

論文読みメモ

Neuroscience-Inspired Artificial Intelligence

2017/12/28 Yasunori Endo

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis,^{1,2,*} Dharshan Kumaran,^{1,3} Christopher Summerfield,^{1,4} and Matthew Botvinick^{1,2}

どんなもの？

神経科学がAIに与えた影響を、Past/Present/Futureの3つの時間軸でまとめたリサーチ論文。

どうやって有効だと検証した？

数々の神経科学の研究がAIに影響を与えてきた歴史が有効性を物語っている。

先行研究と比べてどこがすごい？

ここまで詳しく調べてある論文はすごい。

議論はある？

神経科学とAIの相互の連携を緊密にし “virtuous circle” を回していこうという提案。

技術や手法の肝は？

新たな手法を提示する論文ではないので割愛。

次に読むべき論文は？

いろいろあるが、特に気になるものは以降のスライドで著者名を出しているのでそれを見るとよい

読んだ論文について

- 本文書は下記論文の精読メモである

Neuron Review

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis,^{1,2,*} Dharshan Kumaran,^{1,3} Christopher Summerfield,^{1,4} and Matthew Botvinick^{1,2}

¹DeepMind, 5 New Street Square, London, UK

²Gatsby Computational Neuroscience Unit, 25 Howland Street, London, UK

³Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK

⁴Department of Experimental Psychology, University of Oxford, Oxford, UK

*Correspondence: dhcontact@google.com

<http://dx.doi.org/10.1016/j.neuron.2017.06.011>

用語の定義

- AI(人工知能): Artificial Intelligence
 - 知的な機械を作ろうとするあらゆる研究
 - 機械学習
 - 統計
 - etc.
- NS(神経科学): Neuroscience
 - 脳の振る舞いやメカニズムに関するあらゆる研究分野
 - cognitive neuroscience
 - systems neuroscience
 - psychology
 - etc.

AIとNSの関係について

- もともとAI・NS・心理学が密接に関わっていた(50's - 80's)
- 双方の研究が進んでいくに連れて、分化が進んでいった
- NSを研究すると、AI研究にとって何が嬉しいのか
 - 新しいアルゴリズムやアーキテクチャの発想のもととなる
 - 既存のAIアーキテクチャの検証ができる
- AIが必ずしもNS的に妥当である必要はないが、一つの指標となる

Marr's 3 Levels of Analysis

- Marrは脳機能を3つのレイヤーに分けて考えた
- AIの文脈では、上位2層を重視している
 - 脳機能のメカニズムが注目点であり、アルゴリズムレベルでわかればそれで十分

computational level

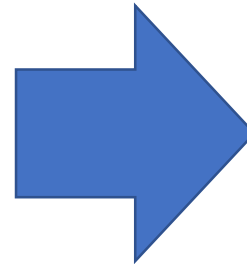
脳はいったい何を計算しているか？

algorithmic level

具体的な計算をどのように実現しているか？

implementation level

脳の特定のアルゴリズムを物理的にどう実現しているか？



これらの *in silico* で
の実現を目指す

完全な脳の再現を目指す Blue Brain
プロジェクトといった例もある

論文の構成について

- From NS to AI...
 - The Past: AI研究の起源を探る
 - The Present: 現在のAI研究とNSとの関連とは？
 - The Future: NSのもつポテンシャルとは？
- From AI to NS...

The Past

- Deep Learning
- Reinforcement Learning

The Past (1): Deep Learning

- ANN(Artificial Neural Network)は、言葉の通りNSが起源
 - 1943 McCulloch et al.: ニューロンによる論理演算
 - 1949 Hebb: オートエンコーダ的サマシング?
 - 1958 Rosenblatt: パーセプトロン
 - 1985 Rumelhart et al.: backpropagation under PDP movement
- その当時のAI研究は…
 - シリアルな論理・記号処理に注力していた
 - しかし、記号処理は現実の複雑な問題に対処するには不適當だった

Parallel Distributed Processing

- Proposed in 1986 by Rumelhart et al.
- 脳が確率的かつ並列に情報を処理しているという知見
- ニューロンのようなシンプルな演算ユニットを組み合わせ学習させることで、人間の認知と振る舞いが生まれると考えた
- PDPの研究がAIに影響を与えた例は多い
 - Distributed (i.e. as vectors) representation of word/sentences
 - Hubel et al., 1959 視覚野V1の研究が、CNNに影響を与えた

