowledge expressed in the ontology. The approach is illustrated in a infectious disease scenario, in particular, it uses the Basic Formal Ontology (BFO) [2] and the Infectious Disease Ontology (IDO) [3].

Keywords

Medical Knowledge, Knowledge Representation, Ontologies, Basic Formal Ontology (BFO), Business Process Management (BPM)

1. Introduction

Expert systems have been used in Medicine successfully since 1990s [4, 5]. An important aspect of the design of such systems regards the acquisition and representation of medical knowledge, a problem that has been addressed by different communities in Computer Science. In medical informatics, a lot of research has been conducted in the development of formalisms to capture medical knowledge. Since late 1990s, many Knowledge Representation (KR) formalisms have been developed, including ontologies, semantic web related formalisms and logics [6]. The focus of medical informatics community is on the representation of (medical) *declarative knowledge* that capture general, background medical knowledge, such as diseases, drugs and treatments [1].

Another strain of research in medical informatics considers the formalization of narrative clinical guidelines into computer interpretable clinical guidelines (CIG) using formalisms such as document models, decision trees, probabilistic models, task-network models [7]. Many domain-specific languages to model CIGs have been proposed, such as Asbru, PROforma, GLIF, EON and GUIDE [8]. More recently, boostered by business demands, Business Process Management (BPM) techniques have been applied to healthcare organizations, in an attempt to streamline

Proceedings of the International Conference on Biomedical Ontologies 2023, August 28th-September 1st, 2023, Brasilia, Brazil

*Corresponding author.
evellin@ufg.br,evellinc@gmail.com (E. Cardoso)
0000-0001-6242-662X (E. Cardoso)

medical operations, thus achieving business goals of improving efficiency and reducing costs. Medical processes have been captured as business processes using process languages such as BPMN and DECLARE. Business processes and CIGs capture (medical) *procedural knowledge* that consists of sequences of actions that must be followed by healthcare providers in certain circumstances [1].

Although promising, the synergies between declarative and procedural knowledge have been little explored in both medical informatics and BPM communities [1]. To tackle this problem, this paper proposes an approach that combines declarative and procedural knowledge. The approach uses ontologies as a knowledge artifact that expresses medical declarative knowledge. This ontology is used as a starting point for an approach that derives a BPMN procedural specification from the knowledge expressed in the ontology. The approach is illustrated in a infectious disease scenario, in particular, it uses the Basic Formal Ontology (BFO) [2] and the Infectious Disease Ontology (IDO) [3].

The rest of the paper is structured as follows: Section 2 briefly introduces the BFO and IDO ontologies and procedural models in the medical sector, briefly introducing BPMN syntax. Section 3 presents the approach for deriving a BPMN procedural model starting from the ontologies, illustrating it in the scenario of infectious diseases. Section 5 concludes the paper and outlines future work.

2. Preliminaries

2.1. Ontologies

Ontologies are a knowledge representation formalism that represent (or strive to represent) reality in such way a group of stakeholders understand the terms that compose a certain domain of discourse, and can thus learn about such domain [2]. They may be classified as *top-level ontologies* (or formal) that contain highly general categories and relations of reality common to all domains, defining concepts such as "process", "material object", etc., or *domain ontologies* that capture a basic set of universal concepts pertinent to a specific scientific domain (e.g., geography, medicine or law).

In this paper, (medical) declarative knowledge is represented as ontologies. The *Basic Foundational Ontology (BFO)* [2] is the top-level ontology chose due to its wide accceptance as an international standard ISO/IEC 21838–2, while the Infectious Disease Ontology (IDO) [3] is the chosen domain ontology. IDO extends the Ontology for General Medical Science (OGMS), which in its turn extends BFO.

2.2. Procedural Models in Healthcare

The BPM discipline is concerned with the formalization and analysis of the activities conducted by an enterprise to produce goods and services to its customers [9]. In healthare organizations, medical processes can be divided as *administrative processes* (concerned with administrative practices like handling of medical order and lab procedures) or *knowlege-intensive processes* which are concerned with the intensive usage of domain specific knowledge in diagnose and treament processes.

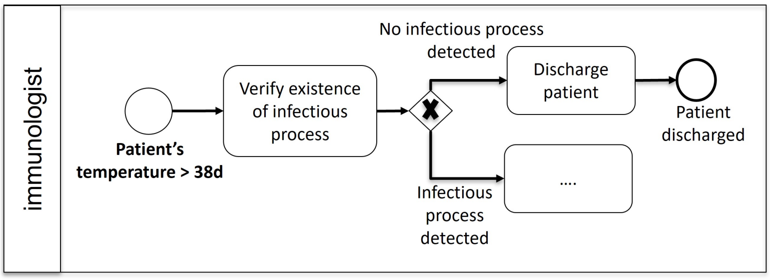

Figure 1: A BPMN specification of knowledge-intensive process in healthcare

Both types of processes can be typically captured by a procedural process language, such as the Business Process Modeling Notation (BPMN). Fig. 1 depicts a knowledge-intensive business process in healthcare. Rounded rectangles denote activities (e.g., *verify the existence of infectious process*), while circles represent events (*patient's temperature > 38 degrees*). Diamonds represent decisions (*presence or absence of infectious process?*), while arrows denote diferent results of decisions.

