自律移動システムの航法における基盤モデル・世界モデルの概観

～航法における Physical AI～

A Survey of Foundation Models and World Models in Navigation for Autonomous Mobile Systems
~ The Era of Physical AI ~

丹羽雄一郎　　Yuichiro Niwa

測位航法学会　個人会員
Member, Institute of Positioning, Navigation and Timing

１．まえがき

近年、自律移動システムの領域では、サイバー空間で発展した AI モデルを物理的な身体性を持つシステムへ適用する「Physical AI」への進化が加速している。特に、環境の物理法則を学習し未来を予測する「世界モデル（World Model）」や、知覚から行動制御までを一気通貫で処理する「視覚-言語-行動（VLA）モデル」の実装が急速に進展している。　昨年の本大会において、生成 AI を用いた航法技術の概念とその可能性について議論した。本稿ではその続編として、実際に提案・実装され始めた具体的な基盤モデルアーキテクチャを対象に包括的なサーベイを行い、これらの技術がいかにして AI を物理世界に接地（Grounding）させ、真の Physical AI を実現するかについて考察する。

２．Physical AI と航法の系譜

NVIDIA の CEO である Jensen Huang は、2024 年 3 月の GTC および 2025 年 1 月の CES において、「知覚 AI (Perception AI)」、「生成 AI (Generative AI)」、「AI エージェント (AI Agent)」に続く AI 進化の最終段階として「フィジカル AI (Physical AI)」を提唱した。これは、デジタル空間でのデータ処理やコンテンツ生成にとどまらず、ロボットのような物理的身体を持ち、現実世界において知覚・推論（Reason）・計画（Plan）・行動（Act）を一貫して実行できる自律型システムを指す。Physical AI は第 4 次 AI ブームにおける最新概念と捉えられがちだが、その基礎となる「身体性を持つ知能」の研究開発は 45 年以上前から行われてきた。その代表例が、カーネギーメロン大学の金出武雄教授らによる野外自動運転車「Navlab」(1988 年) である。Navlab は、カメラや LiDAR による障害物および走行可能領域の検知に加え、INS（慣性航法装置）、車輪オドメトリ、GPS を組み合わせた自己位置・姿勢推定機能を備えていた。これは、古典的な経路探索アルゴリズムと論理的 AI によるセンサーフュージョンを駆使して航法を実現した、Physical AI の先駆的な事例と言える。

３．世界モデルと航法

現代 AI 分野において「世界モデル (World Model)」という用語を広く定着させたのは、David Ha と Jürgen Schmidhuber による論文『World Models』(2018 年) である。同論文では Atari のゲーム「CarRacing」を題材に、車を模したエージェントが、ランダム生成されるトラックを高速で走破できるよう学習するタスクが扱われた。これは AI モデルがシミュレーション内で「車の航法」そのものを学習していたことを意味する。しかし、世界モデルの概念自体はこの論文が初出ではなく、ロボティクスや認知心理学の分野では第 1 次 AI ブームの頃より検討されていた。例えば、自律移動ロボット「Shakey」(1971 年) の行動計画システムにおける世界モデルは、一階述語論理の論理式集合によって記述されていた。

定義上、世界モデルとは「エージェントが環境のダイナミクス（動作原理）を内部でシミュレーションし、予測するために学習した環境の内部表現」である。航法の文脈においては、外界センサによる幾何学的モデル、GNSS や INS による物理計測情報、あるいはそれらを圧縮・抽象化した「潜在空間」の情報がこれに相当する。エージェントはこれを用いて、実世界で試行錯誤することなく、環境との相互作用や未来の状態変化を脳内（モデル内）で予測・計画することが可能となる。

４．Physical AI と世界モデルの関係性

物理的身体を持つ「Physical AI」と、環境シミュレータとしての「世界モデル」の統合は、現在の AI 研究において「身体性を持つ汎用知能（Embodied AGI）」を実現するための最重要課題とされている。
深層学習の父の一人である Yann LeCun は、テキストベースの LLM（大規模言語モデル）の限界を指摘し、Physical AI が物理法則の支配する実世界で機能するためには、世界モデルの搭載が不可欠であると強く主張している。

５．航法の基盤モデル

基盤モデル（Foundation Model）とは、大量のデータによる事前学習を経て、追加の学習なし（ゼロショット）またはわずかな調整で多様なタスクに対応できる AI モデルを指す。LLM がその代表格であるが、近年では画像・動画生成モデルや、ロボットの動作生成に特化した「ロボット基盤モデル」も構築されており、これらは Physical AI の中核技術となっている。
これらのモデルは、学習過程において暗黙的に世界モデルを獲得していると考えられている。航法タスクにおいても例外ではなく、近年の自律移動システム（Physical AI）は、大量の動作データを学習することで、人間が明示的にアルゴリズムを設計しなくとも、自律的に世界モデルを習得し、移動指示などの行動生成が可能となっている。このように、汎化性能が高く、想定外の環境変動や入力に対しても堅牢（ロバスト）に機能することが、航法における基盤モデル活用の最大の利点である。

発表資料は、https://speakerdeck.com/godel　に掲載予定