2024 IEEE International Conference on Web Services (ICWS) | 979-8-3503-6855-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICWS62655.2024.00046

HORAE: A Domain-Agnostic Modeling Language for Automating Multimodal Service Regulation*

[Work-in-Progress Paper]

Yutao Sun, Mingshuai Chen, Kangjia Zhao, Jintao Chen

***Abstract*—Artificial intelligence is rapidly encroaching on the field of service regulation. This work-in-progress article presents the design principles behind HORAE, a unified specification language to model multimodal regulation rules across a diverse set of domains. We show how HORAE facilitates an intelligent service regulation pipeline by further exploiting a fine-tuned large language model named RuleGPT that automates the HORAE modeling process, thereby yielding an end-to-end framework for fully automated intelligent service regulation.**

***Index Terms*—Service regulation, specification languages, compliance checking, multimodal learning, violation recognition.**

I. INTRODUCTION

The rapid advancements in the realm of artificial intelligence (AI) – particularly breakthroughs in deep neural networks and the swift rise of (multimodal) large language models (LLMs) – have triggered a recent surge of interest in *intelligent service regulation*. Employing AI in the field of service regulation may substantially improve the degree of automation and accuracy, thereby yielding a significant cost reduction. However, current AI-based regulation methods predominantly adopt a *plug-and-play* approach: As illustrated in Fig. 1a, regulation industries encompass a wide spectrum of *scenarios* (e.g., healthcare and financial services) and data *modalities* (e.g., text, images, and videos). A common practice is to train a distinct model that caters to a specific combination of a scenario and a modality, e.g., model for urban management in the images modality [1] and model for e-commerce in the text modality [2].

The plug-and-play method, however, suffers from two major issues: (i) *significant resource wastage*: the training and deployment of multiple large-scale AI models tailored for various regulation scenarios and modalities necessarily incur a model proliferation and thereby substantial carbon emissions [3]; and (ii) *confined flexibility and efficiency*: every time a new scenario (resp. modality) is introduced, it has to be (manually) adapted to all the existing modalities (resp. scenarios) to enable a full-fledged regulation architecture (cf. Fig. 1).

In response to these challenges, we propose HORAE – a unified specification language to model *multimodal* regulation rules in a *domain-agnostic* fashion. HORAE leverages the *zero-shot understanding* capability of large language models [4] to translate regulation rules from any scenario (aka, domain) and modality into a structured *intermediate representation* (IR); see Fig. 1b. This representation dissects complex behavior patterns across different domains into a set of fine-grained, readily-detectable events and actions. Consequently, the downstream recognition models and algorithms – being agnostic to specific domains and modalities – can utilize a unified rule interface to undertake and discharge the regulation tasks.

We show that HORAE facilitates an intelligent service regulation pipeline by further exploiting a fine-tuned large language model coined RuleGPT to automatically convert regulation rules written in natural language to the intermediate representation of HORAE. A formal semantics can be developed for HORAE to enable rule-consistency checking and quantitative violation recognition (via, e.g., constraint-solving techniques), see Fig. 1b, thereby yielding an *effective end-to-end framework for fully automated intelligent service regulation*.

Our contributions and future plans are as follows:

- We introduce HORAE as a unified specification language to model cross-domain, multimodal regulation rules. We envisage that, with well-designed semantics, HORAE can facilitate core regulation tasks such as rule modeling, consistency checking, and quantitative violation recognition.
- We plan to collect a benchmark dataset covering a wide range of regulation domains and modalities, and thence create a fine-tuned LLM (using, e.g., LoRA [5]) called RuleGPT to automate the modeling process in HORAE.
- We present an end-to-end intelligent service regulation framework leveraging HORAE and RuleGPT. This framework sheds light on fully automated service regulation admitting effective domain-modality unification.

Below, we first describe the end-to-end framework of HORAE-steered intelligent service regulation in Sect. II and then present detailed design principles of HORAE in Sect. III.

