

# HUAJIAN XIN

Email: [H.Xin-3@sms.ed.ac.uk](mailto:H.Xin-3@sms.ed.ac.uk)  
GitHub: [github.com/xinhjBrant](https://github.com/xinhjBrant)  
Homepage: <https://xinhuajian.wordpress.com>  
Google scholar: <https://scholar.google.com/citations?hl=en&user=E5M9x8wAAAAJ>

## EDUCATION

- Ph. D. Student** | *Artificial Intelligence* Sep. 2024 – Apr. 2027 (expected)  
School of Informatics, the University of Edinburgh  
• Supervisor: Wenda Li  
Edinburgh, United Kingdom
- Bachelor of Philosophy** | *Logic* Sep. 2019 – Jun. 2023  
Department of Philosophy, Sun Yat-sen University  
• GPA: 91% (ranking 2/32)  
Guangzhou, China  
• Selected Coursework:  
    *Logic*: Mathematical Logic, Naive Set Theory, Proof Theory, Model Theory, Computation Theory, Modal Logic,  
    Non-classical Logic, Formal Semantics, Informal Logic  
    *Mathematics*: Mathematical Analysis, Linear Algebra, Abstract Algebra (Lattice Theory and Universal Algebra),  
    Probability theory  
    *Computer Science*: Data Structure and Algorithm, Artificial Intelligence (Logic Programming and Machine Learning)  
    *Philosophy*: Metaphysics, Epistemology, Philosophy of Language, Philosophy of Mind  
• Honor:  
    *National Scholarship* (2021, top prize for university students in mainland China, funded by the central government)  
    *Honors Graduate of Sun Yat-sen University* (2023)
- Undergraduate (Minor)** | *Mathematics and Applied Mathematics* Feb. 2021 – Jun. 2021  
School of Mathematics, Sun Yat-sen University  
Guangzhou, China  
• Selected Coursework:  
    Real Variable Function, Complex Variable Function, Mathematical Statistics

## RESEARCH AND WORKING EXPERIENCES

- Intern** | **DeepSeek AI** Jan. 2024 – Aug. 2024  
*Supervisor*: Chong Ruan and Daya Guo  
Beijing, China
- Research Assistant** | **Sun Yat-sen University** Sept. 2022 – Jan. 2024  
*Supervisor*: Xiaodan Liang and Zhengying Liu  
Shenzhen, China

## SELECTED PUBLICATIONS AND PREPRINTS

- DeepSeek-Prover-V1.5: Harnessing Proof Assistant Feedback for Reinforcement Learning and Monte-Carlo Tree Search** [arXiv](#) & [GitHub](#) & [Huggingface](#) & [X.com](#)  
• Introduced informal chain-of-thought augmented whole-proof generation, reinforcement learning from proof assistant feedback (RLPAF), and the intrinsic-reward-driven Monte Carlo tree search algorithm (RMaxTS).  
• Set a new state-of-the-art in theorem proving for Lean 4, achieving a 63.5% pass rate on the high school-level miniF2F benchmark and a 25.3% pass rate on the undergraduate-level ProofNet benchmark.
- DeepSeek-Prover-V1: Advancing Theorem Proving in LLMs through Large-Scale Synthetic Data** [arXiv](#)  
• Developed a data synthesis pipeline that automatically formalizes natural language math problems into 8 million formal statements with proofs in Lean 4.  
• Set a new state-of-the-art in theorem proving for Lean 4, achieving a 50.0% pass rate on the miniF2F benchmark.
- Proving Theorems Recursively** [NeurIPS 2024 Poster](#)  
• Introduced a recursive, level-by-level theorem proving method, improving on traditional step-by-step approaches by focusing on high-level proof sketches and deferring intermediate conjectures to later stages.
- LEGO-Prover: Neural Theorem Proving with Growing Libraries** [ICLR 2024 Oral](#)  
• Developed a modular theorem-proving approach with an integrated skill library, enabling large language models to propose, prove, and store lemmas for reuse in proving target theorems.  
• Set a new state-of-the-art in theorem proving for Isabelle, achieving a 47.1% pass rate on the miniF2F benchmark.
- MUSTARD: Mastering Uniform Synthesis of Theorem and Proof Data** [ICLR 2024 Spotlight](#)  
• Developed a data synthesis framework that prompts the language model with specific keywords to generate and solve mathematical problems in natural language, followed by translating the problems and solutions into Lean 3 to verify their correctness.