3. The Approach

This section depicts the approach of building process models based on the knowledge extracted from ontologies. As ontologies are commonly structured into *foundational ontologies* $\mathcal{F}$ and *domain ontology* $\mathcal{D}$, our approach starts from a set of *ontologies* $\mathcal{O} = \langle \mathcal{F}, \mathcal{D} \rangle$, where $\mathcal{F}$ is the set of foundational ontologies and $\mathcal{D}$ the set of domain ontologies. The approach is composed by the following steps:

Step 1. The knowledge engineer starts by investigating the concepts of $\mathcal{O}$ for a preliminary understanding of the domain.

Step 2. As the size of ontologies may be significant, the knowledge engineer selects the subset $\mathcal{D}'$ from $\mathcal{D}$ that is relevant for her modeling purposes. $\mathcal{D}'$ may be sub-ontologies or even portions of $\mathcal{D}$.

Step 3. The knowledge engineer explores the semantics of each particular concept to grasp about the domain.

Step 4. The knowledge engineer builds the procedural representation. To perform such step, s/he considers the following sub-steps (not necessarily in this order):

- One must take into account that activities are performed to react to events happening in the world. The hint is to look for the relevant events $\{e_1, e_2, ..., e_n\}$ taking place within the domain
- To express such events in the procedural specification, consider that the occurrence of events is captured by changes in the states of entities of the domain (concepts of the ontology). By searching for the concepts in the ontology and which changes in the state of these concepts are relevant, one can express the relevant events.

- Order the relevant events. Include the events and the activity that has to be performed to address the event (change), including them all the procedural specification.

Step 5. As procedural and declarative knowledge present a complementary view of reality, it is possible that all knowledge required to build the procedural specification is not found in the declarative specification (and vice-versa). In this case, the knowledge engineer will have to complement the procedural specification with the help of domain experts.

Next section illustrates our approach in a infectious disease scenario.

4. Applying the Approach in an Infection Disease Scenario

This section illustrates the approach from Sec. 3 applied in a infectious disease scenario [3]. The infectious disease scenario represents the domain of infectious diseases. It includes different biological scales (gene, cell, organ, organism and population), complementary disciplinary perspectives (biological, clinical, epidemiological), and successive phases of an infectious process (host, reservoir, vector, pathogen). With the recent breakthrough of Covid-19, the availabillity of such ontology is important to integrate heteregenous data sources. Such integration will enable the establishment public policies to contain the disease by public health organizations using statistical data.

Step 1. Starting the approach, the set of ontogies $\mathcal{O}$ related to the IDO have been investigated for a basic understanding of the infectious disease domain. Three ontologies (BFO, OGMS and IDO) have been identified. In this way we have $\mathcal{O} = \langle \mathcal{BFO}, \mathcal{D} \rangle$, where $\mathcal{D} = \langle \mathcal{IDO}, \mathcal{OGMS} \rangle$.

Step 2. Multiple possibilities for selecting parts $\mathcal{D}'$ of $\mathcal{D}$ exist, given that IDO captures many distinct dimensions (biological scales, disciplinary perspectives and different phases). The disciplinary perspective of *epidemiology* has been chosen.

Step 3. The semantics of each particular concept relative to epidemiology has been investigated at Table 5 and section "Epidemiology and surveillance" from [3].

Step 4. Fig. 2 depicts the BPMN specification built in our approach (the three traces inside the first activity denotes that the activity is performed multiple times, in this case, for each geographic region).

This specification has been built by analyzing the semantics of three IDO concepts (*infectious disease incidence*, *infectious disease pandemic* and *infectious disease epidemic*) that are highlighted in boldface in BPMN. On top of identifying the concepts, we have identified the relevant events.

After understanding the semantics of the three concepts, following with the identification of the events, if the incidence of an infectious disease in a population (in a certain geographic region) is above a certain threshold1 (*infection disease incidence* > *threshold1*), this may indicate the existence of an *epidemics* in that region. With that, the public administrator has to check the existence of epidemics in other regions as well. If other regions have not surpass the acceptable threshold1 of infectious disease, then the number of regions with infectious disease is below threshold2 (*#regions* < *threshold2*), no signal of epidemics is found and the surveillance process finishes.

