



NPO法人全脳アーキテクチャ・イニシアティブ

HP: https://wba-initiative.org/

Twitter: @wba_meetings, @wba_initiative

Facebook: https://www.facebook.com/WBA.Initiative/

基本理念

ビジョン: 人類と調和した人工知能のある世界

ミッション:

全脳アーキテクチャのオープンな開発を促進

価値観:

まなぶ: 関連する専門知識を学び、拡める

みわたす: 広く対話を通じて見識を高める

つくる: 共に作り上げる

<u>目次</u>

発刊にあた	ːり	3
2020年度を	を支えてくださった皆様	4
2020年度	の活動方針と予算	5
2020年度	か活動実績	6
財務状況···		9
2021年度	の活動に向けて	10
私とWBAI		12
おわりに		16
参考文献		17
別表 1		19
別表 2		20

発刊にあたり

高度な知能は、良くも悪くも人間の福祉に大きな影響を与えるでしょう。そして、比較的近い将来、人間のあらゆる知的能力を凌駕する汎用人工知能(AGI)が実現されると私たちは考えています。このような状況の中、私たちは「人類と調和した人工知能がある世界」というビジョンを掲げました。このビジョンを実現するためには、世界の特定の組織がAGIの開発技術を独占してしまうリスクを低減する必要があります。そこで、私たちは、脳型からのAGIのオープンな開発を推進しています。

また、脳科学の知見を活用し、脳の構造に基づいた共同開発を行うことで、AIの開発が加速されることは以前から期待されていました。しかし、その具体的な方法論は、長らく明らかなものではありませんでした。しかし、WBAIの約5年間の成果として、BRA駆動型の開発が生み出されまし



NPO法人全脳アーキテクチャ・ イニシアティブ代表 山川宏

た。これは、脳の情報処理を反映したAGIの外部設計仕様書であるBRA(Brain Reference Architecture)を 設計し、それを参照しながらソフトウェアを開発・評価するという開発手法です。

言うまでもなく、私たちは脳全体を反映したAIを作ろうとしていますので、BRAのデータを拡大して、ほぼ脳全体をカバーすることが当面の目標です。そのためには、標準化されたBRAのデータを広く普及させ、脳科学と情報学の接点である人間科学の分野に共同研究者を参入させたいと考えています。そのために、学術的な成果の発表や人材育成などを通じて、活動を推進し、周知していきたいと考えています。

中長期的には、BRAを利用して、AGIをはじめとする脳を使ったソフトウェアを開発する様々な組織が生まれる状況を作りたいと考えています。そのためには、BRAのデータを使った開発を実践し、脳型AIの開発における技術的課題を抽出し、それに対応するノウハウを蓄積していきたいと考えています。またAGIがもたらす危険な側面にも配慮し、それをBeneficial で安全なものとするための議論も引き継続してゆくつもりです。

2020年度を支えてくださった皆様

会員は正会員(社員)と賛助会員からなります。さまざまな企業をはじめとする賛助会員の方に財政的な ご支援をいただきました。

賛助会員一覧

※各カテゴリー内の順番は入会順です。

創設替助会員

※5年間限定のため2020年度までに終了

株式会社ドワンゴ

パナソニック株式会社

トヨタ自動車株式会社

株式会社IPパートナーズ

株式会社東芝

替助会員

後藤健太郎様

株式会社IGPIビジネスアナリティクス &インテリジェンス

匿名希望様

上林厚志様

マテリアライザー合同会社

株式会社TOPWELL

株式会社三菱総合研究所

上野山勝也様

川上量牛様(2020年度入会)

パナソニック株式会社コネクティッド ソリューションズ社様(2020年度 入会)

顧問

- 銅谷 賢治(沖縄科学技術大学院大学神経計算ユニット教授)
- 北野 宏明(特定非営利活動法人システム・バイオロジー研究機構代表)
- 冨田 勝(慶應義塾大学環境情報学 部教授)
- 森川 博之(東京大学先端科学技術 研究センター教授)
- 中島 秀之 (札幌市立大学学長)
- 岡ノ谷一夫様(東京大学広域科学 専攻生命環境科学系認知行動科学 講座教授)

役員

- 代 表(理事):山川宏
- 副代表(理事):松尾豊(東京大学 大学院丁学系研究科教授)
- 副代表(理事): 髙橋 恒一(理化学研究所生命システム研究センターチームリーダー)
- 理事:荒川 直哉(事務局長兼務)
- 監事:上林厚志(竹中工務店・ 2019年7月より)

正会員

正会員(社員) 16名 (2021年3月末時点)

2020年度の活動方針と予算

当法人では、脳型AGIの開発を促進するという目的に即して、教育事業と研究開発事業をすすめていますが、 当初の試みを通じて、直接的に脳型AGIの開発を行いうる神経科学と情報技術の両面に精通した人材育成が必 ずしも容易でないことが明らかになりました。そこで2017年度以降は、神経科学分野の知識を、開発要求仕 様書である脳参照アーキテクチャ(Brain Reference Architecture: BRA)¹として整理し、AI/ML専門家に 提供するという開発方法の構築を進めています。2020年度においては、その記述方法の整備などを行いまい した。以下、事業分野ごとに方針を記します。

