

人工汎用知能の世界

もしある日、私たちが AI に「TikTok や YouTube のようなアプリを作って」と一言命令するだけで、何百万行ものコードを含む完全なプロジェクトを提供してくれるようになったら、それは本当に素晴らしいことでしょう。

2022 年 11 月末に ChatGPT がリリースされた時、それはコードスニペットの修正や幅広い質問に答えるといったタスクを実行することができました。ニューラルネットワークの概念を説明したり、様々なプログラミング言語での「Hello, World!」の例を提供することに優れていました。それは、直接的な答えを提供できる非常に強力な検索エンジンのようなものでした。

そして、検索機能も備えていました。ChatGPT に指示して、ウェブサイトからリンクのリストを取得し、それらについてのドキュメントを作成することができました。

その後、ChatGPT 3.5、4、4o、o1-mini、o1 などのバージョンがリリースされました。

今では、ウェブサイトにはダークモードを追加するようなリクエストを処理できます。HTML、CSS、またはスクリプトを更新するための必要なコードと手順を提供し、ダークモードのトグルを追加する提案もできます。ダークモードの実装には CSS の変更が含まれ、サイトが Markdown を使用している場合はそれも更新する必要があります。

まるで AI が、単なるコードの断片だけでなく、機能全体や機能性を実装できるかのようです。

これらの機能を組み合わせることで、アプリケーションを作成することができます。ですから、ある日、AI ツールにターミナル、ブラウザ、ToDo リスト、タスクアプリ、カレンダー、コード共同作業ツール、またはミーティングアプリを作成するように指示すると、プロジェクト全体のコードを提供してくれるかもしれません。

次に、タスクをより複雑にすることができます。例えば、AI に既存の YouTube コードを統合し、OpenAI、Claude、または Deepseek の API を使用して、YouTube に AI 機能を追加するように依頼することができます。これには、スマートアシスタントの追加、現在の翻訳を AI による翻訳に置き換える、AI による検索機能の強化、さらには日本の生活に関する 100 の面白いショートビデオを YouTube に提供するようにリクエストするなど、専用のショートビデオをキュレーションすることも含まれます。

では、これはアプリです。しかし、もっと高度なタスク、例えばオペレーティングシステムの作成についてはどうでしょうか？ AI に、新しい完全にオープンソースのオペレーティングシステムを設計するように指示することができます。そのシステムは、新しいデザイン、基本的なアプリ、ターミナル、コマンドライン、そしてスケジューラを持ち、Oberon に似たもので、プロセス間の相互作用に文字列ではなくデータ構造を使用するものです。

次は何をしましょうか？ AI に最新の Mac をデザインし、そのオペレーティングシステムを更新

するよう依頼することもできます。

そして、次は何でしょうか？AIに家全体の設計と更新を指示し、私たちの活動、最新の知識、そしてニーズに基づいてすべての電気製品をカスタマイズし、より良い生活環境を作り出すことができます。

そして、次は何でしょうか？AIに、市民の行動と最新の知識に基づいて、彼らの生活を向上させるために都市全体を設計するよう依頼することもできます。

そして最後に、次は何でしょうか？私たちはAIに、地球を改善するよう指示することができます。利用可能なすべての知識と情報を使って、すべての人の生活を向上させるのです。

このエッセイのタイトルを考えるのに苦労しています。「人工汎用知能の世界」と呼ぶことにしましょう。

Deepseekによると、「人工汎用知能（AGI）とは、人間の知能に匹敵するレベルで、幅広いタスクにおいて理解、学習、知識の応用が可能な人工知能の一種を指します。」顔認識、言語翻訳、チェスのプレイなど特定のタスクのために設計された特化型AIとは異なり、AGIは人間が行うことができるあらゆる知的タスクを実行することができます。

AIの未来を考える際に、把握すべき基本的なポイントは2つあります：アルゴリズムと計算能力です。AIアルゴリズムは主に微積分、バックプロパゲーション、トランスフォーマー、GPT、そしてマルチヘッド潜在アテンションに関わっています。

デジタルの世界では、XからYへのマッピングが存在します。ここで、Xはテキスト、画像、動画、音声、コード、さらには任意のバイトデータなど、何でもあり得ます。同様に、Yもこれらのいずれかになり得ます。

コンピュータは本質的にAGI（人工汎用知能）を理解しているわけではありません。それは単に人間が作り出した定義に過ぎず、機械にとってはそれほど重要なことではありません。

AIの物理世界への応用は、自動運転やロボット工学などの分野を含むでしょう。デジタル世界がXをYにマッピングできるなら、物理世界もそれに続くでしょう。例えば、ロボットは食材を料理に変えたり、レゴを組み立てたり、家を飾り付けたり、床をタイル張りしたり、エアコンを設置したり、IKEAの家具を組み立てたりすることができます。

既に産業用ロボットは実用化されています。日本ではFANUC、川崎重工業、安川電機などが有名な企業です。

では、なぜ家庭にもっとロボットが普及していないのでしょうか？消費者向けロボットは多機能で、複数のタスクをこなせる必要があります。例えば、料理ロボットは材料をかき混ぜたり炒めたりするだけかもしれませんが、ユーザーは材料を準備し、後片付けをする必要があります。

将来、ロボットは家庭、店舗、学校、オフィス、映画館、観光地など、現在人間が働いているあらゆる場所に存在するようになるでしょう。

クラウドには世界モデルが存在し、そのサイズは非常に大きく、100 ペタバイト程度になる可能性があります。参考までに、1 ペタバイトは 1,024 テラバイト、1 テラバイトは 1,024 ギガバイトです。Llama 3 70B モデルの 1 つのバージョンのファイルサイズは 21.1 GB です。

世界のロボットは、行動を起こすためにクラウド上のこの世界モデルを参照する必要があります。ロボットが効果的にタスクを実行できる限り、100 ミリ秒や 1 秒のネットワーク遅延は許容されます。