The Droid You're Looking For: C-4PM, a Conversational Agent for Declarative Process Mining*

Yago Fontenla-Seco[1,*], Sarah Winkler[2], Alessandro Gianola[2], Marco Montali[2], Manuel Lama[1] and Alberto Bugarín-Diz[1]

[1]Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS), Universidade de Santiago de Compostela, Spain.

[2]Free University of Bozen-Bolzano, Bolzano, Italy

Abstract

The effective presentation of process models to non-expert users in a way that allows them to understand and query these models is a well-known research challenge. Conversational interfaces, with their low expertise requirements, offer a potential solution. While procedural models like Petri nets are not ideal for linguistic presentation, declarative models, specifically the Declare specification, provide a more suitable option. This paper introduces C-4PM, the first conversational agent for declarative process specifications. C-4PM facilitates tasks such as consistency, conformance, and model checking through a conversation-driven modular pipeline. The feasibility of the tool was assessed through a preliminary evaluation on a healthcare process.

Keywords

Process Mining, Conversational Agent, Declare, NLP, NLG

1. Introduction

Process model understandability and interpretability is widely recognized as an open challenge by the Process Mining community [1, 2, 3, 4]. Highly technical knowledge is needed to understand process models, particularly in complex environments like healthcare [3, 5, 6]. Process models are most typically presented as visualizations using imperative or procedural representations, such as Petri nets or directly-follows graphs (DFGs). DFGs are the most commonly used due to their simplicity and ability to highlight process issues. However, experts caution that DFGs can be prone to misinterpretation, especially when viewed with filtered data [7]. In contrast to imperative process modeling, declarative process specifications, such as those obtained with the Declare language [8], seek to balance flexibility and control by characterizing a process with a set of constraints on the allowed and forbidden action patterns. While this results in compact specifications, the

The BPM 2023 Demos & Resources Forum. 21st International Conference on Business Process Management, Utrecht, Netherlands, September 10-15, 2023

*This research was funded by the Spanish Ministry for Science, Innovation and Universities (grants PRE2018-083719, TIN2017-84796-C2-1-R, PID2020-112623GB-I00, PDC2021-121072-C21) and the Galician Ministry of Education, University and Professional Training (grants ED431C2018/29, ED431G2019/04). All grants were co-funded by the European Regional Development Fund (ERDF/FEDER) program.

*Corresponding author.

hidden dependencies [9] emerging from the interplay of constraints make it very difficult to understand which traces are accepted or rejected by a declarative specification [10].

Making (declarative) process mining accessible to non-expert users calls for developing more effective methods that enhance process model understanding, mitigate misinterpretations, and promote accurate insights. This is especially relevant in complex environments like healthcare, where users possess the domain expertise required to pose relevant queries over their processes, but lack expertise in process mining. For this purpose, we present C-4PM,[1] the first conversational agent for declarative process mining. Based on an event log (and optionally a Declare specification), C-4PM has the ability to discover the declarative specification of a process and perform multiple reasoning tasks, such as consistency, conformance and model checking. Interacting with the user via a text interface, C-4PM takes as input a question expressed in natural language, identifies the user intention and, exploiting a Large Language Model (LLM), translates the relevant parts of the input into an intermediate abstract logical representation that can be appropriately used for the reasoning task identified. Finally, C-4PM expresses the result back in natural language via the LLM. The motivation to use declarative process specifications comes not only from the capacity of expressing Declare patterns directly in natural language, but also because these patterns cover the majority of relationships between activities that are important in practice. Declare supports the specification of constraints on the temporal evolution of a process, without explicitly indicating how process instances should be routed and thus being more flexible than imperative models [8]. Furthermore, as it is grounded in Linear Temporal Logic on Finite Traces (LTLf), this allows to perform reasoning tasks based on this logic. Indeed, C-4PM delegates all reasoning tasks to dedicated tools for Declare and LTLf, using the LLM only for linguistic expression.

Related work. Although the potential of Natural Language Processing (NLP) and Natural Language Generation (NLG) for Business Process Management (BPM) was recognized a decade ago [11, 12], the practical implementation of process mining tools with natural language interfaces has been limited. NLG has been used to explain event logs and processes to users [13, 14], and to verbalize BPMN models [12], but without explicitly supporting reasoning and querying. NLP techniques have also been used to extract BPMN models from textual descriptions of processes [15]. More recently, a conversational interface for process mining was proposed; however, restricted to existing process mining tools (based on imperative models) and lacking reasoning support [1].