教育事業

教育事業は、脳型のAGIの研究開発に必要な人工知能、神経科学、認知科学、機械学習などの異なる複数の専門性を同時に備えた学際的な人材を育成する事業です。2020年度も前年度の活動を継続し、勉強会の実施、および当法人の活動の進展状況をアピールするためにシンポジウムの実施、国内外の外部学術イベントへの参加・協力、外部学術団体との協力・情報交換などを継続することとしました。また、外部の研究者・技術者およびグループとの協力により人材育成、巻き込みをはかることを継続することとしました。WBAへの興味を持つ研究者・技術者・学生の増加をはかるため、ハッカソンを開催することとしました。また、例年通りWBAI奨励賞を授与し、脳型AGIの(国内外の)技術開発の促進において、波及効果の高い開発成果を残した者を評価することで、コミュニティの活性化をはかることにしました。

研究開発事業

研究開発事業の目標は、全脳アーキテクチャ・アプローチによる研究を支援することです。2020年度は、全脳アーキテクチャ関連の研究開発を行う機関や研究者の次のような取り組みの支援をすることにしました。 具体的には、脳の解剖学的構造を記述したBIFと、それに整合的な機能の仮説的な機序を記述したHCDからなるBRA形式に従って、脳型AIの開発に役立つ形での神経科学知見の蓄積を進めることとしました。これにより特定のタスクにおける脳型AIの作成を促進することを期待しました。またそうして構築された脳型AIの実装を評価するために、BRAを活用しながら脳型AGIの神経科学的妥当性を評価する技術の開発も進めることとしました。

さらに、引き続き全脳アーキテクチャ構築のための統合ソフトウェアプラットフォーム開発も引き続き行う こととしました。

当法人が開発促進する脳型AGIをより Beneficial なものとすることにむけた活動も継続することとしました。

2020年度の予算

予定収入は約370万円(前期繰越金約757万円を合計すると約1127万円)でした。支出では、管理費に約97万円、事業費にイベント開催費用、研究開発費を含む約277万円、計約373万円を予定を予定しました。実績については以下の「2020年度の活動実績」と「財務状況」を参照ください。

参考: https://wba-initiative.org/wiki/brain reference architecture

¹ BRAは、ヒトなどの脳全体の情報処理を再現する静的な参照モデルであり、機能部品および部品間の結合を網羅的に表した構造情報と、それらの機能的な意味付けを含むものです。

本年度(2020年度)の活動実績

2020年度の活動方針に沿った教育事業、研究開発事業等を実施しました。

教育事業

教育事業の目標は、長期的に全脳アーキテクチャ・アプローチによる研究開発を行える人材を増加させることです。2020年度は、第5回全脳アーキテクチャ・シンポジウムおよび2回の全脳アーキテクチャ勉強会を開催しました。また、全脳アーキテクチャ・アプローチに関わる新たな企画としてはじめての「WBAレクチャー」を開催しました。財政面で勉強会を支えるため、本年度から勉強会を学生以外の参加者に対して有料化することとしました。またコロナ禍という事情もあり、イベントをリモート開催としました。有料イベントは記録動画を原則 Vimeo で有料配信(シンポジウムは無料配信)、無料イベントは YouTube で配信しています。

全脳アーキテクチャ勉強会

全脳アーキテクチャ勉強会は当法人の創設以前から行われてきた活動です。2020年度は勉強会を2回開催しました。

- 第30回:2020年4月20日「汎用AIと共生インタラクション」² 柳澤琢史氏(大阪大学)、今井倫太氏(慶應義塾大学)間瀬健二氏(名古屋大学)による講演と討論 遠隔開催:Zoom Webinar (参加者約200名) (一般参加費500円)
- 第31回:2020年10月23日「予測する脳と主体性の現象学」³ 吉田正俊氏(北海道大学)、田口茂氏(北海道大学)、大羽成征氏による講演と討論 遠隔開催:Zoom Webinar(参加者約130名)(一般参加費1,000円)

第5回全脳アーキテクチャ・シンポジウム4

2020年10月19日

テーマ:「見えてきた知能研究の本丸」

大森隆司(玉川大学)、岡ノ谷一夫(東京大学)、荒川直哉 (WBAI)、山川宏(WBAI)、市瀬龍太郎 (国立情報学研究所)、松尾豊(東京大学)、銅谷賢治(OIST)、高橋恒一(理研)、浅川伸一(東京 女子大学)、一杉裕志氏(産業技術総合研究所 人工知能研究センター)による講演と討論

会 場: Zoom Webinar (参加者約140名) (一般参加費1,000円)

主催: NPO法人 全脳アーキテクチャ・イニシアティブ

協 賛:新学術領域人工知能と脳科学の対照と融合

後 援:脳情報動態を規定する多領野連関と並列処理、「富岳」成果創出加速プログラム:脳結合データ 解析と機能構造推定に基づくヒトスケール全脳シミュレーション、人工知能学会

² https://wba-initiative.org/14903/

³ https://wba-initiative.org/17986/

⁴ https://wba-initiative.org/17373/

なお、本シンポジウムにおいて、動物のもつ多様な常識的推論の研究開発を促進するAnimal AI Olympicsを企画・開催したMatthew Crosby氏とBenjamin Beyret氏、脳型人工知能の基盤となる深層 生成モデルの研究開発において貢献のあった鈴木雅大氏にWBAI奨励賞、当法人における勉強会やシンポジウムの実行委員長を歴任しWBA勉強会実行委員会の創設時から貢献いただいた生島高裕氏に活動功労賞を授与しました。

IJCAI-20でのチュートリアル

2020年1月8日

金井亮太(アラヤ)、渡邉正峰(Mind in a Device)、山川宏(WBAI)による IJCAI-20のチュートリアル「Conscious AI: Significance and Development」 を実施しました [1]。参加者は20名ほどでした。

