

Golems, Cyborgs, and Artificial Intelligence

Project Assistant Professor, The University of Tokyo **Arisa Ema**

Science and technology research pursues the truth and structure of the universe, life, human beings, and society. Engineering and constructive approaches are taken as an exploration approach. It seeks to "understand by constructing." As a result, human civilization has opened new horizons, and in return, it has encountered various incidents and accidents.

In science and technology studies (STS), these problems are often compared to Frankenstein's monster, golems, and zombies. The so-called "Golem" book series by Harry Collins and Trevor Pinch discuss and cover a variety of incidents and cases, including the Chernobyl disaster, AIDS, Climategate, and vaccines. These discussions also question the structure of society and the responsibility of the experts who created the "golems."

The search for artificial intelligence (AI) and brain science that enables "human-like" flexible behavior and communication also raises questions at the interface between science, technology, and society. "Human-like" in brackets is to emphasize that no actual humans are involved. To be precise, what is composed by technology is "reconstructed human intelligence, skills, and characters based on fragmentary information and data." Assuming we are a collection of information or data, we can build "human-like" AI. As we live, we leave our footprints in the real world and in digital spaces. Some data are biologically assigned, while others are socially

assigned. Aspects such as hobbies and personalities may be said to be decentralized information formed between relationships among people. By scraping these pieces of information and data together,

Frankenstein-monsters, golems, and zombies are created. However, collecting information does not necessarily make one an individual organism.

On the other hand, we should not forget that, as Andy Clark points out, we ourselves are "natural-born cyborgs." In addition to appearance, such as wearing glasses to supplement vision, we already rely on digital spaces for our memories and information processing. Scheduling and community networks are stored in a digital space. Research, calculation, communication, and shopping can no longer be done without the support of information services and devices. Living in a modern society means living as a cyborg, and becoming a grainy zombie after death, constantly sprinkling the material to create a golem using parts of you.

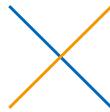
To protect the dignity and value of human beings while advancing technology, multi-stakeholder discussions in various fields such as philosophy, law, and society are required. What kind of society would we like to live in? This is not a question that can be answered easily. Therefore, we must continue dialogue with the general public.





松尾 豊

東京大学大学院 工学系研究科 教授



田中 啓治

理化学研究所 脳科学総合研究センター
シニアチームリーダー



今回の対談は緊急事態宣言下のため、テレビ会議で行われました。同じく過去にあった非常時下の行動が今に繋がっている田中先生のお話をお送りします。

松尾 前回僕がバトンを受け次、田中先生ということですけど、僕の時は子どもの時から話をしたので、僕もそんな感じで聞いていきたいなど。先生はどういう子どもさんでしたか？

田中 スポーツはまあまあ、勉強は親が熱心だったんで割とできる方で、学級委員長なんかもやらされて大人しくはないけども優等生でしたね。あんまり面白くない子でした。東京文京区の白山で育ち、小中学校は国立大学の附属に行きました。

優等生だった高3の12月、東大紛争で受験がなくなり慌てて考えた進路。大学1年、授業が始まったのは12月だった

田中 1969年文京区の高校卒業時、東大紛争真っ只中で。東大と筑波大が受験がなく大阪大学に行きまして。修士卒業まで6年間大阪で過ごしました。そこでちょっと東京の優等生から道を踏み外しまして(笑)、やや乱れた。大して乱れなかったけど乱れました。

松尾 (笑)。あんまりあれですか、大学に行かずにという？

田中 大学に行っても4月から12月まで授業なんかなかったですが、それよりもね。東大理科I類を受験するつもりでいたら、高3の12月くらいに受験がなくなると言われたして。他大学はどこも学科で受けなきゃいけない、私は学科を選ぶ準備ができてなかった。当時、エリート立場にいる人たちは特にエリートにいることに対する疑いを持つべきというエモーションがあって、高3にもなって自分がどんなことやるかも考えずにいたことに愕然としたんです。こんなじゃ私の人生大したことないな、もうちょっと意識的に生きなきゃなと。

