MATHEMATICS RESEARCH CENTRE

Logic Meets AI

Characterizing the expressive power of transformers through logic

Veeti Ahvonen

5 June 2026

WHY THEORY

Background & Motivation

AI models now solve a wide range of problems across many domains. But while applied AI thrives, the theoretical foundations have lagged behind.

Building those foundations could help resolve:

1

Reproducibility crisis

Results that are hard to reproduce

2

Technical debt

Models grown by trial and error, costly to maintain

3

Architectural choices

Design decisions made by intuition

4

Explainability

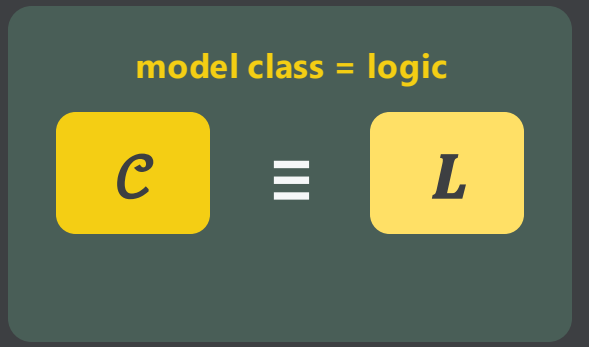
Black-box behavior that is hard to interpret or trust

THE CORE QUESTION

Using Logic to Understand AI

"What can these AI models actually solve?"

Lets us represent a computing model in a precise **logical formalism**, and in ideal cases the resulting formula is **directly interpretable**, turning a black box into something we can read and explain.



Every model has an equivalent formula — and vice versa

WHY IT HELPS

- Architectural choices
- Formal verification
- Automated reasoning
- Explainability

OUR RECENT WORK

Characterizing Transformers

Transformers are the backbone of today's large language models (Vaswani et al. 2017 NIPS).

THEY POWER

ChatGPT | Claude | Copilot & more

OUR CONTRIBUTION

A novel logical characterization of transformers via modal logic.

THE RESULT

Transformers ≡ Logic

In the simplest case, both compute a function $\{0,1\}^* \to \{0,1\}$

INPUT — A WORD

OUTPUT — ONE BIT

Is the word accepted?

Transformer

Operates on a word ***w*** through self-attention layers that aggregate ***global information***

≡

Propositional logic with global modality

$\langle G \rangle \geq 3\ \mathrm{red}$

"at least 3 nodes of w are red"

THEOREM 1 On words, transformers and propositional logic with global modality are equally expressive

→ *Generalizes from words to graphs.*

A., Funk, Heiman, Kuusisto, Lutz, ***"Expressive Power of Graph Transformers via Logic," AAAI 2026.***