Business Process Model Information Extraction from Documents Using In-Context Learning

Patrizio Bellan

Fondazione Bruno Kessler,
Free University of Bozen-Bolzano

Supervisors:

Mauro Dragoni

Chiara Ghidini

The Problem

Process model Extraction from Text *can be regarded as the specific problem of finding algorithmic functions that transform textual descriptions of processes into structured representations of different expressivity, up to the entire formal process model diagram.*

The goal of textual descriptions of business processes is that of being easy to understand and use, but the actual exploitation of the information they contain is often hampered by having to manually analyze unstructured information. But, the extraction of process model information from textual documents is a research area that still lacks the ability to scale to a variety of real-world texts.

The Challenges

- NO Data Available
- Challenging Entities

Factual event

The concert of Pink Floyd was in Venice

↑

Extraction of instances

VS

Concept event

The salesman arrange a road show to talk to potential investors

↑

Extraction of Concepts (e.g. Activity) and Conceptual relations

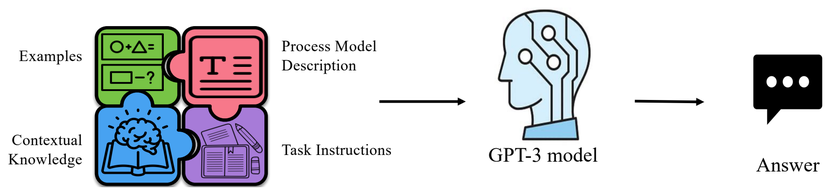
The Solution

Pre-trained language models have shown promising results concerning the knowledge extraction tasks from the models themselves. However, using prompt-based learning strategy to extract process information from such language models has not yet been investigated. In my research, I am investigating a prompt-based in-context learning strategy to extract, from unstructured process descriptions, conceptual information that can be converted into their equivalent process model diagram. Such a strategy is performed in a multi-turn dialog fashion, by posing questions to a language model. Each question targets the extraction of a specific process model information.

I use *in-context learning and the GPT-3 model to solve the data and conceptual entities challenges.*

In-context learning

In-context learning is a prompt-based technique. A Language Model "learns" to do a task simply by conditioning on a prompt (the input) consisting of input-output examples.



Experimental Settings

NO In-context learning

RAW

+ Process description
+ Task Instructions

DEFS

+ Process description
+ Task Instructions
+ Contextual Knowledge

In-context learning

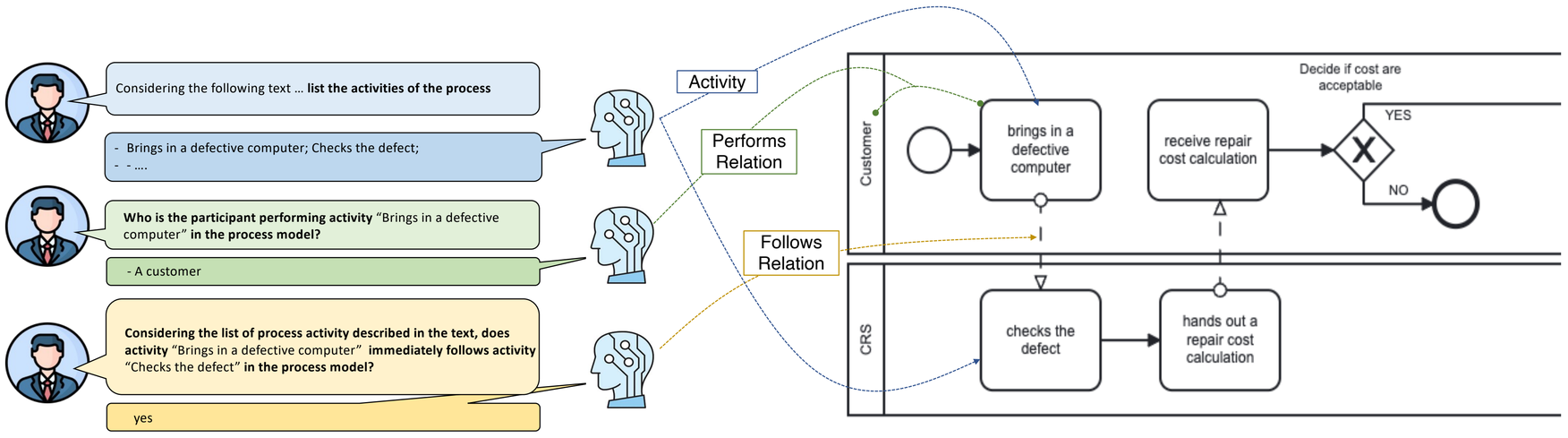
2SHOTS

+ Process description
+ Task Instructions
+ Examples

DEFS+2SHOTS

+ Process description
+ Task Instructions
+ Contextual Knowledge
+ Examples

A customer brings in a defective computer and the CRS **checks the defect** and hands out a repair cost calculation. If the customer decides that the costs are acceptable, the process continues, otherwise, she takes her computer home unrepaired. The ongoing repair consists of two activities, which are executed, in an arbitrary order. ….

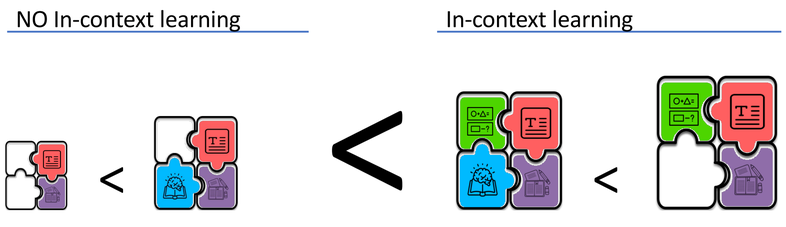


Results

We tested the proposed approach on a subset of documents selected from the PET Dataset

NO In-context learning

In-context learning



Process Graph Construction

Graph Edit Distance measure between Gold-Standard graph and the Extracted one

	NO In-context learning		In-context learning	
Text ID	Raw	Defs	Defs+2Shots	2Shots
doc-1.2	31.0	33.0	13.0	9.0
doc-1.3	20.0	32.0	42.0	39.0
doc-3.3	12.0	14.0	30.0	17.0
doc-5.2	30.0	12.0	22.0	21.0
doc-10.1	19.0	19.0	4.0	6.0
doc-10.6	19.0	19.0	4.0	2.0
doc-10.13	15.0	15.0	13.0	5.0
Average	21.0	18.7	11.2	7.5

Reference

Bellan, P., Ghidini, C., Dragoni, M. (2022). Assisted Process Knowledge Graph Building Using Pre-Trained Language Models. AI*IA
Bellan, P., Ghidini, C., Dragoni, M., Ponzetto, S.P., & Aa, H.V. (2022). Process Extraction from Natural Language Text: the PET dataset and annotation guidelines. NL4AI@AI*IA
Bellan, P., Dragoni, M., & Ghidini, C. (2022). Extracting Business Process Entities and Relations from Text Using Pre-trained Language Models and In-Context Learning. *EDOC*.
Bellan, P., Dragoni, M., & Ghidini, C. (2022). Experiment Maker: A Tool to Create Experiments with GPT-3 Easily. *EKAW*.
Bellan, P., Dragoni, M., & Ghidini, C. (2021). Process Extraction from Text: state of the art and challenges for the future. *ArXiv, abs/2110.03754* (in preparation)
Bellan, P., Dragoni, M., & Ghidini, C. (2020). A Qualitative Analysis of the State of the Art in Process Extraction from Text. *DP@AI*IA*.