

ロボットはなぜお茶を持ってきてくれないのか？

～ 大規模AI時代におけるロボット開発 その課題と展望 ～

中京大学 工学部 機械システム工学科

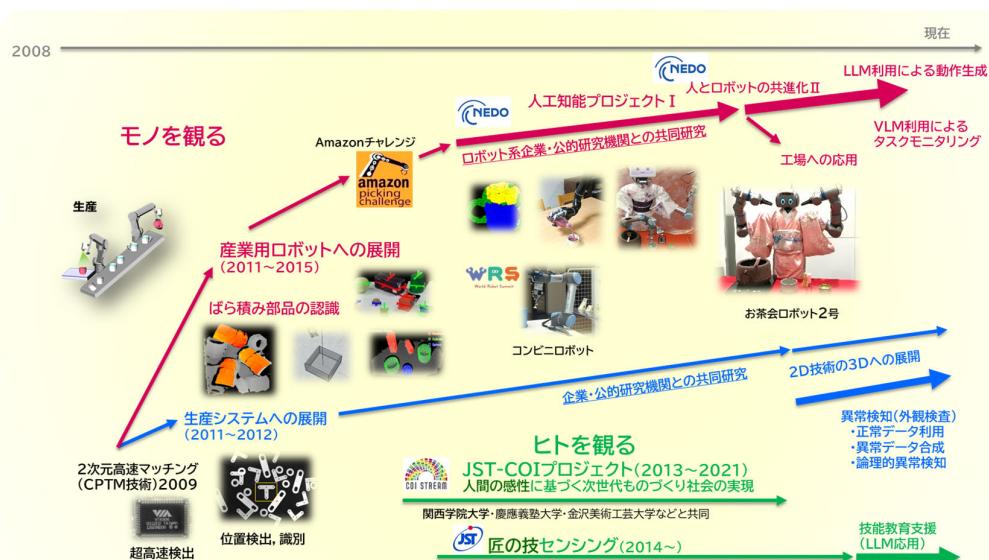
橋本 学

mana@isl.sist.chukyo-u.ac.jp

Advanced Sensing & Machine Intelligence Group,
Chukyo University

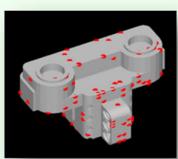
橋本の略歴

- 1987~2007 三菱電機(株)に勤務。生産技術研究所、先端技術総合研究所等にて画像認識、知能ロボットの研究に従事
- 2008~ 中京大学 教授、現在、大学院 工学研究科長、ヒューマン・ロボティクス研究センター長を兼務
- 研究分野 知能ロボティクス、異常検知(画像検査)、自動化システム、物流システム、ヒューマンセンシング
「現場の課題にサイエンスで応える」
- 産官学連携 外観検査、画像照合、ロボティクス、動作分析などで企業様との共同研究、研究生受け入れ、NEDO等の国プロ、技術コンサルティング



最近の主な研究事例

3次元画像認識
ロボット応用



3次元点群処理



把持パラメータ決定



把持位置自動決定



LLM利用動作生成

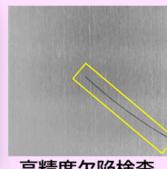


自動プログラム生成

2次元画像認識



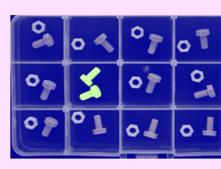
超高速画像マッチング



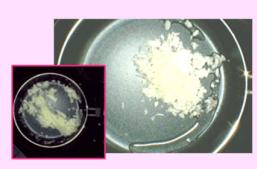
高精度欠陥検査
(検出率99.7%)



外観検査のための
リアルデータ拡張

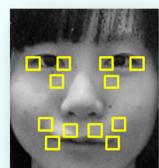


“論理的異常”検知



調理タスクモニタリング

ヒューマンセンシング



表情からの満足度推定



手指動作の認識



熟練技の分析



LLMを利用した技能教育支援システム

AIが作った文章

右手は、ねじを支える役割があるので、動かさず固定しましょう。これにより、ねじ位置が安定し、作業が早くなります。

本日の話題

1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

実用に近い技術の例① Amazon 物流ロボット国際大会



半自動化された配送センター
「自動走行する棚」

完全な無人化のために……

Amazon Robotics Challenge
(国際的なロボット競技大会)



競技で使用されたアイテムの例

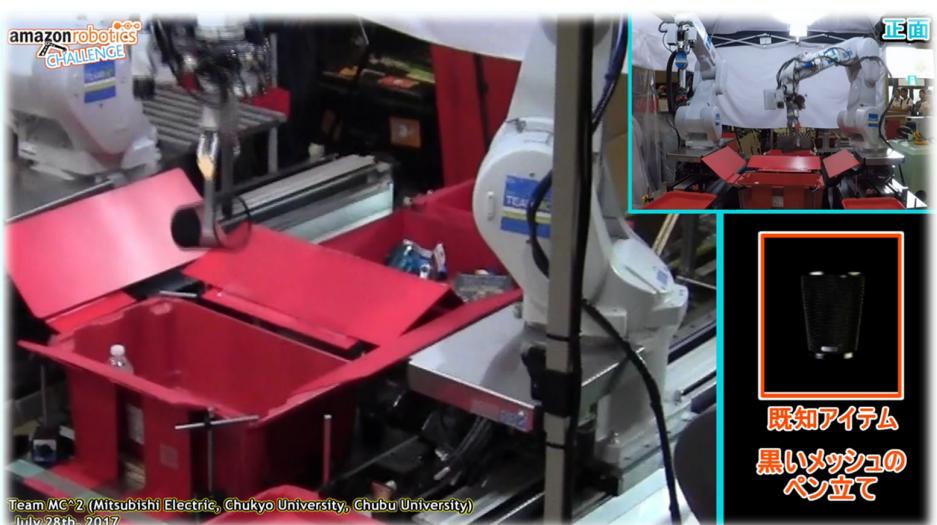


Advanced Sensing and Machine Intelligence, Chukyo Univ.

ESD21 豊田工業大学講演会 Nov. 21, 2025

7

競技のようす



Advanced Sensing and Machine Intelligence, Chukyo Univ.

