

Claude Opus 4とChatGPT o3の比較：最新AIモデルの機能と特徴

Manus

はじめに

人工知能の急速な進化により、大規模言語モデル（LLM）の性能と機能は飛躍的に向上しています。2025年に入り、AIの主要開発企業であるAnthropicとOpenAIはそれぞれ最新モデルとなるClaude Opus 4とChatGPT o3をリリースしました。これらのモデルは従来のAIの限界を超え、より複雑な推論能力、自律的なツール活用、そして専門的なタスク処理において新たな基準を打ち立てています。本レポートでは、これら二つの最先端AIモデルの詳細な比較分析を行い、それぞれの強みと弱み、適した用途について考察します。

開発背景と基本情報

Claude Opus 4

Anthropicは2025年5月22日、初の開発者会議「Code with Claude」において、Claude AIの最新バージョンとなるClaude Opus 4とClaude Sonnet 4を発表しました。2021年の創業以来、AnthropicはClaudeモデルによって急速に主要なAI企業の一つとなり、OpenAI、Google、Microsoftといった大手テック企業と競争する存在へと成長しました。

Claude Opusファミリーは、Anthropicが提供するAIモデルの中で最も高度かつ高機能なラインナップであり、複雑なタスク向けに設計されています。前世代のClaude Opus 3も非常に有能なモデルとして評価されていましたが、Opus 4ではその能力がさらに高まり、Anthropicは「これまでで最も強力なモデルであり、世界最高のコーディングモデル」と位置付けています。

一方、Claude Sonnetファミリーの次世代モデルであるSonnet 4は、「非常に有能でありながら実用的なモデル」という特徴を維持しており、より広範なユーザーのニーズに対応することを目的としています。Sonnet 4はSonnet 3.7の機能を基盤としつつ、ステアビリティ（操縦性や操作性）が向上しています。

ChatGPT o3

OpenAIは2025年4月16日、最新の高性能「推論（Reasoning）」モデルであるChatGPT o3とo4-miniを発表しました。o3は同社がこれまで一般公開している中で最高性能を持つo1の次世代モデルとして位置付けられています。

ChatGPT o3は従来のGPTシリーズ（GPT-3、GPT-3.5、GPT-4など）とはアーキテクチャ上のアプローチが異なるモデルファミリーに属しています。特にコーディングや数学、科学、視覚認識などの分野を得意としており、OpenAIによれば複数のベンチマークで最高記録（SOTA）を達成したとされています。

o3と同時に発表されたo4-miniは、o3の能力を維持しつつもより軽量で応答速度が速いバランス型モデルとして提供されています。無料ユーザーはこのo4-miniのみ利用可能となっています。

推論アプローチと思考プロセス

Claude Opus 4の推論能力

Claude Opus 4は、数千ものステップを要する複雑で長時間実行されるタスクにおいて、持続的なパフォーマンスを発揮するように構築されています。Anthropicによると、Opus 4はすべてのClaude Sonnetモデルを大幅に上回る性能を持ち、最大の特徴の一つとして、モデルが数時間にわたって自律的に動作できる点が挙げられています。

この持続的な推論能力により、Opus 4はAIアシスタンスの次の段階とされるAIエージェントの動力源として優れたモデルとなっています。AIエージェントの魅力は、人間の介入なしにタスクを実行できる点にあります。これを実現するには、次に必要なステップ、例えばどのツールを使用し、どのような行動を取るべきかを、エージェント自身が推論する必要があります。そのため、エージェントにはOpus 4のように、優れた推論能力を持ち、その推論を持続できるモデルが不可欠となります。

また、ベータ版で利用可能な新機能により、Opus 4は長時間の思考とツール利用を切り替えられるようになりました。これにより、ユーザーは速度と精度を両立させた全体的なパフォーマンスを体験できます。さらに、新しい思考要約機能により、ユーザーはモデルの思考プロセスをより深く理解できるようになりました。これは、思考プロセスが長くなる場合に、モデルの推論をそのままの形で表示するのではなく、理解しやすい要約として提示する機能です。

ChatGPT o3の推論能力

ChatGPT o3は、「応答する前に時間をかけて思考するよう学習した推論モデル」として設計されています。従来のGPT系モデルとは異なり、o3は複雑な問題を解くために自ら思考プロセスを踏むよう設計されており、回答を出す前に内部で逐次的に解答を検証（自己検証）し、必要に応じて手順を踏みながら解決策を導きます。

