

LLM's and Semantic Faithfulness in Formal Methods

Ana Pires (PG61130)

Miguel Carvalho (PG61153)

Renato Garcia(PG61542)

Contexto

LLMs conseguem traduzir linguagem natural em especificações formais (LTL). Porém...

- Ferramentas atuais geram fórmulas sintaticamente válidas...
- ... mas não capturam a intenção real do engenheiro
- Em contextos críticos como o ferroviário, isto é perigoso
- Um requisito pode parecer correto e ainda assim falhar - **Falha Silenciosa**

Contexto

- Engenheiros têm dificuldade em transformar diagramas e descrições semi-formais em modelos formais matemáticos
- A tradução de requisitos de LN para especificações formais é um passo intensivo em mão-de-obra e propenso a erros
- Embora LLMs produzam código rapidamente, elas falham em "edge cases", o que torna a verificação formal essencial

Métodos Formais na Indústria

01.

- A Adoção de métodos formais na indústria é travada porque aumenta o ciclo de desenvolvimento de software em até 30%

02.

- A necessidade de treino rigoroso para engenheiros escreverem especificações formais é o principal "gargalo"

03

Métodos Formais na Indústria

03.

- Req2Spec: Já converte 71% dos requisitos da BOSCH (automóvel) em especificações formais.

04.

- NL2LTL: a tradução de Linguagem Natural para LTL já atinge 94.4% de precisão em contextos controlados

03

Objetivos

Explorar o uso de representações intermédias que funcionem como ponte entre a linguagem natural e a especificação formal.

01.

Desenvolver uma Prova de Conceito (PoC) que adapta uma arquitetura de duas fases ao contexto ferroviário.

02.

Resolver o problema da “falha silenciosa.”



Conceitos Base

Arquitetura em duas fases

- **Fase 1** : converter o requisito em linguagem natural numa representação estruturada intermédia
- **Fase 2**: gerar a fórmula formal

REQ2LTL - (Usado em contexto aeroespacial) Framework que propõe a arquitetura de duas fases.

OnionL - Representação intermédia usada pelo REQ2LTL

LTL - descreve propriedades sobre o comportamento de sistemas ao longo do tempo (sequência de estados).

Fidelidade Semântica - tradução captura corretamente o significado e a intenção - não apenas a forma.

Projeto Métodos Formais de Programação

Milestone 1





Análise SWOT

■ Pontos Fracos

- Priorizam validade sintática sobre fidelidade semântica
- Falham em especificações complexas e de maior âmbito
- Outputs frequentemente requerem pós-processamento manual

■ Pontos Fortes

- LLMs atingem até 94.4% de precisão em NL→LTL
- Ferramentas como Req2Spec e SpecGen mostram resultados promissores
- CoT melhora coerência e raciocínio lógico

■ Ameaças

- Escassez de datasets industriais anotados e públicos
- Ferramentas de verificação formal com formatos incompatíveis
- LLMs opacos dificultam confiança em sistemas críticos de segurança

■ Oportunidades

- Arquiteturas de duas fases com representação intermédia
- Raciocínio neuro-simbólico para maior fiabilidade
- Integração de human-in-the-loop para reduzir erros



■ 01

- Estudo do estado da arte:
 - Métodos Formais na Indústria
 - Influência de LLM's no Software Formal
 - LTL
 - NL2spec, Req2LTL, NL2LTL

■ 02

- Recolha inicial de Requisitos
 - Transformação nas respetivas fórmulas LTL corretas

■ 03

- Implementação do PoC

■ 04

- Testes e Avaliação

Plano



- Reunião Semanal com os orientadores
 - Professor Hugo Macedo
 - Zhenze Wu (Critical Software)
 - Segundas-feiras às 10H
- Reunião de Grupo às Quintas-Feiras
- Rotatividade no uso das ferramentas
- Escalar as perguntas para as reuniões semanais serem produtivas

Plano



Conclusão

O foco tem sido a compreensão aprofundada do problema e a exploração das ferramentas existentes - NL2LTL, NL2Spec e Req2Spec - pelo que ainda não existe uma calendarização detalhada definida. À medida que o grupo consolidar o conhecimento sobre as ferramentas e o domínio, será possível estabelecer um plano de trabalho mais concreto para as fases seguintes.

Obrigado