ESD21 豊田工業大学講演会 Nov. 21, 2025

8

実用に近い技術の例② World Robot Summit 組立てロボット大会



実用に近い技術の例② World Robot Summit 組立てロボット大会

ロボット v.s. 人間

ロボットによる組み立て



3分20秒

人間による組み立て



1分02秒

本日の話題

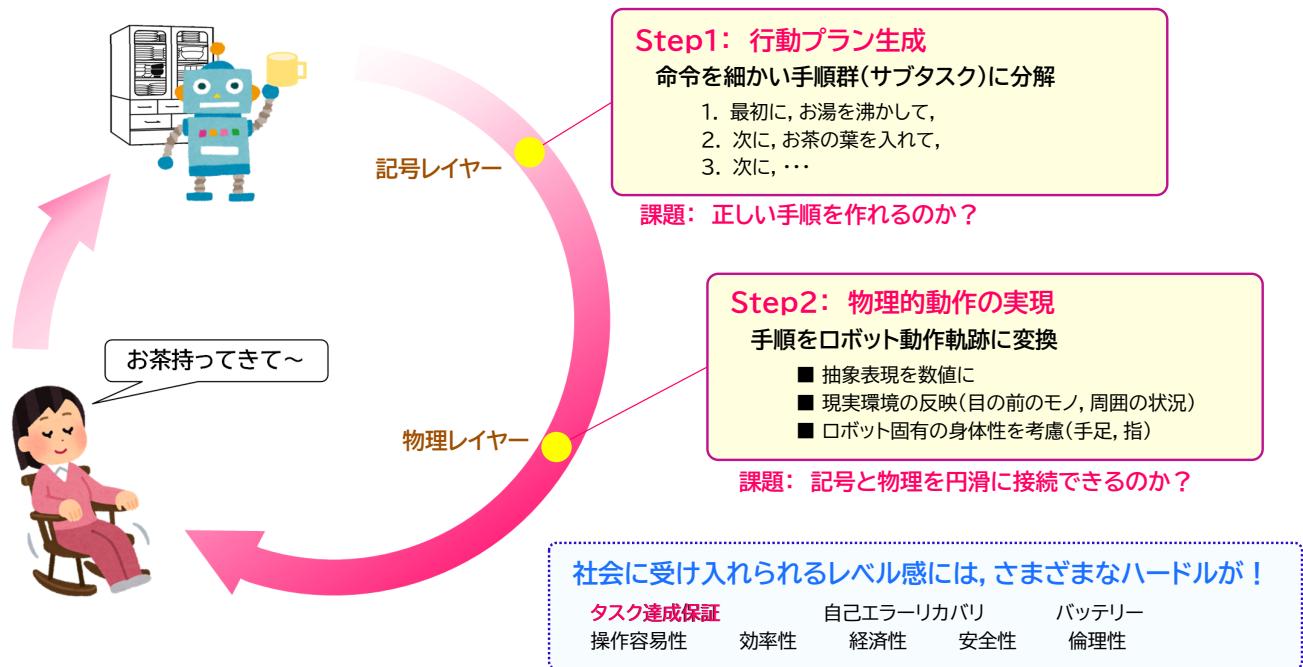
1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

開発中のお茶会ロボット



双腕ロボットによるお茶会デモシステム

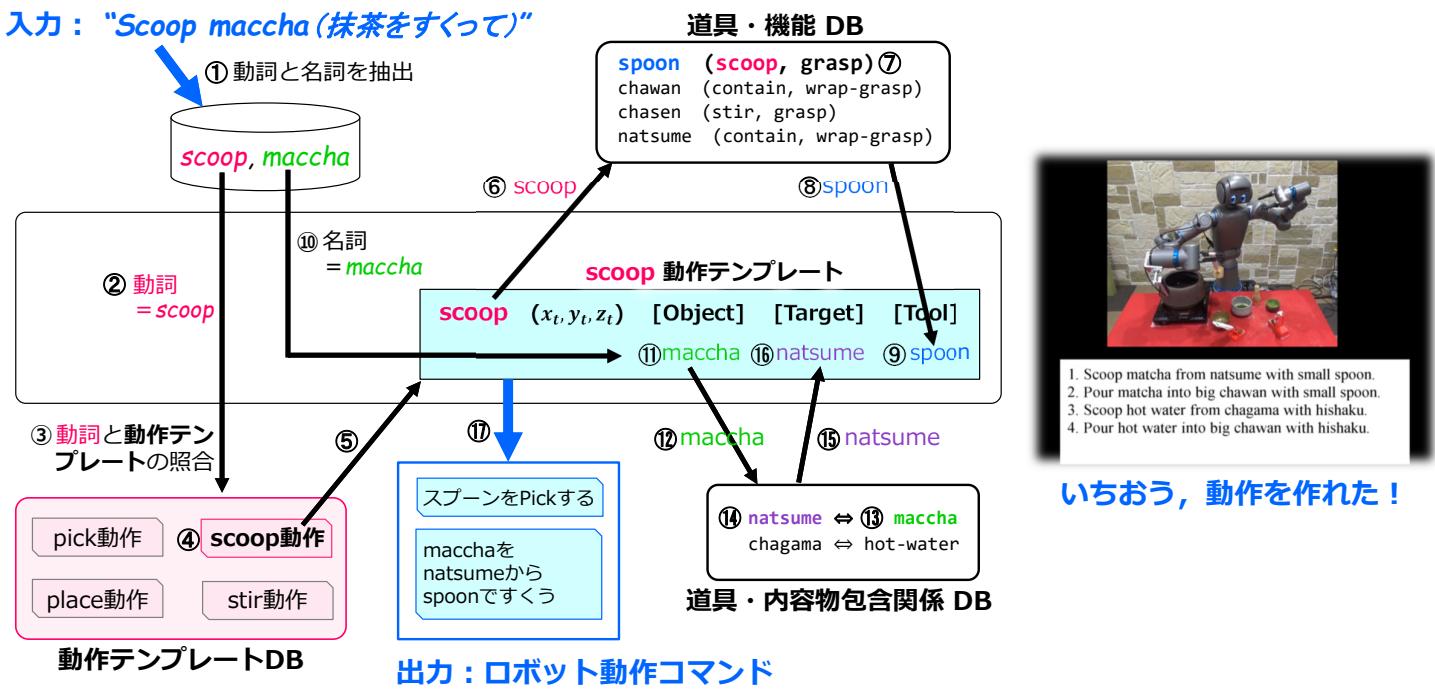
お茶会ロボットは、なぜ難しいのか？



本日の話題

1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

予備検討 シンプルな指示文からの動作生成 (2022年開発)



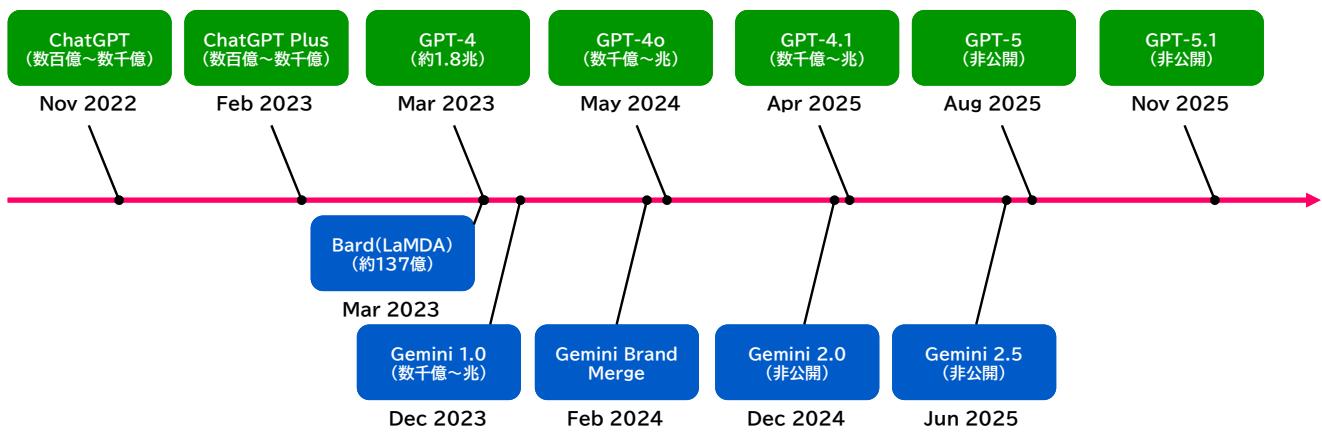
Advanced Sensing and Machine Intelligence, Chukyo Univ.

