

自律型AIによる科学的発見の最前線:「Aristotle Agent」と「Get Physics Done (GPD)」の深層解析

Gemini 3.1 pro

第1章: 次世代人工知能におけるパラダイムシフトと形式的推論の台頭

人工知能(AI)の進化は、自然言語処理やパターン認識の領域において目覚ましい成果を上げてきたが、科学研究の最前線で求められる「厳密性」と「形式的証明」の壁に直面していた。大規模言語モデル(LLM)は確率論的なトークン生成に依存しており、幻覚(ハルシネーション)を伴う性質上、ミッションクリティカルな工学や理論物理学、純粋数学における推論エンジンとしては本質的な限界を抱えていた。この限界を突破する新たなパラダイムが急速に形成されつつある。それが、論理的真理を形式言語によって検証する「数学的超知能(Mathematical Superintelligence: MSI)」の追求と、物理世界のシミュレーションおよび実験ワークフローを自律的に統合・実行する「エージェント型物理学者」の誕生である。

本報告書は、この新たな技術潮流を牽引する2つの革新的なAIシステム、すなわちHarmonic社が開発した世界初の自律型AI数学者「Aristotle Agent」と、Physical Superintelligence PBC (PSI)が公開したオープンソースのAI物理学者「Get Physics Done (GPD)」について、そのアーキテクチャ、哲学、技術的基盤、そして科学的探求における認識論的意義を網羅的に解析するものである。これらのシステムは単なる計算ツールではなく、科学的仮説の立案から検証、そして論文執筆に至るまでの一連のプロセスを自律的に推進する知的エージェントとして機能し、人類の科学的発見の速度を根本的に変革する可能性を秘めている。

第2章: Aristotle Agent — 幻覚なき数学的超知能(MSI)の実現と哲学的基盤

2.1 開発の背景と「数学的超知能」のビジョン

「Aristotle Agent」は、金融プラットフォームRobinhoodの最高経営責任者(CEO)であるVlad Tenevと、AI研究者のTudor Achimによって設立された先進的なAI研究機関Harmonicによって開発された¹。Harmonicのミッションは、人間を凌駕する数学的能力を持つAI「数学的超知能(MSI)」を構築することである³。彼らのアプローチの根底には、数学があらゆる推論と問題解決の絶対的な基礎であり、数学的推論能力を極めることが、結果として物理学やソフトウェア工学といった広範な領域における汎用的かつ厳密なAIの実現に直結するという確信が存在する²。

従来の言語モデルベースのアプローチとは異なり、Aristotle Agentは「幻覚のない(

hallucination-free) 出力を保証するためにゼロから設計されている⁵。これは、単に尤もらしいテキストを生成するのではなく、出力結果が完全に形式的に検証可能であることを意味する。定量的推論の領域において、同モデルは論理の破綻がないことをシステムレベルで保証しており、確率的な推測を証明可能なソリューションに置き換えることで、正確性が極めて重要となる実世界の高リスクなアプリケーションにおける最も信頼できるモデルとして位置付けられている¹。

2.2 アリストテレス的哲学とAIエージェントの基本原則

「Aristotle(アリストテレス)」という命名は、単なる歴史的偉人へのオマージュにとどまらず、このAIエージェントの深層的な設計思想そのものを体現している。古代ギリシャの哲学者アリストテレスは、エージェント(行為者)の動因について、非理性的な欲求(appetitive desires)と、それを正しい善へと導く理性(reason)の二元的な相互作用を提唱した⁶。デカルト的形而上学が理性を意志の中核としたのに対し、あるいはヒュームが理性を単なる欲求充足の道具とみなしたのとは対照的に、アリストテレスの枠組みでは、理性が情動や直感的な衝動を操舵し、システム全体をエラーから守りながら目的へと導く中核的役割を果たす⁶。

この哲学的枠組みは、Aristotle Agentのニューロシンボリック(Neuro-symbolic)なAIアーキテクチャに見事なまでに符合している。大規模言語モデル(LLM)の持つ直感的・確率論的な生成能力やパターン認識能力は、アリストテレス的な「欲求」や「ヒューリスティックな情動」に相当し、時に幻覚や論理的飛躍(エラー)を引き起こす。しかし、それを定理証明支援系(Lean 4)という厳密な形式論理、すなわちアリストテレス的な「理性」が強固に操舵し、絶対的な真理へと導くのである⁴。

さらに、アリストテレス哲学における「能動知性(Agent Intellect)」と「受動知性(Patient Intellect)」の概念も、このシステムの動作原理を説明する上で重要である⁸。ギリシャ哲学において、能動知性は不動であり同一性によって対象を認識する絶対的な論理基盤であるのに対し、受動知性は運動を伴い、離散的な操作を通じて対象を追跡する⁸。Aristotle AgentにおけるLean 4の不変の公理系(能動知性)と、MCTS(モンテカルロ木探索)を通じて動的に探索空間を移動する探索アルゴリズム(受動知性)の組み合わせは、まさにこの古代の認識論の現代的な実装と言える。また、アリストテレスが『自然学』や『形而上学』において因果関係(質料因、形相因、動力因、目的因)を体系的に分析したように⁷、このエージェントは数学的命題の因果や構造を包括的かつ体系的に解き明かす能力を備えているのである。

