

情報科学技術分野における戦略的
重要研究開発領域に関する検討会

2024-6-11

AIとシミュレーション

鶴岡 慶雅
（東京大学／産総研）

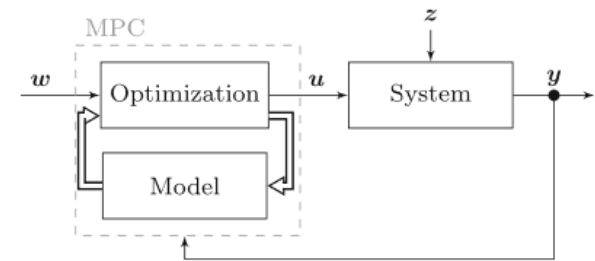
制御と強化学習

- 動的なシステムを適切に制御して望ましい状態を達成する
 - ロボット、自動車、ドローン、空調、プラント、インフラ、鉄道、etc.

- 制御理論的アプローチ

- モデル予測制御

- 有限時間先までの最適制御の繰り返し
 - システムの数理的モデルが使えることを前提



(Schwenzer et al., 2021)

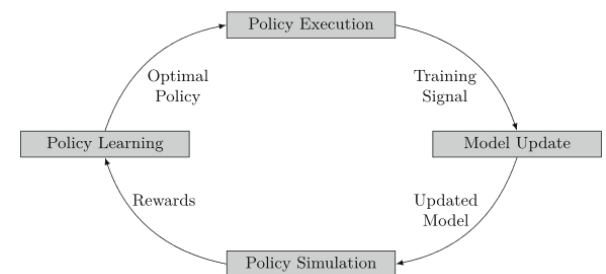
- AI的アプローチ

- 強化学習

- システムとの相互作用を繰り返して方策を学習

- モデルベース強化学習

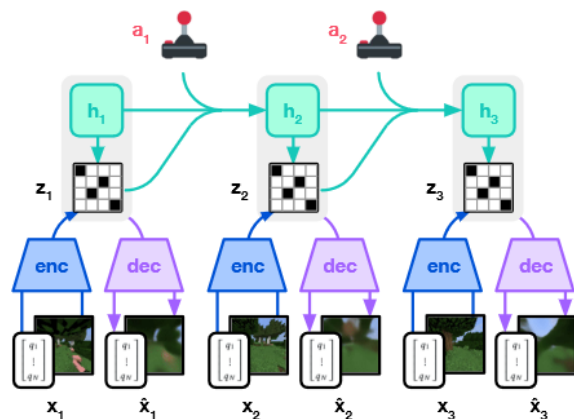
- システムのモデルを観測データから学習
 - モデルを利用して方策を学習



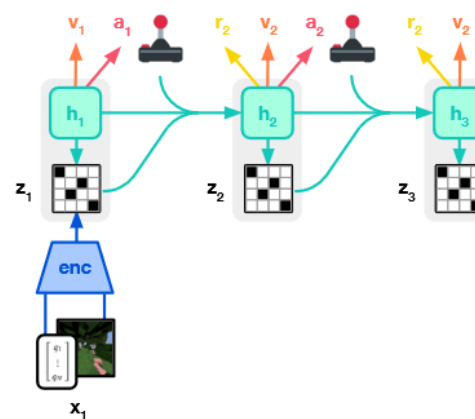
(Polydoros et al., 2017)

世界モデル

- 世界モデル(world model)の例
 - 状態遷移確率 $P(s'|s, a)$
 - 状態 s で行動 a をとったときに状態 s' に遷移する確率
- DreamerV3 (Hafner et al., 2023)
 - 世界モデルを利用したモデルベース強化学習



