



Institute of Mathematics
Academia Sinica

Department of Mathematics
National Taiwan University



Lakeside Lectures



Speaker:

Dawsen Hwang (Google DeepMind)

Title:

Achieving IMO Gold: Robust Mathematical Reasoning with LLMs

Date: November 3 (Mon.), 2025

Time: 14:00-15:00

Venue: Room 202, Astro-Math Building (NTU Campus)

Abstract: In 2025, an advanced version of Gemini with Deep Think achieved a historic gold-medal standard at the International Mathematical Olympiad (IMO), the world's most prestigious competition for young mathematicians. This breakthrough represents a significant leap in AI reasoning, as our model solved five of the six problems end-to-end in natural language—producing rigorous, human-readable proofs directly from the problem statements, all within the official competition time limit.

This talk will provide high-level ideas and design principles about the LLM and how it achieves robust mathematical reasoning. A critical challenge in advancing mathematical AI is that existing evaluation benchmarks are often too simplistic or focus only on correct short answers, failing to assess the deep, multi-step reasoning required for Olympiad-level mathematics. We will present how we overcame this issue by establishing a new north-star for our research: a suite of rigorous evaluation benchmarks designed to reward the entire problem-solving process. This talk will detail the design of these benchmarks, and explore how they, in turn, guided the high-level research and design principles that were fundamental to achieving gold-level performance.

Refreshment: 13:30

Organizers:

Olivier Hénnot(NTU), Chun-Ju Lai(AS), Chao Ming Lin(NTU), Colin McSwiggen(AS), Masao Oi(NTU)