

Avviso di Seminario

Title: Can Reinforcement Learning Teach Small Language Models to Reason for Text2SQL?

Where: Sala riunioni sez. di Informatica & Automazione, 1° piano ex-OMI, stanza 1.14

When: Mercoledì 15 aprile 2026 alle ore 11:30

Speaker: Prof. Paolo Papotti (EURECOM)

Abstract

The ability to interact with complex databases using natural language (NL) is a key step in democratizing data access, a long-standing goal in the enterprise world. While Large Language Models (LLMs) have shown remarkable promise in translating NL questions into SQL queries (Text2QL), their performance stalls when faced with the complexities of real-world enterprise databases. This talk will report a promising solution to enhance the reasoning capabilities of LLMs for this task. Our "Think2SQL" methodology investigates various strategies for improving LLM performance, including Zero-Shot Learning (ZSL), Supervised Fine-Tuning (SFT), and Reinforcement Learning (RL). RL, using rewards crafted around SQL execution accuracy, significantly boosts the performance of small models, achieving results comparable to those of frontier LLMs on complex datasets. Finally, we will highlight the path forward for Text2SQL systems capable of navigating the nuances of human language, such as ambiguity, in a real-world enterprise context.

Short biography of the speaker: Paolo Papotti is a Professor at EURECOM (France) since 2017 and the holder of a Chair of Artificial Intelligence at the 3IA Institute since 2024. He got his PhD from Roma Tre University (Italy) in 2007 and had research positions at the Qatar Computing Research Institute (Qatar) and Arizona State University (USA). His research is focused on data management and NLP. He has authored more than 150 publications and his work has been recognized with best paper awards (CIKM 2024, ISWC 2024), best demo awards (SIGMOD 2015, DBA 2020, SIGMOD 2022), and Google Faculty Research awards (2016, 2020)..

For further information please contact Prof. R. Torlone (riccardo.torlone@uniroma3.it)