On the contrary, if a number of regions that population has infectious disease surpass a certain threshold (*#regions* > *threshold2*), this indicates a *pandemics.* In this case, the public administrator proceeds with the monitoring and two situations may happen. One is when

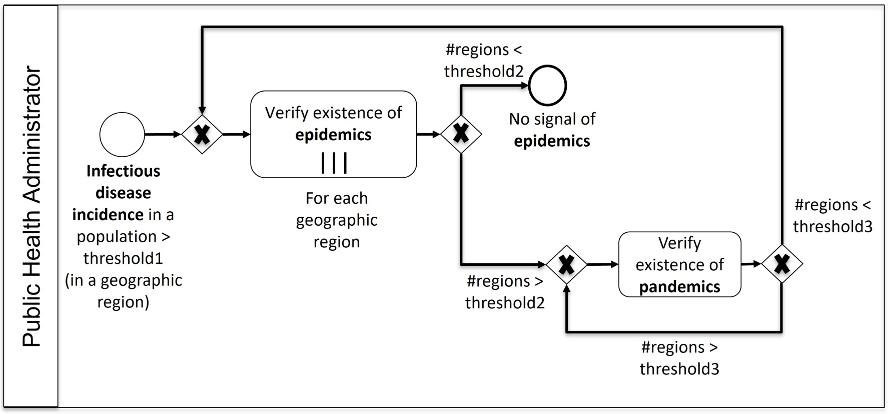

Figure 2: Procedural Specification (BPMN) built from an infectious disease ontology [3]

we continuosly have a pandemics (*#regions > threshold3*). The second one happens when the number of regions drop and we no longer have a pandemics, but we may still have an epidemics (*#regions < threshold3*).
Step 5. Notice that the ontology captures only knowledge about the domain, but not the real-time data relative to the incidence of infectious diseases in population or the thresholds for epidemics and pandemics. In this case, public admnistrators have to provide such numbers.

5. Conclusion

This paper proposed a preliminary approach to derive a BPMN procedural specification from the knowledge expressed in an ontology. The approach is illustrated in an infectious disease scenario, using BFO (top-level ontology) and IDO (domain specific ontology). As the integration between declarative and procedural knowledge is generally challenging in many domains (not only Medicine), I hope that this preliminary approach provides an initial insight on how to solve this problem.

In particular, an advantage of this approach in the infectious disease scenario is the possibility of integrating data to the overall approach presented in this paper. Two types of data may be integrated, together with two distinct monitoring perspectives. One perspective considers the usage of the values of incidence of infectious diseases in population and the thresholds for epidemics and pandemics. With these values, it is possible to use epidemic simulators in public health monitoring [10]. The second monitoring approach considers the usage of process mining [9] to monitor the BPMN procedural process, and thus, to monitor in which stage of the process we are (if we have a epidemics, pandemics or no abnormal public health event).

As future work, I consider this integration with data, together with the usage of more concepts from the ontology to derive more information for the procedural specification. Further, I also want to explore the synergy between declarative and procedural knowledge in the other way round, such as, can we derive the ontology from a procedural specification?

References

[1] S. Bragaglia, F. Chesani, P. Mello, M. Montali, Conformance Verification of Clinical Guidelines in Presence of Computerized and Human-Enhanced Processes, Springer International Publishing, 2015, pp. 81–106.

[2] R. Arp, B. Smith, A. D. Spear, Building Ontologies with Basic Formal Ontology, The MIT Press, 2015.

[3] S. Babcock, J. Beverley, L. Cowell, B. Smith, The infectious disease ontology in the age of covid-19, 2021. doi:10.31219/osf.io/az6u5.

[4] P. Windyga, D. Almeida, G. Passariello, F. Mora, J. Coatrieux, Knowledge-based approach to the management of serious arrhythmia in the ccu, Medical and Biological Engineering and Computing 29 (1991) 254–60. doi:10.1007/BF02446707.

[5] A. Batarekh, A. D. Preece, A. Bennett, P. Grogono, Specifying an expert system, Expert Systems with Applications 2 (1991) 285–303. doi:https://doi.org/10.1016/0957-4174(91)90036-E.

[6] D. Riaño, M. Peleg, A. ten Teije, Ten Years of Knowledge Representation for Health Care (2009–2018): Topics, Trends, and Challenges, Artificial Intelligence in Medicine 100 (2019).

[7] M. Peleg, Computer-Interpretable Clinical Guidelines: A Methodological Review, Journal of Biomedical Informatics 46 (2013) 744–763.

[8] D. Isern, A. Moreno, Computer-Based Execution of Clinical Guidelines: A Review, International Journal of Medical Informatics 77 (2008) 787–808.

[9] W. M. P. van der Aalst, Business Process Management: A Comprehensive Survey, ISRN Software Engineering 2013 (2013).

[10] W. Hogan, M. Wagner, M. Brochhausen, J. Levander, S. Brown, N. Millett, J. DePasse, J. Hanna, The apollo structured vocabulary: an owl2 ontology of phenomena in infectious disease epidemiology and population biology for use in epidemic simulation, Journal of Biomedical Semantics 7 (2016) 50. doi:10.1186/s13326-016-0092-y.