CRDS 研究開発戦略センター
Center for Research and Development Strategy

国立研究開発法人科学技術振興機構

大学見本市 2024 Innovation Japan
JST研究開発戦略センターセミナー

テクノロジー、サステナビリティ、Well-beingの交差点
～イノベーションの動向を掴む～

「生成AIとロボット研究の新たなフロンティア」

2024年8月23日

JST研究開発戦略センター（CRDS）
システム・情報科学技術ユニット
茂木 強

発表内容

1. **はじめに**
   - AI技術の進展とロボット研究
   - 生成AIとは何か
   - 生成AIとロボットの関係
   - ロボットに学ばせる
2. **ロボットの最新動向**
   - 米国における生成AIのロボット適用事例
   - 中国におけるヒューマノイド市場投入
   - 日本における発達するロボットの研究
3. **今後の展望**
   - 今後の方向性
   - 社会的インパクト

©2024 CRDS

はじめに

©2024 CRDS

AI技術の進展とロボット研究

第１段階：機械学習・深層学習の適用（2010～）

- ヒトの視覚システムのモデル化から始まり、文字認識、画像認識・音声認識など、さまざまな感覚情報からの認識・判断能力が人間を上回るレベルにまで達した。

第２段階：基盤モデル・生成AIの適用（2020～）

- 大規模言語モデルと画像・動作情報を組み合わせることで、行動計画や動作生成などの運動情報を出力できるようになり、これまで困難だったタスクの自動化の可能性が高まっている。

第3段階：ロボット基盤モデルの構築（今後）

- ロボットに特化した基盤モデルや現場環境のモデル化（世界モデル）の研究が進んでいる。近い将来、人間のような知能・身体性・対話能力を備え、共に多様なタスクを担えるロボットの実現が期待できる。

