

JSAI2020 企画セッションKS09

次世代AI研究開発(2)さらなる進化に向けて 趣旨説明

2020年6月9日

科学技術振興機構(JST) 研究開発戦略センター(CRDS)

福島俊一

https://researchmap.jp/toshikazu_fukushima
toshikazu.fukushima@jst.go.jp



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

本企画セッションの開催趣旨

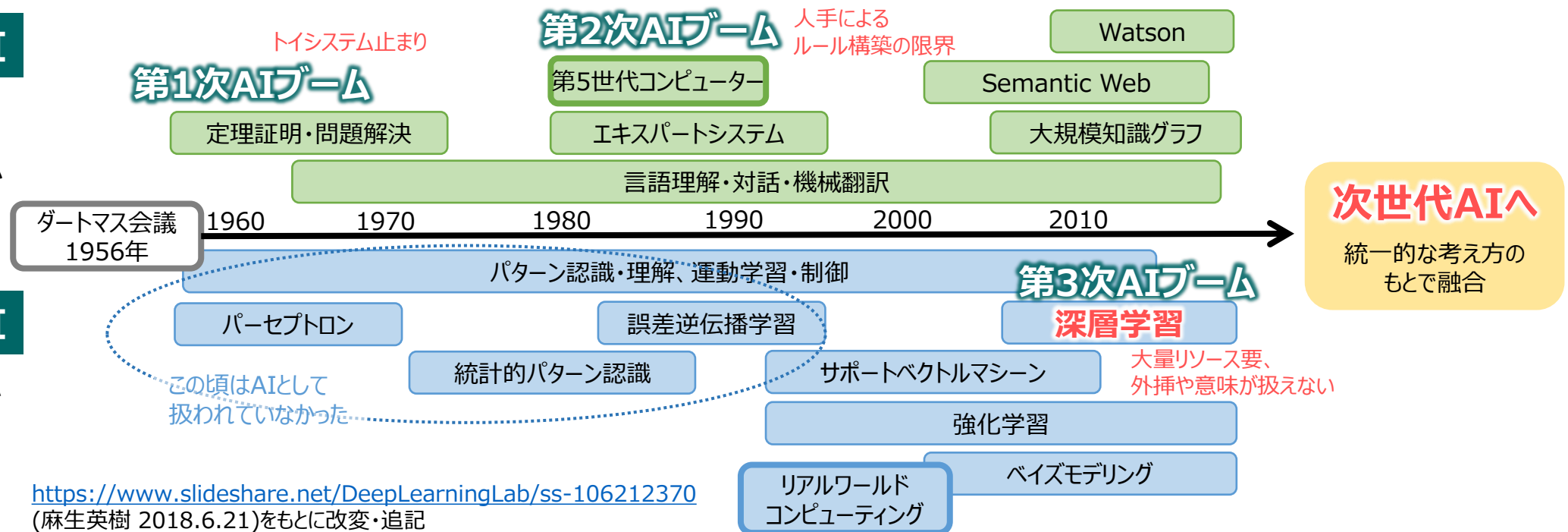
深層学習によって高精度なパタン認識・生成が可能になっただけでなく、柔らかな記号処理も可能になってきました。これまで、データから学習する「即応的知能」と記号・論理を用いる「熟考的知能」という人間の知能の2つの側面が、別々のAIシステムとして技術発展を遂げてきましたが、それらが統一的な考え方のもとで融合する可能性が見えてきたのです。本企画セッションでは、このようなAI進化の方向性について、招待講演2件とパネル討論によって論じます。