II. INTELLIGENT SERVICE REGULATION FRAMEWORK

Fig. 2 depicts an overview of our end-to-end framework of HORAE-steered intelligent service regulation. This framework consists of the following four main steps:

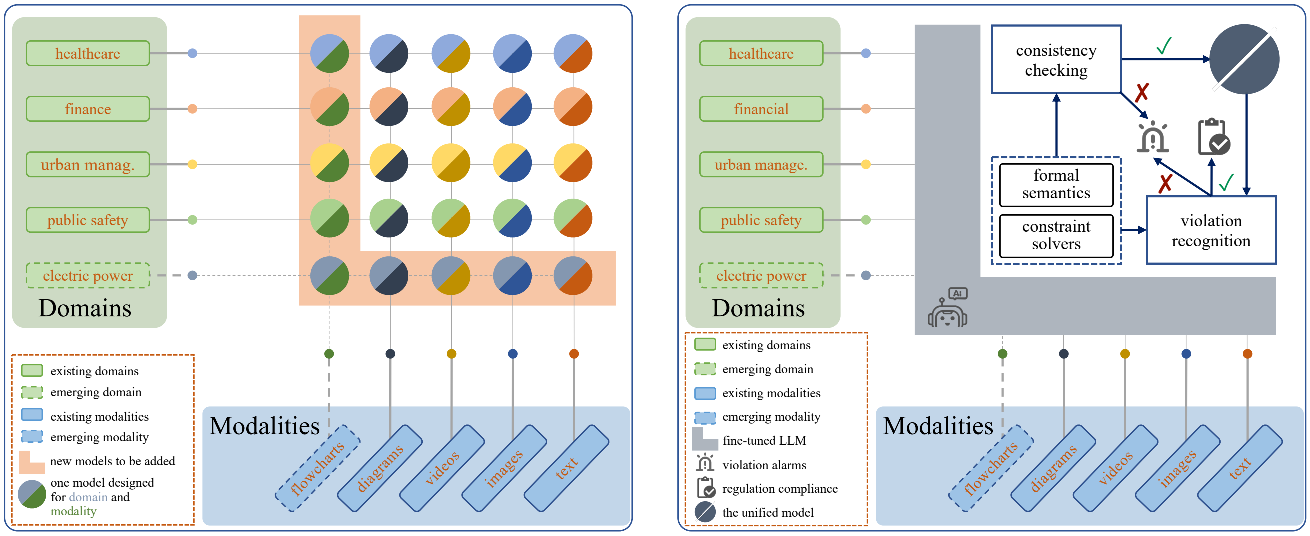
1) *Multimodal Rule Alignment*: This initial step aligns (preprocessed) multimodal regulation rules – leveraging existing multimodal models – to the text modality such that

* HORAE (/ˈhɔːriː/) refers to – in Greek mythology – the goddesses of order who guarded the gates of Olympus (Homer, *The Iliad*).

This work has been partially funded by the National Key R&D Program of China under grant No. 2022YFF0902600, by the ZJNSF Major Program under grant No. LD24F020013 and by the ZJU Education Foundation's Qizhen Talent program. (Corresponding author: Mingshuai Chen)

Y. Sun, M. Chen, K. Zhao, J. Chen are with the Zhejiang University, Hangzhou 310027, China (e-mail: {ytsun, m.chen, konkaz, chenjintao}@zju.edu.cn).

2836-3868/24/$31.00 ©2024 IEEE
DOI 10.1109/ICWS62655.2024.00046



(a) The plug-and-play approach

(b) The HORAE architecture

Fig. 1: The plug-and-play approach vs. the HORAE architecture. Conventional plug-and-play methods are often confined to distinct models for specific domains and data modalities, thus requiring extensive retraining and resource expenditure. In contrast, HORAE acts as a unified specification language to model multimodal regulation rules in a domain-agnostic fashion.

rules of different formats can later be interpreted through a *unified medium*, i.e., HORAE rules.

2) *Rule Dataset Construction*: The aligned textual rules are formed into a (semi-)structured rule dataset, which consists of a *multimodal lexicon* – key terms associated with multimodal media, e.g., images – and a *multimodal rule library* – rules in a format that is close to HORAE.
3) *Rule Modeling and Checking*: The (semi-)structured rule dataset is then translated into HORAE using our fine-tuned RuleGPT. As per the formal semantics of HORAE, we can check the qualitative and quantitative *consistency* of the rule library to detect potential conflicts before deploying it to downstream regulation tasks.
4) *Violation Recognition*: The downstream recognition tasks are discharged by multimodal models and algorithms, which assess the *violation probabilities* of basic events in the rule library. These violation probabilities contribute to an overall likelihood of rule violation (computed by a probability calculation engine).