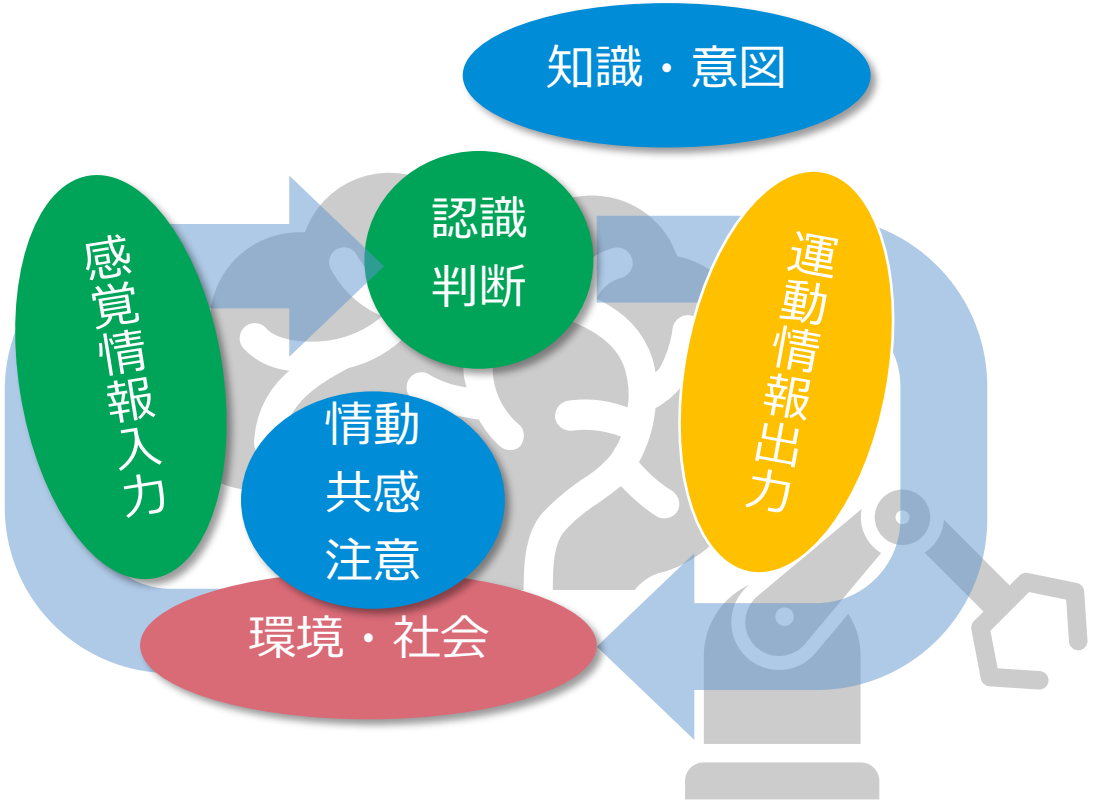


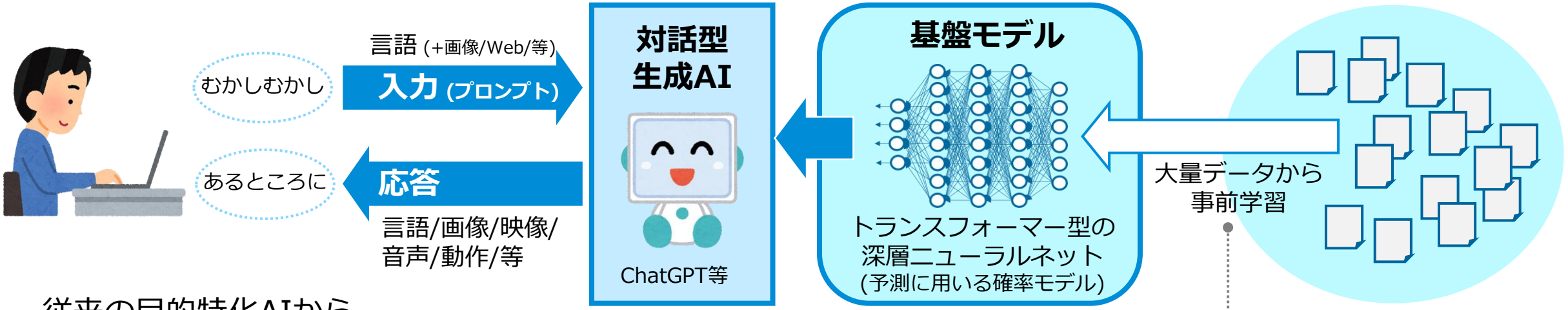
図　ロボットの構成とAI技術の適用範囲

（研究開発の俯瞰報告書 システム・情報科学技術分野 （2019年）「ロボティクスと社会」をもとに作成）

©2024 CRDS

対話型生成AIはどのようなものか

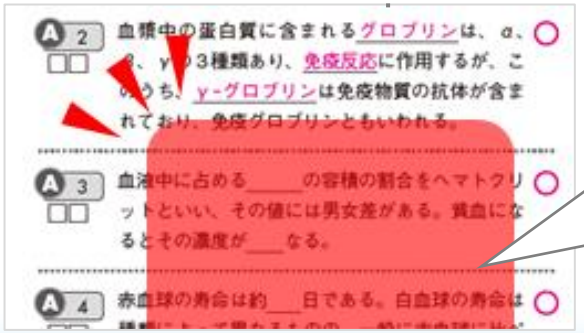
- 大量データから言葉のつながり傾向を学習した確率モデルに基づき、入力の続きを予測して応答
- 現在主流のトランスフォーマー型の深層ニューラルネットでは、幅広いコンテキスト情報を参照しつつ、予測のために注目すべき箇所を適切に選択できるようになり、予測精度が大幅に向上



従来の目的特化AIから
道具としての汎用AIへ

- まるで人間のような自然な対話
- 様々なタスクで、専門的な知識・能力を備えたような応答

世界中のテキストをもとに、自分で穴埋め問題を作って解く練習をひたすらしたら、聞かれたことに対して、たいてい続きを答えられるようになった！



事前学習：
穴埋め問題を自分で作って解く練習(自己教師あり学習)

©2024 CRDS

生成AIとロボットの関係

■ より多様でより大量のマルチモーダルデータを学習したロボット基盤モデルを構築し、実世界におけるロボットの行動計画・動作生成を柔軟化・ロバスト化する試みが急速に進展

Google

① **PaLM-SayCan：自然言語による曖昧な要求**に対して、何ができるか、ロボットが行動を選択して実行

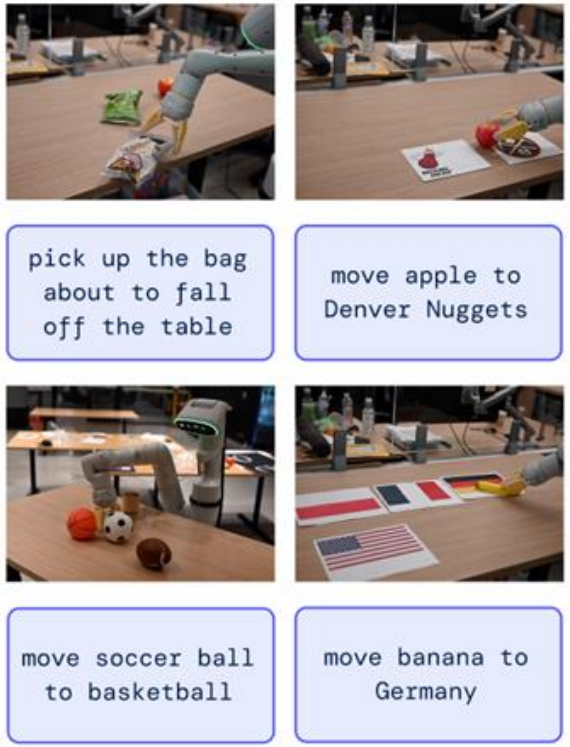
② **RT-1：実機を用いた大規模学習**によってロボット動作生成の汎用性を高めた