ディープラーニングの発達とNS

- NS continued inspiring AI after PDP period
 - Deep Belief Network by Hinton et al., 2006
 - ImageNet dataset
 - Inspired by research on human language by Deng et al., 2009
 - Invention of “dropout” technique by Hinton et al., 2012
 - ポアソン分布的に発火するニューロンから着想を得た
 - Hinton の論文にはぱっと見書いてなさそうだったけどな(◡•.̣•◡)

The Past (2): Reinforcement Learning

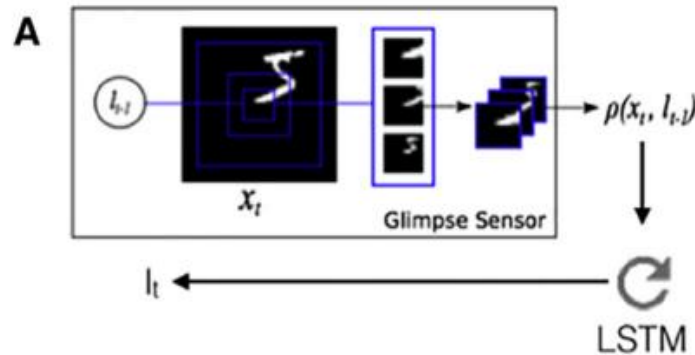
- Temporal-difference (TD) method
 - 条件付け実験(パブロフの犬)と密接に関連
- AIの進歩のコアとなっている

The Present

- Attention
- Episodic Memory
- Working Memory
- Continual Learning

The Present (1): Attention

- 多くのCNNモデルが入力全体を同様に扱う一方、霊長類の視覚は場所や物体に注目しながら扱っている
- Mninh et al., 2014: “glimpses” architecture がクラッタをうまく処理できた



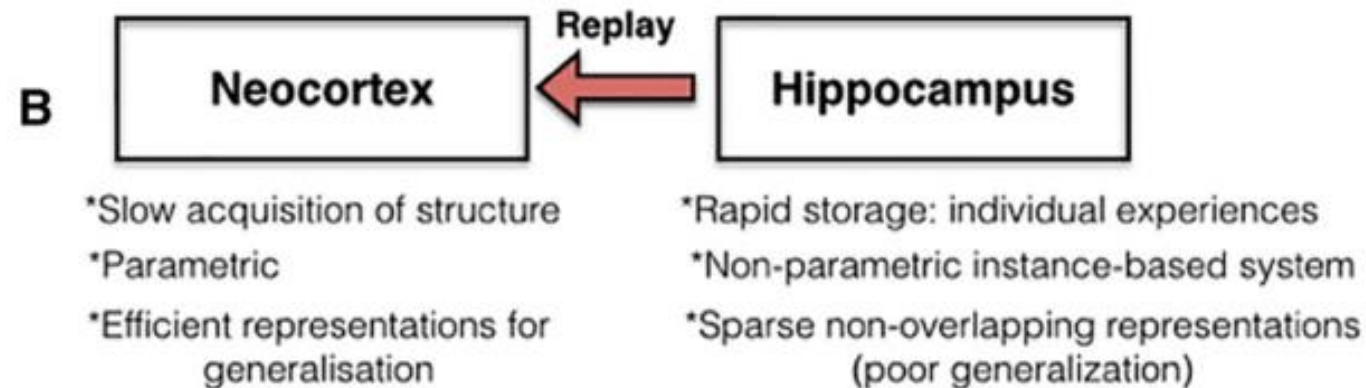
- attention メカニズムを適用することにより、複数物体認識 (Ba et al., 2015) などで単純なCNNよりも高性能を実現

Attention in memory

- attention は知覚だけでなく、脳の内部の記憶を読む際にも使われている
- このメカニズムが機械翻訳等に応用されている
- “DRAW” をはじめとする生成モデルにも役立っている
 - “mental canvas” に、部分ごとでインクリメンタルに画像を生成するモデル

The Present (2): Episodic Memory

- エピソード記憶は “instance-based” & “one shot” な記憶
 - 海馬での記憶が “one shot” な記憶に相当
 - 睡眠・休憩中に海馬から新皮質で記憶が再生され、固定される
 - Favor events that lead to high levels of reinforcement



Episodic Control

- 経験に基づいた急速な振る舞いの変化
- 海馬では、rewarded action sequences can be internally re-enacted from a rapidly updateable memory store