熟考的AI

シンボリズム
言語・論理系
演繹型が中心

即応的AI

コネクショニズム
知覚・運動系
帰納型が起点



次世代AIの方向性・期待

現状の課題

- ① 学習に大量の教師データや計算資源が必要



- ② 学習範囲外の状況に弱く、実世界状況への臨機応変な対応ができない



- ③ パターン処理は強いが、意味理解・説明等の高次処理はできていない

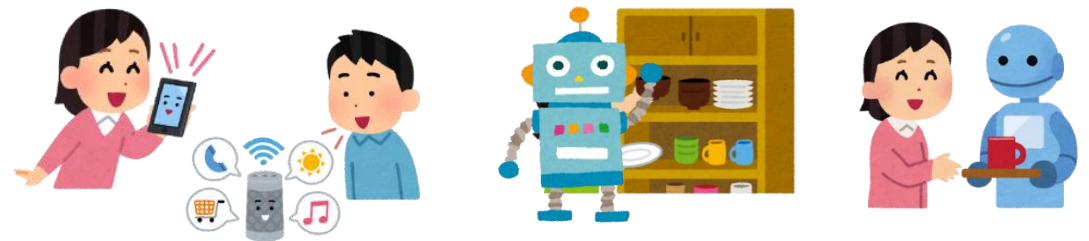


次世代
AI

即応的知能と熟考的知能を統一的に扱う

- 深層学習と知識・記号推論の融合
- 人間の知能からのインスパイア

- ① 教師なしでも知識獲得・成長できる
② 実世界で安全性・頑健性を確保できる
③ コンテキストに応じ適切な推論・アクション可能



- ✓ 解釈性や対話性が高まり、人間との親和性が向上
- ✓ AIシステムの安全性や開発効率も向上
- ✓ 言語・精神疾患の理解 等へもつながる

注目動向-1

NeurIPS 2019でのYoshua Bengio講演: From System1 Deep Learning to System2 Deep Learning

<https://videos.neurips.cc/video/slideslive-38922304>

System1 (即応的知能)	System2 (熟考的知能)
● 高速 ● 直感的 ● 無意識的 ● 非言語的 ● 習慣的	● 低速 ● 論理的 ● 意識的 ● 言語的 ● 計画・推論

現在の深層学習

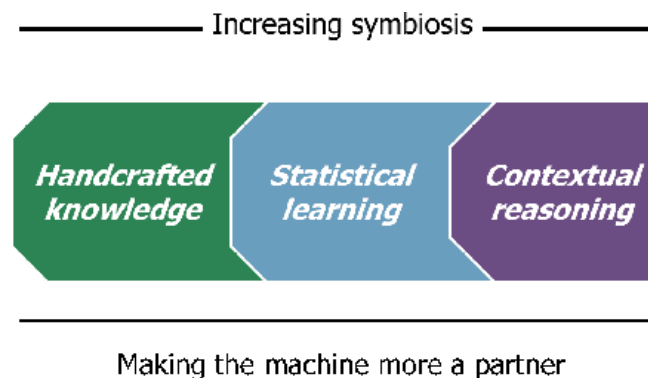
将来の深層学習

System1 & 2はDaniel Kahneman (行動経済学の創始者、2002年ノーベル経済学賞)が、有名な著書「Thinking, Fast and Slow」において、人間の思考には2種類のモードがあると述べたもの。

AAAI 2020では、2018年チューリング賞の3人(Godfathers of deep learning)とKahnemanの対談も行われた。

DARPA (米国国防高等研究計画局): “AI Next Campaign”

<https://www.darpa.mil/work-with-us/ai-next-campaign>



DARPAが2018年に発表、次世代AI研究に20億ドル以上の大型投資を実施

第1の波 : Handcrafted Knowledge (エキスパートシステム)
専門家の知識をキャプチャして関心のある状況に適用

第2の波 : Statistical Learning (深層学習)
人間が提示するデータセットから一般的な知見を得る

第3の波 : Contextual Reasoning (文脈を理解し合理化する機械)
人間と機械の共生を促進
人間が把握・理解・行動する以上の速度でデータを生成することで、
・複雑でタイムクリティカルな環境におけるより良い意思決定を促進する
・大規模、不完全かつ矛盾した情報を共有することを可能にする
・安全かつ高度な自律性にて、無人のシステムが重要なミッションを実行できるようにする