Our preliminary implementation of the proposed framework indicates that the above steps suffice to produce highly accurate outcomes in a fully automated manner in various domains.

III. DESIGN PRINCIPLES OF HORAE

HORAE serves as the basis of our intelligent service regulation framework by modeling a set of (multimodal) regulation rules in a structured, domain-agnostic fashion. The following principles are employed throughout the design of HORAE:

- *Generality*: HORAE must accommodate regulation rules across different domains and multilingual texts; it shall abstract away domain-specific elements while encoding common regulation patterns.
- *Structuration*: HORAE shall remain (semi-)structured to effectively store, check, and manipulate regulation rules while resolving potential ambiguities.
- *Automation*: HORAE shall feature simple patterns and structures to allow for an automated translation of rules written in natural language to those in HORAE via LLMs.
- *Quantification*: HORAE shall encode both qualitative and quantitative aspects of rule satisfaction, i.e., whether a rule is satisfied and the likelihood of it being satisfied.

In line with these principles, we design the syntax (as per generality, structuration, and automation) and formal semantics (concerning quantification) of HORAE.

A. Syntax

Our syntactic analysis over a collected multilingual benchmark set of regulation rules reveals the following observations:

- *Independency*: Regulation rules that are ostensibly disparate in grammatical structure may well encode semantically similar regulation intentions. Hence, HORAE's syntax shall be independent of any specific natural language grammar and optimized to admit the most diverse set of intentions with as few grammatical categories as possible.
- *Rule Types*: Regulation rules typically fall into categories of *enforced*, *recommended*, or *forbidden* behaviors. HORAE is thus expected to provide a simple mechanism to specify these types for regulation rules.
- *Behavioral Composition*: Regulation rules often describe behaviors as combinations of sub-behaviors using logical connectives. HORAE supports this compositionality by maintaining an abstract layer of *basic events*, which can be logically assembled to describe complex behaviors.
- *Temporality*: Temporal properties are crucial in many regulation domains, especially where timing constraints

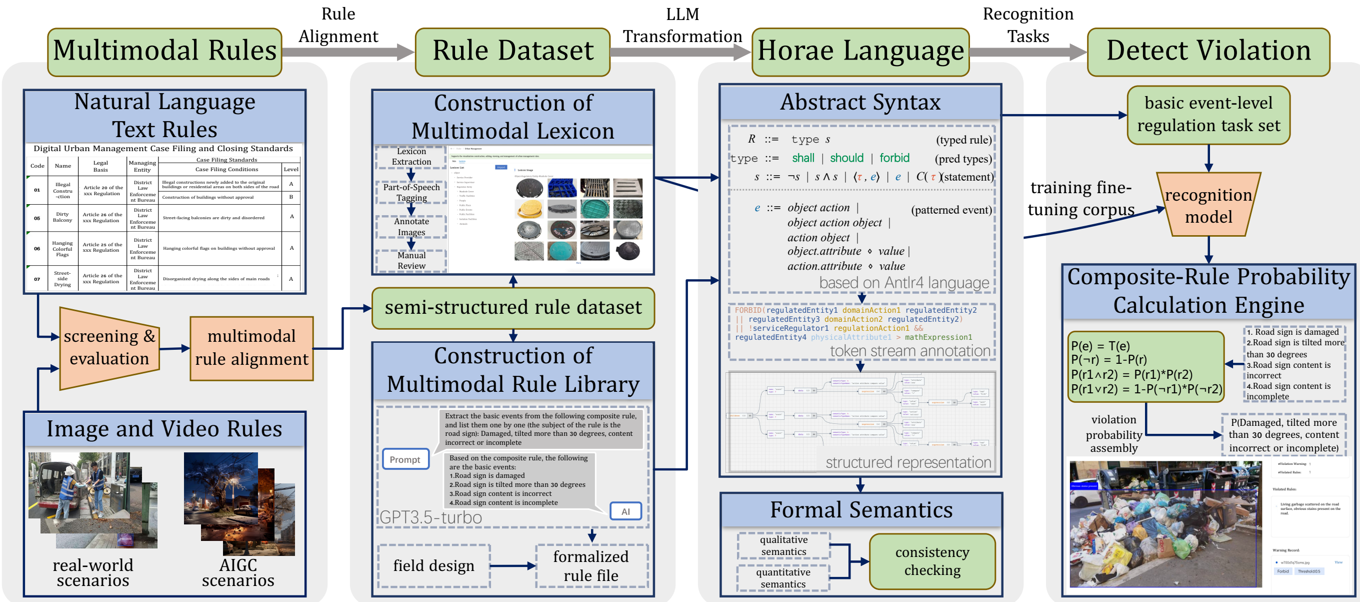


Fig. 2: The end-to-end framework of HORAE-steered intelligent service regulation.

