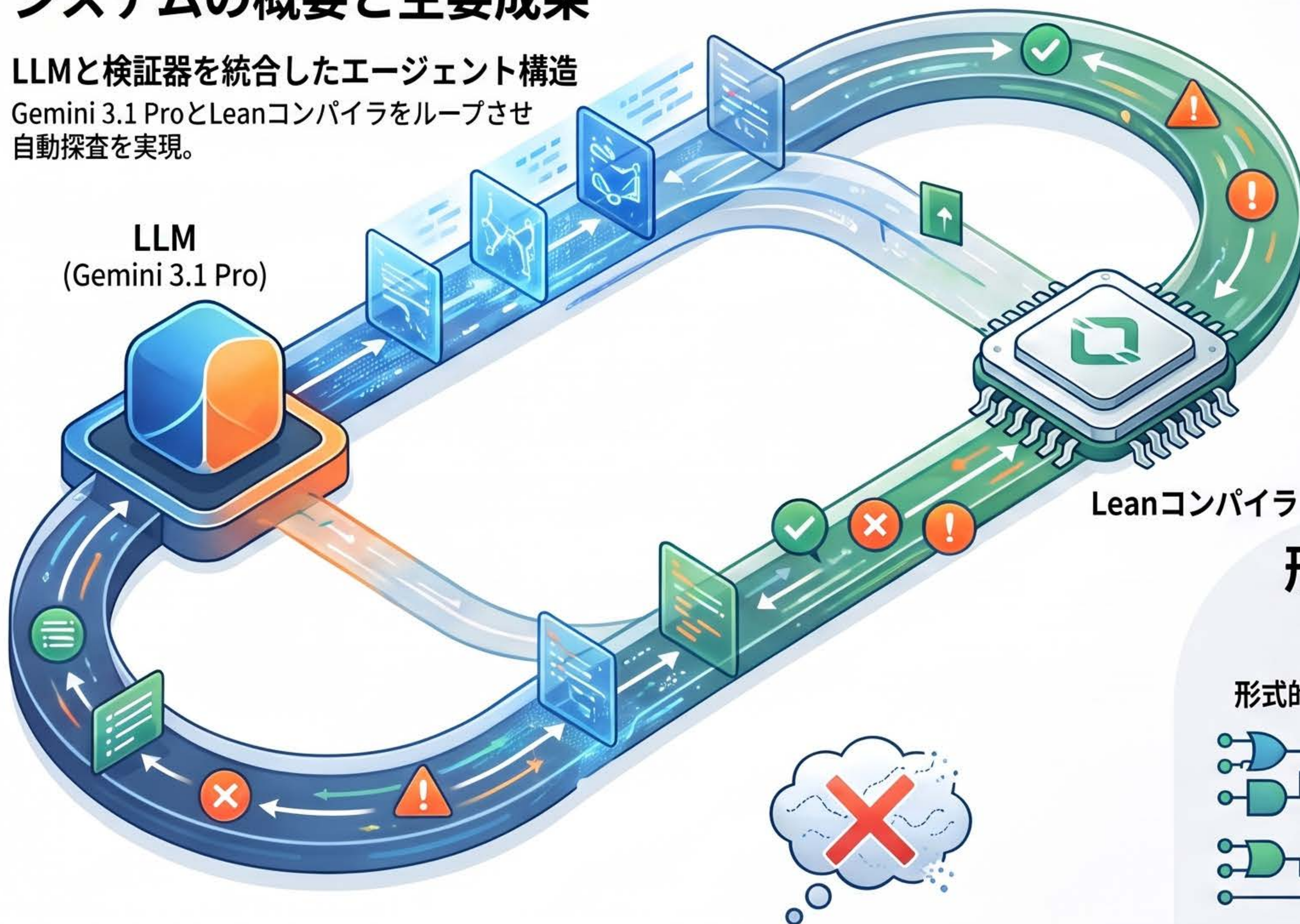


AlphaProof Nexus : AIと形式検証が切り拓く数学研究の未来

システムの概要と主要成果

LLMと検証器を統合したエージェント構造
Gemini 3.1 ProとLeanコンパイラをループさせ
自動探査を実現。



ハルシネーションの不活性化と排除
実在しない補題はsorry検知や検証失敗
によって確実に弾かれる。

Erdős問題



9件解決

353件探索
約2.55%成功率

機械検証可能な
Lean証明

OEIS問題



44件解決

492件探索