注目動向-2

東大公開シンポジウム「深層学習の先にあるもの－記号推論との融合を目指して－」

第1回: 2018年1月22日



第2回: 2019年3月5日



講演者:

中島秀之、松尾豊、谷口忠大、麻生英樹、
尾形哲也、山川宏、國吉康夫、浅井政太郎

第1回の開催趣旨説明文:

「現在深層学習 (Deep Learning) がAIに革命をもたらしています。深層学習は機械学習のための大変強力な道具ですが、それだけでAIシステムの全てが実現できるわけではありません。ボトムアップの機械学習とトップダウンの推論システムの融合により強力なAIシステムが構築できると考えていますが、その具体的手段はまだ確立されていません。現在、その融合に関心を持つ研究者たちによるワークショップ形式の議論を公開で行うことにより、日本のAI研究の未来像を示したいと考えています。」

第2回の講演動画(東大TV)

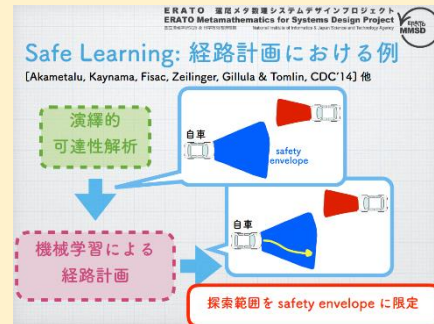
https://today.tv/contents-list/2018FY/beyond_deep_learning

融合型AIの研究例-1

戦略プロポーザル「第4世代AIの研究開発—深層学習と知識・記号推論の融合—」からの抜粋
<https://www.jst.go.jp/crds/pdf/2019/SP/CRDS-FY2019-SP-08.pdf>

① 帰納・演繹の疎な組み合わせ

- 演繹的に作られたシステムの枠内で、機械学習によって探索やスコアリングを最適化
- あるいは、機械学習モジュールの振る舞いを演繹的に検証
- DeepMind AlphaGo、統計的機械翻訳、Safe Learning等



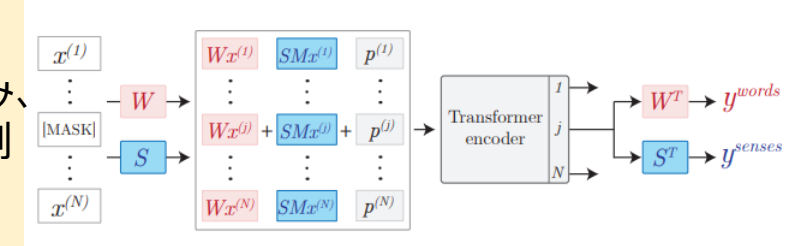
<http://group-mmm.org/~ichiro/talks/20180924okayama.pdf>

② 深層学習への知識の組み込み

SenseBERT(AI21 Labs)

- BERTに、WordNetを組み込み、コンテキストから単語の意味予測を可能に

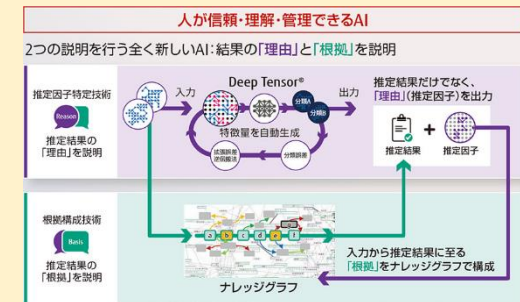
<https://arxiv.org/pdf/1908.05646.pdf>



Deep Tensor + ナレッジグラフ(富士通研)