作業記憶モデラソン

2020年7月~9月

第5回WBAハッカソンに先駆け、作業記憶のモデルを募集した6。

応募が1件あり、評価にあたって生物学的妥当性の検討が進展したため、応募者に重要功労賞を授与⁷ した。

第1回WBAレクチャー

2021年2月7日

テーマ:「認知機能の脳構造に沿った分解手法ー脳機能の体系的理解を目指してー」⁸

山川 宏(WBAI)、布川 絢子(WBAI)、田和辻 可昌(早稲田大学)

会 場: Zoom Webinar (参加者約190名) (無料)

第5回WBAハッカソン~『今、ここ』を超えて

2021年5月~8月

作業記憶を用いた課題を解く全脳アーキテクチャエージェントを作成することをテーマとします。

2020年度中に WBAI のパートナーである Cerenaut の協力により準備を終了し、2021年8月末までオンラインで Cerenaut との共催で開催中です。⁹

研究開発事業

研究開発事業の目標は、全脳アーキテクチャ・アプローチからの脳型AGIの研究開発を促進することです。

開発方法論

私たちは2018年度より脳型AIの開発方法論の議論を進めてきました。現在は、図1に示すように脳参照アーキテクチャ(BRA)を設計するという作業と、それを活用してソフトウエアを開発するという作業へと分割しすすめるBRA駆動開発という形になりました。 BRAは、脳のメゾスコピックレベルの解剖学的構造である脳情

⁵ https://ijcai20.org/tutorials/; https://wba-initiative.org/en/12120/

⁶ https://wba-initiative.org/15968/

⁷ https://wba-initiative.org/18192/

⁸ https://wba-initiative.org/17902/

⁹ https://wba-initiative.org/18626/

報フロー(Brain Information Flow: BIF [2][3][4]) に対して、それに整合的な機能的な計算機能の仮説である仮説的コンポーネント図(Hypothetical Component Diagram: HCD)を付加した静的な参照モデルです(図1)。BIF における Circuit は、脳内のメゾスコピック・レベルのネットワークシステム内の任意の部分ネットワークを表します。特定の細胞タイプの神経細胞グループに対応する Uniform circuit からの出力として、Circuit 間には Connection が設定されます。

仮説的コンポーネント図は、興味の対象となる脳領域(ROI)のメゾスコピック・レベルの解剖学的構造と整合するように ROIが担う機能を組織化したもので、コンポーネント間の依存関係を記述した有向グラフであるUMLのコンポーネント図の一種です。しかし、必ずしも脳の真実と一致しているという保証がないという意味で仮説的です。

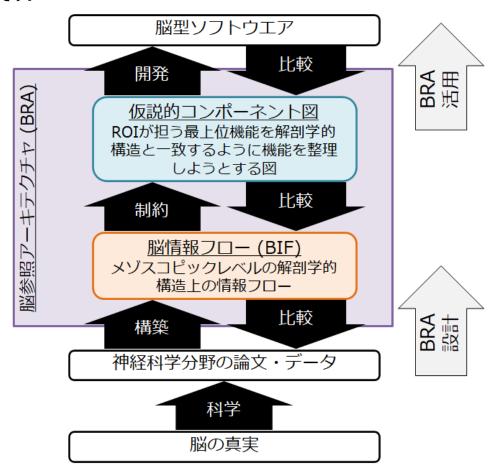


図1: 脳参照アーキテクチャ (BRA) とBRA駆動開発

BRA駆動ソフトウエア開発において直接的に利用されるのは仮説的なコンポーネント図です。ほとんどの脳領域において、コンポーネント図のように組織化された機能の記述を神経科学知見から直接的に得ることはできません。そこで、WBAI は、神経科学知見(主に解剖学的構造~BIF)に整合的な形で仮説的なコンポーネント図を作成する Structure-constrained Interface Decomposition(SCID法)の提案によりBRAの整備を進めており、その全体像を整理しました[5]。

全脳アーキテクチャ・アプローチによる研究を支援するためのソフトウェアなどの研究インフラストラクチャを整備する活動を他の組織と連携・協力しながら行っています。2020年は、以下に述べるような研究開発において共同の活動を行いました。特に新学術領域研究「人工知能と脳科学の対照と融合」とは確率的生

成モデルの利用について連携をしています [6][7][8]。また、新学術領域研究「脳情報動態を規定する多領野連関と並列処理」とも神経情報学の基盤整備(下記)などのテーマで共同の研究を行っています。

神経情報学の基盤整備

BRAの基盤は脳全体におけるメゾスコピックレベルの解剖学的な構造です。そこで、哺乳類において最も充実したメゾスコピックなコネクトームの1つである Allen Institute の Mouse Connectome から信頼性の高い接続情報を抽出するための分析を行っています。

BRA設計 (SCID法)

個別の脳器官に対するSCID法を用いたBRA設計としては、大脳基底核ループについて2019年以来BIF の記述を行ったものに注意と対応づいた仮説を付与し [9]、嗅内野の経路積分機能についての仮説 [10] および眼球運動のBRAの構築 [11][12] を提案しました。さらに海馬周辺の Papez 回路の調査も進めました [13]。

脳の広域の機能としては、新皮質内で流通する信号の一般的な意味についての仮説 [14][15] の検討、前障(claustrum)による注意機能 [16]、AGIに関わる標準化された脳器官についての議論 [17] を進めました。