松尾 高校の頃から今やられてるようなことに対するの興味はもうあったんですか？

田中 いや、受験がなくなるという頃慌てて考えだして(笑)。どの学科を第1志望にして北海道から九大までどこを受けるか考えるほんの短期間にWienerのサイバネティクス、それからバイオニクスに出会って。面白いなと思って、大阪大学基礎工学部生物工学科の受験を決めたのが、私の人生の非常に重要な出発点なんです。阪大は東京の進学校ではそんなに有名でなく僕にはちょっとした選択でもありました。

松尾 授業がない間はどんなことを？

田中 ちょっとスポーツやったり、生物工学科は当時非常にアンポピュラーな分野だったんで面白い同級生が全国から集まっていた。楽しかったですね。授業がない仲間仲間達といろんなこと勉強したり議論したり遊んだり、それが今の原点ですね。高3から大学1年の経験が。

環境づくりから始まった学部・修士時代

松尾 3期生と書かれてますが、これは学科のことですか？

田中 そうですね。当時は大学院がそんなに強調される前で、学科ができる時に学部学生を募集して、その人たちが研究室に入るのにあわせて教授が赴任してきて。非常に顕著な活躍をされた方としては、当時大学院生でおられた大阪大学の生物物理の柳田先生。だけど学部学生が中心で、開設時は教授が2、3人ぐらいであとから新しい若い教授が赴任して来られたんですね。すると教授の先生たちを学部学生が「よくいらっやいました」と歓迎するんですね(笑)。こういうちょっと普通と違う状況で育ち、非常に面白い経験でした。東京の小中高校と比較的ルールに乗った世界からほとんどルールがない、何もなくて新しいものを作るというところに放り込まれて、物事には前提はないんだという持論を、身をもって二十歳前後に経験しました。

松尾 それは重要な経験ですね。卒論はありますよね。研究室に配属されてという感じなんですか？

田中 ええ。普通は学部生の卒論なんてのは大事じゃないですが、できたばかりで院生があまりいない時に入りましたから比較的よく先生たちに面倒見ていただいて。神経科学の実験研究室に入った時、院生が1人もおらず僕たちが一番上で、いろんな実験も4年生でも相当にやらせていただいた。実験セット1つしかないから、昼間先生が実験後、夜は僕ともう1人が交代で使って、実験を朝までやり続ける。4年生にはもうすごい恵まれた環境で。甘やかされて育ちました。

松尾 へえ～そうですか。どんな研究だったんですか。

田中 サイバネティクスやバイオニクスに憧れて、最初は松尾先生に近いような脳の情報処理の理論やろうと思ったんですけど。卒業研究やる頃に向けて脳の研究やるには、あまりにもデータが少ないんじゃないかと。しばらく実験研究、5年ぐらい経って理論研究に進もうと思って、網膜の実験をやりました。網膜の情報処理は、下等動物では出力細胞の1つ手前にある網膜中のアマクリン細胞がいろんなことをやるおかげで非常に複雑な働きをしている。それがよく分かってなかったので、細胞内記録用の尖った電極で、コイとカエルの目玉だけ取り出したEyecup Preparationで、光刺激を与えてアマクリン細胞がどういう反応するかという比較的単純な実験でしたね。

松尾 面白いですね。何か新しいことは見つかったんですか？

田中 学部の研究からは見つけるに至りませんでした。しかし僕にとって非常に良かったのが、網膜と脳の神経回路は同じぐらい複雑なネットワークを構成していると当時言われ、実験しやすいおかげで神経科学の研究では比較的進んでいたんですね。だけど生物学的な研究が多く、系統的にまとめるのは比較的少なかった。生物工学科の先生には生物の実験系・理論系だけでなく、情報科学、制御理論、生物物理の方がいらして、卒論ではその中で自分の研究の位置づけを手書きで一所懸命書きました。ワープロもない時代で、たくさん情報をまとめて整理して研究を位置づけるという、サイエンスで大事なことを学部の時勉強させていただいた。

