FSU | FLORIDA STATE UNIVERSITY

STUDENT SEMINARS @ FSU Computer Science

ATOMIC: Attention Reallocation to Mitigate Irrelevant Context in Small Language Models

Bolin Shen

Presentation Abstract

Small language models have recently gained increasing attention due to their strong cost efficiency, computational efficiency, and competitive performance across a wide range of reasoning tasks. However, when reasoning inputs contain irrelevant context, SLMs are substantially more vulnerable to distraction than large language models, leading to severe performance degradation. Through a fine grained analysis of inference time attention, we identify a severe underlying issue: SLMs exhibit a markedly weaker ability to distinguish relevant context from irrelevant context, as evidenced by consistently lower AUROC scores when treating attention over RC and IC as a binary discrimination problem. As a result, SLMs often assign high attention to distracting tokens in irrelevant context and erroneously incorporate them as evidence during reasoning, which disrupts effective evidence aggregation. To address this limitation, we propose ATOMIC. Our approach first performs task conditioned context level filtering to suppress irrelevant information, then introduces an Evidence Likelihood Score to amplify attention on critical evidence tokens within relevant context, and finally reallocates attention during answer inference to further reinforce key evidence. Extensive experiments across multiple mathematical reasoning benchmarks demonstrate that our method achieves state of the art performance.

Speaker's Profile

I am a first-year Ph.D. student in Computer Science at Florida State University, where I am fortunate to work under the supervision of Prof. Yushun Dong. Before that, I completed my master's degree at the University of Michigan – Ann Arbor.

My research mainly focus on strengthening Small Language Models (SLMs), especially by improving their memory mechanisms and reasoning abilities to achieve efficient yet powerful intelligence. In parallel, I investigate Graph Neural Networks (GNNs) with an emphasis on model extraction attacks and defenses mechanisms. Ultimately, my goal is to build secure, robust, and trustworthy machine learning systems that contribute to the broader vision of Responsible AI.

Seminar Time: 2026-03-03 Tuesday, 11:45 AM - 12:45 PM

Location (In-Person Only): Room 353, LOVE Building

Contact FSU Student Seminar Organizers

Dr. Yushun Dong

Email: yd24f@fsu.edu