

Online Plausibility Checks for Patient Pathways

Medical ontologies in online process mining.

Emmanuel Helm

Department of Medical and Bioinformatics
University of Applied Sciences Upper Austria
4232, Hagenberg, Austria
emmanuel.helm@fh-hagenberg.at

Georg Buchgeher · Lisa Ehrlinger

Software Competence Center Hagenberg GmbH
4232, Hagenberg, Austria
georg.buchgeher@scch.at
lisa.ehrlinger@scch.at



Problem Statement

A patient pathway describes all events recorded during the treatment of one specific patient. In the field of PODS4H, data quality issues (e.g., completeness, timeliness, or plausibility) remain a major challenge [1]. In an online setting (cf. Online Process Mining, as introduced in [2]), the *completeness* of the patient pathway is unknown. Hence, other quality checks, e.g., the verification of guideline compliance, become more arduous.

Objective

In this poster, we discuss how to measure the plausibility of an online patient pathway, i.e., an ongoing patient encounter. We want to validate if the currently observed pathway remains plausible whenever new events are recorded.

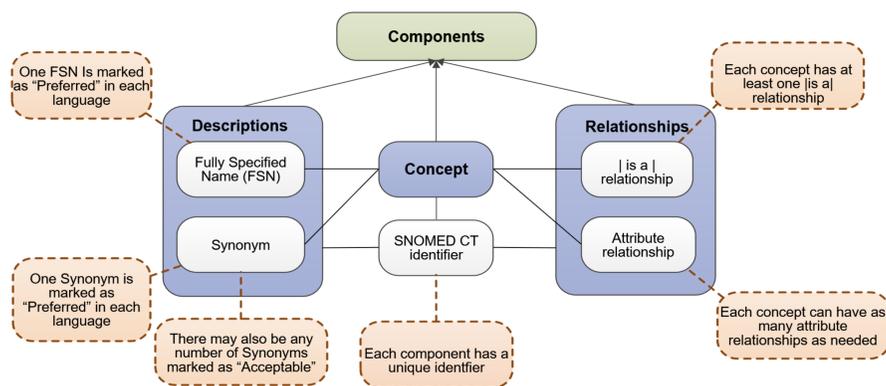


Figure 1. SNOMED CT logical model, from [4].

Idea

We utilize the Shapes Constraint Language (SHACL [6]) to continuously validate patient pathways. The pathway, that is, the *online* process instance, is modelled as continuously growing Resource Description Framework (RDF [7]) data graph. Di Ciccio et al. [5] already provide a XES RFD vocabulary to enable the translation (<http://semantics.id/ns/xes>). This data graph is then validated against a shapes graph.

Di Ciccio et al. [5] discovered declarative constraints in existing event logs and translated them into SHACL constraints. We want to extend this approach and add constraints based on a publicly available Web Ontology Language (OWL [8])-based medical ontology. We plan to use parts of SNOMED-CT as shapes graphs. This approach is based on previous work in the industrial domain [3].

“SNOMED CT is a core clinical healthcare terminology that contains concepts with unique meanings and formal logic based definitions organized into hierarchies. SNOMED CT content is represented using three types of components:

- Concepts representing clinical meanings that are organized into hierarchies.
- Descriptions which link appropriate human readable terms to concepts.
- Relationships which link each concept to other related concepts.”

– SNOMED CT Starter Guide [4]

W3C defines the Shapes Constraint Language as:

“A language for validating RDF graphs against a set of conditions. These conditions are provided as shapes and other constructs expressed in the form of an RDF graph. RDF graphs that are used in this manner are called ‘shapes graphs’ in SHACL and the RDF graphs that are validated against a shapes graph are called ‘data graphs’. As SHACL shape graphs are used to validate that data graphs satisfy a set of conditions they can also be viewed as a description of the data graphs that do satisfy these conditions.”

– <https://www.w3.org/TR/shacl>

Discussion

We discuss this vision and expected challenges with the PODS4H community, specifically:

1. How to model the data graphs, i.e., the online process instances?
2. How to construct the shapes graphs?
3. How can SNOMED CT support the construction of the shapes graph? Is the expressiveness of SNOMED CT sufficient for this use case?
4. Is it feasible to use SNOMED CT as shapes graph and the XES event log as data graph or vice versa? Open World Assumption vs. Closed World Assumption?
5. Should we use SNOMED CT Expression Constraint Language (ECL) instead of SHACL? Could we utilize an approach similar to Di Ciccio et al. [5].

References

- [1] Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O.A., Sepúlveda, M., Helm, E., et al., Process Mining for Healthcare: Characteristics and Challenges, Journal of Biomedical Informatics, Volume 127, 2022.
- [2] van der Aalst, W. Process Mining: Data Science in Action. Vol 2. Springer, 2016.
- [3] Buchgeher, G., Ehrlinger, L., et al.: Representing Technical Standards as Knowledge Graph to Guide the Design of Industrial Systems (Extended Abstract). 18th International Conference on Computer Aided Systems Theory, 2022.
- [4] SNOMED CT, Starter Guide. IHTSDO, June 2022.
- [5] Di Ciccio, C., Ekaputra, F.J., Cecconi, A., Ekelhart, A. and Kiesling, E., Finding Non-compliances with Declarative Process Constraints through Semantic Technologies, CAISE Forum 2019, pp. 60-74. Springer, Cham.
- [6] W3C, Shapes Constraint Language (SHACL), 2017. Online: <https://www.w3.org/TR/shacl>
- [7] W3C, RDF Primer, 2004. Online: <https://www.w3.org/TR/rdf-primer>
- [8] W3C, OWL 2 Web Ontology Language Document Overview (Second Edition), 2012. Online: <https://www.w3.org/TR/owl2-overview>

Acknowledgements

The research reported in this paper has been funded by BMK, BMDW, and the State of Upper Austria in the frame of the COMET Programme managed by FFG.