ESD21 豊田工業大学講演会 Nov. 21, 2025

15

2022年 黑船来航

実用レベルのLLM(大規模言語モデル)が登場



LLMを利用した行動プランの作成

GPTへの問い合わせ

「お茶をもってきて」という命令を、手順に分解してください。

GPTの答え

1. お茶の葉を取得する
2. ティーポットを確保する
3. お茶の葉をティーポットに入れる
4. 水を沸かす
5. 沸騰した水をティーポットに注ぐ
6. お茶を淹れる
7. お茶をカップに注ぐ
8. お茶を提供する

【参考】

PaLM-SayCan (Google2022)のデモ

「オレンジジュースをこぼしたので、代わりの飲み物を持ってきて」



- ①まず、倒れたカップをゴミ箱に捨て、
- ②次に、スポンジを持ってきて拭き、
- ③最後に新しい飲み物を持ってくる

という長いレンジの行動を生成できた。

Long-horizon instruction (長い地平の指示文) を解釈、実行できた。

検証実験 本当に LLM は使えるのか？

実験目的

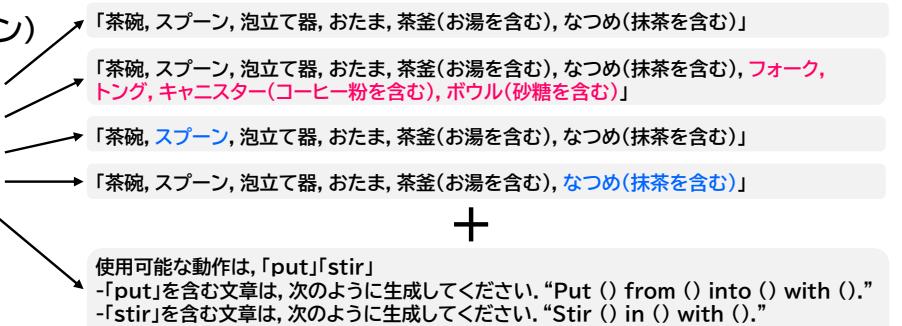
多様な質問文、条件・制約、ターゲットを与えた結果、GPT が生成した動作手順の問題点を分析

■ GPTに与える質問文(3 パターン)

1. ○○を作るための手順を教えてください。
2. ○○を作るための手順を生成してください。
3. ○○を作るための手順をロボットが実行可能な形で生成してください。

■ GPTに与える条件・制約(9 パターン)

1. シーンの設定
 1. 過不足なく道具がある状況
 2. 必要のない道具も入っている状況
 3. 不足している道具がある状況
 4. 不足している材料がある状況
2. 動作・文構造の指定
3. シーン + 動作・文構造の指定



■ ターゲット(4 パターン)

1. 抹茶
2. コーヒー
3. カレーライス
4. みそ汁

GPTによる不適切な応答例（実験 #72）

(実際のGPT応答の日本語訳)

GPTに与えたクエリー

- User（指示文）
 - カレーライスを作るための動作手順を1文ずつ生成してください。

- System（条件・制約）
 1. 指定された動作（put、stir）のみを使用してください。
 2. 指定された道具と材料（お皿、おひつ（ご飯入り）、鍋（カレー入り）、ボウル（福神漬入り））のみを使用してください。
 3. 必要であれば、以下のように「put」を含む文章を作成してください。Put (ingredient) from (tool) into (tool) with (tool)。
 4. 必要であれば、以下のように "stir" を含む文を生成してください。Stir (ingredient) in (tool) with (tool)。

GPTからの応答

1. しゃもじで皿におひつから米を入れる。
2. 福神漬をボウルからスプーンで皿に入れる。
3. 皿に入れた福神漬をお玉でかき混ぜる。

【考察】

1. お米と福神漬しか使ってないので、そもそも、カレーライスにならない。
2. お玉は、福神漬を扱うには大きすぎる。
3. 福神漬をかき混ぜてはいけないし・・・

GPTによる不適切な応答例（実験 #34）

(実際のGPT応答の日本語訳)

GPTに与えたクエリー

- User（指示文）
 - コーヒーを作る手順を一言ずつ教えてください。

- System（条件・制約）
 1. 指定された道具と材料（カップ、スプーン、泡立て器、おたま、ヘラ、フォーク、ナイフ、トング、茶釜（お湯入り）、容器（コーヒー粉入り）、なつめ（抹茶粉入り）、ボウル（砂糖入り））のみを使用する。
 2. 指定された動作（put、stir）のみを使用する。
 3. 以下のように「入れる」を含む文章を作成してください。（道具）から（道具）に（道具）で（材料）を入れる。
 4. 以下のように「かき混ぜる」を含む文が必要なら作ってください。（道具）の中の（材料）を（道具）でかき混ぜる。

GPTからの応答

1. 容器からコーヒー粉をお玉で茶釜に入る。
2. 茶釜の中のコーヒーを泡立て器でかき混ぜる。

【考察】

1. お玉は、コーヒー粉をすくう道具としては不適切。
2. コーヒーを「茶釜」の中に作ろうとしている。
3. 泡立て器は、コーヒーを混ぜる道具としては不適切。

GPTによる不適切な応答例（実験 #97）

(実際のGPT応答の日本語訳)

GPTに与えたクエリー

- User (指示文)
 - 味噌汁を作るための動作手順を1文ずつ生成してください。

- System (条件・制約)
 1. 指定された動作 (put, stir) のみを使用してください。
 2. 指定された道具と材料 (茶碗, スプーン, 泡立て器, お玉, フライ返し, フォーク, ナイフ, トング, 茶釜 (お湯入り), 容器 (コーヒー粉入り), なつめ (抹茶粉入り), ボウル A (白味噌入り), ボウル B (赤味噌入り), ボウル C (わかめ入り)) のみを使用してください。
 3. 必要であれば、以下のように「put」を含む文章を作成してください。 Put (ingredient) from (tool) into (tool) with (tool).
 4. 必要であれば、以下のように "stir" を含む文を生成してください。 Stir (ingredient) in (tool) with (tool).

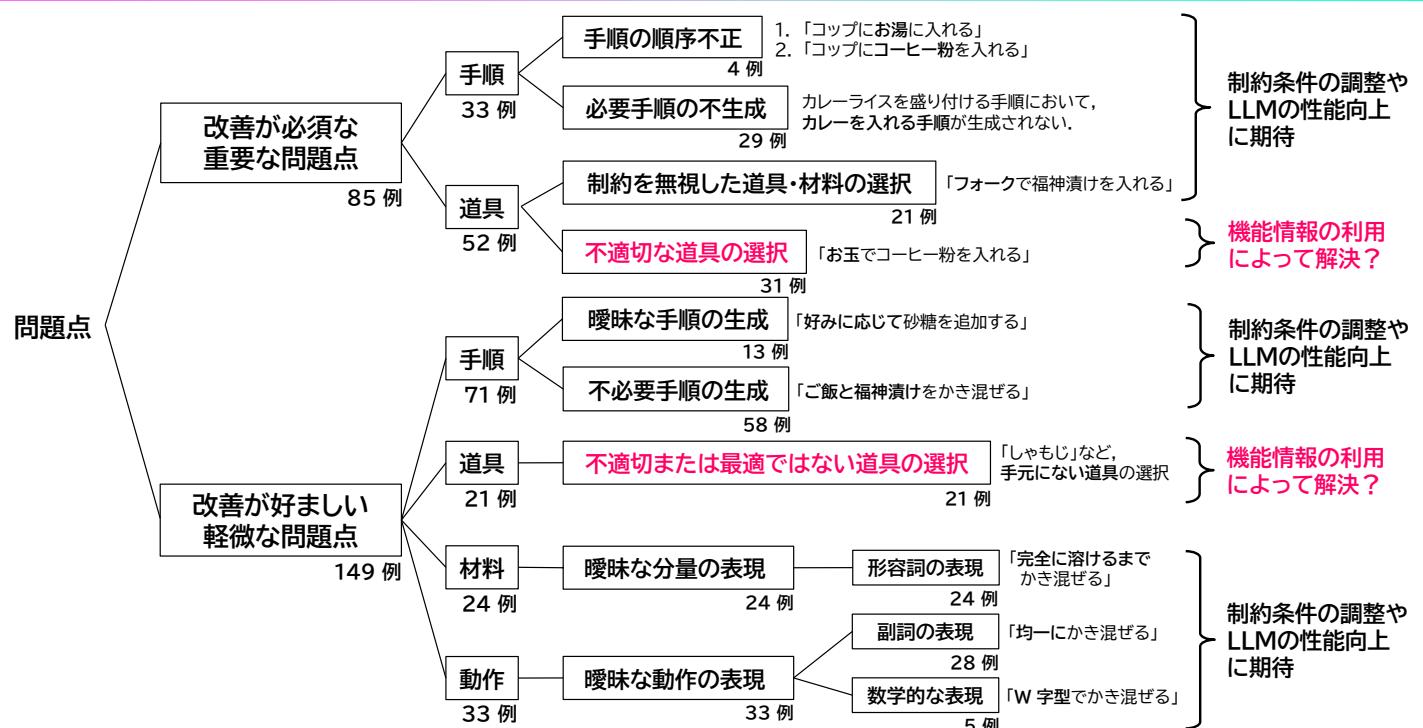
GPTからの応答

1. お玉でボウル A の白味噌を茶碗に入れる。
2. 泡立て器で味噌をかき混ぜる。
3. お玉でボウル B の赤味噌を茶碗に入れる。
4. 泡立て器で味噌をかき混ぜる。
5. トングでボウル C のわかめを茶碗に入れる。
6. スプーンで茶碗の中のわかめをかき混ぜる。