ロボット実機13台で17カ月、700以上のタスクをカバーする13万エピソードの動作データを収集・学習、700種類のタスクで97%の成功

③ **RT-2：Web上のテキストと画像も学習**することで、RT-1モデルで未学習だった物体も操作可能

④ **RT-X：**世界33研究機関が参加する**史上最大のオープンソースロボットデータセットプロジェクト**

RT-2において創発されたロボット動作生成例(一部)
https://doi.org/10.48550/arXiv.2307.15818



OpenAI+Figure AI

● **ChatGPT+Figure01：言語で対話、視覚や記憶から説明、計画して行動、推論して理由説明**等をデモ

「何か食べ物をもらえますか？」
「テーブルにある食べられるものはリンゴだけだったので、それをあなたに渡しました」

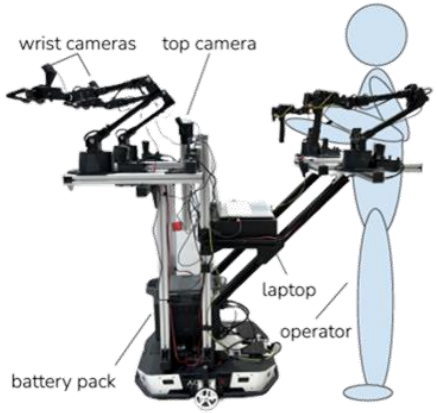
https://www.youtube.com/watch?v=Sq1QZB5baNw

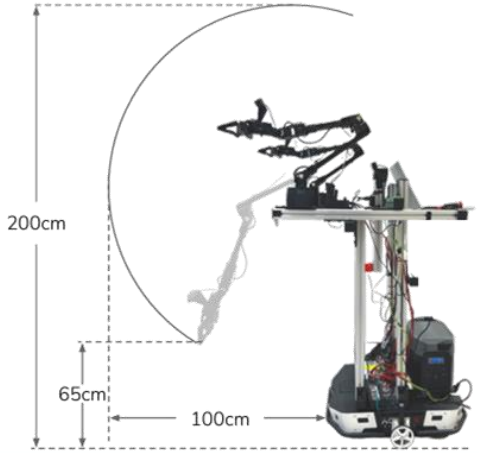
①Google+Everyday Robots (2022年8月)、② Google+Everyday Robots (2022年12月)、
③Google DeepMind (2023年7月)、④Google DeepMindほか: Open X-Embodiment発表(2023年10月)、
その後、Google DeepMindは AutoRT、SARA-RT、RT-Trajectoryも発表 (2024年1月)

