

introduction to world model

世界模型（World Models）作為人工智慧領域的核心概念，正在重新定義機器如何理解、預測和與環境互動。本報告深入探討了世界模型的完整發展歷程，從 1990 年代的理論基礎到 2025 年的商業應用，為讀者提供這一前沿領域的全面視角。

歷史起源與理論基礎

Dyna 算法：世界模型的奠基之作

世界模型的概念可以追溯到 1990 年 Richard S. Sutton 提出的 Dyna 算法。Sutton 在 GTE 實驗室工作期間，開發了這一開創性的模型化強化學習（Model-Based Reinforcement Learning, MBRL）架構。

Dyna 算法的核心創新在於將三個關鍵組件整合為統一框架：

試錯學習（Trial-and-Error Learning）：智慧體通過與真實環境互動來學習最優策略。這一過程基於強化學習的基本原則，智慧體根據獲得的獎勵信號調整其行為。

世界模型學習（World Model Learning）：系統同時學習環境動態的內部模型，該模型能夠預測在

給定狀態下執行特定動作的結果。這個預測模型成為後續規劃的基礎。

規劃機制（Planning）：利用學習到的世界模型，智慧體可以在不與真實環境互動的情況下，通過內部模擬來評估不同動作序列的潛在結果，從而制定更優的策略。

Dyna 算法的理論基礎建立在**動態規劃（Dynamic Programming）**理論之上，特別是與**時序差分學習（Temporal-Difference Learning）**和強化學習的緊密聯繫。這種整合方法使得智慧體能夠同時從真實經驗和模擬經驗中學習，大幅提高了樣本效率。

Sutton 在其理論框架中提出了****放鬆規劃（Relaxation Planning）****的概念，這是一種增量式規劃方法，可以隨時中斷和恢復。與傳統的深度搜索不同，放鬆規劃通過一系列淺層搜索（通常只有一層

深度) 來逐步改進策略, 最終達到與任意深度傳統搜索相同的效果。

早期發展與挑戰

在 Dyna 算法之後的十餘年間, 模型化強化學習領域面臨著可擴展性問題。早期的方法主要依賴表格化表示 (Tabular Representations) 或簡單的函數近似, 這限制了它們在高維狀態空間中的應用能力。儘管如此, 這一時期的研究為後續的發展奠定了重要的理論基礎。

研究者們逐漸認識到, 有效的世界模型需要能夠處理高維觀察 (如圖像)、複雜動態系統, 以及不確定性建模等挑戰。這些認識為後來的深度學習時代世界模型研究指明了方向。

深度學習時代的突破

Ha & Schmidhuber 的開創性工作

2018 年標誌著世界模型研究的重大轉折點。*David Ha 和 Jürgen Schmidhuber 發表的論文「World Models」*不僅重新定義了世界模型的概念, 更提供了在高維觀察環境中實現世界模型的具體技術路徑。

技術架構詳解

Ha 和 Schmidhuber 的世界模型框架包含三個核心組件, 這一設計深受人類認知系統的啟發:

視覺組件 (Vision Component, V): 採用變分自動編碼器 (Variational Autoencoder, VAE) 架構,

負責將高維觀察 (64×64×3 RGB 圖像) 壓縮為低維潛在表示 (32 維潛在向量 z)。VAE 的使用確保了潛在空間的連續性和可解釋性, 同時引入的隨機性有助於提高模型的泛化能力。

記憶組件 (Memory Component, M): 使用混合密度網絡-循環神經網絡 (Mixture Density Network-RNN, MDN-RNN) 來建模環境的時序動態。該組件預測下一個時間步的潛在向量分布

$p(z_{t+1} | z_t, a_t, h_t)$, 其中 h_t 是 RNN 的隱藏狀態。混合高斯分布的使用使模型能夠處理環境中的隨機事件和多模態結果。

控制器組件 (Controller Component, C): 採用極簡的線性模型設計，僅包含 867 個參數。控制器接收來自視覺組件的當前觀察編碼和記憶組件的隱藏狀態，輸出相應的動作。這種設計哲學將模型的複雜性集中在世界模型部分，而保持策略組件的簡潔性。

關鍵創新與貢獻

夢境訓練 (Dream Training): Ha 和 Schmidhuber 證明了智慧體可以完全在由其世界模型生成的「夢境環境」中進行訓練，並將學習到的策略成功遷移到真實環境中。在 VizDoom 的「Take Cover」任務中，智慧體在虛擬環境中學習躲避火球，並在真實環境中表現出色。

