NL2TL: Transforming Natural Languages to Temporal Logics using Large Language Models

Yongchao Chen
MIT / Harvard University
yongchaochen@fas.harvard.edu

Rujul Gandhi
MIT
rujul@mit.edu

Yang Zhang
MIT-IBM Watson AI Lab
Yang.Zhang2@ibm.com

Chuchu Fan
MIT
chuchu@mit.edu

Abstract

Temporal Logic (TL) can be used to rigorously specify complex high-level specification for systems in many engineering applications. The translation between natural language (NL) and TL has been under-explored due to the lack of dataset and generalizable model across different application domains. In this paper, we propose an accurate and generalizable transformation framework of English instructions from NL to TL, exploring the use of Large Language Models (LLMs) at multiple stages. Our contributions are twofold. First, we develop a framework to create a dataset of NL-TL pairs combining LLMs and human annotation. We publish a dataset with 28K NL-TL pairs. Then, we finetune T5 models on the lifted versions (i.e., the specific Atomic Propositions (AP) are hidden) of the NL and TL. The enhanced generalizability originates from two aspects: 1) Usage of lifted NL-TL characterizes common logical structures, without constraints of specific domains. 2) Application of LLMs in dataset creation largely enhances corpus richness. We test the generalization of trained models on five varied domains. To achieve full NL-TL transformation, we either combine the lifted model with AP recognition task or do the further finetuning on each specific domain. During the further finetuning, our model achieves higher accuracy (>95%) using only <10% training data, compared with the baseline sequence to sequence (Seq2Seq) model.[1][2]

1 Introduction

Temporal Logic (TL) has been widely used as a mathematically precise language to specify requirements in many engineering domains such as robotics (Tellex et al., 2020), electronics design (Browne et al., 1986), autonomous driving (Maierhofer et al., 2020). TL can capture the complex spatial, temporal, and logical requirements of both human languages and environmental constraints, and can be transformed into executable actions or control inputs for robots (Gundana and Kress-Gazit, 2022; Raman et al., 2013; Boteanu et al., 2016; Patel et al., 2020; Gopalan et al., 2018).

Unlike many robotics works that try to use end-to-end black-box models to infer robotic behaviors directly from natural language (NL) (Ahn et al., 2022), using structured TL as the intermediate has a twofold benefit – the TL can be used for direct planning, and the TL representation can be used to identify specific sources of failure and provide automatic feedback to a non-expert user (Raman et al., 2013). However, TL has a steep learning curve. Communicating one's goals and constraints through NL is much more intuitive to a non-expert. Therefore, a model able to transform NL instructions into TL is a missing but crucial component for interactive robots and engineering designs.

Currently, there is no general tool to perform automated translations between TL and NL that takes the following requirements into consideration:

- **Cross-domain generalization.** Although TL is used in many engineering domains, current NL-to-TL approaches largely constrain their training data to a single domain. These datasets mostly lack plentiful corpus richness of NL-TL and have their own specified formats of Atomic Propositions (AP). Then the models fail to generalize to other domains (Gopalan et al., 2018), even though the structure of TL itself is not dependent on the domain and should be generic.

- **Variability of NL instructions.** Past work often constructs synthetic data algorithmically, requiring limited forms of the NL input. Real-world NL utterances cannot be encoded into

[1]Datasets and Codes are available at https://github.com/yongchao98/NL2TL

[2]Project Page is available at https://yongchao98.github.io/MIT-realm-NL2TL