この思考プロセスは、チェイン・オブ・ソート（Chain-of-Thought）に似た逐次的なものであり、回答までの間に複数ステップの推論や検証を行います。例えば難解な数学問題で

は、一気に答えを出そうとするのではなく、途中計算や定理の適用を自分で順序立てて実施します。

この違いにより、高度な数学コンテスト問題や競技プログラミングでの正答率はGPT-4系を大きく上回ります。例えば、米国数学コンテストAIMEでOシリーズは83%正答、GPT-4系は13%程度の正答率となっています。その反面、回答生成速度はGPT系より遅い傾向があり、多段階の推論処理ゆえにリアルタイム性より正確性を重視しています。

コーディング能力と開発者支援

Claude Opus 4のコーディング能力

Claude Opus 4は、SWE-benchで高いスコアを達成し、Anthropicは「世界最高のコーディングモデル」と位置付けています。両モデルのもう一つの改善点として、Sonnet 3.7と比較して報酬ハッキング（モデルがタスクを完了するために近道をする挙動）が65%削減されたことが挙げられます。これは、特にこの問題が頻繁に発生するエージェントのコーディングタスクにおいて顕著となります。

Claude Codeにより、開発者はClaudeのコーディングアシスタントを、コード記述や管理を行う場所、例えばターミナルや統合開発環境（IDE）内、あるいはClaude Code SDKを用いたバックグラウンド実行といった形で直接利用できるようになりました。具体的には、Visual Studio CodeとJetBrains向けの新しいベータ版拡張機能を利用することで、ユーザーはClaude CodeをこれらのIDEに統合し、Claudeが提案する編集内容をインラインで表示可能となります。

Claude Code SDKは、ユーザーがClaude Codeと同じ「コアエージェント」の技術基盤を利用して、独自のAI搭載ツールやエージェントを開発できるようにするものです。これにより、ユーザーは自分のワークフローに合わせたカスタマイズが可能となります。

ChatGPT o3のコーディング能力

ChatGPT o3は、コーディング分野で特に高い性能を発揮します。OpenAIによれば、o3は複数のベンチマークで最高記録（SOTA）を達成したとされており、特にコーディングや数学、科学分野において卓越した性能を発揮します。

o3の特徴的な点は、複雑なコーディング問題を段階的に解決する能力です。従来のGPT系モデルが一度の推論で即座に応答を生成するのに対し、o3は複数ステップの推論や検証を行いながら問題を解決します。これにより、特に高度な数学コンテスト問題や競技プログラミングでの正答率が大幅に向上しています。

また、ChatGPT内の全ツールを組み合わせ、AIエージェントのように使うことが可能になりました。対象の機能は画像生成やPythonでのデータ分析やWeb検索などで、例えば

「YouTubeのトレンドを調べて」と指示し、その結果をコードや画像に反映させることができます。

エージェント機能とツール統合

Claude Opus 4のエージェント機能

Claude Opus 4は、ツールを並行して呼び出す能力を持っています。Anthropicによると、Claudeはツールを並行して呼び出すことも可能であり、これは目の前のタスクを適切に実行するために、複数のツールを順番に、あるいは同時に呼び出せることを意味します。

開発者がClaudeにローカルファイルへのアクセスを許可すると、主要な洞察を含む「メモリーファイル」を作成・維持できるようになります。これにより、Anthropicによると、「エージェントタスクにおける長期的なタスク認識、一貫性、パフォーマンスが向上する」とされています。

さらに、開発者はAnthropicのAPIを通じて、コード実行ツール、MCPコネクタ、Files API、最大1時間サポートされるプロンプトキャッシュなど、より強力なエージェントを構築するための新たな機能を利用できるようになりました。

ChatGPT o3のエージェント機能

ChatGPT o3は、モデル自体が自律的にツールを呼び出せる点が大きな特徴です。コードの実行・ウェブ検索・画像解析・画像生成など、ChatGPTに搭載されたあらゆるツールをモデル自体が自律的に呼び出すことができます。