生物学的妥当性評価

脳型ソフトウエアが、Uniform Circuit程度までの粒度において、神経科学が捉える脳の真実をどれだけ 再現しているのかを見積もる技術が生物学的妥当性評価です。脳に近い設計空間内でAGIを探索する WBAアプローチでは、この技術を確立し、ソフトウエアを評価し続ける必要があります。そこで、 BRAを前提とした場合における生物学的妥当性の評価について検討を進め [18][19]、BRAの既存神経 科学知見への整合性を見積もる適切度と、BRAに対する脳型ソフトウエアの再現性を見積もる忠実度に 分割して行う評価方法を見いだしました。

機械学習の開発

以前からの研究を発展させ、隠されたマクロな時系列を見出しながら、その空間における強化学習についての研究を行いました [20]。

OpenなAI開発コミュニティの形成

Slack (開発者用SNS) 上のコミュニティで議論や情報交換を行っています。

AGIをBeneficialなものとするための活動

開発を促進しているAGIが人類にもたらす影響の大きさから、AGI を Beneficial で安全なものとし、かつ民主化するための活動を行っています。2020年度は、汎用AIが投げかける人間の本質への投げかける問いについてのインタビュー [21] や知能爆発についての発表 [22] を行いました。

活動とボランティア

本年度における活動は、事務局の業務委託職員を除くと、基本的に手弁当の正会員とボランティアによって 行われました。とりわけ、全脳アーキテクチャ勉強会とWBAシンポジウム、WBAレクチャーの開催はボラン ティア主導で行われました。

財務状況

2020年度の貸借対照表(別表1)と活動計算書(別表2)を文書末に示します。

決算収入(当期)は約360万円。支出は管理部門で約82万円、事業部門で約161万円、計約243万円でした(約117万円の黒字)。

経常収益のほとんどは賛助会員年会費から来ています(創設賛助会員2社からは2019年度まで5年分の年会費を前払いいただいていました)。雑収入は、勉強会とシンポジウムの参加者から徴収した参加費を含みます。

事業費のうち外注費は事務局費用の50%を按分したものです。業務委託費は、従来ほとんどを外部資金に頼っていた研究開発促進のための調査費用を一部WBAIが負担したものです。奨励金・賞金は、WBAI奨励賞2件と作業記憶モデラソンにおける重要功労賞の副賞として支払われました。2020年度は勉強会のためにZoomのWebinarを契約したため、通信費が多くなっています。

管理費のうち、事務局の人件費は外注費の形で支出しています(事業費と50%按分)。地代家賃は、ガルム合同会社のご厚意で無料となっています。支払報酬は経理事務を一部委託している会計事務所に支払っているものです。

2021年度の活動にむけて

脳型のAGIを個人の力量で開発できる脳科学と情報技術の両方に精通した人材を多く育成することは難しいという教訓を踏まえ、WBAIはBRA駆動開発を誕生させました。これは、膨大に蓄積しつつある神経科学知見を用いて外部設計仕様となるBRAを設計し、それに基づいてソフトウェアを開発・評価するという脳型ソフトウェアの開発手法です。このため、脳全体にわたるAGIソフトウェアの開発を推進するWBAIにおける、当面の技術目標は、脳全体のBRAデータを構築することになりました。

昨年までにBRAデータ形式が概ね標準化され、BRAを活用した開発の定式化が大きく進みました。そこで本年度は、これらの技術を普及させ、神経科学と情報学の接点に興味を持つ研究者の参入を促したいと考えています。そのためには、学術的な成果の発表や人材育成を通じて、活動を推進・普及させていきたいと考えています。そこで、引き続き本年度においても、教育事業と研究開発事業を継続し、以下などの活動をおこなう予定です。

教育事業(人材育成事業)

本事業の目標は、全脳アーキテクチャ・アプローチからの研究開発に必要な、人工知能、神経科学、認知科学、機械学習などの異なる専門性を同時に備えた学際的な人材を長期的に育成・増加することにあります。 本年度も例年の活動を継続し、複数回の勉強会を開催すること、年次のシンポジウムを開催することで一般の関心を喚起します。

ハッカソン(オンラインの国際コンペティション)を開催することで、WBAへの興味を持つ研究者・技術者・学生の増加をはかります。国内外の外部学術イベントへの参加・協力、外部学術団体との協力・情報交換などを行うことを通じ、当法人の進展状況をアピールしてゆきます。

外部の研究者・技術者およびグループとの協力により人材育成、巻き込みをはかります。

WBAI奨励賞の募集を行い、脳型AGIの(国内外の)技術開発の促進において、波及効果の高い開発成果を 残した方々を評価することで、コミュニティの活性化をはかります。

研究開発事業

本事業の目標は、全脳アーキテクチャ・アプローチによる外部での研究の先導と促進によりオープン・プラットフォーム上で民主的なAGI研究を加速することです。つまり、私たちは既存のWBA研究に関わる研究機関等と競合しないように活動を進め、情報交換などを行ったり、オープン・プラットフォームを利用した具体的な機械学習モデルの実装などといった研究の活性化をすすめます。

2020年度、私たちは、脳の解剖学的構造を記述したBIFと、それに整合的な機能の仮説的な機序を記述したHCDからなるBRA形式に従って、脳型AIの開発に役立つ形での神経科学知見の蓄積を進めます。これをもちいて特定のタスクにおける脳型AIの作成を促進します。またそうして構築された脳型AIの実装を評価するために、BRAを活用しながら脳型AGIの神経科学的妥当性を評価する技術の開発も進めます。