are significant. HORAE supports temporality by allowing *timestamped events* and t*emporal constraints*, providing a natural means of modeling time-sensitive regulation rules.

Based on these observations, we propose to model a *regulation rule R* in HORAE that adheres to the following syntax:

$$R \quad ::= \quad \texttt{type}\ s \qquad \text{(typed rule)}$$

$$\texttt{type} \quad ::= \quad \textsf{shall} \mid \textsf{should} \mid \textsf{forbid} \qquad \text{(predefined types)}$$

$$s \quad ::= \quad \neg s \mid s \wedge s \mid \langle \tau, e \rangle \mid e \mid \mathcal{C}(\tau) \qquad \text{(statement)}$$

$$\begin{aligned} e \quad ::= \quad & \textit{object action} \mid & \text{(patterned event)} \\ & \textit{object action object} \mid \\ & \textit{action object} \mid \\ & \textit{object.attribute} \diamond \textit{value} \mid \\ & \textit{action.attribute} \diamond \textit{value} \end{aligned}$$

This abstract syntax consists of a *top-level* and a *bottom-level* grammar, as indicated by the dashed line therein. The former combines (timestamped) basic events via logical connectives into a regulation rule of certain type, whilst the latter assembles fine-grained sentence patterns and components into such basic events. Slicing basic events into smaller, detectable ingredients improves the precision of downstream recognition models.

B. Formal Semantics

An unambiguous formal semantics of HORAE is crucial for representing, interpreting, and reasoning about a typically large set of regulation rules. In particular, it provides a mechanism to check the consistency of a rule library to detect potential conflicts before deploying it to downstream regulation tasks.

To reason about regulation rules *qualitatively*, we can interpret basic events as Boolean values and timestamps as non-negative real numbers. A rule library is considered *qualitatively consistent* if there exists an interpretation where all rules evaluate to true. Such qualitative consistency can be checked using state-of-the-art satisfiability modulo theories (SMT) solvers. This approach ensures that the rule library is logically sound and free of contradictions.

We further plan to extend the qualitative semantics to its *quantitative counterpart* to characterize the likelihood of rule satisfaction. Such quantitative aspects are particularly crucial for intelligent service regulation since the underlying recognition models inherently produce imprecise results (measured by certain confidence factors).

IV. CONCLUSION

This work-in-progress paper presents an end-to-end intelligent regulation framework steered by a domain-agnostic modeling language for multimodal service regulation. Future work includes formalizing the semantics of HORAE and fine-tuning and optimizing RuleGPT, thereby ultimately achieving full automation of the intelligent service regulation framework.

REFERENCES

[1] A. Kaginalkar, S. Kumar, P. Gargava, and D. Niyogi, "Review of urban computing in air quality management as smart city service," *Urban Climate*, vol. 39, p. 100972, 2021.

[2] M. A. Raji, H. B. Olodo, T. T. Oke, W. A. Addy, O. C. Ofodile, and A. T. Oyewole, "E-commerce and consumer behavior: A review of ai-powered personalization and market trends," *GSC Advanced Research and Reviews*, vol. 18, no. 3, pp. 066–077, 2024.

[3] A. S. Luccioni, S. Viguier, and A. Ligozat, "Estimating the carbon footprint of BLOOM, a 176B parameter language model," *J. Mach. Learn. Res.*, vol. 24, pp. 253:1–253:15, 2023.

[4] J. Wei, M. Bosma, V. Y. Zhao, K. Guu, A. W. Yu, B. Lester, N. Du, A. M. Dai, and Q. V. Le, "Finetuned language models are zero-shot learners," in *ICLR*. OpenReview.net, 2022.

[5] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-rank adaptation of large language models," in *ICLR*. OpenReview.net, 2022.