- 深層学習により推定結果を出力するとともに、推定に強く影響した因子をもとにナレッジグラフを参照して推定根拠を出力し、説明可能性を向上

<https://blog.global.fujitsu.com/jp/2018-12-27/01/>



③ 演繹世界への変換

科学的法則式発見(阪大)

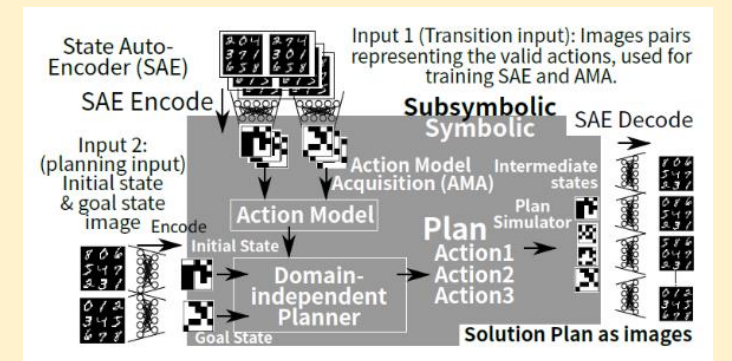
- 能動的観測を通して客観性・普遍性・再現性・健全性・簡潔性・数学的許容性を検証することで、帰納的な実験式を法則式へ

[鷲尾・元田2000]「スケールタイプ制約に基づく科学的法則式の発見」(人工知能15-4) 他

一階述語論理変換(IBM)

- 画像をオートエンコーダで一階述語論理に変換し、初期状態からゴール状態に至る計画を既存ソルバを用いて生成
- 初期状態とゴール状態も画像で与える

<https://arxiv.org/pdf/1902.08093.pdf>



融合型AIの研究例-2

戦略プロポーザル「第4世代AIの研究開発—深層学習と知識・記号推論の融合—」からの抜粋
<https://www.jst.go.jp/crds/pdf/2019/SP/CRDS-FY2019-SP-08.pdf>

④ シミュレーションと帰納的手法の組み合わせ

データ同化

- ある未知のパラメータを含んだ方程式に基づいたシミュレーションの結果を、実際の観測値を用いて補正

シミュレーションと機械学習の融合

- 希少事象、複雑系、特殊・極限状態等、データが取れない/とても取りにくい世界における設計・計画等を高度化・最適化

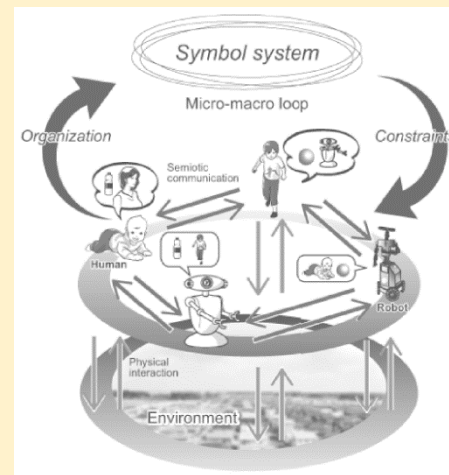


[NEC-産総研連携研究室]

⑥ 記号創発ロボティクス

- 身体を基盤とした主体が環境と相互作用することで概念や言語を獲得していく過程・原理を、機械学習やロボティクスの技術を用いて実現することを通して探る研究分野

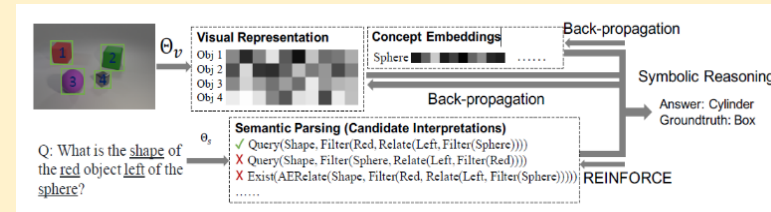
[谷口2016]「記号創発問題—記号創発ロボティクスによる記号接地問題の本質的解決に向けて—」(人工知能 31-1)