さらに引き続き全脳アーキテクチャ構築の基盤となる統合ソフトウェアプラットフォーや、当法人が開発 促進する脳型AGIをより Beneficial なものとすることにむけた活動も継続する予定です。

なおこれら研究開発活動の多くの部分が外部研究機関の研究資金により進められていますが、年度もその 一部については直接当法人が研究費を支出する予定です。

年度の予算

予定収入は約211万円で、会費収入176万円のほか有料化した勉強会の収益を含みます。2020年度の会費収入は323万円でしたので、予定される会費収入は大幅に減少していますが、これは今まで当法人を支えていただいた創設賛助会員3社が5年の期限を迎えたためです。当期予定収入と前期繰越金約916万円を合計すると約1127万円となります。

支出では、管理費に約83万円、イベント開催費用、研究開発費を含む事業費に約222万円、計約305万円を予定しています(当期予定収入との関係では94万円の赤字になります)。なお、2020年度の予算では管理費に約97万円、事業費に約277万円の計約373万円の支出を予定していました。

私とWBAI

公立はこだて未来大学 佐藤直行



公立はこだて未来大学の佐藤直行と申します。3回 目の勉強会から参加しています。

私は脳科学の背景からWBAIに参加しています。脳科学と言っても広い分野で、対象は分子からネットワークレベルまで様々、用いる実験計測技術も様々です。私の専門は計算論的神経科学(

computational neuroscience)という、脳の働きを計算論として理解しようする分野です。特に、海馬や大脳皮質の働きに興味があり、心理実験や脳計

測実験(脳波や機能的脳画像)をしたり、記憶や認知に関わる神経回路のシミュレーションをしたりしています。計算論的神経科学では、脳を理解するために「脳を創って理解する」のがひとつの方法で、脳の原理的な理解や、脳型計算機の実装を目指しています。これはWBAIの目的「人間のような汎用人工知能を創る」とも合致すると思います。一方、脳科学としては、脳がどう働いているかを知りたいので、若干の違いはあるかもしれません。

コンピュータと違って、脳は「自律分散的な制御」がひとつの特徴です。コンピュータは中央処理装置が全ての演算を実行しますが、脳は異なるモダリティの情報処理(例えば、視覚、聴覚、体性感覚など)が異なる脳部位で実行・保持されていて、それらのネットワークとして(なぜか)統合的に機能しているようです。このため、どこかの一部が損傷しても全体がフリーズしたりはしません。例えて言えば、試合中のサッカーチームで、メンバーが一人二人欠けても試合をそれなりに続行できるような感じです。脳の構成要素もそのように、ネットワークの中で協調的に働く仕組みを何か内在しているように思われます。

WBAの作業仮説は「各モジュールは個別の機械学習アルゴリズムを実行している」ことです。「全体としてうまく働くために、複数の脳部位がどのように連動すればよいのか」はWBAとしても脳科学として根幹をなす問題で、私もWBAIの活動を通じて勉強したいと考えています。

WBAIの良さのひとつは、分野の全く違う方々が、新しい事柄にチャレンジしようという心意気で、同じ場に集まって気軽に議論ができることだと思っています。私の興味は脳寄りですが、勉強会の参加者の方々話をすると、新しい人工知能を作りたい、やっぱり脳を創りたい、事業に生かしたい、興味があったので勉強したい、何も知らないので来てみた、など本当に様々です。どの意見も刺激になります。今後とも、皆様と一緒にWBAIを盛り上げていきたいと思います、よろしくお願いいたします。

東京大学 松尾 豊



人間の脳が行っていることは情報処理であり、何らかのアルゴリズムで実現されている。それを計算機で行うことも可能である。ただし、進化のなかで絶妙に最適化されており、また、言語・教育・文化などと一体になって共進化しており、さらには、人間の身体や環境に依存する部分も大きいため、人間とそっくりな情報処理ができる可能性はかなり低い。人間の脳が行っている情報処理は、おそらくいくつかの主要な要素で構成されるはずである。そのひとつが、深層学習のような深い階層を持つネットワー

クにより、関数近似を行う仕組みである。また、別のひとつは強化学習であり報酬に応じて行動を強化する仕組みである。

ただし、これまでの学術研究のなかで、根本的に間違った定式化からスタートしているものも少なくない。例えば、これまでに一般的に定式化されている強化学習の問題設定は、人間の脳の主要な要素という点ではおそらく間違っている。教師なし学習という問題設定も長らく間違っていたが、最近は「自己教師あり学習」という形で正しい方向に修正されてきた。そもそも特徴量を見つけること(すなわち表現学習)が重要であるということも、深層学習が広く理解されるまでは、問題設定としては長らく無視、あるいは軽視されてきた。こういう問題設定のそもそもの間違いに、研究者(あるいは研究コミュニティ)は無自覚である場合が多い。

キーワードだけ出しておくと、他の主要な要素として考えられるのは、世界モデル、蒸留、想像、万能 チューリングマシンなどである。特に、言語を用いることが、情報処理においてどのように根本的な変化 をもたらしたかが、その説明の主要部分になると思われる。

深層学習をはじめとする最新の技術で実装可能である(あるいは近い将来に実装可能になると思われる)こと、それが人間の知能をどう作り上げているかの妥当な説明があること、それが脳の各機能と対応づいていること、進化的に妥当な説明が存在すること、などの要件を満たすアルゴリズムの全体像、すなわち知能のアーキテクチャを見つけることができれば、多くの計算機科学的・脳科学的・哲学的・心理学的な疑問に答えることができるようになるだろう。それを見つけたいというのがWBAIの活動であり、(その予想が妥当であるかどうかというより、一研究者としての心的態度として)それが近い将来明らかになるのではと考え、研究を行っている。