(a) World Model Learning



(b) Actor Critic Learning

生成モデルに基づく世界モデル

- GAIA-1: 自動運転のための生成的世界モデル



<https://wayve.ai/thinking/introducing-gaia1/>

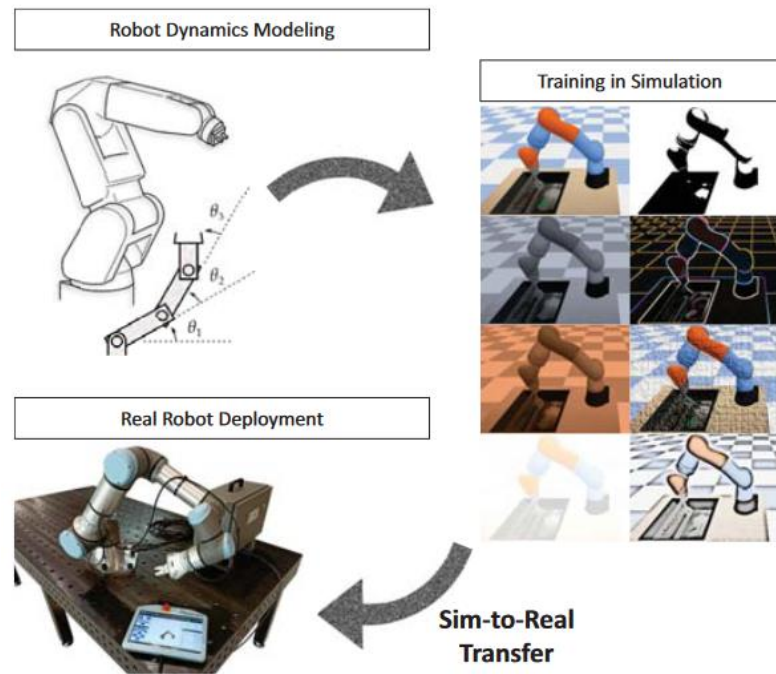
Hu et al., GAIA-1: A Generative World Model for Autonomous Driving, arXiv:2309.17080, 2023

シミュレーション

- システムに関する知識を利用して動的な特性を明示的にモデリング
 - 典型的には
 - 微分方程式(モデリング) → 数値積分(シミュレーション)
 - 内部状態も含めてモデリング
- シミュレータの例
 - ロボティクス: MuJoCo, Isaac gym, Brax, etc.
 - 自動運転: CARLA, CarMaker, CarSim, etc.
 - プラント: DWSIM, OpenModelica, Plant Simulation, etc.

Sim-to-real transfer

- シミュレータ環境で学習したエージェントを現実世界のタスクに利用
- Sim-to-real gap (reality gap)
 - 現実とシミュレーションの違い
 - それによって引き起こされる性能低下、操作の失敗
- 対策
 - ドメインランダム化
 - システム同定
 - 敵対的・ロバスト強化学習
 - 転移学習
 - 知識蒸留
 - メタ学習、メタ強化学習
 - Human-in-the-loop correction
 - etc.

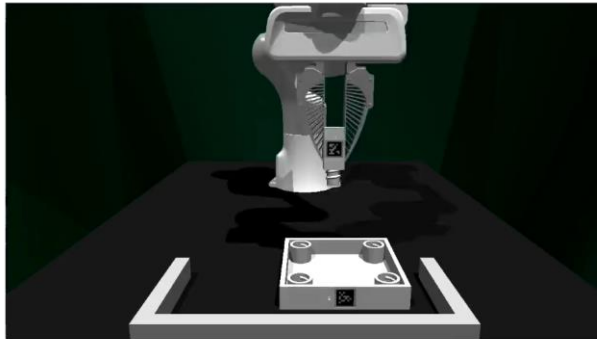


(Zhao et al., 2020)

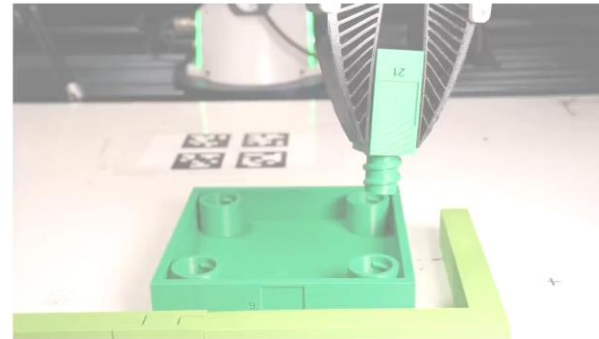
Human-in-the-loop correction

- TRANSIC (Jiang et al., 2024)
 - シミュレータで学習した方策を実世界に適用
 - 必要に応じて人間が介入して操作を訂正
 - 訂正データから残差方策を学習し、元の方策に統合