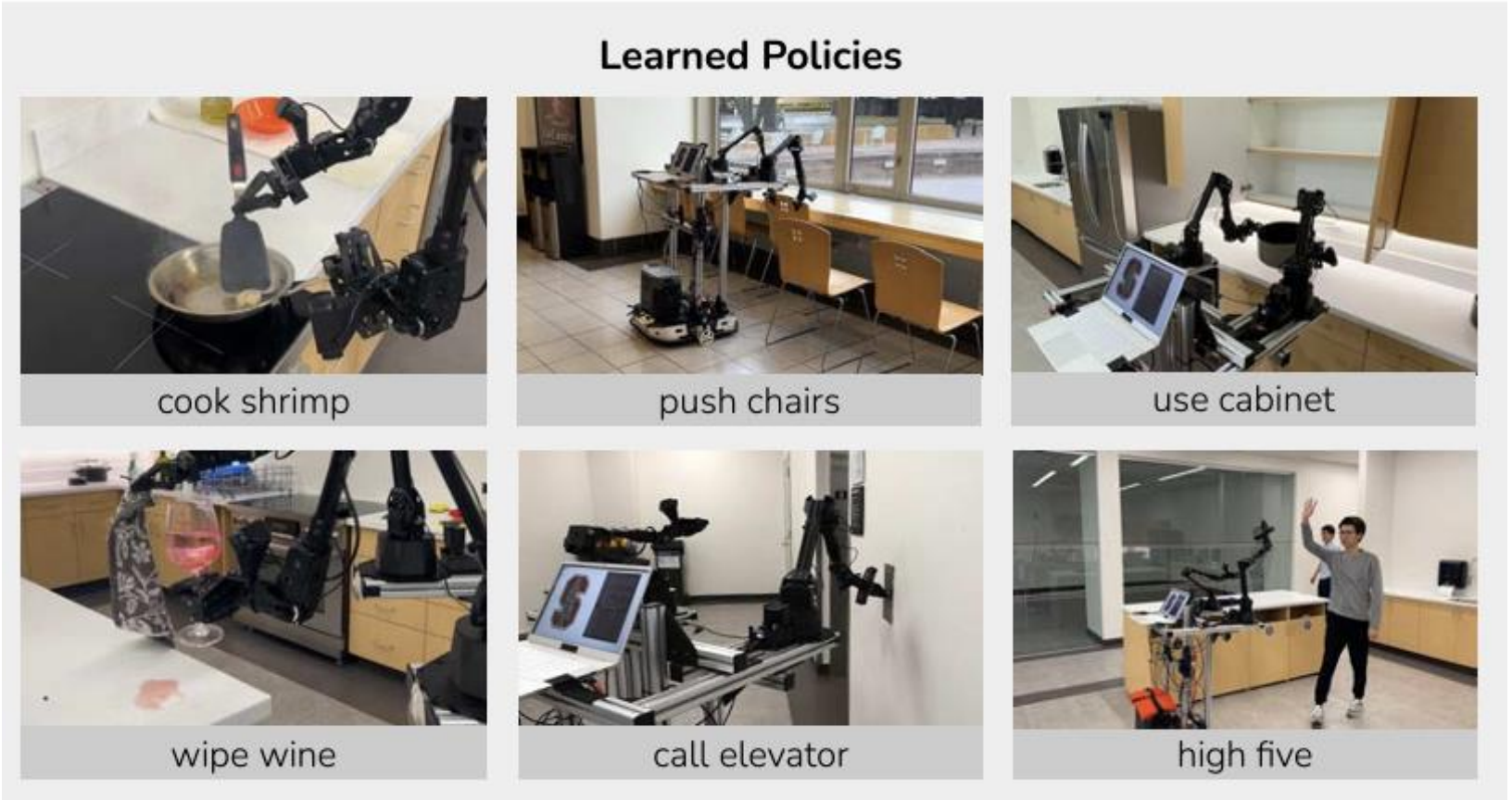
©2024 CRDS

ロボットに学ばせる Stanford大 Mobile ALOHA

- 双腕ロボットの動作を人がテレオペレーションすることで学習する模倣ロボット、比較的少ない試行で学習した後は、ロボットが自律的に実行するようになる
- 低価格化することで普及を促進し大量の学習データを集めることが可能







出典：プロジェクトのホームページより
https://mobile-aloha.github.io/

©2024 CRDS

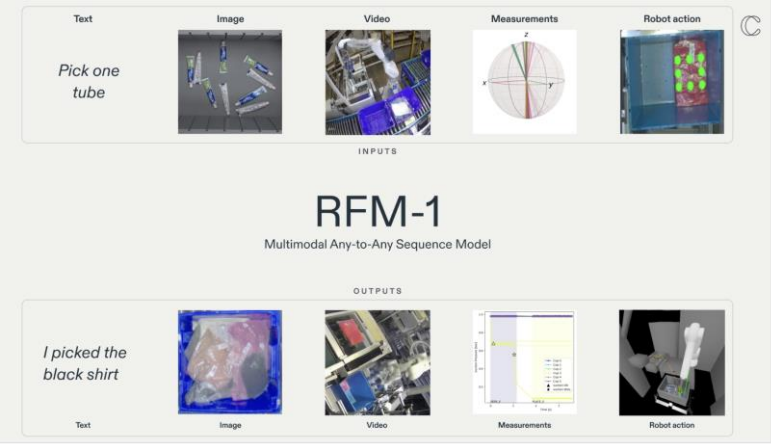
ロボットの最新動向

©2024 CRDS

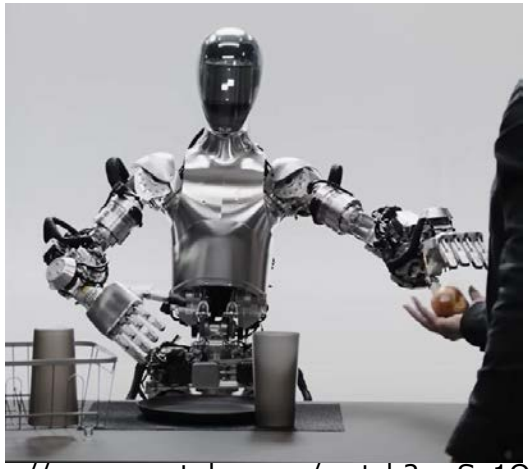
最新ロボット事例紹介（米国）

- マルチモーダルデータを学習したロボット基盤モデルによる行動計画・動作生成の実証が進む中、ロボットでは、ヒューマノイド（人型ロボット）で基本的な動作が可能になりつつある
- 実用化には、環境適応性、リアルタイム性、コスト、安全性など課題が残されている