WBAIは、山川氏との議論がきっかけで2015年にNPOとして組織され、私自身も副代表として活動してきた。人類が誕生して以来の長い歴史の中で、知能の仕組みという人類史上で最も重大で深遠な謎が解けるかもしれない瞬間に、我々が研究ができる状況にあるということは大変な幸運である。この幸運に感謝し、WBAIが社会に大きな貢献ができることを祈念している。

WBA勉強会実行委員 生島高裕



数理先端技術研究所の生島と申します。私が活動に参加し始めたのは、第12回 全脳アーキテクチャ勉強会 ~ 脳の学習アーキテクチャ ~ (2016年01月14日)からだと思います。最初は勉強会のレポート作成などでお手伝いしていました。その後、実行委員のメンバーになり、第1回シンポジウム「加速する人工知能、加速する世界」(2016年5月18日)では実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきました。ここでは実行委員長を努めさせていただきます。

勉強会全体は、脳科学、神経科学とDeep Learning、強化学習のような機械学習、AI系の話がベースになっております。しかし、その中でも、海馬、前頭前野など言語、推論、意味、価値、創造性の話など最先端の話が多くあり、他の勉強会とは差別化されていると思っています。また精神疾患、社会性などがテーマとしてあり、今後のAI発展に伴う、人間との双対的な現象出現の示唆も感じられます。

シンギュラリティに向かって、経済、統治、社会の未来についてもAIがベースとして認知されるようになってきています。この数年では、貨幣の消失、デジタルレーニン主義、信用スコア、AI倫理など活発に議論されています。すべての現象の発端は第3次AIブームによるものでしょう。それらの課題に対するマイルストーンの年として2030年は重要な時期であると思います。

さて、1956年のダートマス会議(The Dartmouth Summer Research Project on Artificial Intelligence)の7つの課題、自律、自然言語、ニューラルネット、計算複雑性、自己改善、抽象、創造性、このテーマに関してどこまで進んでいるでしょうか?楽しみです。

個人的興味としましては、やはり人類が出現し繁栄している直接の原因、自我、創造性の解明にあると思っています。

自我に関してはエネルギーの概念に影響を受けているフロイトの力動論が今後どのように説明されていくかに興味があります。局所論(意識、前意識、無意識の3階層モデル)、構造論(エス、自我、超自我)はわかりやすい人間の心のモデルとなっています。武谷三男の三段階論的過程での現象論的段階を経てモデル的段階へは達していますが、本質論的な説明には至っていない現状です。

発達心理学はAIの研究方針として、個体発生的に系統発生的な知性の創発の指針を与えてくれました。ただ精神疾患も含めた異常系を理解するには力動論的な現象論的知見がAIの研究方針に役立つのではないかと思っています。

また、観測技術の進歩はいつも科学のパラダイムシフトを後押ししました。大きくはビッグバン直後の「第一世代星」観測技術、小さくはナノレベルでの分子構造観測技術があります。もちろんその観測結果からの本質論追求への理論的考察が最も重要で、今後納得のいく精緻な心の理論が構築されると思っています。

その追及にはフリストンの自由エネルギー原理のように物理の最小作用の原理からの説明、複雑ネット ワーク的なスケールフリー 性、クラスター性、スモールワールド性はもとより、意識に関する統合情報理 論のように情報の多様性・情報の統合からの説明は必要不可欠だと思います。その知見から心のシステム 論が構築されることを期待しています。

現在のオペレーティングシステムは意識をカーネルモード、無意識をソフトウエアモードと考えますとアバウトな心の表現はできています。しかし、エス、自我、超自我に対応するものが実装されていないわけです。特に超自我である倫理、本能であるエスはどのように実装すべきかわかっていません(人間もわかっていませんが…)。しかも作り込みではなく創発させることを考えますと課題は山積みと思います。創造性については科学だけでなく、アートの世界など全ての分野でいつも言われることです。個人的は数学におけるポアンカレの科学と方法で述べられた数学的審美眼は特に興味深く思われます。これは様々な分野で使われる価値関数、効用関数、評価関数と言われるもので一種の目標を表す美意識です。それと現実の差を表す関数が損失関数、誤差関数、予測誤差でこれらを最小にしようとして脳活動は常に行われているのでしょう。

敵対的生成ネットワークにおける生成ネットワークに対する識別ネットワークが評価関数に対応するものと思っています。これはアートの世界におけるクリエータ・オーディエンス問題です。また、蒸留は生成ネットワークが新たな組み合わせを作るときのコストを下げるための手段と考えられます。蒸留は数学、物理の理論における説明性の向上に関係していると思われ、常に難解な理論はよりわかりやすい表現へと進化します。18歳までに超重力理論の理解ができるコースもあると聞きます。それはある意味、オントロジー構造の最適化です。この人類における創造性のテクニックがAIに応用されているということ自体が素晴らしいことと思います。

シンボリズムの観点からは推論の解明と実装が大きな問題ですね。類推、演繹、帰納、アブダクション、 一般化、抽象化、双対性のうち類推、演繹が若干進んでいるのが現実でしょうか?