From Simulation to Reality



Simulation

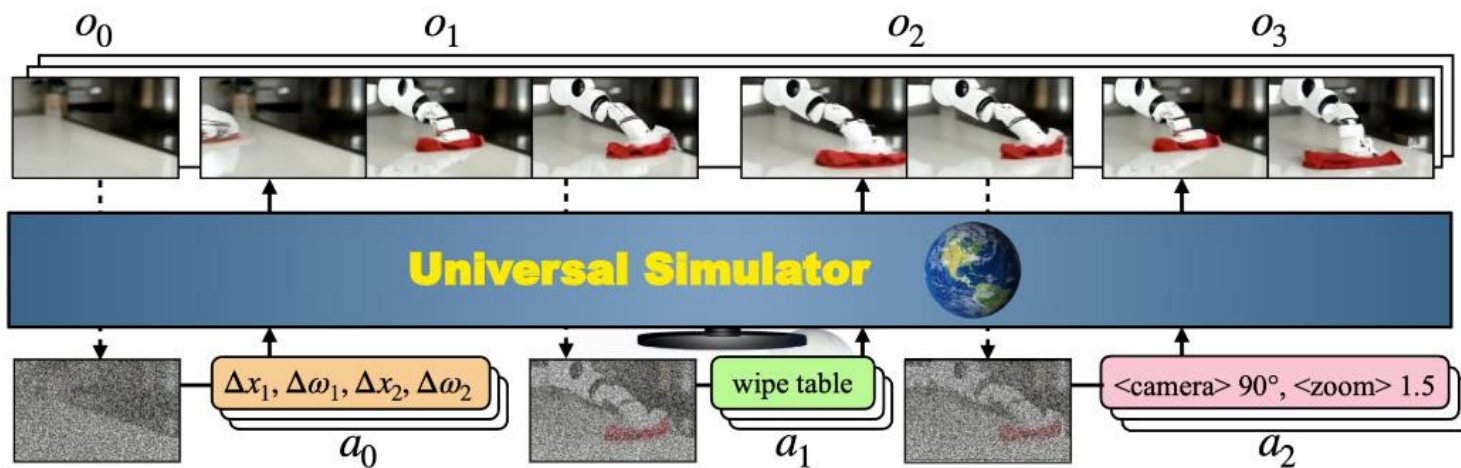


Real World

汎用実世界シミュレータの学習

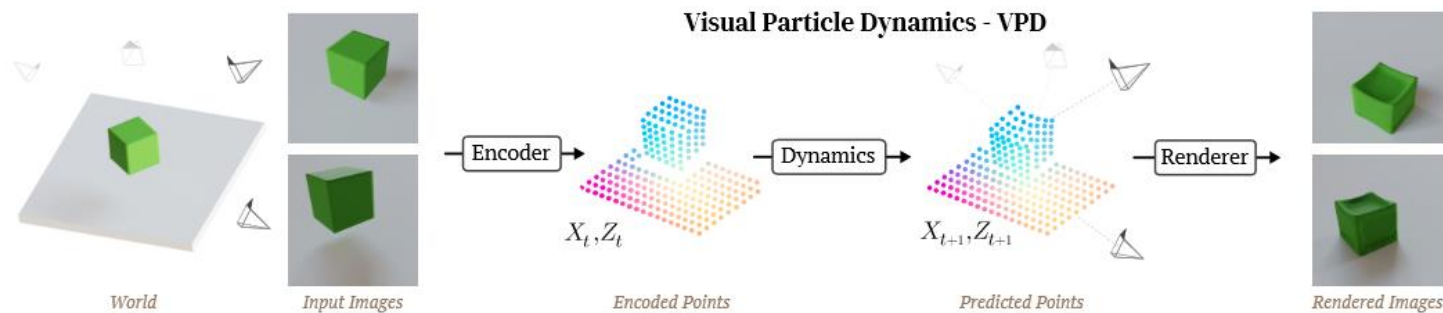
- Universal Simulator (UniSim)
 - 拡散モデルによる動画生成
 - 静止画とアクション(位置指定、言語、カメラ角度など)を入力

Action: [Wash hands](#)

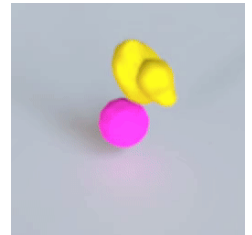


3次元物理シミュレータの学習

- 複数視点のRGB-D動画から3D点群表現による物理シミュレータを学習
 - オブジェクトの位置や形状、姿勢等の物理的な情報なしで学習



視点変換の例



今後

- Sim-to-real transfer
 - シミュレータの精度や速度の向上により、シミュレータを利用した事前学習はますます重要に
 - しかし reality gap は常に存在
 - 対策手法の研究開発
 - ドメインランダム化、システム同定、ロバスト強化学習、転移学習、知識蒸留、メタ強化学習、human-in-the-loop、etc.
- シミュレータの(半)自動構築
 - 大量の実世界データと生成モデルにより世界モデルを学習
 - 内部状態のモデリング
 - 数理的モデルとの統合