約8.94%成功率

機械検証可能な
Lean証明

Erdős問題9件・OEIS問題44件を解決

未解決問題に対し、自然言語ではなく機械検証可能なLean証明を出力。

形式検証の真価と人間の役割

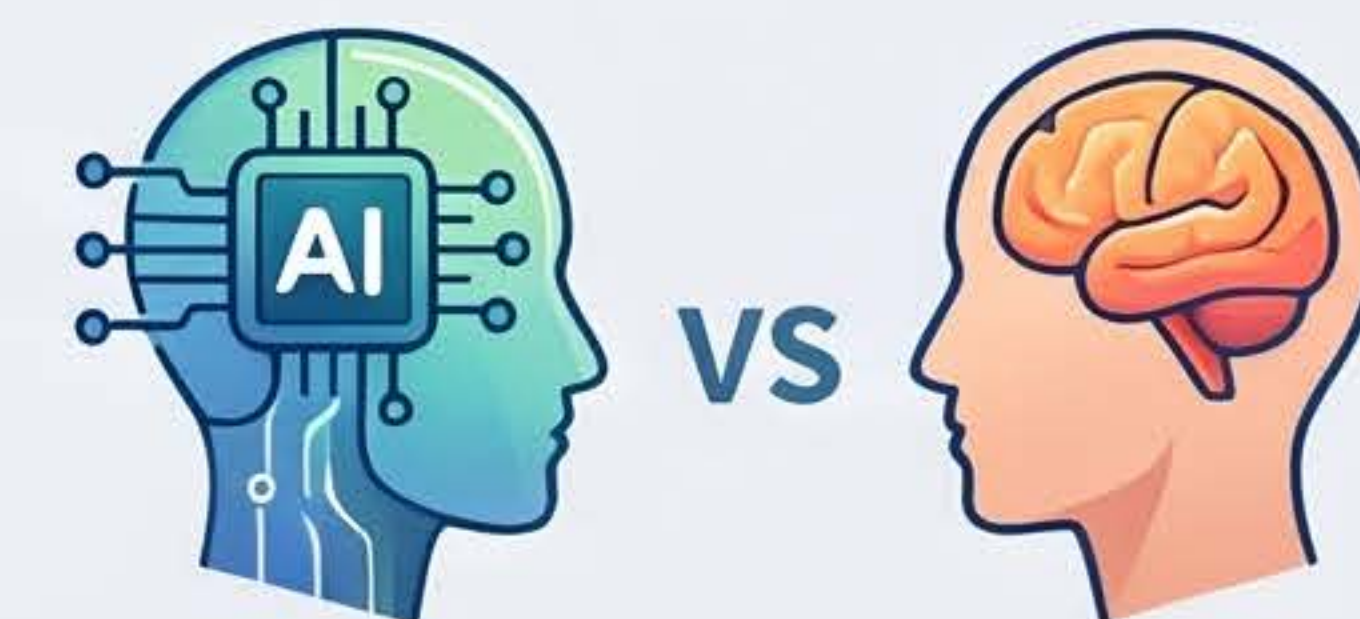
形式検証の真価



形式的正しさと意味的妥当性の分離

Leanは論理の正しさを保証するが
問題設定の意味論的正しさは保証しない。

人間とAIの新しい共同研究モデル



AI: 大容量探索と
証明検証

人間:
問題定義と理解

AIが大容量探索と証明検証を担い、
人間は問題定義と理解を担う。