【考察】

1. 2種類のお味噌を混ぜているのはよいのだが・・・
肝心な「お湯」を入れる手順が抜けている。

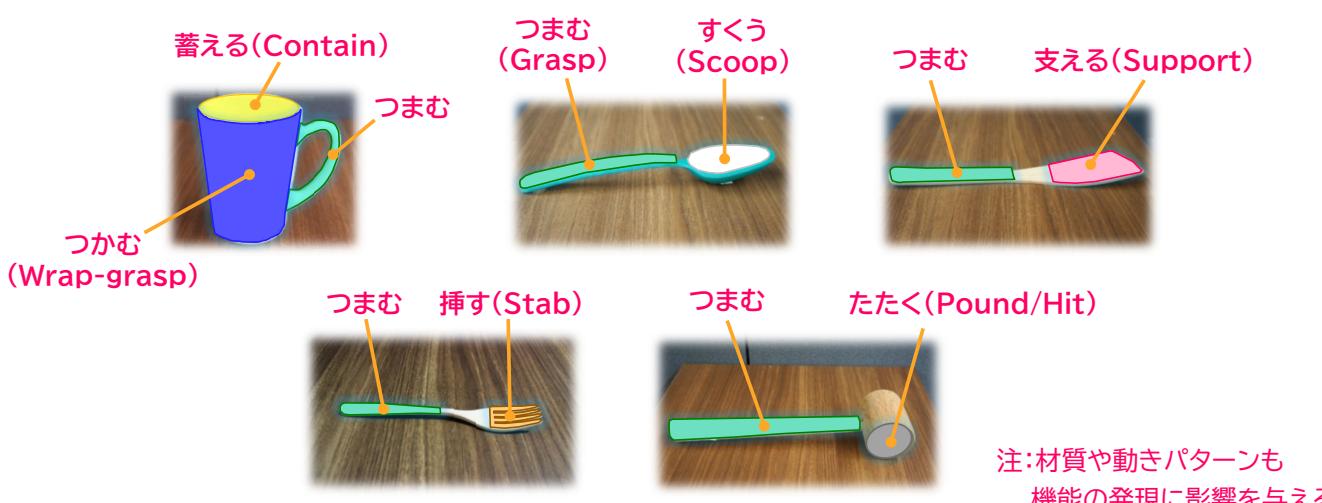
問題点の整理と解決の方向性



道具が持っている“機能”とは？



道具は“形”に“機能”を持っている

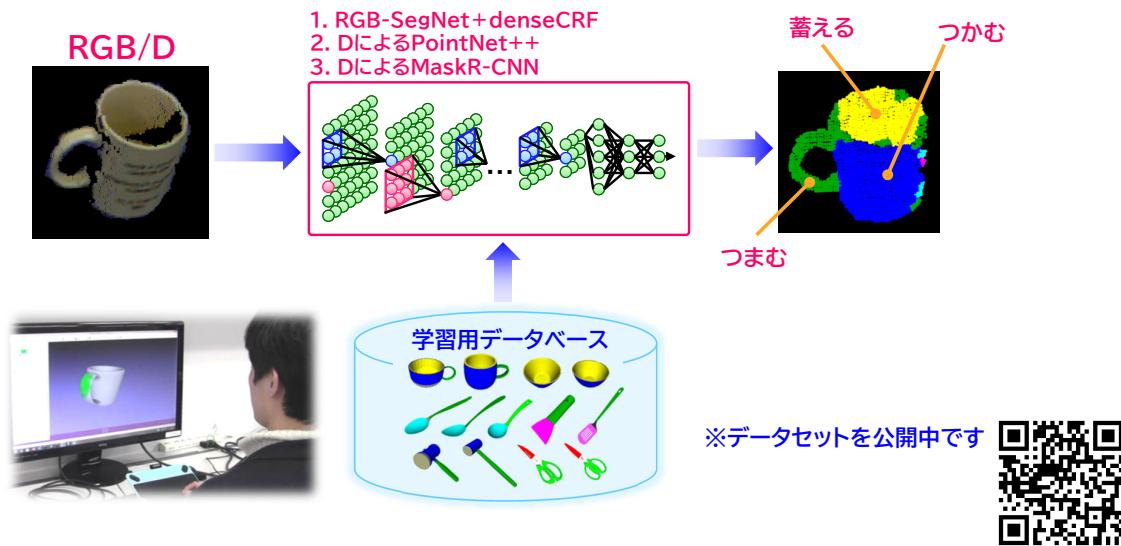


Affordance, Signifierとその認識に関する研究例

- [1] Yamanobe, et al. A Brief Review of Affordance in Robotic Manipulation Research, Journal of Advanced Robotics, 2017.
- [2] Hermans, J.M. Rehg and A.F. Bobick: "Affordance Prediction via Learned Object Attributes," ICRA2011.
- [3] A.Myers, C.L.Teo, C.Fermuller, and Y.Aloimonos: ``Affordance Detection of Tool Parts from Geometric Features", ICRA2015.
- [4] A.Nguyen, D.Kanoulas, D.G.Caldwell, and N.G.Tsagarakis: ``Detecting object affordances with Convolutional Neural Networks", Proc. IROS2016.
- [5] A. Roy and S. Todorovic: ``A Multi-Scale CNN for Affordance Segmentation in RGB Images", ECCV2016.
- [6] A. Nguyen, et. al, ``Object-Based Affordances Detection with Convolutional Neural Networks and Dense Conditional Random Fields", IROS2017.