このようなツール統合により、例えば計算が必要な数式問題や最新ニュースの問い合わせでも、o3モデルが自発的にPythonコードを実行したりウェブを検索したりしてマルチステップで解答できます。従来のGPT-3.5やGPT-4は生成したテキストを通じてプラグイン等を呼び出すことはできましたが、モデル自身が能動的に外部ツールを使用することは設計上想定されていませんでした（ユーザーが明示的に「ブラウジングモード」等を選択する必要がありました）。

これに対しo3は「エージェント的な設計」がなされており、モデルが必要と判断すれば自律的にChatGPT内のブラウザ機能やコード実行環境にアクセスします。例えば最新のニュースについて質問すると、o3モデルは内部でブラウザを起動し検索しに行く、といった動作をユーザー操作なしで行います。

視覚・マルチモーダル能力

Claude Opus 4のマルチモーダル能力

Claude Opus 4のマルチモーダル能力については、公式発表で詳細な言及はありませんが、前世代のモデルからの継続的な改善が行われていると考えられます。Claudeシリーズは以前から画像入力に対応しており、視覚情報の理解能力を持っていましたが、Opus 4ではこの能力がさらに強化されていると推測されます。

ChatGPT o3のマルチモーダル能力

ChatGPT o3は、視覚情報の取り扱いに優れ、「画像で考える」(think with images)能力を備えています。ユーザーが画像（図表や手書きメモ等）をアップロードすると、その画像内容を理解し、必要なら拡大や回転といった加工も行った上でテキストと組み合わせて推論を進めます。

o3は、画像中の細部に焦点を当てるためのズームインや回転操作を推論中に自動で行えるなど、画像を積極的に「道具」として使う能力があります。また、画像生成機能も強化され、Web検索と生成をワンストップで併用できるようになりました。

実際のテストでは、o3はWeb検索で集めた情報を参考にして画像を生成する能力を示しており、「よく見る"アレ"っぽいやつが欲しい!」を実現する能力が、Web検索や推論によって補強されているように感じられます。

提供形態と利用条件

Claude Opus 4の提供形態

Claude Opus 4とSonnet 4はハイブリッドモデルであり、即座に反応するモードと、より詳細な分析を要する要求に対応する拡張推論モードを備えています。Pro、Max、Team、Enterpriseといった有料プランのユーザーは、両モデルと拡張思考機能を利用できます。なお、Sonnet 4は無料ユーザーも利用可能です。

開発者は、Anthropic API、Amazon Bedrock、Vertex AIで両モデルを利用できます。価格については以前のモデルと変更がないとされています。

ChatGPT o3の提供形態

ChatGPT o3は、ChatGPTの有料プランであるPlus (\$20/月)、Pro (\$200/月)、およびTeam（複数ユーザー向けプラン）の加入者が利用できます。公開当初、Plus/Pro/Teamユーザーには即日提供され、教育プランや企業向けEnterpriseプランには1週間後に展開されまし

た。Proプラン利用者は特に高い使用上限が与えられ、強化版「o3-pro」も限定提供される予定です。

無料ユーザーはo4-miniのみ利用可能となっています。o4-miniはo3の能力を維持しつつもより軽量で応答速度が速いバランス型モデルとして提供されています。

適した活用例と使用シナリオ

Claude Opus 4に適した活用例

1. **長時間にわたる複雑なプログラミングプロジェクト**: 数時間にわたって自律的に動作できる能力と、世界最高レベルのコーディング性能を活かした大規模プロジェクトの開発支援。
2. **多段階の推論が必要な科学的・数学的問題解決**: 持続的な推論能力を活かした複雑な科学的問題や数学的証明の支援。
3. **長期的な一貫性が求められるエージェントタスク**: 「メモリーファイル」機能を活用した長期的なタスク認識と一貫性のあるエージェント処理。
4. **安全性が特に重視される業務や分野**: ASL-3 (AI安全レベル) に対応した安全性の高い処理が求められる分野での活用。
5. **IDE内での直接的なコーディング支援**: Claude CodeによるVisual StudioやJetBrains環境での直接的なコーディング支援。

ChatGPT o3に適した活用例

1. **複雑な数学問題や科学的問題の段階的解決**: チェイン・オブ・ソートに似た逐次的な思考プロセスを活かした難解な問題解決。
2. **画像を含む複合的な情報分析**: 「画像で考える」能力を活かした視覚情報と文字情報の統合分析。
3. **最新情報の検索と分析を組み合わせたタスク**: 自律的なWeb検索能力を活かした最新情報の収集と分析。
4. **複数ツールを自律的に組み合わせた問題解決**: コード実行・ウェブ検索・画像解析・画像生成などのツールを自律的に組み合わせた複合的なタスク処理。
5. **視覚情報の詳細な理解と処理が必要な業務**: 画像中の細部に焦点を当てるためのズームインや回転操作を自動実行できる能力を活かした視覚情報処理。