典型的なマイルストーンは、数学においての証明の検証が自動化されていないことです。宇宙際タイヒミュラー理論など予想問題の証明は大変なため人類も苦労しています。数学論文の証明部分の検証の自動化はいつできるでしょうか?個人的にはこの後、美意識に基づく予想問題生成ができることにより機械が創造性を勝ち得ることと思っています。フィールズ賞を機械が取ってこそシンギュラリティの日と思っています。

このような夢のある課題について、是非勉強会でもテーマにしていただければと思います。

全脳アーキテクチャの今後の発展を期待しております。

おわりに

全脳アーキテクチャアプローチによる脳型AGI開発は、外部設計仕様であるBRAを脳科学的知見で制約しながら設計し、BRAを参照しながらソフトウェアを開発・評価するBRA駆動開発という形で具現化されてきました。この開発手法により脳型AGIを完成させるためには、少なくとも以下の5つのマイルストーンを達成する必要があると考えています。

- 1. <u>全脳の領野を含むBIFの完成</u>:複数の哺乳類の脳科学的知見を統合して、脳のほぼ全体をカバーするキメラ型の脳情報フロー(BIF)を構築する。
- 2. <u>典型的機能に対するHCDの完成:</u> 概ね全てのBIFに対して、少なくとも1つ以上の仮想コンポーネント図(HCD)を構築する。
- 3. <u>WBA全体の初期実装:</u> 脳のほぼ全ての領野がBRAに基づいて実装され、統合されたソフトウェアを作成する。
- 4. <u>アーキテクチャ探索の自動化:</u> 脳型AIソフトウェアを仮想環境で実行することで、テストを自動化し、様々なアーキテクチャを効率的に比較・検討することができるようにする。
- 5. 人や動物が解決できる典型的な認知課題に対応できるように、HCDのカバー範囲を拡大する。
- 6. <u>WBAシステムの完成:</u> 脳の神経回路に対応するほぼすべての部品を組み立てたソフトウェアにより、 人間の典型的な作業を脳と同じ方法で解決できることを実現する。

こうしたマイルストーンを踏まえ、当面は、第1、第2項目を含む脳全体のBRAデータの構築に注力するとともに、第3項目の実現可能性を評価していきます。この段階では、脳科学と情報学の接点に関心のある研究者を巻き込むことで、その研究開発を加速できるでしょう。そこでまずは、学術的な成果の発表や人材育成を通じて、BRAを活用した開発を広めていきたいと考えております。こうした活動を着実に継続することにより、近い将来、脳型AGIの開発に有用な技術基盤が構築され、後半のマイルストーンに着手できるとでしょう。脳型AGIの完成が見えてくると、安全性への配慮がより重要になってきます。そこで、いずれ到達する脳型AGIをBeneficialで安全に実現するための議論も継続してゆきます。

私たちは、上記ような展望のもと、今後も継続的な取り組みを行っていきたいと考えていますので、ご理解とご支援をいただければ幸いです。

発表論文など

- [1] Hiroshi Yamakawa, Imagination Architecture and Consciousness in the Brain, in *IJCAI-PRICAI 2020 Tutorial T11 'Conscious AI: Significance and Development'*, January 8th 2021. https://wba-initiative.org/en/12120/
- [2] 佐々木明, 荒川直哉, 山川宏, 脳の機能仮説についての構造化データの構築, , 信学技法 ニューロコンピューティング 研究会 (NC) , NC2020-11, pp. 9 13, 2020年10月29-30日.
- [3] Mei Sasaki, Naoya Arakawa, Hiroshi Yamakawa, Construction of a whole brain reference architecture (WBRA), *International Symposium on Artificial Intelligence and Brain Science*, P-31, 10-12, October, 2020. http://www.brain-ai.jp/symposium2020/posters/
- [4] Naoya Arakawa & Hiroshi Yamakawa, The Brain Information Flow Format, *The 1st Asia-Pacific Computational and Cognitive Neuroscience (AP-CCN) Conference*, September 26-27, 2020. Presentation: https://www.youtube.com/watch?v=0fqdCVCIKC0&t=16s
- [5] Yamakawa, H. (2021). The whole brain architecture approach: Accelerating the development of artificial general intelligence by referring to the brain. In arXiv [cs.AI]. arXiv. http://arxiv.org/abs/2103.06123
- [6] 谷口忠大, 山川宏, 全脳に学び確率的生成モデルで描く統合的認知アーキテクチャ, 新学術領域研究人工知能と脳科学の対照と融合NEWSLETTER, vol.8, 2020年10月.
- [7] Taniguchi, T., Yamakawa, H., Nagai, T., Doya, K., Sakagami, M., Suzuki, M., Nakamura, T., & Taniguchi, A. (2021). Whole brain Probabilistic Generative Model toward Realizing Cognitive Architecture for Developmental Robots. In arXiv [cs.AI]. arXiv. http://arxiv.org/abs/2103.08183
- [8] Taniguchi, A., Fukawa, A., & Yamakawa, H. (2021). Hippocampal formation-inspired probabilistic generative model. In arXiv [cs.AI]. arXiv. http://arxiv.org/abs/2103.07356
- [9] Yamakawa, H. Attentional Reinforcement Learning in the Brain. *New Gener. Comput.* 2020. doi:10.1007/s00354-019-00081-z
- [10] Ayako Fukawa, Takahiro Aizawa, Hiroshi Yamakawa and Ikuko Eguchi Yairi, Identifying Core Regions for Path Integration on Medial Entorhinal Cortex of Hippocampal Formation, *MDPI Brain Science*, 10(1), 28, 2020.
- [11] 田和辻可昌, 荒川直哉, 山川宏, 脳幹視覚系神経核における抑制と加算機能に関する仮説の提案 ~ BIF形式に基づく滑動性追従眼球運動と衝動性眼球運動の統一的記述, 信学技法 ニューロコンピューティング研究会(NC), NC2020-10, pp. 9 13, 2020年10月29-30日.
- [12] Yoshimasa Tawatsuji, Naoya Arakawa, Hiroshi Yamakawa, Knowledge representation for neural circuits subserving saccadic eye movement based on a Brain Information Flow description, *International Symposium on Artificial Intelligence and Brain Science*, P-45, 10-12, October, 2020. http://www.brain-ai.jp/symposium2020/posters/
- [13] 山川宏(WBAI/東大), 内観における Papez 回路の役割について, 人工知能学会第15回汎用人工知能研究会, SIG-AGI-015-03, 2020.8.29. http://id.nii.ac.jp/1004/00010725/
- [14] Hiroshi Yamakawa, "Understanding the computational meaning of the neocortical interrarea signals", CSHL Meeting: From Neuroscience to Artificially Intelligent Systems (NAISys), November 9 - 12, 2020.
- [15] Yamakawa, H. Revealing the computational meaning of neocortical interarea signals. *Frontiers in Computational Neuroscience*. (2020)