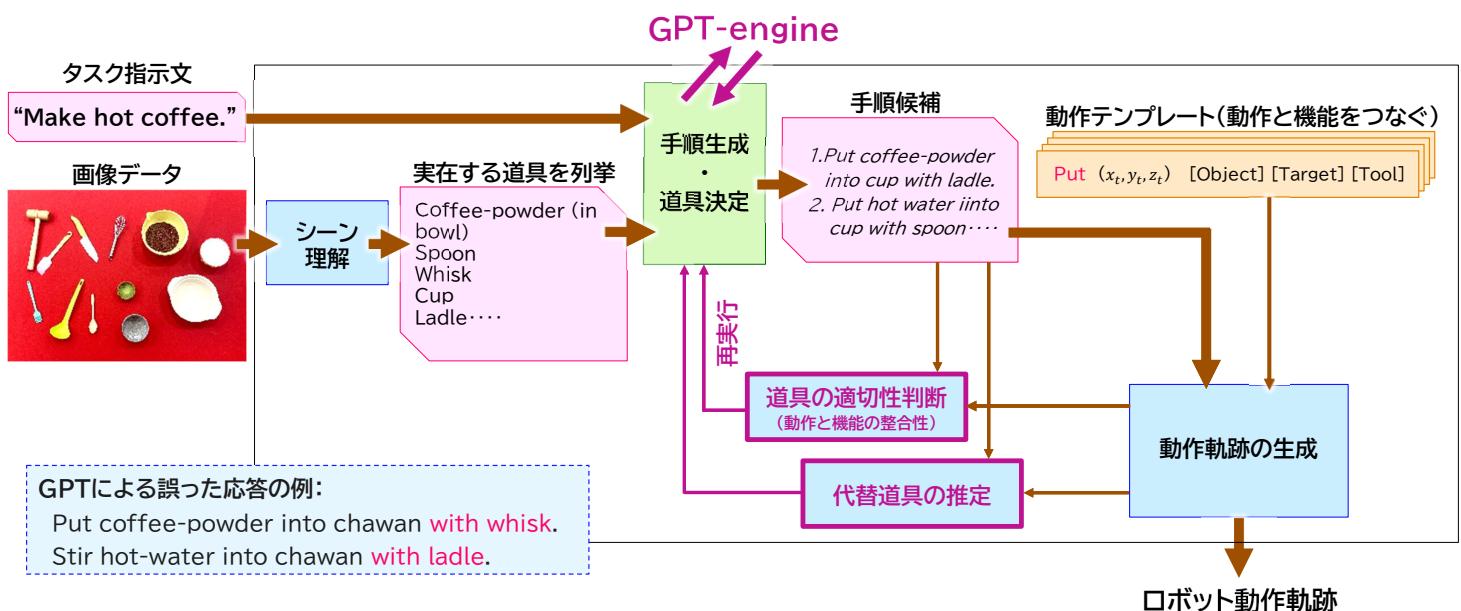
機械学習に基づく“機能”の認識

- 部位ごとに“機能”的な意味をもつラベル名を与える → “機能認識”と考える



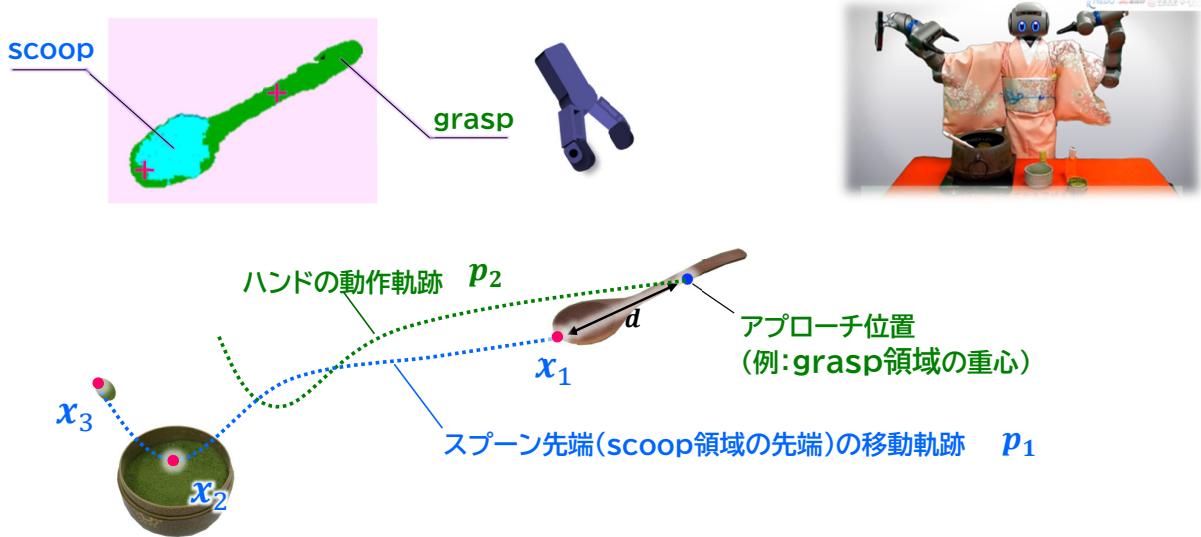
研究事例① “機能”情報を利用した手順のエラー修正

- シンプルなタスク指示文をもとに、GPTエンジンを用いて手順候補を生成
- 機能情報を用いて、GPTの誤りや不足な点を自動修正し、よりよい手順生成を実現



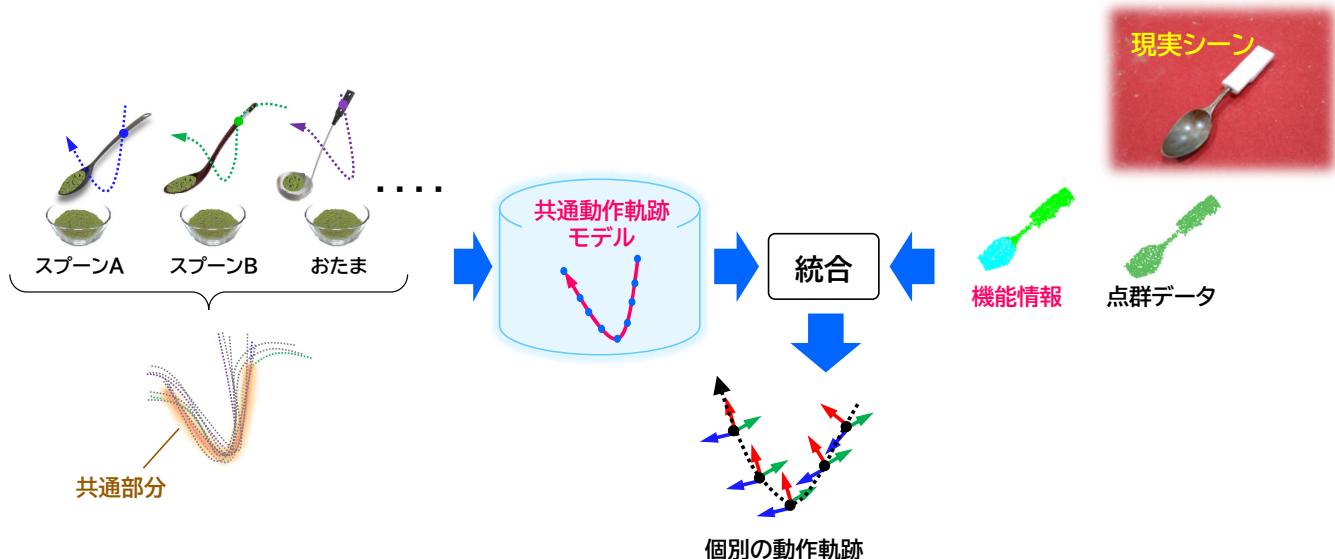
1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

“機能情報” 利用の基本 ……汎用的なロボット動作の生成



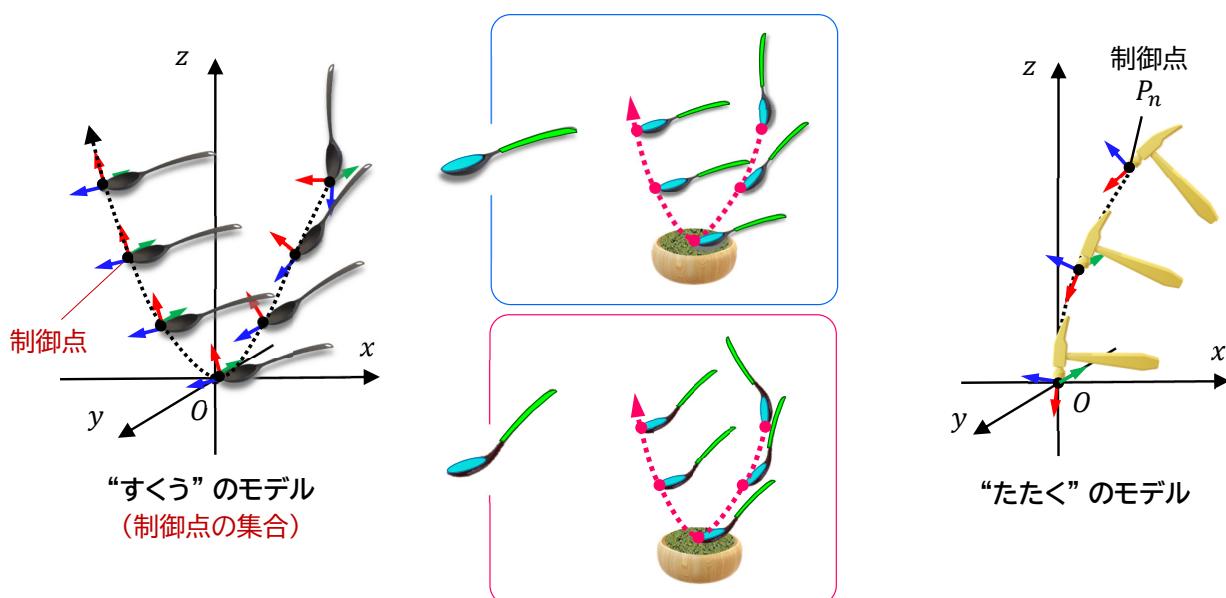
研究事例② 道具の種類ごとに「動き」をモデル化……“共通動作軌跡モデル”