⑤ ニューロ系とシンボル系・推論系の融合

Neuro-Symbolic Concept Learner (MIT)

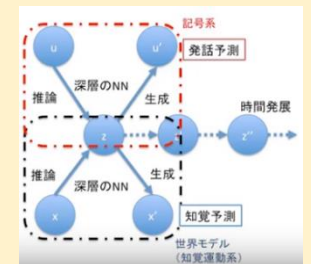
- 子供が環境と少数の教師情報やQAから学ぶようなシステムとして、画像・空間に関する教師あり学習と、QAの強化学習という2系統を統合、マルチタスクのカリキュラム学習



<https://arxiv.org/pdf/1904.12584.pdf>

知覚運動系と記号系の2階建てモデル(東大)

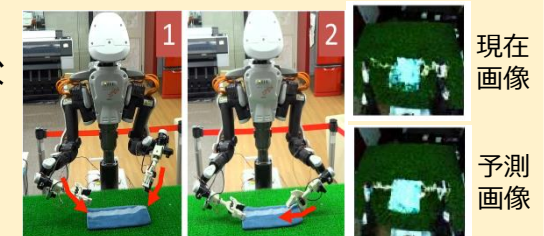
- 上側は発話予測、下側は知覚予測、上下で潜在構造を共有し、2つの予測問題をマルチタスク学習



[松尾2019]「意味理解と創造」(深層学習の先にあるものシンポジウム2)
https://todai.tv/contents-list/2018FY/beyond_deep_learning/01

予測学習(早大)

- 環境とアクションにより、どのような結果になるかを予測、実際の結果を観測し、予測した結果と実際の結果の差異から学習



[尾形2017] <https://ieeexplore.ieee.org/document/7762066>

文科省 令和2年度 戦略目標「信頼されるAI」

https://www.mext.go.jp/b_menu/houdou/2020/mext_00487.html

人間中心社会 with Trusted Quality AI [AI戦略2019]

人間を考える、人間を知る、人間から学ぶ

- ・「人間中心AI社会原則」、ELSI考慮
- ・人間・社会の理解(人文・社会科学)
- ・脳情報処理の研究成果・知見

人間の主体的意思決定の問題 (フェイク・改竄、思考誘導等)

- フェイク検知、改竄防止
- 主体的意思決定支援 等

AIシステムの信頼性・安全性の問題 (脆弱性・解釈性、差別・品質問題等)

- AI品質評価、安全設計
- 説明可能AI、公平性保証 等

現在のAI技術の限界 (大量学習、常識等)

- 深層学習(即応的AI)のさらなる進化
- 知識・記号推論(熟考的AI)との統合
- 実世界環境で発達・成長 等

[AIに対する社会的要請への対応]

- ・AI技術そのもののさらなる発展・革新も見据え、かつ、社会からの要請(人間中心のAI社会原則)を満たす、人々が信頼できるAIの実現

▶ データ信頼性確保、集合知醸成 (主体的意思決定支援)

▶ AIシステムの信頼性・安全性確保 (AIソフトウェア工学/機械学習工学)

[AI技術そのものの発展・革新]

▶ 現在のAI技術の限界克服 (次世代AIアーキテクチャ)

- ・大量教師データがなくとも、実世界環境中で効率的かつ効果的に処理・判断
- ・文脈・常識等を理解し、人間とより適切にコミュニケーション、迅速にアクション実行

→ t

現在募集中のJSTプログラム

■ JST戦略的創造研究推進事業**CREST「信頼されるAIシステム」**研究提案募集

https://www.jst.go.jp/kisoken/boshuu/teian/top/ryoiki/ryoiki_c10.html

研究総括：相澤彰子(国立情報学研究所 コンテンツ科学研究系 教授／所長補佐)