溫度參數調節: 通過調節 MDN-RNN 的溫度參數，研究者可以控制生成環境的不確定性水平。較高的溫度使虛擬環境更加困難和不確定，從而防止智慧體利用世界模型的缺陷。實驗顯示，在溫度為 1.15 的設定下訓練的智慧體在實際環境中達到了 1092 ± 556 的最佳成績。

進化策略優化: 控制器的訓練採用協方差矩陣適應進化策略 (CMA-ES) 而非梯度下降，這種選擇使得系統可以避免傳統強化學習中的信用分配問題，同時便於並行化實現。

PlaNet: 向潛在動態的演進

2019 年，Danijar Hafner 等人提出的 Deep Planning Network (PlaNet) 進一步推進了世界模型的發展。PlaNet 的核心創新在於引入了潛在動態模型 (Latent Dynamics Model)，該模型結合了確定性和隨機性轉換組件。

PlaNet 提出的潛在**過度射擊 (Latent Overshooting)** 目標函數是一個重要的技術創新。這個**多步驟變分推理目標**確保模型能夠準確預測多個時間步的未來狀態，這對於有效規劃至關重要。

與 Ha 和 Schmidhuber 的工作相比，PlaNet 在以下幾個方面實現了顯著改進：

更精細的抽象表示: PlaNet 將環境抽象為關鍵特徵，使得規劃過程更加高效

純模型化方法：完全依賴學習到的動態模型進行決策，不需要與真實環境的額外互動

連續控制優化：特別針對需要精確連續控制的任務進行了優化

Dreamer 系列：世界模型的成熟化

Dreamer：通過潛在想像學習行為

2020 年，Hafner 團隊發布了 Dreamer 算法，這標誌著世界模型從概念驗證向實用系統的重要轉變。

Dreamer 的核心思想是「通過潛在想像學習行為」(Learning Behaviors by Latent Imagination)。

Dreamer 整合了 PlaNet 的潛在動態建模與強化學習，使智慧體能夠在潛在空間中模擬詳細的軌跡

並基於想像的場景優化策略。與傳統方法相比，**Dreamer 避免了在真實環境中的試錯學習，而**

是依靠內部模擬來改進策略。

DreamerV2：離散世界模型的突破

2021 年的 DreamerV2 實現了兩個關鍵技術突破，使得世界模型首次在 Atari 環境中超越了

Rainbow 等頂級模型化自由算法：

離散變分自動編碼器 (Discrete VAE)：放棄了傳統 VAE 中單峰正規分布的強假

設，改用 OneHotCategoricalDistribution。這種設計能夠更好地表示系統的隨機轉換，**特別適用**

於具有離散狀態轉換的環境。

KL 平衡 (KL Balancing)：這一技術創新解決了變分學習中的訓練不穩定問題，確保了模型訓練的收斂性。

DreamerV3：多樣化領域的掌控

2023 年發布的 DreamerV3 代表了世界模型發展的重要里程碑，實現了以下突破性成就：

統一超參數框架：DreamerV3 能夠在不同環境域中使用相同的超參數配置實現

異性能，這證明了其作為通用算法的潛力。實驗涵蓋了從連續控制到離散動作、從低維輸入到高

維圖像的各種環境。

Minecraft 鑽石收集：DreamerV3 成為首個能夠在 Minecraft 中從零開始收集鑽石

的算法，這一成就展示了世界模型在複雜、開放式環境中的能力。

Symlog 預測：引入 Symlog 函數來處理不同規模的獎勵和價值函數，解決了訓練中的數值不穩定問題。Symlog 變換公式為： $\text{symlog}(x) = \text{sign}(x)\ln(|x|+1)$ ，其逆變換為 $\text{symexp}(x) = \text{sign}(x)(\exp(|x|)-1)$

模化效益：DreamerV3 證明了模型參數規模與性能的正相關關係，從 8M 到 200M 參數的模型顯示了一致的性能提升趨勢。

商業化應用與突破

SORA：視頻生成作為世界模擬器

2024 年，OpenAI 發布的 Sora 代表了世界模型在視頻生成領域的重大突破。Sora 的技術報告明確提出了「視頻生成模型作為世界模擬器」的概念，這一觀點重新定義了世界模型的應用範疇。