- 道具固有の使い方(動作)を、道具のカテゴリごとに共通モデルとして構築
→ 共通モデルにその場でセンシングした情報を組み合わせて、個別の軌跡を生成



研究事例② 道具の種類ごとに「動き」をモデル化……“共通動作軌跡モデル”

- 道具の機能属性を認識し、**機能領域の軌跡(座標 & 姿勢)**を生成する。



研究事例② 動作生成実験の結果(すくう)

		使用道具(スプーン)				
		1	2	3	4	5
対象物	1	10/10	9/10	9/10	10/10	10/10
	2	6/10	9/10	7/10	6/10	10/10

		使用道具(お玉)				
		1	2	3	4	5
対象物	1	9/10	3/10	10/10	8/10	8/10
	2	8/10	8/10	9/10	9/10	8/10



いずれも1つの“すくう”モデルを利用して動作生成(成功率 83.0%)

研究事例② 動作生成実験の結果(注ぐ)

		使用道具(スプーン)				
		1	2	3	4	5
対象物	1	10/10	10/10	8/10	9/10	10/10
	2	9/10	10/10	10/10	10/10	9/10

		使用道具(お玉)				
		1	2	3	4	5
対象物	1	7/10	5/10	5/10	8/10	5/10
	2	7/10	7/10	5/10	4/10	9/10

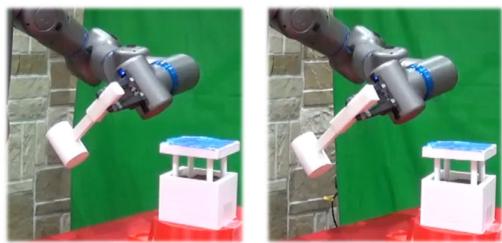


いずれも1つの“注ぐ”モデルを利用して動作生成(成功率 78.5%)

研究事例② 動作生成実験の結果(たたく, 塗る)

たたく動作

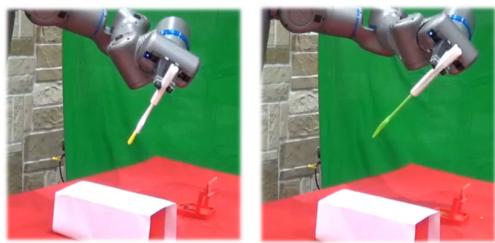
		使用道具(ハンマー)				
対象物		4/10	8/10	7/10	7/10	7/10



いずれも1つの“たたく”モデルを利用して動作生成(成功率 66.0%)

塗る動作

		使用道具(刷毛)				
対象物		8/10	8/10	10/10	8/10	9/10



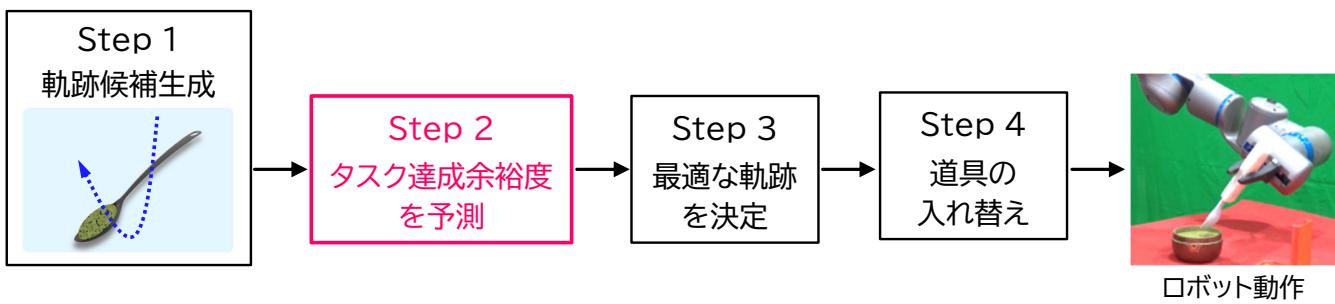
いずれも1つの“塗る”モデルを利用して動作生成(成功率 86.0%)

研究事例③ 環境に存在する事物との干渉問題の解決



研究事例③ 環境に存在する事物との干渉問題の解決

- Step 1. 手順中の「動詞」をもとに、仮の軌跡候補を生成（最適な軌跡とは限らない）
- Step 2. タスク達成余裕度（環境と干渉しない？本当に実行できる？）を予測
- Step 3. Step1, 2を繰り返し、最適な軌跡を決定
- Step 4. 必要に応じて、適切な道具に入れ替える