松尾 なるほど。修士も同じ研究をされたんですか？

田中 当時研究室の教授は塚原仲晃先生って（後に）日航機事故で亡くなった先生ですけれども。網膜はずっとやるにはどうかなと思ったんで方向を変えて。小脳の研究が当時日本では伊藤正男先生の影響もあって盛んで、修士は小脳の研究をしました。別な先生について。四足歩行の基本的な時間パターンは脊髄の中の回路でできていて、それを制御している小脳の役割は何か、実は大事なものは外乱が入った時にそれを元のリズムに修復するというのを研究しようとした。大変な実験でしたが、楽しかった。当時ミニコンピューターが入りだして生物工学科4階にPDP12が1台だけあって、窓越しに僕たちの2階の実験室まで線を這わせ、ベルトコンベヤー上を歩くネ

コの小脳と脊髄の下行路の神経細胞の記録を取った。歩行は周期が一定じゃないから、どこらへんで神経細胞が活動するかの解析は単純じゃなく、周期をノーマライズするプログラムを学科で初めて作ってオンライン解析をしたんです。先生は医学部出身でそういうプログラムはできなかったけど、僕たちは生物工学科でちょっとでも習ったからできて、その後の研究に大いに役立ちました。

松尾 今でもおかしくない本当にすごく先端的な研究ですね。

田中 今はもう当たり前ですからどうってことは。5年後に本当は理論研究に戻るはずが、とてもじゃないけど神経科学の実験研究の進展ってそんなに速くないので、それ以来ずーっと実験研究を続けて、その間に頭も鈍っちゃって（笑）だけど、情報処理機械との関係はそもそも出発点からあるんです。

卒後の研究は神経科学の発展と共に

松尾 田中先生は修士阪大でその後東大に行かれ、NHK研究所、理研といろいろ移ってこられました。その中で最初の「こういうところから研究しよう」という思いからどんな経緯で今の研究に至っているんでしょうか。

田中 ちょっと修正すると私は阪大修士卒業後、いきなりNHK放送科学基礎研究所に研究員で就職する機会に恵まれました。東大医学部の伊藤先生のところに研究生で所属し、時間をかけて博士号をとらせていただきました。NHKの基礎研と後に合併された技術研究所には14年間いて、理研に移った。修士卒業後1975年から今の2020年の間ってというのは、神経科学、脳科学が爆発的に発展した時期だったんですね。1970年にもうかなり研究ありましたが、概していえば神経細胞の研究だったのがネットワークの研究になり、脳全体の研究になり、そして心理学と結びついて認知科学とも対話するようになっていった、神経科学が発展していく過程があったんですね。その中で私自身の興味もどんどん展開して、最初は心理学的なことに興味がなくて、小規模ネットワーク中での情報処理の様子に興味があったんですが、そのうちに行動している動物の中で回路がどう動くかに興味に移り、最初はネコ、次にマカク属サル、1996年からヒトの研究も行って。するとヒトのことを勉強しなきゃいけない、さらに心理学的な考察もしなきゃいけない。大脳連合野に入ってくると認知科学の話になり、人工知能の先生方の考え方も勉強しなきゃいけない。そういうふうに展開してきました。それで特に最初は小脳の研究、NHK研究所時代は第一次視覚野の研究、それから視覚系のもう少し高次の研究もやり、さらに前頭前野の研究が中心になってきて研究を拡張しました。理研は恵まれた環境で可能性はいろいろあったので、神経科学の発展や興味に合わせて重要と思われることに向いてきた。

松尾 なるほど。NHKは福島先生とかおられましたか。

田中 私が3階、福島先生が2階におられて、ネオコグニトンの研究されてました。当時はそこまで重要な方が同じグループに居られるとは分からなかったですね。

松尾 ネオコグニトンは目立ってたわけではなくて？

田中 そうですね、その当時も有名でしたが、やはりNHKの研究所であって人工知能の研究所ではなかったので、福島先生はある程度グループを持ってましたがそれでも3、4人かな。

松尾 そうですか。理研に行かれると甘利先生もおられる？

田中 私は1989年に理研に行ったので、まだおられなかったですね。甘利先生は、独特のIntuitionで実験の話題もばつと理解してコメントする。だから独特の重要な役割、日本の理論研究と実験研究のインタラクションを作るっていう。日本は甘利先生や福島先生のおかげで脳科学の実験面と理論面のインタラクションの歴史は比較的豊かで、私もお陰を被りました。