2. C-4PM: Overview and Features

C-4PM is based on the Rasa Framework [16] and is divided in two main components: the Rasa Stack (or Rasa Open Source) and the Action Server (based on the Rasa SDK). In the Rasa Stack, the Natural Language Understanding (NLU) module handles user intent classification and entity extraction from the user input; and the Rasa Core module is in charge of the dialogue management i.e.: deciding the next action in the conversation based

[1]Repository with source code, documentation and demonstration: https://github.com/Yagouus/c-4pm

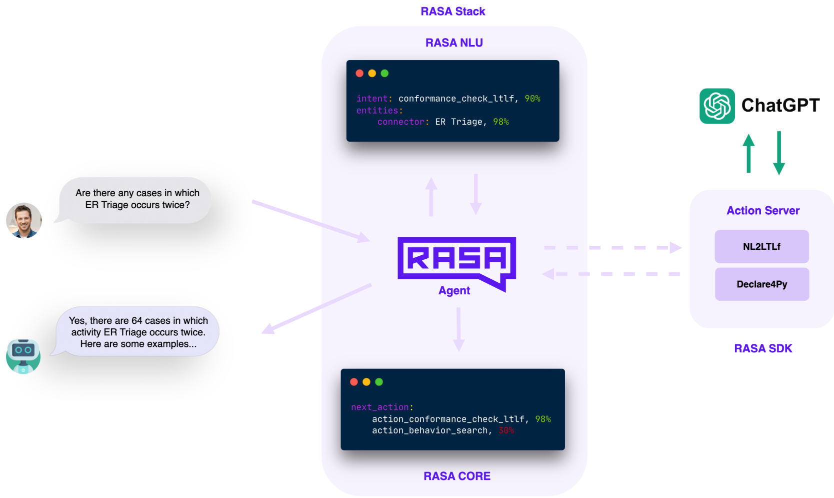

Figure 1: Architecture of C-4PM. Shadowed in purple are the two main components: the Rasa Stack, in charge of the Natural Language Understanding and dialogue management; and the action server, where the reasoning tasks and Natural Language Generation of appropriate responses are carried out.

on the context and user input. Once an action is decided by the dialogue manager, the Action Server runs the appropriate reasoning tasks and generates the response that will be delivered to the user. All components work together to enable effective communication and interaction in the conversational system.

Reasoning is supported by the NL2LTL [17] and Declare4Py [18] frameworks. NL2LTL is in charge of translating the natural language input of the user to an appropriate LTLf formula that can be used by Declare4Py for multiple reasoning tasks. When the NLU module identifies that the user intent is to perform a reasoning task in which some LTLf formula has to be extracted from the input, the NL2LTL module is called. NL2LTL can operate with two different engines, namely a RASA implementation and a GPT one, in our particular case the GPT engine is used as we have extended the LTLf formulas originally supported by NL2LTL and we found that adapting the prompt used with the GPT engine yields more accurate results that those obtained by modifying the original RASA engine. Declare4Py is a Python library based on the Declare language that is designed to analyze processes using declarative, constraint-based models (process specifications). Backed by Declare4Py, C-4PM has the ability to automatically discover the declarative specification of a process based on event log data using the approach presented in [19] and perform all reasoning tasks supported by Declare4Py, namely: conformance, consistency, model, and query checking[2].

Running example To better illustrate how C-4PM operates we will use an example based on the Sepsis event log [20]. This log contains 1050 traces describing the treatment process of sepsis patients in a Dutch hospital. For instance, users can ask:[3]

[2]Even though supported by Declare4Py, multi-perspective Declare in C-4PM is proposed as future work.
[3]All tasks with exemplary queries can be found in the project's repository: https://github.com/Yagouus/c-4pm

- "Explain the model of the process", for the system to explain the discovered Declare specification in natural language to grasp a general idea of it (specification description).
- If the discovered (or given as input) Declare specification allows for any trace at all (consistency checking) as a way of checking if the specification is correct and aligns with the event log e.g.: "Does the model allow for any behavior?".
- Examples of traces that conform (or not) to the specification to get a general idea of the frequent behavior in the process can be retrieved by asking "Can you give me traces from the log that conform to the model?", i.e., for conformance checking.
- To know whether some behavior is supported by the model, e.g., whether it is possible that patients experience first admission to normal care, and later on to intensive care, one may ask "Is it possible that activity Admission IC occurs after activity Admission NC has taken place?". This constitutes a form of model checking.
- Knowing that cases in which the above behavior happens are considered problematic [20, p.299], a user might ask "Are there cases where Admission IC happens after Admission NC?".