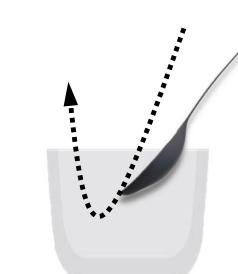


研究事例③ 環境に存在する事物との干渉問題の解決

$$\text{タスク達成余裕度} = \text{干渉余裕度 } M_I + \text{実行余裕度 } M_E$$

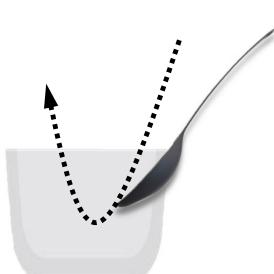
1. 干渉余裕度 M_I

道具と容器が干渉しない度合い



容器と干渉しない

$$M_I = \text{大}$$

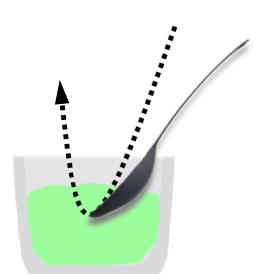


容器と干渉する

$$M_I = \text{小}$$

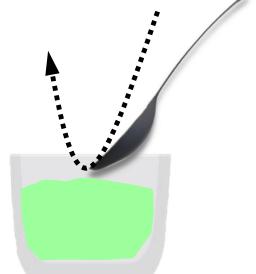
2. 実行余裕度 M_E

目的のタスクを実行できる度合い



すぐえる

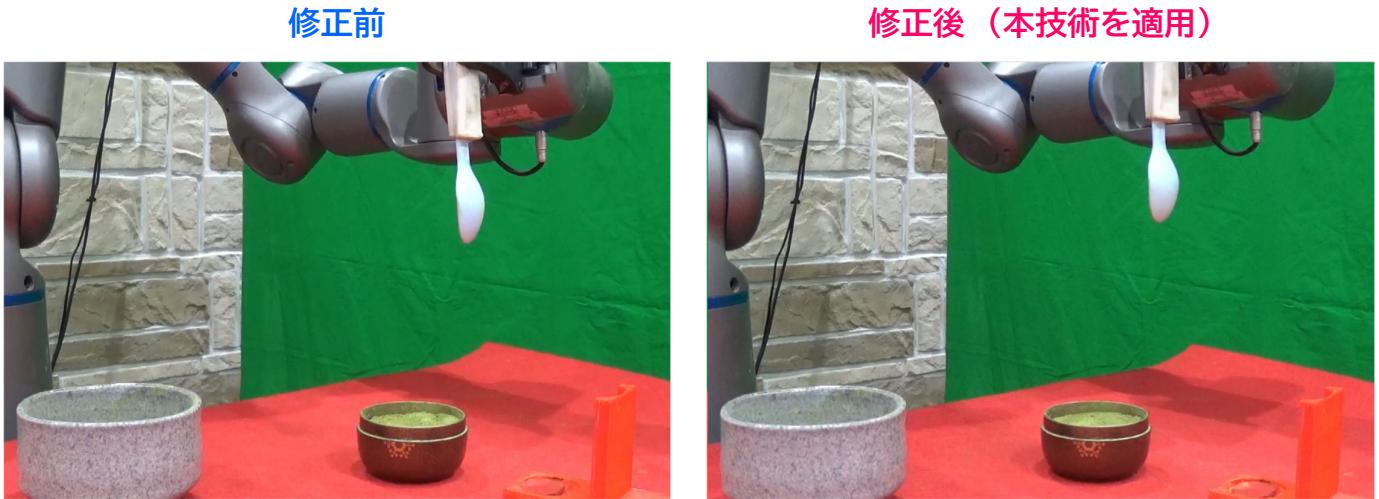
$$M_E = \text{大}$$



すぐえない

$$M_E = \text{小}$$

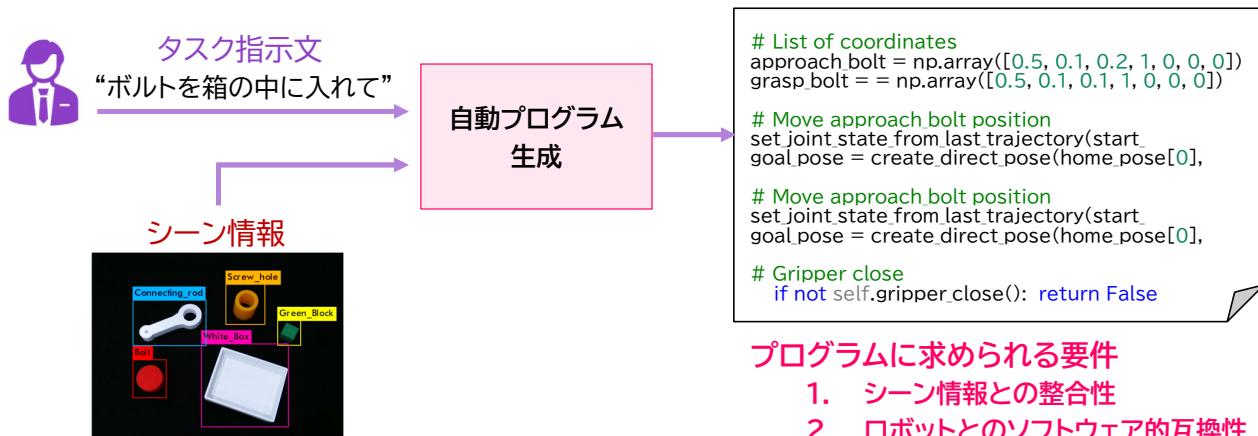
研究事例③ 実験結果



研究事例④ 生成AIを用いたロボットプログラミングの自動化

研究の目的：自然言語による指示文をもとに、実ロボットで動作可能なプログラムを生成

- 入力：タスク指示文、シーン情報(RGB-D 画像)
- 出力：対象物の3次元上の位置と姿勢をもとに動作するロボットプログラム



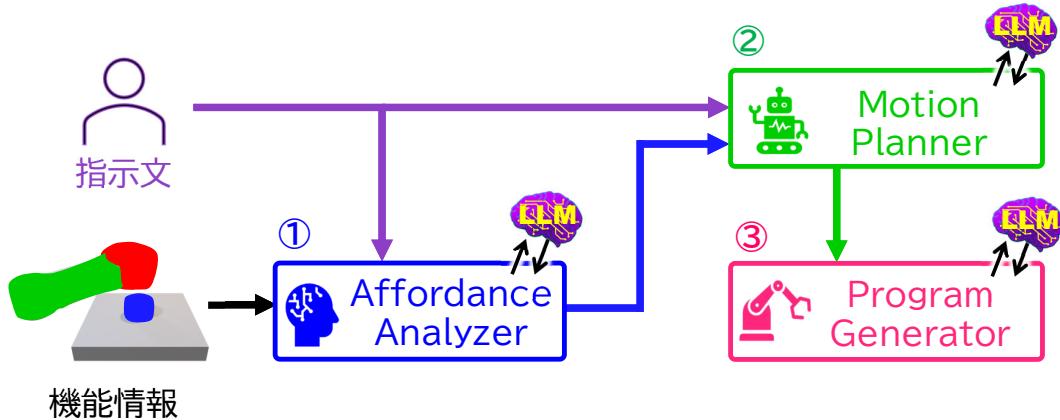
研究事例④ 生成AIを用いたロボットプログラミングの自動化

アイデア：段階的にLLMを適用し、ロボット動作プログラムを生成

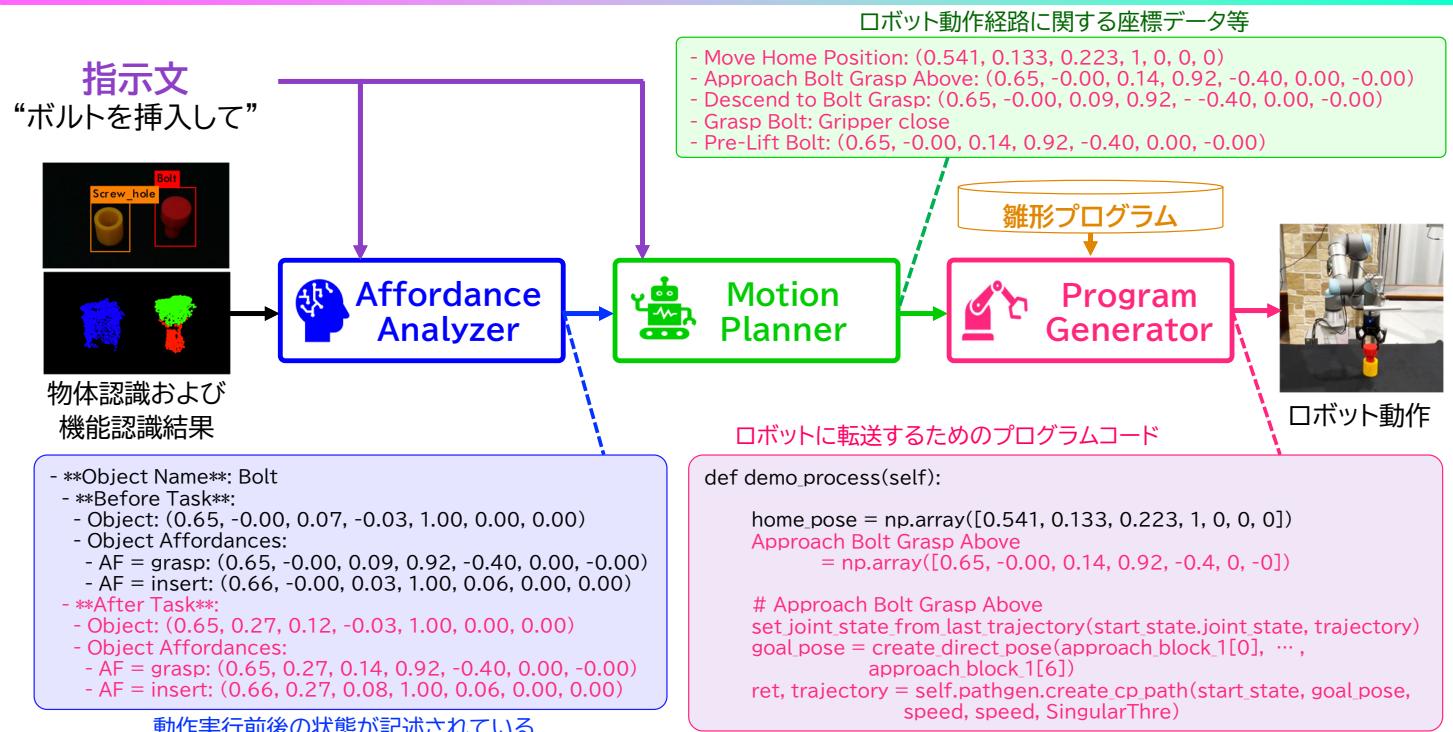
Step ① Affordance Analyzer: 機能情報を利用してシーンを理解、対象の目標状態を決定