技術架構創新：Sora 採用 Transformer 架構操作視頻和圖像的時空補

丁，能夠處理可變持續時間、解析度和長寬比的視覺數據。這種統一表示方法實現了大規模生成模型的高效訓練。

新興模擬能力：Sora 展現了多項令人驚嘆的新興能力：

3D 一致性：能夠生成具有動態攝像機運動的視頻，當攝像機移動和旋轉時，人物和場景元素在三維空間中保持一致

長程連貫性和物體永續性：能夠在長視頻中維持時間一致性，即使物體被遮擋或離開畫面也能保持其特性

世界互動性：能夠模擬影響世界狀態的動作，如畫家在畫布上留下持續的筆觸

數位世界模擬：能夠模擬人工過程，包括視頻遊戲，可同時控制 Minecraft 中的玩家並高保真度渲染世界及其動態

Google DeepMind 的 Genie 系列

Genie 2 的發布標誌著互動式世界生成的重要進展。與 Sora 專注於視頻品質不同，Genie 2 引入了用戶驅動的世界操作新維度。

Genie 2 採用**自回歸潛在擴散模型架構**，在大型視頻數據集上訓練，逐幀生成視頻。其處理流程包括：自動編碼器將視頻幀壓縮到潛在空間；基於 Transformer 的自回歸模型根據先前幀和智慧體動作預測下一個潛在幀；潛在擴散過程精煉生成真實的視頻幀；解碼潛在表示為視覺幀。

2025 年推出的 Genie 3 實現了實時互動能力，同時在一致性和真實性方面相比 Genie 2 有顯著改進。這標誌著世界模型從離線生成向實時互動應用的重要轉變。

NVIDIA Cosmos：物理 AI 的世界基礎模型

NVIDIA Cosmos 代表了世界模型商業化應用的最新高度。作為專為物理 AI 開發構建的平台，Cosmos 提供了最先進的生成世界基礎模型（World Foundation Models, WFM）、高級分詞器、護欄和加速數據處理管道。

技術規模與能力：Cosmos WFM 基於 9,000 兆 tokens 和 2,000 萬小時的機器人和駕駛數據訓練，涵蓋現實世界的幾乎所有場景。這種大規模訓練使模型能夠生成物理基礎的合成數據，用於訓練和評估現有模型。

開放模型授權：Cosmos 採用開放模型授權方式，使 5 個人和小型初創企業也能夠利用高級物理 AI 技術。這一策略有望加速整個生態系統的發展。

產業合作夥伴：包括 1X、Agile Robots、Agility、Figure AI、Foretellix、Uber、Waabi、XPENG 等在內的全球物理 AI 領導企業已成為 Cosmos 的首批採用者。

Meta 的導航世界模型

Meta 的導航世界模型 (Navigation World Model, NWM) 專注於視覺運動能力智慧體的導航任務。 NWM 採用條件擴散 Transformer (Conditional Diffusion Transformer, CDiT) 架構，在人類和機器人 智慧體的自我中心視頻多樣化集合上訓練，規模達到 10 億參數。

NWM 的核心優勢在於其動態約束整合能力。 與具有固定行為的監督導航策略不同， NWM 能夠在規劃過程中動態整合約束條件，這使其成為下一代導航系統的靈活而強大的工具。

在未知環境中， NWM 利用其學習到的視覺先 驗從單一輸入圖像想像軌跡，展示了強大的泛 化能力。實驗證明其在從零開始規劃軌跡或對 外部策略採樣的軌跡進行排名方面都表現出色。

當前挑戰與限制

技術挑戰

不確定性建模與部分可觀察性：當前的世界模型在處理不確定性和部分可觀察環境方面仍面臨重大挑戰。 如 Reddit 討論中提到，建模不確定性和部分可觀察性是當前面臨的主要技術難題之一。

可遷移表示學習：創建能夠跨不同任務和環境遷移的表示是另一個關鍵挑戰。Ha 和 Schmidhuber 在原始論文中已經指出， VAE 可能編碼對任務不相關的觀察部分，這限制了模型在新任務上的重用能力。

真實性與抽象性的平衡：如何在內部模擬中實現真實性與抽象性的適當平衡是一個持續的研究問題。過於真實的模型可能導致計算複雜性過高， 而過於抽象的模型則可能遺漏關鍵資訊。

理解與解釋性問題

缺乏真正理解：如 VE3 分析指出，大型語言模型在世界建模方面面臨的一個根本問題是缺乏真正的

理解。它們基於數據模式而非概念的真正理解來生成答案，這可能導致錯誤，特別是在需要深度理解或推理的場景中。

黑盒性質：隨著模型變得越來越複雜，世界模型的可解釋性成為一個重要問題。如 Verity AI 指出，未來的 AI 世界模型可能變得無法解釋，因為其決策過程將類似於人類直覺，而我們也無法完全解釋人類直覺。