2.3 アーキテクチャの深層: 形式的検証と非形式的推論の高度な統合

Aristotle Agentの技術的ブレイクスルーは、形式的検証(Formal Verification)と非形式的推論(Informal Reasoning)の高度な統合にある。自然言語のみに依存するシステムとは異なり、AristotleはLean 4で証明を記述し、その数学ライブラリであるMathlibに照らして検証することを必須としているため、ギャップや不健全な公理のないソリューションが保証される⁴。アーキテクチャは主に以下の表に示す3つの高度なサブシステムから構成されている⁴。

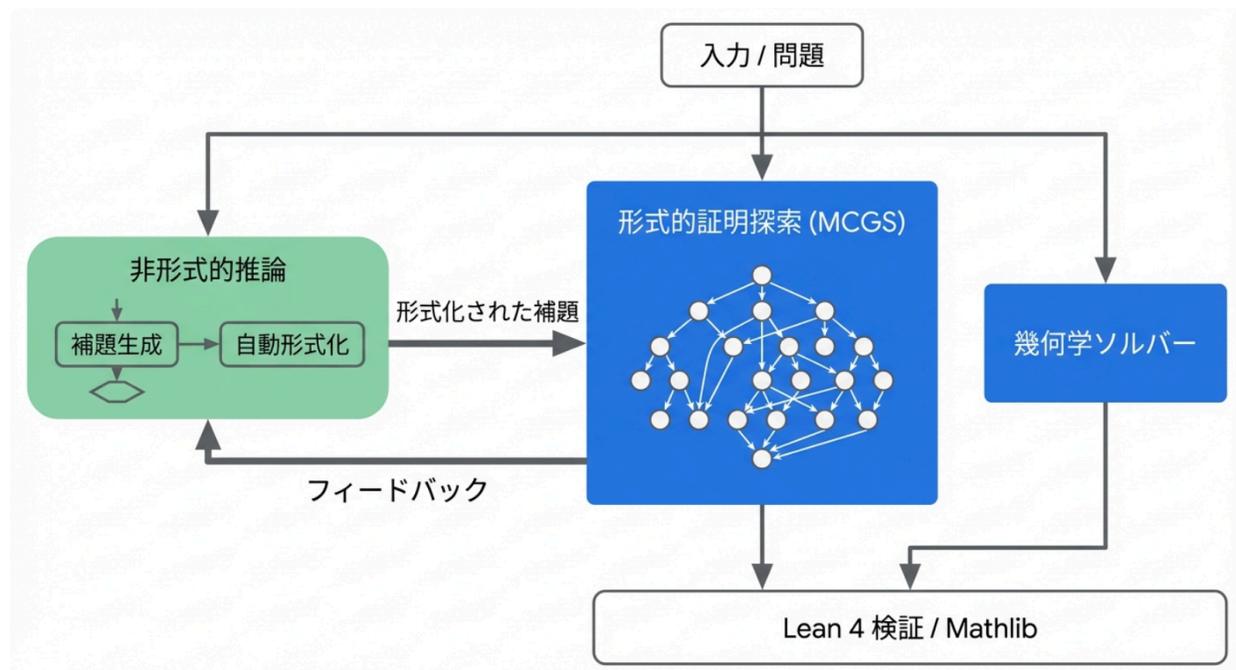
サブシステム名	コア技術とメカニズム	果たす役割と特徴

Lean証明探索アルゴリズム	並列化モンテカルログラフ探索 (MCGS)、PUCT、トランスフォーマー・ポリシー	Lean証明スケッチを読み込み、未証明のゴールを解決する。等価な状態を統合して木からグラフへ変形し、計算効率を最大化する。
補題ベースの非形式的推論	LLMによる自然言語推論、自動形式化 (Autoformalization)、動的思考予算	数学命題を非形式的に証明し、小さな補題に分解。Leanコードへ変換し、形式検証のフィードバックに基づいて反復修正する。
専用幾何学ソルバー	AlphaGeometryベースのアルゴリズム	平面幾何学の問題をLeanの外部で専用に処理し、図形的な推論を効率的かつ正確に解決する。

1. Lean証明探索アルゴリズム (Lean Proof Search Algorithm) の革新性 探索の核となるのは、高度に並列化されたモンテカルログラフ探索 (Monte Carlo Graph Search: MCGS) である⁴。従来のモンテカルロ木探索 (MCTS) を拡張したMCGSは、同一のゴール式、ローカルコンテキスト、変数名を持つLeanの論理状態を同等と見なすことで、探索空間を単純なツリー構造から相互接続されたグラフへと変形し、計算の重複を排除して効率を飛躍的に高めている⁴。探索戦略にはPUCT (Predictor Upper Confidence bound applied to Trees) の変法が用いられている。数学の証明においては、単一のアクションが成功すれば状態が証明される「OR条件」と、アクションの結果生じるすべてのサブゴールが証明されなければならない「AND条件」が複雑に絡み合う。このアルゴリズムは、これらの条件を動的に評価し、最も有望なアクションと最も解決が困難な状態を優先的に探索することで、計算リソースを最適化している⁴。

2. 補題ベースの非形式的推論システム (Lemma-Based Informal Reasoning System) このシステムは、「隠された思考の連鎖 (Hidden Chain of Thought)」、「非形式的なコメント」、「形式的なLeanコード」という3つの異なるレイヤーの出力を共進化させる⁴。特筆すべきは「動的な思考予算 (Dynamic thinking budget)」の割り当て機能である。これにより、システムは特定の推論プロセスがゴール達成に極めて重要であることを自己認識し、そこにより多くの計算リソースと推論ステップを割り当てることができる⁴。この非形式的推論モジュールが生成した仮説や補題は、即座に自動形式化 (Autoformalization) され、Lean 4の厳密な環境でテストされる。

