

記号的AIシステム

(応募者の今までの専門分野)

- ・記号入力(考える対象問題のよく整理された説明)
→使う前に人の手による詳細な説明が必要で、使うのは大変
- ・論理操作(古典論理,ベイズ論理,時制論理)に基づく
→正しさが保証された思考
- ・思考の過程を説明可能なシステム(理解しやすい,安心できるAI)
- ・将棋/囲碁のように将来のことを考えて意思決定
→目前にある誘惑には騙されない/より賢く、信頼できる
- ・宇宙船・社会インフラなど、ミスの許されない対象に使える

これからの研究/抱負

未踏分野

記号/非記号複合AIシステム

(neural-symbolic hybrid)
カメラ/マイクを通して世界を理解
論理に基づいて行動できる柔軟で信頼できるシステム

外界

カメラ/マイク
論理による判断



東京大学大学院 博士3年 浅井 政太郎

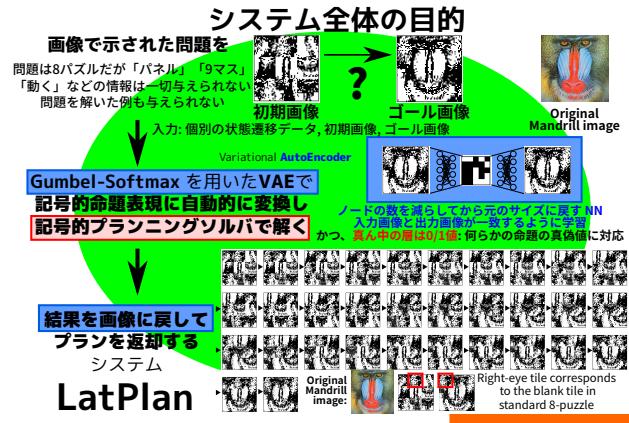
非記号的AIシステム

(近年もてはやされている分野)

- ・ニューラルネット(NN)・深層学習・強化学習
- ・画像・音声・文章を直接入力
→写真などをそのまま使えるので使いやすい
- ・データからの学習に基づく
→正しさがない/学んだ勘にもとづいて行動
- ・説明困難(大抵うまくいくが失敗もする、理由は不明)
- ・学習結果に基づく反射的/直後の/近視眼的行動決定
→乱雑な環境で生き残る虫/動物のような柔軟単純知能



外界を記号表現に変換するNN+記号推論による意思決定



プロトタイプシステム(完成済み)のアーキテクチャ



短期(5年間)/長期の研究計画

短期目標: LatPlan の基礎的性質/拡張性/応用性の研究

1年目: LatPlan基礎の完成 → 2年目~ : LatPlanの適用範囲の拡大

アーキテクチャの妥当性を証明

少数データから汎化できるPDDL生成機

現在のLatPlanの命題生成:
→汎化能力あり。少数の画像から学習して
全画像を命題に変換可能

現在のLatPlanの状態遷移モデル(PDDL)生成:
→汎化能力なし

→問題を解くには全ての状態遷移が必要

Domain Acquisition (Yang07, Konidaris14 etc.)
→汎化能力ありのPDDL生成機

画像で示された解の検証の自動化

現状は手作業で確認/学習記の誤差に依存

実世界タスクでの有用性を実証

State of the Art 画像認識ネットワークの利用
→実世界写真データへの適用

推論空間の表現力の拡張

一階述語論理 $\text{pred}(\text{obj1}, \text{obj2})$ の学習
→R-CNNによるオブジェクト認識

多様なデータ種別への拡張

→音声用オートエンコーダにより
音声に対する論理推論
→テキスト用オートエンコーダにより
テキストに対する論理推論

4年目~ 実運用を見据えた 非古典的モデルへの発展

不確実性を扱う推論機の使用

PO/MDPプランナによるロバストな意思決定

並列処理を扱う推論機の使用

スケジューリング問題

階層的タスクネットワークの自動生成
より複雑な問題設定への適用

長期目標: Latplan を記号/非記号 複合システムの「業界標準」に確立

State of the Art とは?

- ・ハイブリッド車といえば「トヨタのプリウス」 第一人者は強大
- ・CPUといえば「Intel の x86」 な産業インパクトを得られる
- ・スマホといえば「Apple の iPhone」

産業応用

言葉を真の意味で理解するAI



現状の会話人工知能(ボット)は反射的返答
会話を「理解しているように見せている」だけ

- ・Apple Siri、Microsoft りんな

→人の会話を理解(=論理的に考えられる)して
対話するシステム HAL2000: 映画2001年宇宙の旅に登場した人工知能。会話やカメラを通して状況を理解/問題を解決でき、チェスも強い↑

安全/柔軟/人と協調出来る産業用機械

今までの産業用機械: 柔軟でない

→人が一から十まで指示して同じ動作を繰り返す

→工場に人が紛れ込んだら?

→巻き込み事故か 異常検知で全工場がストップ

複合システムがあれば?

カメラから外部を理解でき、かつ論理的な機械

観測・計画・実行による自律行動が可能

→作業方法(計画)は自分で考えてくれる

人間は作業の目的だけを提示(あれやつて!)

想定外(人, 自然現象)に対して自力で対応

→人間との協調行動が可能なロボット