靜態知識基礎：當前的世界模型主要依賴於訓練數據，無法像人類那樣持續學習和適應。這種限制使得模型難以構建強健且無偏見的世界模型，特別是在動態或不熟悉的環境中。

可擴展性與實用性限制

計算資源需求：現代世界模型，特別是像 NVIDIA Cosmos 這樣的大規模模型，需要巨大的計算資源進行訓練和部署。這限制了其在資源受限環境中的應用。

控制環境外的脆弱性：儘管世界模型在受控環境中表現出色，但在開放式、不可預測的真實世界環境中的表現仍有待驗證。

資料依賴性：世界模型的品質嚴重依賴於訓練資料的品質和多樣性。如果訓練資料缺乏多樣性或存在偏見，模型的世界理解將受到相應限制。

未來發展方向

邁向人工通用智慧

世界模型作為 AGI 的基石：Yann LeCun 等研究者認為，世界模型是實現人工通用智慧的關鍵組件。LeCun 提出了包括自監督學習、世界模型和認知架構在內的三大關鍵技術，這些技術的整合被視為 AGI 發展的必要條件。

體現智慧的發展：如體現 AGI 綜述所指出， 體現智慧（Embodied AI）與 世界模型的結合正在推動 AI 系統向更 類人的互動能力發展。 體現 AGI 被定 義為能夠在真實世界任務中表現出人 類水準能力的體現 AI 形式。

技術融合與創新

多模態整合：未來的世界模型將更好地整合視覺、聽覺、觸覺等多種感知模式，形成更全面的世界表示。NVIDIA Cosmos 已經在這個方向上邁出了重要步伐。

因果推理整合：如 TuringPost 分析指出， 世界模型未來發展的一個 重要方向是整合因果 AI。這種整合對於實現真正的 AGI至關重要。

層次化與混合表示：未來的世界模型可能採用層次化、 多層次和連續/離 散混合表示的架構，以更好地處理複 雜的真實世界場景。

應用領域擴展

自動駕駛革命：世界模型在自動駕駛領域的應用正在快速發展。 從 GAIA-1 到 DriveDreamer ，再到 MUVO 等專門針對駕駛場景的世界 模型正在推動自動駕駛技術的進步。

機器人技術突破：如 Towards AI 分析所述，世界模型正在推動機器人技術向更智慧、 更自主的方向發展。機器人將能夠預測、規劃和適應複雜環境，實現真正的自主性。

科學發現與探索：世界模型在科學發現中的應用潛力巨大。通過模擬複雜系統，研究者可以探索新的科學現象並測試理論假說。

結論與展望

世界模型的發展歷程展現了人工智慧領域從理論探索到實際應用的完整演進過程。從 1990 年 Sutton 的 Dyna 算法奠定的理論基礎，到 2018 年 Ha 和 Schmidhuber 的開創性框架，再到近年來 Dreamer 系列的技術突破和商業化應用的蓬勃發展，世界模型已經成為現代 AI 系統的核心組件。

當前的發展趨勢顯示，世界模型正在朝著更大規模、更多模態、更高智慧的方向演進。NVIDIA Cosmos、OpenAI Sora、Google Genie 等商業化產品的成功展示了世界模型的巨大應用潛力。同時，導航世界模型、遊戲世界生成等專門化應用也證明了這一技術的廣泛適用性。

然而，世界模型的發展仍面臨諸多挑戰。不確定性建模、真實世界泛化、計算效率、可解釋性等問題需要持續的研究投入。特別是在邁向人工通用智慧的路徑上，世界模型需要與其他 AI 技術深度融合，形成更加完整的智慧系統架構。

12 展望未來，世界模型有望成為實現真正智慧機器的關鍵技術。通過持續的技術創新和跨領域合作，我們正在逐步構建能夠真正理解和互動世界的 AI 系統。這不僅將推動人工智慧技術的革命性發展，更將為人類社會帶來前所未有的變革機遇。

在這個充滿挑戰與機遇的時代，世界模型研究正站在歷史的轉折點上。從學術研究到商業應用，從理論探索到實際部署，世界模型正在重新定義智慧機器與物理世界互動的方式。這一發展趨勢預示著我們正在進入一個 AI 系統不僅能夠感知世界，更能夠真正理解和預測世界的新時代。