Aristotle Agentにおける形式的・非形式的推論の統合アーキテクチャ



非形式的な推論エンジン（言語モデル）が補題と証明戦略を生成し、並列化されたモンテカルログラフ探索（MCGS）を通じてLean 4の形式検証環境で評価される反復的プロセス。

このアーキテクチャの統合により、Aristotle Agentは単なる計算機を超え、論理的推論と直感的飛躍を組み合わせた真の「数学的協働者」として機能することが可能となっている。

第3章：Aristotle Agentの実証的成果とインフラストラクチャの極致

3.1 ProofBenchにおける圧倒的優位性と数学界への衝撃

Aristotle Agentのアーキテクチャの優位性は、外部機関による厳格なベンチマークテストや実際の数学研究において実証されている。特に、Vals AIが開発した形式数学ベンチマーク「ProofBench」における結果は、現在のAI業界における地殻変動を示すものである。

ProofBenchは、100の公開データと100の非公開テストデータから構成され、モデルは自然言語の定理を与えられた上で、非形式的な証明のヒントなしに、Lean 4の証明チェッカーで完全に受理される形式的証明をゼロから生成しなければならない¹⁴。確率論、確率解析、測度論、関数解析、可換代数、代数幾何学、数理論理学など、大学院レベルの高度な現代数学を網羅するこの過酷なテストでの成功は、推論の連鎖全体にわたって論理的正確性を維持できることを明確に証明するものである。

14。

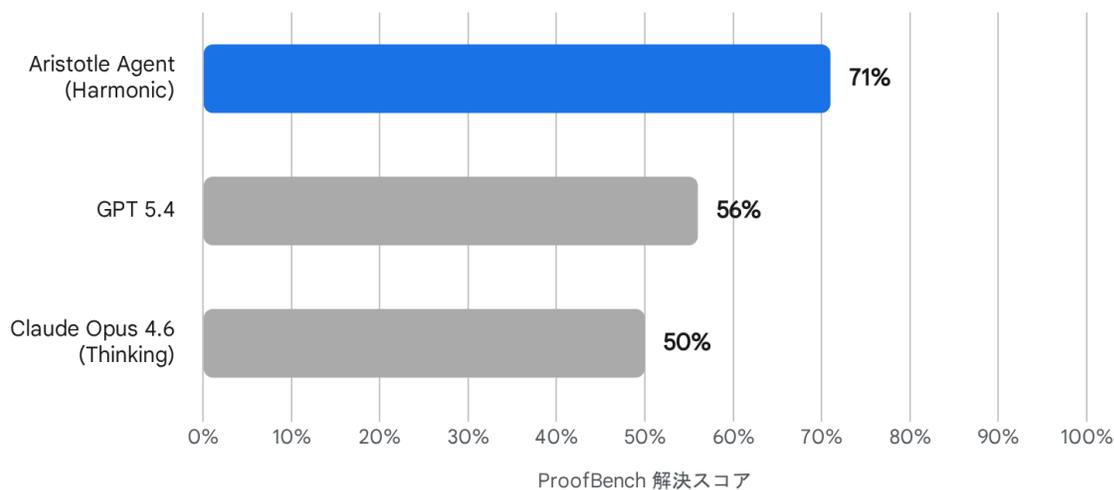
以下の表は、主要なAIモデルのProofBenchにおけるスコアを比較したものである。

順位	AIモデル名	開発・提供元	ProofBench スコア	備考
1	Aristotle Agent	Harmonic	71%	専用の形式数学エージェント。圧倒的な首位。
2	GPT 5.4	OpenAI	56%	汎用基盤モデルにおけるトップ。
3	Claude Opus 4.6 (Thinking)	Anthropic	50%	高度な推論能力を持つ汎用モデル。
23	MiniMax M2.7	MiniMax	3%	比較対象となる他の最新モデルの一例。

このデータが示す通り、Aristotle Agentは最先端の汎用基盤モデルであるGPT 5.4に対して15ポイントもの圧倒的な性能差をつけている¹⁴。汎用言語モデルが形式数学において限界を迎えつつある中、専用に構築されたエージェント型アーキテクチャがいかに優れているかを定量的に実証している。

ProofBench（形式数学ベンチマーク）におけるAristotle Agentの圧倒的リード

15% の性能差 (対トップ汎用基盤モデル)



Vals AIによるProofBenchベンチマーク結果。Aristotle Agentは、最先端の汎用基盤モデルであるGPT 5.4を15ポイント上回り、形式的数学証明タスクにおいて第1位にランクされている。

データソース: [Vals AI](#)

3.2 IMO 2025と「エルデシュの問題」の単独証明

理論的ベンチマークにとどまらず、Aristotle Agentは歴史的な数学的課題においても画期的な成果を挙げている。同システムは、2025年国際数学オリンピック(IMO)の問題において、6問中5問について正しい形式的解決策を提供し、金メダル相当のパフォーマンスを達成した⁴。

さらに学术界に衝撃を与えたのは、約30年間にわたり未解決であったErdos Problem #124(エルデシュの問題)をAristotle単独で証明し、Leanで検証可能な証明を生成したことである¹⁸。この問題は「整数べき乗の集合の完全なシーケンス」に関するものであり、数十年間数学者のアプローチを退けてきた。著名な数学者であるテレンス・タオ(Terence Tao)は、この過程におけるAIの振る舞いを詳細に分析している。ChatGPTのような汎用LLMが生成したアイデアには軽微なエラーや論理的ギャップが含まれていたが、Aristotleはこれらのギャップを自動的に修復し、Leanによる検証済みの証明を完全に構築することに成功したのである¹⁵。