製品名	Covariant AI社Covariant RFM-1	米国 Figure AI社 Figure 01	米Agility Robotics社 Digit
特長	ロボット基盤モデル 大量のロボット操作データを学習し、人間のような推論能力をもつ	自律汎用二足ロボット 5本指による把持、荷重運搬 人間との自然なインタラクション	二足歩行型の自律ロボット 階段昇降や段差乗越、物の運搬 3本指のグリッパー
応用分野	物流倉庫や製造工場などで、ピッキング作業の自動化	製造、出荷、物流、倉庫、小売	倉庫内物流、建設現場、災害救助、製造業、農業、介護



出典：Covariant AI社のホームページ



出典：https://www.youtube.com/watch?v=Sq1QZB5baNw

出典：Agility社ホームページ 仕様資料

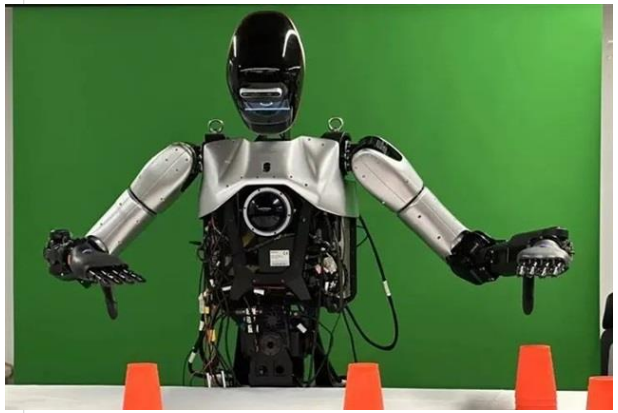
©2024 CRDS

最新ロボット事例紹介（中国）　ヒューマノイドの台頭

- 2024年上海で開幕された世界人工知能会議（WAIC）ではヒューマノイドが多数展示
- 安価なヒューマノイドが大量に市場に出回ることで進化が加速されることが期待される

製品名	Unitree社 G1	Humanoid Robot Innovation Center　Qinglong（健康龍）	Fourier Intelligence-Robotics社 GR-1
特長	二次開発できる汎用ヒューマノイドロボット 日本国内でも約600万円で販売中	43自由度、多モード移動、大規模言語モデルによるインタラクション	量産型の汎用ヒューマノイドロボット、大規模言語モデル搭載、オープンソース。
応用分野	研究機関や企業がカスタマイズすることで様々な応用が可能	製造業、物流、サービス業での多様なタスク自動化	研究・教育、受付、デモ、保安点検、製造、医療・リハビリなど

出典： Unitree社のホームページより



Qinglong, China’s first full-sized general-purpose humanoid robot, unveiled at World Artificial Intelligence Conference 2024

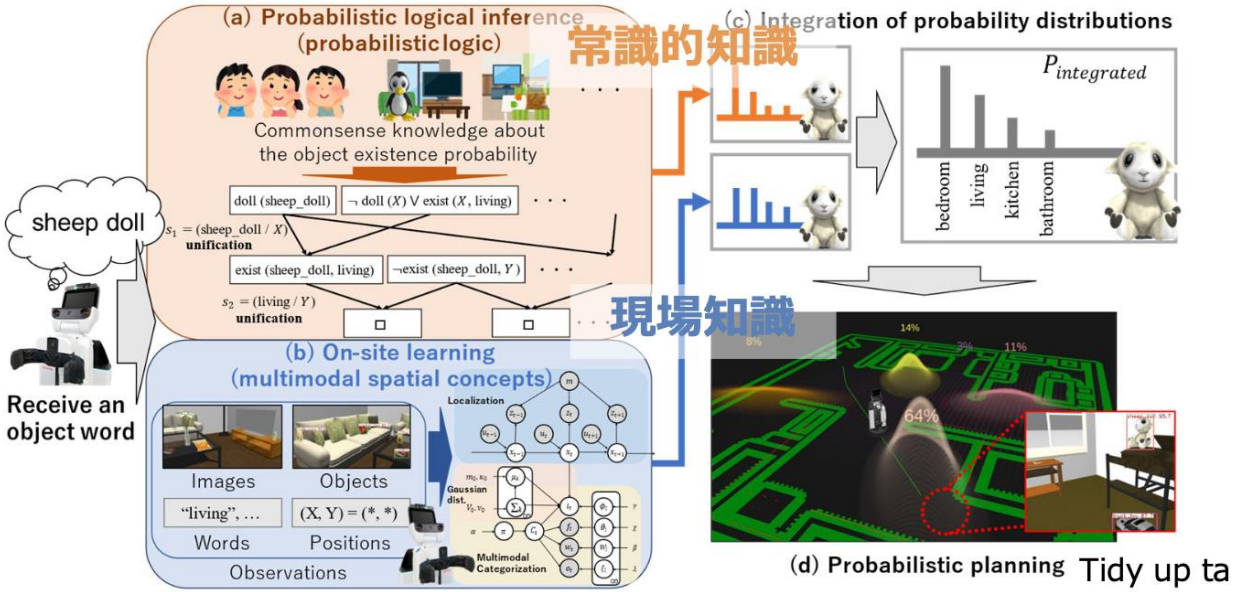
出典： Fourier Intelligence-Robotics 社社のホームページより