In the latter two cases, C-4PM translates the user input to LTLf and to the corresponding Declare constraint $\mathit{response}(\mathit{AdmissionNC}, \mathit{AdmissionIC})$, and checks in the former case whether it is consistent with the model, and in the latter uses it for conformance checking.

3. Evaluation and Future Work

As C-4PM is a proof-of-concept implementation, extensive experiments are left for future work. However, all the described functionality can already be used and tested. A preliminary user validation has been conducted with process mining experts to evaluate its performance. The performed tests have yielded positive results, indicating the potential effectiveness of the tool in addressing the challenges of presenting process models and facilitating the interaction of non-expert users with a declarative process specification; effectively allowing them to perform multiple reasoning tasks in a transparent and understandable manner. While further investigation and evaluation are necessary to fully assess the maturity and reliability of the tool, the initial feedback provides promising indications of its viability and potential.

Future work Future work is focused on enhancing the user experience by exploring new NLU and NLG techniques, potentially migrating towards a more powerful Large Language Model (LLM) NLU based systems that can provide a deeper and more context-aware conversation. Furthermore, expanding the scope of reasoning tasks that the system can handle, such as complex decision-making and logical inference, would contribute to its overall effectiveness. Finally, future research efforts should focus on conducting more comprehensive evaluations to establish the tool's maturity and reliability in a wider range of contexts. Conducting a comprehensive user validation study involving a diverse range of users and scenarios would provide even more valuable insights into the system's usability and performance, guiding further improvements and refinements.

References

[1] L. Barbieri, E. Madeira, K. Stroeh, W. van der Aalst, Towards a natural language conversational interface for process mining, in: PM Workshops, 2022, pp. 268–280.
[2] F. Mannhardt, Responsible Process Mining, Springer, 2022, pp. 373–401.
[3] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: A literature review, Journal of Biomedical Informatics 61 (2016) 224–236.
[4] I. Beerepoot, et al., The biggest business process management problems to solve before we die, Computers in Industry 146 (2023) 103837.
[5] R. Mans, W. van der Aalst, R. Vanwersch, Process mining in healthcare: evaluating and exploiting operational healthcare processes, SpringerBriefs in BPM, 2015.
[6] E. Batista, A. Solanas, Process mining in healthcare: A systematic review, in: Proc. 9th IISA, 2018, pp. 1–6.
[7] F. Zerbato, P. Soffer, B. Weber, Process mining practices: Evidence from interviews, in: Business Process Management, Springer, 2022, pp. 268–285.
[8] C. Di Ciccio, M. Montali, Declarative Process Specifications: Reasoning, Discovery, Monitoring, Springer, 2022, pp. 108–152.
[9] M. Montali, et al., Declarative specification and verification of service choreographies, ACM Trans. Web 4 (2010) 3:1–3:62.
[10] C. Haisjackl, et al., Understanding declare models: strategies, pitfalls, empirical results, Softw. Syst. Model. 15 (2016) 325–352.
[11] H. van der Aa, et al., Challenges and opportunities of applying natural language processing in business process management, in: 27th COLING, 2018, pp. 2791–2801.
[12] H. Leopold, Natural Language in Business Process Models, 2013.
[13] Y. Fontenla-Seco, et al., A framework for the automatic description of healthcare processes in natural language: Application in an aortic stenosis integrated care process, Journal of Biomedical Informatics 128 (2022) 104033.
[14] Y. Fontenla-Seco, M. Lama, A. Bugarín, Process-to-text: A framework for the quantitative description of processes in natural language, in: Trustworthy AI - Integrating Learning, Optimization and Reasoning, 2021, pp. 212–219.
[15] H. van der Aa, H. Leopold, H. A. Reijers, Detecting inconsistencies between process models and textual descriptions, in: Proc. 13th BPM, 2015, pp. 90–105.
[16] T. Bocklisch, J. Faulkner, N. Pawlowski, A. Nichol, Rasa: Open source language understanding and dialogue management, 2017.
[17] F. Fuggitti, T. Chakraborti, NL2LTL – a python package for converting natural language (NL) instructions to linear temporal logic (LTL) formulas, in: AAAI, 2023, pp. 16428–16430.
[18] I. Donadello, et al., Declare4py: A python library for declarative process mining, in: BPM (PhD/Demos), CEUR Workshop Proceedings, 2022, pp. 117–121.
[19] F. M. Maggi, et al., Efficient discovery of understandable declarative process models from event logs, in: CAiSE 2012, 2012, pp. 270–285.
[20] F. Mannhardt, Multi-perspective Process Mining, Ph.D. thesis, Technical University of Eindhoven, 2018.