Aristotleは「完全なエージェント型(Fully Agentic)」として機能し、英語で提示された問題に対してゼロから証明と形式化を自律的に行う。最大24時間にわたり人間の介入なしで稼働し続け、コードリポジトリやLeanプロジェクト内のファイルを直接編集して、大規模な形式化プロジェクトに「そのままライ

ブラリに統合可能なコード(Library-Ready Code)」を提供する能力を持つ¹⁵。

3.3 インフラストラクチャ・エンジニアリングの極致: pbccとpython-memtools

数学的超知能を実環境で稼働させるためには、極めて高度なインフラストラクチャが必要となる。Harmonicのエンジニアリングチームは、計算のボトルネックを解消するために独自のツール群を開発し、オープンソースとして科学コミュニティに還元している。

1. 無限のスケーラビリティを誇るREPLサービス Aristotleと実際のLean証明システムとの対話は、専用のREPL(Read-Eval-Print Loop)サービスを通じて行われる。このサービスは意味論的にステートレス(無状態)に設計されており、高価なGPUの容量に依存することなく、安価なプリエンパティブルインスタンスを用いてGoogle Kubernetes Engine(GKE)上で数十万のCPUコアへと動的にスケーリング可能である⁴。これにより、MCGSの膨大な並列計算を経済的かつ効率的に支えている。

2. pbccコンパイラ: Protocol Buffersの限界突破 データ規模の増大に伴い、AI開発においてはPythonの標準的なProtocol Buffers実装がボトルネックとなる。標準実装はメモリ消費が激しく、特に2GBを超える巨大なメッセージの処理において頻繁にクラッシュするという致命的な問題が発生していた⁴。これを解決するため、Harmonicは事前コンパイル(AOT)アプローチを採用したカスタムコンパイラpbccを開発した。pbccは最適化されたC++テンプレートを生成し、Pythonのネイティブなデータ型(リストや辞書)に直接マッピングする。これにより、煩雑なキャスト処理を不要にし、型ヒント(mypy等)を完全にサポートしながら、2GBのサイズ制限を完全に排除した高速かつメモリ効率の良い大規模データ処理を実現している⁴。

3. python-memtoolsによる深層デバッグ さらに、長時間のデータ生成スクリプトがCPU使用率0%のままハングアップするという難解なバグに直面した際、既存のプロファイラ(py-spy等)では非同期タスクの状態を可視化できなかった⁴。そこで開発されたpython-memtoolsは、Pythonプロセスのメモリ状態の完全なスナップショットを取得し、基本タイプオブジェクト(base type object)のメモリ上の固有シグネチャ(自身のタイプへのポインタを持つ等)を利用して全Pythonオブジェクトを追跡する画期的なツールである⁴。これにより、入力グラフの予期せぬサイクルに起因する非同期タスクのデッドロック(asyncio.Taskの待機ループ)や、Cコードとlibuv間の低レベルなデッドロックを正確に特定し、解決に導いた⁴。

3.4 ソフトウェア工学と物理学への拡張(Verification Problemの解決)

Aristotle Agentのロードマップは、純粋数学の枠組みにとどまらない。Harmonicは現在、この形式的推論アーキテクチャを「厳密な数学的理解が不可欠な依存領域」、すなわちソフトウェア工学や理論物理学へと拡張している²。

現状のAI技術は、高度な技術的コンテンツを生成する能力を持つ一方で、最終的には人間がその正当性を「検証(Verification)」しなければならないという致命的なボトルネック(いわゆるVerification Problem)を抱えている⁴。Aristotleは、Lean 4の証明支援系を用いて公理レベルまで遡ってソリューションを検証することで、出力の正確性に関する「機械的に検証可能な保証(Machine-checkable guarantee of correctness)」を提供する⁴。これは、ミッションクリティカルなソフトウェア開発や、物理学の理論的探求において、人間の検証作業を不要にする画期的な進歩であ

る。ソフトウェアコーディングの検証能力を測るVerinaベンチマークにおいても最先端の成績を収めており、物理的現象の数学的モデリングにおけるエラーの排除に大きく貢献すると期待されている⁴。

第4章：物理世界への拡張：Physical Superintelligence (PSI) と「Solve Everything」

数学の領域でAristotle Agentが論理的真理の自動化を進める一方で、物理的現実のモデリング、シミュレーション、そして実験の領域では、Physical Superintelligence PBC (PSI) が開発した「Get Physics Done (GPD)」が決定的な革新をもたらしている。

4.1 PSIと「Solve Everything」フレームワークの哲学

PSIは、科学技術分野で著名なAlex Wissner-Gross博士らによって設立されたAIスタートアップであり、「Solve Everything (すべてを解決する)」という極めて壮大かつ具体的なミッションを掲げている¹⁹。その目標は、人間の直感や偶然に基づく単発的な発見に依存してきた従来の物理学研究のあり方から脱却し、人工超知能の力を通じて「科学的発見を工業化 (Industrialize scientific discovery)」することである¹⁹。