©2024 CRDS

最新ロボット事例紹介（日本）

- 現場状況をモデル化した「世界モデル」を構築することで多様な環境に適応
- 認知発達ロボティクス：ロボットが人間のように環境との相互作用を通じて学習し、適応する能力に関する研究

大規模言語モデル＋世界モデル：場所概念モデルによる現場学習と常識的知識を 統合した計算論モデルを構築。「これ片付けて」や「羊の人形を取ってきて」といった 発話から具体的な生活物理支援タスクを実現.、常識的知識として**大規模言語モデル**を活用し，現場のオンライン場所概念学習と融合的に利用。



出典：Shoichi Hasegawa他, "Inferring Place-Object Relationships by Integrating Probabilistic Logic and Multimodal Spatial Concepts," IEEE/SICE International Symposium on System Integration (SII), 2023.

予測符号化理論（能動的推論）：乳幼児からの成長のように、他者や環境との相互作用・実世界操作を通じて、自己・環境の認知、言語獲得、行動・推論等の認知機能を発達させていく過程を、予測誤差最小化原理(自由エネルギー原理)によって統一的に説明。大量の教師あり事前学習は不要になり、資源効率も改善され得る。

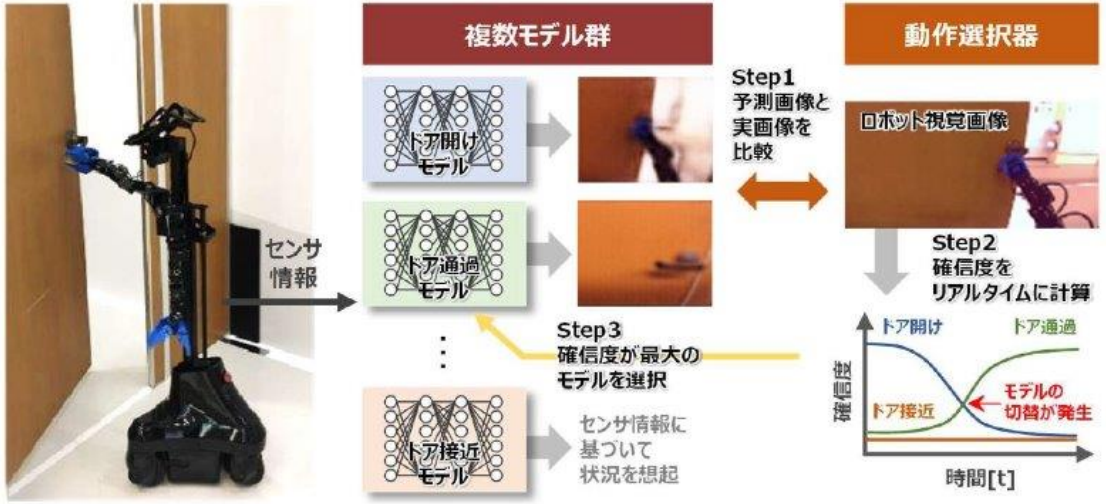


図３ 複数予測モデルのリアルタイム切替技術

１つの予測モデルでは対応しきれない複雑な作業に対応可能

出典：深層予測学習型ロボット制御技術開発 https://www.waseda.jp/top/news/79811

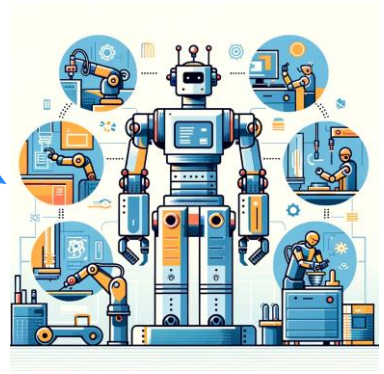
©2024 CRDS

今後の展望

©2024 CRDS

今後の方向性

■ 物理的な身体機能を獲得したAIシステムとしてのロボットの発展の方向性と典型的なタイプ



タイプP.多様な実世界タスクをこなすロボット

ある程度限定された環境のもとで、1台のシステムが様々なタスクを実行できるようにする。タスクに応じた様々な動作、繊細な動作、高速な動作を追求する。人間にはできないようなタスク・動作まで実現。