応募〆切：2020年6月23日(火)正午※厳守

■ 同さきがけ**「信頼されるAI」**研究提案募集

https://www.jst.go.jp/kisoken/boshuu/teian/top/ryoiki/ryoiki_p13.html

研究総括：有村博紀(北海道大学 大学院情報科学研究院 教授)

応募〆切：2020年6月23日(火)正午※厳守

■ 同**ACT-X「AI活用で挑む学問の革新と創成」**研究提案募集

https://www.jst.go.jp/kisoken/boshuu/teian/top/ryoiki/ryoiki_a03.html

研究総括：國吉康夫(東京大学 大学院情報理工学系研究科 教授)

応募〆切：2020年6月23日(火)正午※厳守

■ JST**未来社会創造事業「超スマート社会の実現」**領域

重点公募テーマ**「異分野共創型のAI・シミュレーション技術を駆使した健全な社会の構築」**

<https://www.jst.go.jp/mirai/jp/uploads/application-guideline-r02-c6.pdf#page=6>

テーママネージャー：鷺尾隆(大阪大学 産業科学研究所 教授)

応募〆切：2020年6月30日(火)正午※厳守

戦略目標のもとになった JST CRDSで発行した報告書や企画・開催したワークショップ等

- ① 戦略プロポーザル: 第4世代AIの研究開発—深層学習と知識・記号推論の融合— 2020年3月
<https://www.jst.go.jp/crds/pdf/2019/SP/CRDS-FY2019-SP-08.pdf>
- ② 科学技術未来戦略ワークショップ報告書: 深層学習と知識・記号推論の融合によるAI基盤技術の発展 2020年3月
<https://www.jst.go.jp/crds/pdf/2019/WR/CRDS-FY2019-WR-08.pdf>
- ③ 戦略プロポーザル: AI応用システムの安全性・信頼性を確保する新世代ソフトウェア工学の確立 2018年12月
<https://www.jst.go.jp/crds/report/report01/CRDS-FY2018-SP-03.html>
- ④ 科学技術未来戦略ワークショップ報告書: 機械学習型システム開発へのパラダイム転換 2018年3月
<https://www.jst.go.jp/crds/report/report05/CRDS-FY2017-WR-11.html>
- ⑤ 2019年度人工知能学会全国大会(第33回)企画セッション: 機械学習における説明可能性・公平性・安全性への工学的取り組み 2019年6月
https://www.jst.go.jp/crds/sympo/201906_JSAI/index.html
- ⑥ 戦略プロポーザル: 複雑社会における意思決定・合意形成を支える情報科学技術 2018年3月
<https://www.jst.go.jp/crds/report/report01/CRDS-FY2017-SP-03.html>
- ⑦ 科学技術未来戦略ワークショップ報告書: 複雑社会における意思決定・合意形成を支える情報科学技術 2017年10月
<https://www.jst.go.jp/crds/report/report05/CRDS-FY2017-WR-05.html>
- ⑧ 公開ワークショップ報告書: 意思決定のための情報科学～情報氾濫・フェイク・分断に立ち向かうことは可能か～ 2020年2月
<https://www.jst.go.jp/crds/pdf/2019/WR/CRDS-FY2019-WR-02.pdf>
- ⑨ 研究開発の俯瞰報告書: システム・情報科学技術分野(2019年) 2019年3月
<https://www.jst.go.jp/crds/report/report02/CRDS-FY2018-FR-02.html>

本企画セッションのAgenda

15:20～15:30	趣旨説明	福島 俊一（科学技術振興機構）
15:30～16:00	講演1 「世界モデルと記号処理」	松尾 豊（東京大学）
16:00～16:30	講演2 「確率的深層生成モデルによる実世界自律知能の創成に向けて ～記号創発ロボティクスが生み出す認知アーキテクチャ～」	谷口 忠大（立命館大学）
16:30～17:00	パネル討論	中島 秀之（札幌市立大学） および 上記登壇者