Step ② Motion Planner: 現在のシーンから目標状態へ遷移するための動作経路を決定

Step ③ Program Generator: 動作経路をもとに、実際のロボット動作プログラムを生成

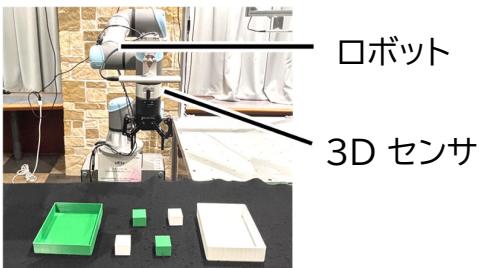


研究事例④ 生成AIを用いたロボットプログラミングの自動化

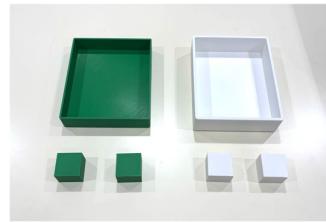


研究事例④ 生成AIを用いたロボットプログラミングの自動化

- 性能評価実験の目的
 - 指示文とシーンから、実行可能なロボット動作プログラムを生成できるかを評価
- 実験方法
 - タスク: 4種類の指示文を使用
 1. Put タスク(2種類)
 2. Insert タスク(2種類)
 - 実験回数: 各タスクに対して、20回ずつ(計80回)
- 実験環境



■ 対象物



Put タスクで使用する対象物



Insert タスクで使用する対象物

デモ動画

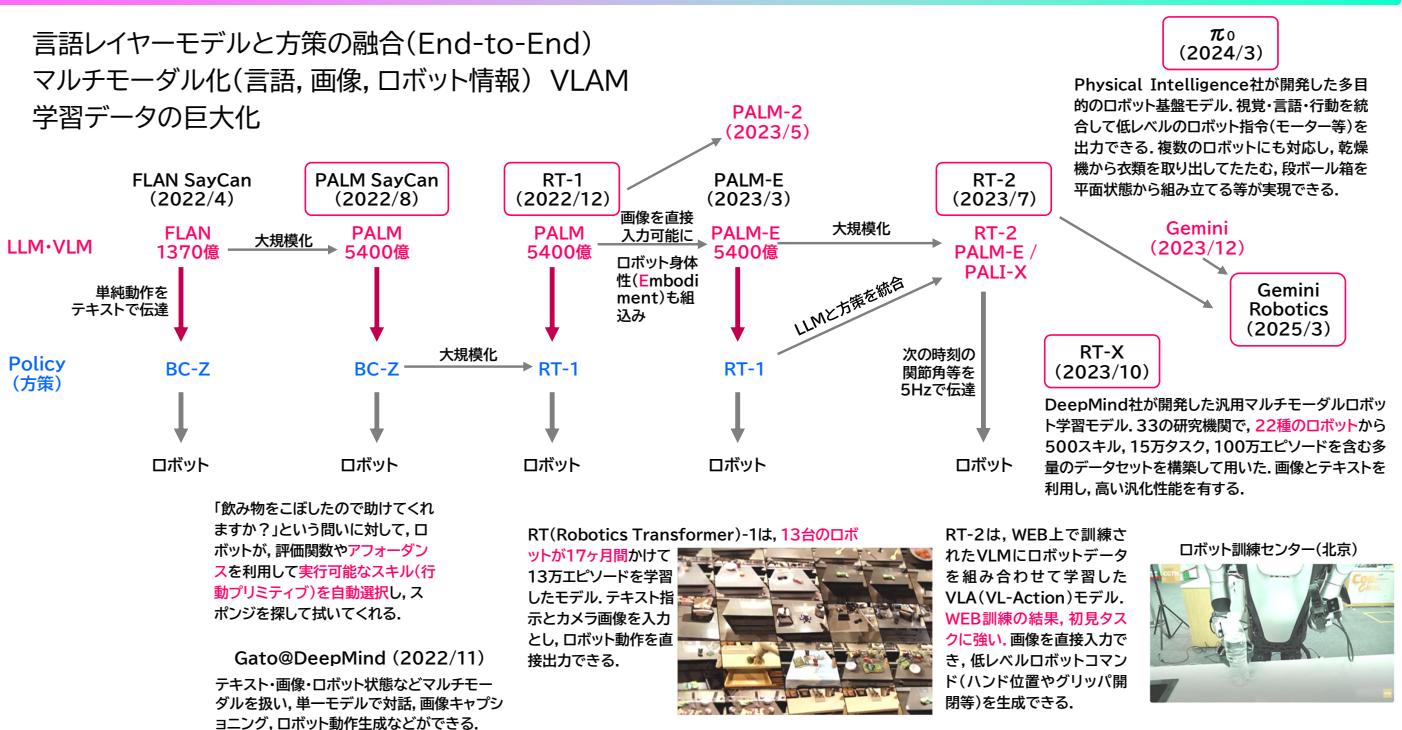
(66秒)

LLMと視覚情報に基づく階層的処理による
ロボットプログラム自動生成

1. イントロダクション いま、ロボットに何ができるか？
2. お茶会ロボットは、なぜ難しいのか？
3. 解決への道筋
 - (1) 記号レイヤーでの行動プラン生成
 - (2) 物理レイヤーへの接続
4. AI時代のロボット研究と期待

ロボット動作生成に関する研究動向 (Googleほか)

- 言語レイヤーモデルと方策の融合(End-to-End)
- マルチモーダル化(言語, 画像, ロボット情報) VLAM
- 学習データの巨大化



まとめ

1. イントロダクション いま、ロボットに何ができるか？

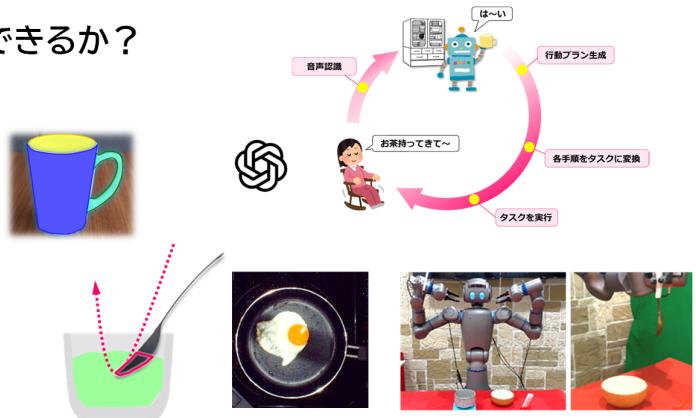
2. お茶会ロボットは、なぜ難しいのか？

3. 解決への道筋

(1) 記号レイヤーでの行動プラン生成

(2) 物理レイヤーへの接続

4. AI時代のロボット研究と期待



連携相談の連絡先（共同研究、学術相談、技術コンサル）

中京大学 工学部 橋本学 mana@isl.sist.chukyo-u.ac.jp

おわり