[想定する適用例] 産業用ロボットの知能化（セル生産、職人技の学習など）

P: Performance　（タスク性能追求型）
H: Humanoid　（人との共生追求型）
A: Adaptive　（環境適応性追求型）

タイプH.実世界で人間と協働・共進化するロボット

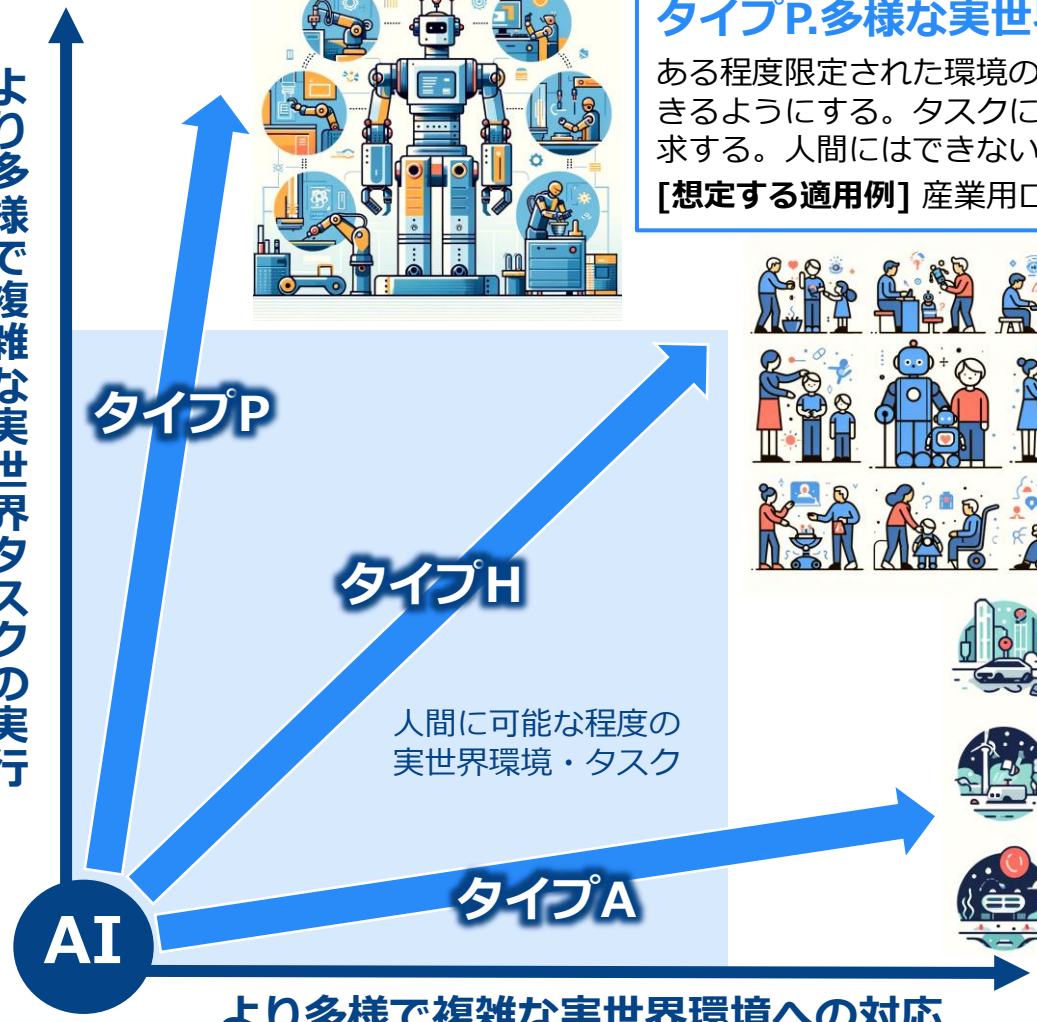
人間のような知能・身体性・対話能力等を備え、人間の代わりに／人間と協働で様々なタスクを担えるようにする。人間自体を理解するためのサイエンス研究面も含み、能動的に学習し、発達・共進化するロボットへと発展させる。

[想定する適用例] 発達・共進化（日本発の認知発達ロボティクス研究）

タイプA.多様な実世界環境で稼働するロボット

タスクの種類は、ある程度限定されているが、環境や条件に多様性・複雑性・変動性・不完全性があっても、ロバストにタスクを遂行できるようにする。人間が立ち入れない環境でさえ稼働可能。