PSIは、新しい物理法則の発見や革新的な理論構築のプロセスを加速させることが、トランジスタやレーザー、原子力エネルギーの開発に匹敵する「物理学の新たな黄金時代」を到来させると予見している¹⁹。この目標を達成するため、PSIは物理的超知能のための「世界初の垂直統合型ファクトリー」を構築中であり、Google、OpenAI、Meta、ハーバード大学、MIT、スタンフォード大学などの最高峰の頭脳を結集している¹⁹。

このビジョンの最初の具体的な具現化であり、科学コミュニティへのオープンソース貢献として公開されたのが、世界初のエージェント型AI物理学者「Get Physics Done (GPD)」である¹⁹。

4.2 GPDの設計思想と対応ランタイム環境

GPDは、単純なプロンプト入力では対応できない「困難な物理学の研究課題 (hard physics research problems)」を自律的に支援するために特化して設計されている²⁵。理論的導出、複雑な数値計算の安定性チェック、シミュレーション環境の構築、そして学術論文の執筆まで、長期的な視野 (long-horizon) を必要とする研究プロジェクト全体を管理・実行する機能を持つ²⁶。

特筆すべきは、人間の研究者が頻繁に陥る表記揺れや前提の不整合をシステムレベルで排除する能力である。GPDは、最大18の異なる物理学分野にまたがって、表記法 (notation)、符号の選択 (sign choices)、および前提条件の一貫性をプロジェクト全体で厳格に維持する²⁶。

GPDは柔軟性を重視しており、単一のLLMに依存せず、以下の表に示す4つの主要なAIランタイムと統合して動作するよう設計されている²⁶。

サポートされるランタイム	コマンドプレフィックス	設定・管理ファイルの場所	特徴
Claude Code	/gpd:help	.claude/settings.json	MCPサーバーを使用し、高度なコード生成と検証タスクを処理。
Codex	\$gpd-help	.codex/config.toml	マルチエージェント機能を有効化し、専門のエージェントロールを登録して動作。
Gemini CLI	/gpd:help	(シェル承認を通じて管理)	experimental.enableAgentsを利用し、自動編集とシェルの直接承認を実行。
OpenCode	/gpd-help	.opencode/opencode.json	プロジェクト内の権限と構成を統合的に管理し、オープンソース環境に最適化。

第5章 : Get Physics Done (GPD) の自律的ワークフローと厳密な検証機構

5.1 4段階のワークフローと階層的コンテキスト管理

GPDの研究プロセスは、物理学者が論文を執筆するまでの実際の思考プロセスを模倣し、以下の4つの明確なステージに構造化されている²⁵。

1. 定式化 (Formulate):

研究の初期段階において、GPDがユーザーに対してターゲットを絞った質問を自律的に行う。これにより、問題の範囲、前提となる物理的仮定、使用する表記法、および最終的な検証のための具体的な目標を明確に定義する。

2. 計画 (Plan):
定義された目標に基づき、具体的なタスク、依存関係、および成功基準からなる段階的なロードマップ (ROADMAP.md) を生成する。
3. 実行 (Execute):
多段階の解析的導出、文献レビュー、数値的チェック、論文の執筆など、タスクの性質に応じた専門のスペシャリスト・エージェントが動的に配備され、並列または直列で実行される。
4. 検証 (Verify):
物理的整合性のチェック (次元解析、対称性の制約、極限ケース、保存則など) や、数値的な収束テストを含む包括的な監査を実行する。

この複雑なワークフローが途中で破綻しないよう、作業は「永続的なコンテキスト構造 (Persistent context structure)」として階層化されている²⁶。最上位の **Project** (研究ワークスペース) の下に、**Milestone** (論文提出等の主要目標)、**Phase** (研究の個別セグメント)、**Plan** (依存関係に基づく Wave 群)、そして最小単位である **Task** が整理される²⁶。これにより、エージェントは数週間におよぶ研究プロセスにおいても、過去の前提や文脈を見失うことがない。

5.2 厳密な物理検証 (Physics Validation) コマンド群

GPD が一般的なコーディング AI と一線を画す最大の理由は、研究結果の科学的完全性と物理的妥当性を監査するための専門的な「物理検証 (Physics Validation)」コマンド群を実装している点である²⁶。これらは、厳格な査読者 (ピアレビューアー) として機能する。

物理検証コマンド	機能詳細と物理学研究における意義
/gpd:dimensional-analysis	導出されたすべての方程式に対し、次元の整合性 (長さ、時間、質量などの単位の一貫性) を体系的に監査する。方程式の根本的な誤りを瞬時に検出する。
/gpd:limiting-cases	質量がゼロに近づく極限、あるいは速度が光速に近づく極限など、特定の極端なシナリオや簡略化された条件下において物理法則が崩れないかを体系的に検証する。
/gpd:numerical-convergence	複雑な数値計算において、グリッドサイズや時間ステップを変化させた際の安定性と精度を保証す

	るための体系的な収束テストを提供する。
/gpd:compare-experiment	プロジェクト内で生成された理論的予測と、既存の実験データや天体観測データとを照合し、理論の現実世界における妥当性を検証する。
/gpd:validate-conventions	プロジェクトの全フェーズにおいて、符号の定義（例：計量の符号関数）や座標系の規約が崩れていないか検証し、長期プロジェクトでの致命的な不整合を防ぐ。
/gpd:regression-check	プロジェクトに新たな変更を加えた後、以前に検証済みの「真実 (truths)」が破壊されていないかを自動的に再評価する。