Episodic Memory in AI

- Deep Q Network (DQN) における “experience replay” が脳における記憶の流れを *in silico* で再現している
 - Replay buffer が海馬に相当
 - 報酬の高いイベントを優先した experience replay がより効果的
- ディープなRLは学習が遅いという問題があったが、 episodic control を実装することによって解決した (Blundell et al., 2016)
- エピソード記憶的なシステムは、少ないサンプルから新しい概念を学習できる可能性がある

The Present (3): Working Memory

- 作業記憶は前頭葉とその近辺に存在し、情報を維持・操作する能力が人間の知性の特性となっている
 - 中央的なコントローラと周辺のドメインに特化したメモリ領域によって構成されていると考えられている
- LSTMとDNCが作業記憶に相当
 - LSTM: Long Short Term Memory
 - DNC: Differential Neural Computer

The Present (4): Continual Learning

- 人間を始めとする知的なエージェントは、異なる時間軸で異なるタスクを学習し、思い出していく必要がある
 - つまり、以前のタスクでの学習を忘れることなく、新しいタスクを学習できるということ
 - 以前のタスクに関する知識を守るメカニズムが見つまっている
- ANNには “catastrophic forgetting” の問題が知られていた
- EWC (Elastic Weight Consolidation) proposed by Kirkpatrick et al., 2017, is a solution
 - ネットワークの重みのうち、以前のタスクで重要な一部分を固定する

The Future

- Intuitive Understanding of the Physical World
- Efficient Learning
- Transfer Learning
- Imagination and Planning
- Virtual Brain Analytics

The Future (1): Intuitive Understanding of the Physical World

- 人間の場合、空間・数・物体感といった概念が乳児の段階にすでによく発達する一方で、ほとんどの AI は実現できていない
- ある ANN は、個々の物体とその関係に分解することで、シーンを解釈することができる
- ディープなRLを用いて、子供が対話的な実験を通して常識を理解する仮定を捉えることができた (Denil et al., 2016)

The Future (2): Efficient Learning

- 人間の認知は、事前知識を元に自在に用いることで数少ない事例から新しい概念を学習できる点が特徴としてあげられる
- この認知機能に関する課題として Lake et al. が 2016 年に “characters challenge” を作った
- 最近の AI アルゴが、確率的モデル・ディープな生成モデルの両面で進歩があった

The Future (3): Transfer Learning

- 人間は経験の汎化と移転にも長けている
 - 車を運転できる人が、別の種類の車にも乗れるなど
- “Progressive Network” が成果を上げている (Rusu et al., 2016)
 - あるゲームに関する学習を利用して、他のゲームの学習時間を短縮
 - シミュレーション環境のロボットの学習を利用して、実世界のロボットの学習時間を短縮
- 移転学習の文脈で、関係性をもとに推論する能力も考えられる
 - 一般的に人間がどのように高次の移転学習をしているかはわかっていない

The Future (4): Imagination and Planning

- 動物は将来の報酬をシミュレートして、柔軟に行動を選択する
 - 内部に環境モデルを構築し、シミュレーションを行っていると考えられる
 - 例えばラットが選択点で止まっている時に、後に通行可能な軌跡を通っている時に観測される海馬の活動と似たような活動が観測される
- ディープな生成モデルがここでも活躍する
 - 現実空間のレイアウトをもとに連続したサンプルを生成できるモデルが提案された (Gemici et al., 2017)
 - シミュレーションの結果をエージェントの行動計画に用いる手法はまだ存在していない
- 海馬でどのように内部モデルが存在しているかはまだ不明

The Future (5): Virtual Brain Analytics

- AIシステムの“ブラックボックス”を、神経科学の手法で解明
 - 単独細胞の記録、イメージング等
- 次元削減によるニューラルネットワーク(DQN)の可視化 (Zahavy et al. 2016)

From AI to NS

- 機械学習の手法を用いた神経画像(fMRI/MEG)の解析
- 脳とTD学習の関連
- CNNと ventral visual stream における情報表現の関連
- Backpropagation と脳における学習の違い
 - フィードフォワード・フィードバックの対称性が脳では観測されない
 - 脳は局所的な信号を用いて学習する一方、BPでは非局所的な誤差信号を用いる

所感

- NS が AI の発展に重要な役割を果たしていることが伝わった
- AI が思いの外幅広く脳機能をカバーしつつあることがわかった
- 引いている論文が非常に多く、筆者のリサーチ力がすごい
 - 読むべき論文が芋づる式に見つかるので、一度読んだほうがいい
- DeepMind 強い
 - 筆者を始め、引かれている論文も DeepMind 発が多かった