Script Induction for Agent-based Simulations

Mehwish Alam & Nils Holzenberger
first.last@telecom-paris.fr

December 4, 2024

Description

Taxpayers occasionally find clever ways to combine existing laws to achieve unexpected tax discounts. These combinations are often referred to as tax loopholes, and come up often enough in tax courts, likely causing a significant tax deficit to governments [1]. While policy-makers can use available historical data to guide their policy choices, unexpected reactions of entities to new policies can hardly be foreseen. We aim to discover tax loopholes by simulating a legal and economic environment and letting agents interact with the environment and possibly one another, guided by reward maximization. This exploration leads to the discovery of new schemes of behavior, including previously unknown tax loopholes. Code ReCivil, a tool for agent-based simulation, relies on agent-based simulation environments, formalized legal rules, reward-oriented exploration and knowledge extraction [2].

To find tax loopholes, Code ReCivil must identify strategies used by agents [3]. It must turn sequences of events produced by the simulation into schemes that humans can make sense of. Here, we propose to turn to Natural Language Processing (NLP) and the rich research area of script induction [4, 5]. By translating event logs into human-readable natural language, we will make it possible to leverage methods from data-driven script induction.

Objectives

Our goal is to study whether it is feasible to induce scripts from event logs by using natural language.

Verbalizer. The first step is to verbalize event logs, the result of existing CodeReCivil project, by automatically translating atomic and structured data into free-form natural language. This first step should take half a month.

Script Induction. The second step will explore script induction on this large-scale NLP dataset. Script induction is an open research area, where Large Language Models (LLMs) have recently shown promising results [6]. During the next 2 months, we will explore LLM-based approaches. Our baseline will be to use LLMs to spell out schemas [7]. Further, since the rule framework is part of the CodeReCivil simulation and is thus known, we will explore methods to give LLMs access to structured rules [8]. The intern will spend the following 2 months to compare these approaches to classical automated script induction [9].

The last month will be spent writing up our results for submission to a tier-A NLP conference. Depending on the outcomes of the project, and how many new challenges we identify, we envision submitting a proposal for a PhD thesis on this topic.

References

[1] Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme. Shelter check: Proactively finding tax minimization strategies via ai. *Tax Notes Federal, Dec*, 12, 2022.

[2] Peter Fratric, Nils Holzenberger, and David Restrepo Amariles. Rules2lab: from prolog knowledge-base, to learning agents, to norm engineering. In *EUMAS 2024*, 2024.

[3] Peter Fratric, Mostafa Mohajeri Parizi, Giovanni Sileno, Tom M. van Engers, and Sander Klous. Do agents dream of abiding by the rules?: Learning norms via behavioral exploration and sparse human supervision. In Matthias Grabmair, Francisco Andrade, and Paulo Novais, editors, *Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law, ICAIL 2023, Braga, Portugal, June 19-23, 2023*, pages 81–90. ACM, 2023.

[4] Roger C. Schank and Robert P. Abelson. Scripts, plans and knowledge. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence*, pages 151–157, 1975.

[5] Noah Weber, Anton Belyy, Nils Holzenberger, Rachel Rudinger, and Benjamin Van Durme. Human schema curation via causal association rule mining. In *Proceedings of the 16th Linguistic Annotation Workshop (LAW-XVI) within LREC2022*. European Language Resources Association, 2022.

[6] Abhilasha Sancheti and Rachel Rudinger. What do large language models learn about scripts? In *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics, *SEM@NAACL-HLT*, pages 1–11. Association for Computational Linguistics, 2022.

[7] Sha Li, Ruining Zhao, Manling Li, Heng Ji, Chris Callison-Burch, and Jiawei Han. Open-domain hierarchical event schema induction by incremental prompting and verification. In *Proceedings of the 61st Annual Meeting of the*

Association for Computational Linguistics (Volume 1: Long Papers), pages 5677–5697. Association for Computational Linguistics, 2023.

[8] Agnieszka Lawrynowicz, Luis Galárraga, Mehwish Alam, Berenice Jaulmes, Vaclav Zeman, and Tomás Kliegr. Neurosymbolic methods for rule mining, book chapter to appear in the "Handbook of Neurosymbolic Artificial Intelligence". *CoRR*, abs/2408.05773, 2024.

[9] Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers, and Clare R. Voss. Connecting the dots: Event graph schema induction with path language modeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 684–695. Association for Computational Linguistics, 2020.