- [16] 布川絢子(WBAI), 矢入郁子, 山川宏(WBAI/東大), 脳の注意機能を担う前障(claustrum)の内部機構, 人工知能学会第15回汎用人工知能研究会, SIG-AGI-015-06, 2020.8.29. http://id.nii.ac.jp/1004/00010728/
- [17] 山川宏, 脳に学んだカノニカル知能機構の構築にむけて 一学習で得た知識を多面的に再利用する仕組みとは一, 認知モデル研究会, 2021年1月9日.
- [18] Hiroshi Yamakawa, Towards a qualitative evaluation of biological plausibility for brain-inspired software, The 1st Asia-Pacific Computational and Cognitive Neuroscience (AP-CCN) Conference, September 26-27, 2020. Presentation: https://www.youtube.com/watch?v=mtXcFHIKbFw&t=9s
- [19] Hiroshi Yamakawa, Naoya Arakawa and Koichi Takahashi, Whole brain reference architecture to evaluate biological plausibility of human-like artificial intelligence, *International Symposium on Artificial Intelligence and Brain Science*, P-30, 10-12, October, 2020. http://www.brain-ai.jp/symposium2020/posters/
- [20] Heecheol Kim, Masanori Yamada, Kosuke Miyoshi, Tomoharu Iwata, Hiroshi Yamakawa, Reinforcement Learning in Latent Action Sequence Space, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020), October, 2020. http://ras.papercept.net/images/temp/IROS/files/0338.pdf
- [21] 山川宏、汎用AIが投げかける人間の本質への根源的な問い, フロネシス, vol.12, no.1, pp.116-121, April, 2020.
- [22] 高橋恒一, 知能爆発はいつどのように起きるのか, シンギュラリティサロン@SpringX, 2021年1月30日.

貸借対照表

(2021年3月31日現在)

特定非営利活動法人全脳アーキテクチャ・イニシアティブ

科目			金額(単位:円)		
I	資産の部				
	1.	流動資産			
		現金預金	9,145,734		
		流動資産合計		9,145,734	
	資産の	含計			9,145,734
I	負債の部				
	1.	流動負債			
		前受金	400,000		
		預り金	6,126		
		流動負債合計		406,126	
	負債合計				406,126
Ш	正味財産の部				
		前期繰越正味財産		7,569,416	
		当期正味財産増減額		1,170,192	
	正味財産合計				8,739,608
	負債及び正味財産合計				9,145,734

活動計算書

2020年4月1日から2021年3月31日まで 特定非営利活動法人全脳アーキテクチャ・イニシアティブ

科目		金額	
I 経常収益			
1. 受取会費等			
正会員年会費	150,000	0.000.000	
賛助会員費 2. その他収益	3, 080, 000	3, 230, 000	
受取利息	76		
雑収入	367, 963	368, 039	
経常収益計	501, 505	500,005	3, 598, 039
Ⅱ 経常費用			3,000,000
1. 事業費			
(1) 人件費			
給料手当	162, 750		
法定福利費			
人件費計	162, 750		
(2) その他経費	005 105		
業務委託費	325, 105		
外注費 講師等謝金	626, 180 72, 000		
広告宣伝費	12,000		
交際費			
会議費			
旅費交通費			
通信費	284, 525		
消耗品費	330		
新聞図書費			
支払手数料	7, 491		
地代家賃	100.070		
奨励金・賞金	130, 072		
寄付金 雑費			
その他経費計	1, 445, 703		
事業費計	1, 110, 100	1,608,453	
2. 管理費			
(1) 人件費			
給料手当	0		
法定福利費	0		
人件費計	0		
(2) その他経費			
業務委託費 外注費	533, 500		
支払報酬料	264, 000		
交際費	204,000		
通信費	4, 520		
消耗品費			
諸会費			
支払手数料	11,622		
地代家賃			
租税公課	4, 592		
雑費	1, 160		
その他経費計	819, 394	010 001	
管理費計 経常費用計		819, 394	9 497 947
栓吊貨用訂 当期正味財産増減額			2, 427, 847 1, 170, 192
前期繰越正味財産額 前期繰越正味財産額			7, 569, 416
次期繰越正味財産額			8, 739, 608
V17914KA22 44-775 F1 FE BH			0, 100, 000