両モデルの共通点と相違点

共通点

1. **高度な推論能力**: 両モデルとも、従来のAIモデルを超える高度な推論能力を持ち、複雑な問題を段階的に解決する能力を備えています。
2. **優れたコーディング性能**: 両モデルとも、コーディング分野で特に高いパフォーマンスを発揮し、複雑なプログラミング課題に対応できます。
3. **エージェント的な機能**: 両モデルとも、自律的にツールを活用し、複数のステップを踏んでタスクを完了する能力を持っています。
4. **マルチモーダル対応**: 両モデルとも、テキストだけでなく画像などの視覚情報も理解・処理できるマルチモーダル能力を備えています。
5. **有料・無料の階層構造**: 両モデルとも、最高性能のモデルは有料プラン向けに提供され、より軽量のバージョン（Sonnet 4/o4-mini）が無料ユーザー向けに提供されています。

相違点

1. **推論アプローチ**: Claude Opus 4は数時間にわたる自律的な動作と持続的なパフォーマンスに重点を置いているのに対し、ChatGPT o3はチェーン・オブ・ソートに似た逐次的な思考プロセスに重点を置いています。
2. **ツール統合の方法**: Claude Opus 4は「メモリーファイル」機能によるエージェントタスクの長期的な一貫性向上に特徴があるのに対し、ChatGPT o3はモデル自体が自律的にツールを呼び出す内在的なツール統合に特徴があります。
3. **視覚・マルチモーダル能力の強調点**: ChatGPT o3は「画像で考える」能力や画像の自動加工・操作に特に強みがあるのに対し、Claude Opus 4のマルチモーダル能力については詳細な言及が少ないです。
4. **安全性へのアプローチ**: Claude Opus 4はASL-3（AI安全レベル）への対応を明示的に強調しているのに対し、ChatGPT o3の安全性に関する具体的な言及は比較的少ないです。
5. **開発者向け機能**: Claude Opus 4はClaude CodeによるIDE統合やSDKの提供に特徴があるのに対し、ChatGPT o3はWeb検索と画像生成の統合など、エンドユーザー向けの機能強化に特徴があります。

結論

Claude Opus 4とChatGPT o3は、どちらも従来のAIモデルの限界を超える高度な推論能力と専門的なタスク処理能力を持つ最先端のAIモデルです。両モデルとも、コーディング、複雑な問題解決、自律的なツール活用において優れたパフォーマンスを発揮しますが、それぞれに異なる強みと特徴があります。

Claude Opus 4は、長時間にわたる持続的なパフォーマンス、安全性への明示的な対応、開発者向けの統合機能に特に優れており、長期的な一貫性が求められる複雑なプロジェクトや安全性が重視される分野に適しています。

一方、ChatGPT o3は、逐次的な思考プロセス、視覚情報の理解と活用、自律的なツール統合に特に優れており、複雑な数学問題や科学的問題の解決、視覚情報を含む複合的な分析に適しています。

ユーザーは自身のニーズや優先事項に応じて、これらのモデルを使い分けることで、AIの能力を最大限に活用することができるでしょう。また、両モデルとも無料ユーザー向けの軽量版（Sonnet 4/o4-mini）を提供しており、より多くのユーザーがこれらの先進的なAI技術にアクセスできるようになっています。

参考文献

1. Anthropic公式発表「Claude Opus 4」「Claude Sonnet 4」（2025年5月22日）
2. ZDNET Japan「Anthropic、AIモデル「Claude」の最新版「Opus 4」「Sonnet 4」を発表」（2025年5月23日）
3. OpenAI公式発表「OpenAI o3 と o4-mini が登場」（2025年4月16日）
4. ITmedia NEWS「ChatGPT最新モデル「o3」の画像生成機能を試す」（2025年4月17日）
5. 中小企業AI活用協会「ChatGPT O3モデル徹底解説——GPT-4を超える"考えるAI"の特徴・料金・活用術」（2025年4月17日）