5.3 定量的解析 (Quantitative Analysis) と論文生成

物理学の実験やシミュレーションにおいては、不確実性の管理が不可欠である。GPDの定量的解析コマンドは、この側面を完全に自動化する²⁶。

- **/gpd:parameter-sweep**: 広範なパラメータ空間を並列実行で探索し、系の振る舞いの全体像をマッピングする。
- **/gpd:sensitivity-analysis**: どの物理パラメータが最終的なモデル出力に最も大きな影響を与えるか、不確実性がどのように伝播するかを定量化する。
- **/gpd:error-propagation**: 多段階の複雑な計算を通じて、初期条件の誤差や測定不確実性がどのように増幅・伝播していくかを厳密に追跡する。

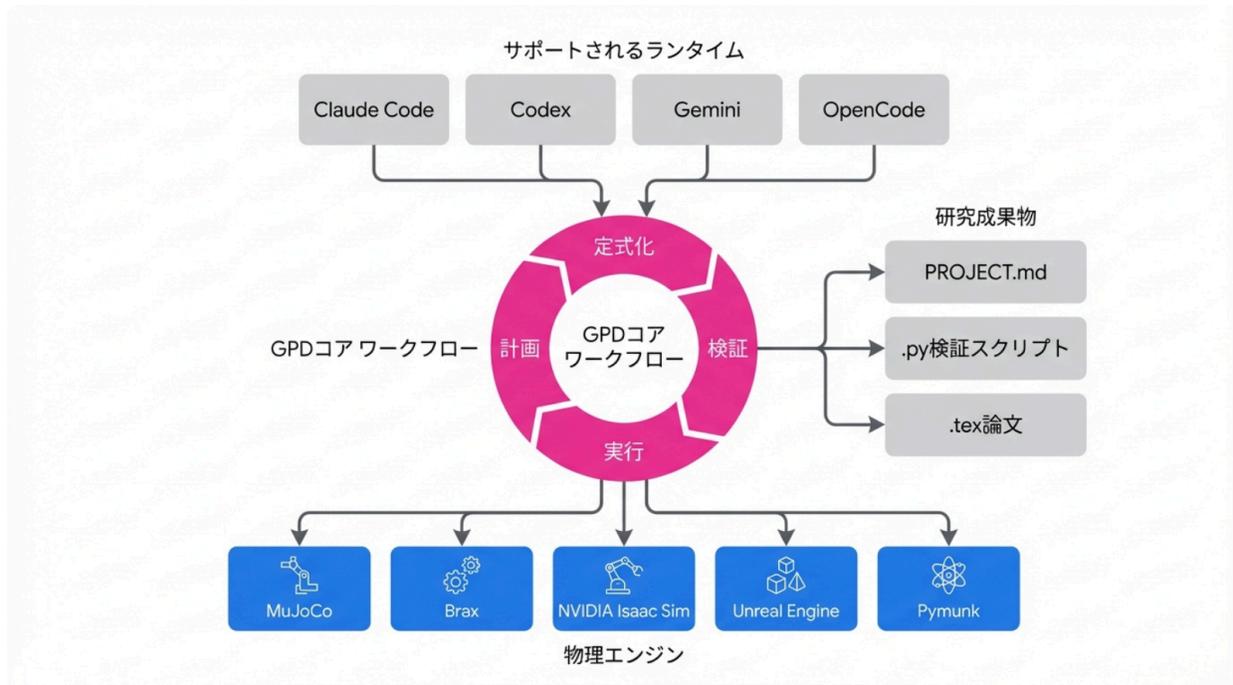
さらに、これらの分析結果をまとめるための「仮説ブランチ機能 (/gpd:branch-hypothesis)」により、複数の代替理論アプローチを並行して展開し、結果を比較することが可能である²⁶。最終段階では、論文執筆コマンド (/gpd:write-paper) が起動し、6段階のピアレビュー模倣 (/gpd:peer-review) を経た上で、arXivへの投稿に最適化されたLaTeX/ZIPパッケージ (/gpd:arxiv-submission) を生成する²⁶。これは、発見から発表までのタイムラグを劇的に短縮する。

第6章：物理シミュレーションとロボティクス環境のオーケストレーション

理論の構築と検証にとどまらず、GPDは現代の高度な計算物理学エンジンやロボティクス・シミュレータと直接連携し、物理現象の再現や実験セットアップの摩擦を劇的に軽減する²³。GPDは単なるエディタ上のAIではなく、計算資源を操作する「オーケストレーター」として機能する。GPDのエコシステムが直接的に支援または連携する主要な物理エンジンは以下の通りである。

連携する物理シミュレーションエンジン	特徴とGPDによる自動化のメリット
MuJoCo (Multi-Joint dynamics with Contact)	Google DeepMindが開発する高性能な汎用物理エンジン。接触を伴う複雑な多関節システムのシミュレーションに特化 ²⁷ 。GPDは煩雑なXMLベースの環境構築やリアルタイム計算用C APIのセットアップを自動化する。
Brax	JAX上に構築された、大規模な剛体シミュレーションのための高速かつ完全に微分可能な(fully differentiable)物理エンジン ²⁷ 。GPDは強化学習やロボティクスのための勾配ベースの最適化ループ構築を支援する。
NVIDIA Isaac Sim	NVIDIA Omniverse上に構築された高忠実度のロボティクス・シミュレーション環境。RTXレイトレーシングによるリアルな物理環境やマルチセンサー(Lidar等)の統合をGPDがサポートする ²⁷ 。
Unreal Engine	高度な3D描画エンジン。AIエージェントが仮想環境を構築し、デジタル入力に対する生物学的細胞の反応を探求するような最先端のニューロコンピューティング研究等においてGPDが統合的に機能する ²⁵ 。
Pymunk	Pythonicな2D剛体力学ライブラリ。ゲーム開発や初期の理論モデルの軽量な実験・デモ環境の構築をGPDが迅速に支援する ²⁷ 。