松尾 将棋の研究時はプロ棋士といろいろやられたんですか？

田中 2011-12年とかに実験を盛んにやりました。その頃絶頂期でものすごい勢いだった羽生先生に3回ぐらい被験者になってもらいました。

松尾 やっぱり違うんですか？羽生先生は他の棋士と。

田中 プロ棋士のデータ20人分とアマチュア20人分の比較は割と簡単ですが、羽生さん1人と他のプロ棋士100人を比較しても、それが本当に羽生さんの能力の原因かは非常に難しいですね。脳活動に特別なことは見つからなかった。ただ言えるのは羽生さんはすごく真面目で研究者肌。ファンクションMRIの実験時は課題を仔細に至るまで非常に真剣にやられる。モニターのため目の像を取っていると羽生さんは問題が出るたび瞳孔がぱっと開くんですよ。大体緊張すると瞳孔が開くんですが、非常に顕著。能力の原因ではないと思います。

松尾 羽生さん凄く負けず嫌いで、とにかく負けると嫌だから回避したいみたいなものがあるかもしれませんね。

田中 羽生さんはとにかく研究しまくってる感じがしましたね。対談してもね、どんどんおひとりでお考えが進んで、結論までどんどんどんどん話が進んじゃうんですよ（笑）。他のプロ棋士の方は、思い込み的な発言をされる方がいらっしやるんですけど、もっとこう整ったというか、ロジックが出てくる。

転換点—将棋の研究で初めて触れた人間独特の話題

松尾 理研でやられている中で一番大きな転換点になった研究や、凄く前進した研究はご自身ではどうですか？

田中 いくつかあるんですけども、せっかく出ましたから、将棋の研究の話を。最初、将棋の研究は持ち込まれて偶然やり始めたんです、実は。最初はなんだその話と思ったのが、やって

る間に面白いと思いだして問題を煮詰めて。それまでのファンクションMRI使って第一次視覚野の眼優位性カラムのイメージングとか、高次野に展開とかの研究は、動物の研究から十分発想できたんですね。動物にも多分あると思うけど、動物の直観の研究ってのは全くなかった。将棋における直観の役割はやっぱり人間独特の話題で、初めてそのことに接する機会だったので、大いに面白かった。特に顕在的な論理的な意識に上る面と、無意識の潜在的な面の関係は今話題になりつつあると思います。意識に上らない潜在的なプロセス、直観というのは初心者では働かないがエキスパートになるとうんと働くということの研究して、人間の能力は言語の能力だけでなく動物OSによってだと感じる機会があったので、松尾先生の動物OSと言語アプリもものすごく面白いなと思ってね。それでお聞きしたかったのが、動物OSの上に言語アプリが載っていると。じゃあ、動物OS部分はヒトとマカク属サルぐらいは同じで、人間の高等動物に比べて優れた部分はその上の言語アプリ部分だけなのか？

松尾 ええ、僕はそれはほぼイエスだと思ってまして。言語アプリの原始的なものはマカク属サルにもあると思いますが、人間の場合一番違うのは他人の発話を予測しようとする報酬関数が別にある。動物はそのまま predictive coding みたいなもので自分が次に何を見るか聞くかを予測する学習、センサ入力、アクチュエータ出力に対しての予測問題設定になっている。一方で人間は相手の発話を予測する別の予測問題が組み込まれると。多分、そこだけが違いでそれによってすごい大きな変化が起こる。発話予測には、例えば相手のメンタルモデルを持っておく必要があり、相手の学習内容を自分もシミュレートしないといけないので、ある種の蒸留、Hinton先生がいう相手のニューラルネットワークを自分にコピーみたいないろんなことが起こると思ってます。逆にいうと自閉症の子が他人に対し興味を持たないとかは多分その報酬関数だけ異常があって、他人に対する発話予測が設定されてない。

田中 でも、そうして言語で表象されたものとしてゴールを共有し、それを実現するため相手がどう考えているかのモデルを自分の中に作る過程は、動物OSを使うんじゃないですか？