[想定する適用例] 屋外ロボットの知能化（農業、インフラ保守点検など）



※上記画像は生成AIを用いて作成した

©2024 CRDS

ロボットの発展がもたらす社会的価値

■ AIが物理的な身体機能を獲得すると、デスクワークから現場作業まで幅広く労働市場に波及、大きな社会・経済的インパクトをもたらすことが期待できる

タイプ	現状の実用レベル	現状の研究レベル	最終的な目標レベル
P Performance	**タスク毎の作り込み** • 強く統制された現場環境 • 個々のタスクのプログラミング **実用例** • 自動車製造ラインの溶接ロボット • 半導体製造でのウェハ搬送ロボット	**タスク毎の事前学習** • 個別タスクの模倣学習と強化学習 • 熟練者技能の学習 **研究事例** • 米）Covariant AI社 RFM-1 • 米）Google社 RT-1/RT-2	**多様な実世界タスクをこなすシステム** • 1台でさまざまなタスクを実行 • 繊細な操作、高速な操作を追求 **事例：汎用型組立ロボット** • 車種やモデルごとに異なる部品の組み立てを1台のロボットで対応
H Humanoid	**単機能サービスロボット** • 限定環境での安全な定型作業 • 限定的話題や定型的対話(案内受付等) **実用例** • 飲食店における配膳ロボット • 病院における搬送ロボット	**対話的サービスロボット** • タスクに応じた行動計画推論 • 自由対話による指示理解・説明 **研究事例** • 日）認知発達ロボティクス • 米）Google社 PaLM-SayCan	**実世界で人間と協働・共進化するシステム** • 人間同様の知能・身体・対話能力 • 能動的に学習し発達する **事例：サービスアシスタントロボット** • 家事やサービス業などで人間の作業を能動的に支援
A Adaptive	**専用機による作業** • 人間による機器の操作 • 限定環境での自動運転 **実用例** • 建設現場における重機ロボット • 農業における収穫ロボット	**限定環境で働くフィールドロボット** • 不整地対応ロコモーション • 移動とマニピュレーションの統合 **研究事例** • 米）Agility Robotics社 Digit • スイス）ETH Zurich ANYmal C	**多様な実世界環境で稼働するシステム** • 多様・複雑・不完全性への頑強性 • 環境変化や想定外への適応 **事例：フィールドロボット** • 農業やインフラ点検作業など高負荷／危険な作業を人間に代わってこなす

©2024 CRDS

まとめ

©2024 CRDS

まとめ

1.AI技術の進展とロボット研究

- 第1段階（2010～）：機械学習・深層学習▶環境の認識・判断能力の向上
- 第2段階（2020～）：基盤モデル・生成AI ▶行動計画や動作生成が可能に
- 第3段階（今後）：ロボット基盤モデル▶人間のような知能・身体性・対話能力へ

2.生成AIとロボットの統合

- ロボット基盤モデルの進化：柔軟な行動計画と動作生成の自動化
- データの多様化と大規模学習：マルチモーダルデータによる高精度な動作生成

3.ロボットの最新動向

- 米国：生成AIを適用したヒューマノイドロボットの能力向上
- 中国：安価で汎用性の高いヒューマノイドロボットの登場
- 日本：認知発達ロボティクスと世界モデルによる現場環境適応の研究

4.今後の展望

- 多様な実世界タスクをこなすロボット
- 実世界で人間と協働・共進化するロボット
- 多様な実世界環境で稼働するロボット

©2024 CRDS

■作成担当■
茂木 強 フェロー （システム・情報科学技術ユニット）

CRDS-FY2024-XR-09
その他報告書

テクノロジー、サステナビリティ、Well-beingの交差点 ～イノベーションの動向を掴む～
「生成AIとロボット研究の新たなフロンティア」

令和6年8月 August 2024
ISBN 978-4-88890-945-7

国立研究開発法人科学技術振興機構　研究開発戦略センター
Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町7 K's五番町
電 話　 03-5214-7481
E-mail crds@jst.go.jp
https://www.jst.go.jp/crds/

本資料は著作権法等によって著作権が保護された著作物です。
著作権法で認められた場合を除き、本資料の全部又は一部を許可無く複写・複製することを禁じます。
引用を行う際は、必ず出典を記述願います。
なお、本報告書の参考文献としてインターネット上の情報が掲載されている場合には、本報告書の発行日の1ヶ月前の日付で入手しているものです。
上記日付以降後の情報の更新は行わないものとします。

This publication is protected by copyright law and international treaties.
No part of this publication may be copied or reproduced in any form or by any means without permission of JST,
except to the extent permitted by applicable law.
Any quotations must be appropriately acknowledged.
If you wish to copy, reproduce, display or otherwise use this publication, please contact crds@jst.go.jp.
Please note that all web references in this report were last checked one month prior to publication.
CRDS is not responsible for any changes in content after this date.

Copyright © 2024 CRDS All Rights Reserved.