GPD (Get Physics Done) のワークフローと物理シミュレーション統合エコシステム



GPDは「定式化・計画・実行・検証」の4段階のワークフローを中核とし、LLMランタイム（Claude, Gemini等）からの入力を受けて、MuJoCo、Brax、Isaac Simなどの高度な物理シミュレーションエンジンの実験設定と実行を自律的にオーケストレーションする。

近年、Physical Intelligence (π) のような企業が汎用ロボティクス基盤モデル (VLAモデル) を用いて、物理世界におけるロボットの複雑な操作 (モラベックのパラドックスの克服) に取り組んでいるように、AIの主戦場はデジタル空間から物理空間へと移行しつつある²⁹。GPDは、こうした現実世界への応用を見据え、その前段階である高度な物理シミュレーション環境での実験を完全に自動化する重要なインフラストラクチャとして位置付けられる。

第7章：結論：形式的推論と経験的シミュレーションの統合が拓く科学の未来

「Aristotle Agent」と「Get Physics Done (GPD)」の登場は、人工知能の応用範囲が自然言語による「模倣的推論」から、形式論理と物理法則に基づく「自律的・科学的推論」へと明確に移行したことを示している。両者は、目的とアプローチこそ異なれど、科学的知見の自動生成という共通の終着点に向かっている。

数学的真理と物理学的真理は、その検証プロセスにおいて根本的に異なるアプローチを要求する。HarmonicのAristotle Agentは、純粋数学の世界において「演繹的确实性 (Deductive Certainty)」

を追求する。Lean 4のような定理証明支援系を用いた検証は、少数の基礎的な公理から論理規則のみを用いて定理を導出するプロセスであり、ここでの真理は絶対的である²。このアプローチは、理論物理学の根幹をなす数学的構造の健全性を保証する上で極めて強力である。対照的に、PSIのGPDは物理世界を対象とするため、「経験的・数値的妥当性 (Empirical and Numerical Validity)」に依存する。物理法則は数学言語で記述されるものの、その適用には次元解析や極限における振る舞い、そして実験データとの照合といった経験的な監査が不可欠である²⁶。

両者の関係性は、アリストテレス哲学における「形相 (Form)」と「質料 (Matter)」の関係になぞらえることができる。Aristotle Agentが事物の純粋な論理構造や数学的形式 (Form) を構築・検証するのに対し、GPDはそれが具体的な物理世界 (Matter) においてどのように振る舞い、エネルギーや力がどのように伝播するかをシミュレートする。理論的基盤の完全無欠な証明 (Aristotle) と、物理環境でのテストとデータ駆動型の調整 (GPD) は、次世代の科学革命の両輪として機能する。

これらエージェント型システムの普及により、研究リソースの配分は劇的に変化する。これまで研究者の多大な時間を奪ってきた、表記の統一、数値計算の収束確認、シミュレータ環境の構築、さらには論文のLaTeXフォーマット化といった周辺作業の多くが自動化されることで²⁶、人間の研究者は、AIに提示するべき「より根源的で創造的な仮説の立案」に専念することが可能となる。

さらに、「Verification Problem (検証問題)」の根本的な解決がもたらされる。これまではLLMが生成した物理モデルや数式に微小なハルシネーションが含まれているリスクがあり、その検証に多大な人的コストがかかっていた。しかし、Aristotleのような形式的証明システムが基礎となる数理モデルの健全性を証明し⁴、GPDがそのモデルに基づく物理シミュレーションと次元・極限解析を実行するというパイプラインが確立されれば、AIが生成した科学的知見に対する「検証のボトルネック」は完全に解消される。

論理的飛躍のない演繹 (Aristotle) と、現実世界を正確に模倣・検証する帰納的シミュレーション (GPD) が融合する時、人類は科学的発見そのものを「工業的スケール」で生み出す、新たな知のパラダイムを迎えることとなるのである。

引用文献

1. Harmonic Launches Aristotle App, Claiming “Hallucination-Free” AI Math Answers, 3月 22, 2026にアクセス、
<https://android.com/2025/harmonic-launches-aristotle-app-claiming-hallucination-free-ai-math-answers/>
2. Why Vlad Tenev and Tudor Achim of Harmonic Think AI Is About to Change Math—and Why It Matters - Sequoia Capital, 3月 22, 2026にアクセス、
<https://sequoiacap.com/podcast/training-data-harmonic/>
3. Solving the AI Reasoning Gap: How Harmonic is Building Mathematical Superintelligence, 3月 22, 2026にアクセス、
<https://www.indexventures.com/perspectives/solving-the-ai-reasoning-gap-how-harmonic-is-building-mathematical-superintelligence/>
4. Harmonic - News - Harmonic Fun, 3月 22, 2026にアクセス、
<https://harmonic.fun/news>

5. Harmonic Announces IMO Gold Medal-Level Performance & Launch of First Mathematical Superintelligence (MSI) AI App - Business Wire, 3月 22, 2026にアクセス、
<https://www.businesswire.com/news/home/20250728394917/en/Harmonic-Announces-IMO-Gold-Medal-Level-Performance-Launch-of-First-Mathematical-Superintelligence-MSI-AI-App>
6. Aldeal: Sentience and Ideology - IEEE Xplore, 3月 22, 2026にアクセス、
<https://ieeexplore.ieee.org/iel7/8964404/10412088/10412092.pdf>
7. Efficient Cause (Philosophy) – Study Guide | StudyGuides.com, 3月 22, 2026にアクセス、
<https://studyguides.com/study-methods/study-guide/cmjig30xsgglv01aabhwc6rvk>
8. AI ODDISSEY -Physics and Artificial Intelligence | by miguel iradier | Medium, 3月 22, 2026にアクセス、
https://medium.com/@miguel.iradier_91782/ai-oddissey-physics-and-artificial-intelligence-ad3b3718022f
9. Acting-arrangement Ontology Introduced - Durham E-Theses, 3月 22, 2026にアクセス、
<http://etheses.dur.ac.uk/15175/1/PEMBERTON000899140.pdf?DDD24+>
10. Alexander of Aphrodisias : a source of Origen's philosophy ? - OpenEdition Journals, 3月 22, 2026にアクセス、
<https://journals.openedition.org/philosant/807>
11. Aristotle: IMO-level Automated Theorem Proving - arXiv, 3月 22, 2026にアクセス、
<https://arxiv.org/html/2510.01346v1>
12. Aristotle: IMO-level Automated Theorem Proving | alphaXiv, 3月 22, 2026にアクセス、
<https://www.alphaxiv.org/overview/2510.01346v1>
13. Aristotle: IMO-Level Theorem Prover, 3月 22, 2026にアクセス、
<https://www.emergentmind.com/topics/aristotle-imo-level-automated-theorem-proving>
14. ProofBench - Vals AI, 3月 22, 2026にアクセス、
https://www.vals.ai/benchmarks/proof_bench
15. Aristotle API, 3月 22, 2026にアクセス、
<https://aristotle.harmonic.fun/>
16. Robinhood CEO Vlad Tenev Touts 'Autonomous Mathematician' As Harmonic Unveils Aristotle Agent - Sahm, 3月 22, 2026にアクセス、
<https://www.sahmcapital.com/news/content/robinhood-ceo-vlad-tenev-touts-autonomous-mathematician-as-harmonic-unveils-aristotle-agent-2026-03-18>
17. Robinhood CEO Vlad Tenev Touts 'Autonomous Mathematician' As Harmonic Unveils Aristotle Agent (UPDATED) - Benzinga, 3月 22, 2026にアクセス、
<https://www.benzinga.com/markets/tech/26/03/51317862/robinhood-ceo-vlad-tenev-touts-autonomous-mathematician-as-harmonic-unveils-aristotle-agent>
18. Aristotle from Harmonic just proved Erdos Problem #124 in Lean all by itself. This problem has been open for nearly 30 years : r/math - Reddit, 3月 22, 2026にアクセス、
https://www.reddit.com/r/math/comments/1paewg5/aristotle_from_harmonic_just_proved_erdos_problem/
19. Physical Superintelligence, 3月 22, 2026にアクセス、
<https://www.psi.inc/>
20. Physical Superintelligence - Dr Alex Wissner-Gross : r/accelerate - Reddit, 3月 22,

- 2026にアクセス、
https://www.reddit.com/r/accelerate/comments/1r5fvjk/physical_superintelligence_dr_alex_wissnergross/
21. Alex Wissner-Gross: A new equation for intelligence | TED Talk and Transcript - Reddit, 3月 22, 2026にアクセス、
https://www.reddit.com/r/ControlProblem/comments/449fz7/alex_wissnergross_a_new_equation_for_intelligence/
 22. \$14-\$129/hr Ai Training For Physics Jobs in Union City, CA - ZipRecruiter, 3月 22, 2026にアクセス、
<https://www.ziprecruiter.com/Jobs/Ai-Training-For-Physics/-in-Union-City,CA>
 23. Get Physics Done (GPD) download | SourceForge.net, 3月 22, 2026にアクセス、
<https://sourceforge.net/projects/get-physics-done-gpd.mirror/>
 24. Get Physics Done (GPD) Reviews - 2026 - SourceForge, 3月 22, 2026にアクセス、
<https://sourceforge.net/projects/get-physics-done-gpd.mirror/reviews/>
 25. AI Updates | A Research Resource - Glia.ca, 3月 22, 2026にアクセス、
<https://glia.ca/AI-updates/>
 26. psi-oss/get-physics-done: The first open-source agentic AI ... - GitHub, 3月 22, 2026にアクセス、
<https://github.com/psi-oss/get-physics-done>
 27. physics free download - SourceForge, 3月 22, 2026にアクセス、
<https://sourceforge.net/directory/?q=physics>
 28. Best Open Source Python Physics Software - SourceForge, 3月 22, 2026にアクセス、
<https://sourceforge.net/directory/physics/python/>
 29. Physical Intelligence (π), 3月 22, 2026にアクセス、
<https://www.pi.website/>
 30. Robotics: why now? - Quan Vuong and Jost Tobias Springberg, Physical Intelligence, 3月 22, 2026にアクセス、
<https://www.youtube.com/watch?v=cGLa8DsOYdk>
 31. Inside the \$5.6B Startup Building Robot Brains (Physical Intelligence) - YouTube, 3月 22, 2026にアクセス、
<https://www.youtube.com/watch?v=b8BDUa-xbyA>