松尾 使いますが、僕のイメージはこの動物OSの方はセンサ入力、アクチュエータの出力があって、それを時系列に将来を予測しながらプランニングしている構造になっている。その学習過程で、どういう状況で何をしたら何が起こるかという世界のダイナミクスに関する情報「世界モデル」を学習する。一方で言語アプリは言葉を入力にして”世界モデルをシミュレータとして発動”させるんです。発動後何が起こるか想像し、また言葉に直して出力している。これを相手の発話も予測しながら自分の発話を作って、というようなループを回して行うので、動物OSと言語アプリで潜在構造が共有されている。

入力となるものと予測したいものはそれぞれ異なっていて、動物OSの場合はセンサ入力を元に将来のセンサ入力の予測問題を解いてるし、言語アプリは発話を入力として将来の相手の発話の予測問題を解いている。潜在構造は一緒に問題が2つというマルチタスク構造をしてるんじゃないかと…

田中 人間の場合は動物OSを使う時に入力と出力のところが違う。動物はセンサなりアクチュエータだけけど人間は言語で表わしたものを入力し出力する。だけどその言語的なものから言語的なものの処理の、人間がOSを使う時、過程はもう潜在的になる？

松尾 ええ、潜在的なんですけどこれはHassabisさんはイマジネーション、僕は絵を描くって言うんですけど、言語を条件とした絵を生成すること、視覚だけじゃないですが、視覚的な絵の生成をやっているのだと思います。

田中 動物OSをヒトは動物があまり使わなかったやり方で使うようになった。そうするとヒトの動物OSは動物と違ってきていてヒト独特の発展を遂げたのでは？

松尾 なるほど、面白い質問です。2個あって一つはまずアーキテクチャー上、活性化が外界からの刺激で起こっているのか、それとも内的な言語による刺激なのか区別する必要がありますので、その機構は必要ですね。もう1個は言語から何か刺激を受けて想像して言語的な反応を返すこと、社会的なコミュニケーションも含めシミュレートして、今度動物OSが学んじゃうんです。動物OSの方がどんどんメタになる感じです。

田中 なるほど。“エキスパートの直観”は潜在的で説明できないんです。あとで解析して時間をかければこういうことだと分析できますが、その時には意識に残ってない。だけどそれが膨大な情報処理を含んでますよね。僕は動物になかったものが人間は発展したと思うんです。脳科学者なのでいうと、人間の脳で前頭前野の一番前の所にある10野というのがマカク属サルではちっちゃいんだけど、人間ではぐあーっと広がっている。10野は、入力層の4層、小さい細胞がいっぱいある層が非常によく発達していて、入力層にそういう小さい細胞が沢山あると次元を拡大して問題を簡単な方法で解くのに非常に有利。小脳で顆粒細胞でぶわあーっと大きい問題を展開しておき、後で中間層で問題整理する。大脳皮質も4層で入力次元を拡大するんで有利なんだけど、10野もそう。その4層で次元を拡大した世界は、松尾先生や高次元科学の丸山先生の話ではあまり高次元になると人間は顕在的に反応できない。すると4層がうんと発達した10野は顕在的な過程じゃなくて潜在的な高次元処理をやってるんですよ。だけどヒトでなんでこんなに発達したか。やっぱり言語でドライブして言語に戻していくという処理が人間でうんと仕組みがあって、前頭極は特にいろんなところから総合的な入力を受けているところなんです。言語処理のためにはそういう総合的な情報が

必要。言語アプリで動物OSを使うために独特な動物OSに有利な構造を持った動物OSが、ヒトの脳では前頭極の発達だったんじゃないかと思ってるんです。

松尾 多分、言語の入力をし始める時点で何かすごいいろんなことが起こるんですよ。Disentanglement、物事を要素に分解し言語で表すと、現実世界にない組み合わせが指定可能になるので、一角獣みたいに現実にはないが想像できるという空間も探索でき、それによりすごく広がる。一方、進化的には多分そんな難しい変化を一気に起こせないの、報酬関数が1個追加されただけであとは雪崩式に起こったのではと。

田中 確かに、サルとマカク属サルとの共通祖先から人間が分かれてきたのは動物の進化の歴史の中では本当に短いところで起きたのは仰る通りだけど、やっぱりヒトとサルの脳はすごく違うから、その中で進化は相当にある。脳の変化があり、動物OSが言語アプリを制限するのは当然だけど、言語アプリ側が動物OSを変えていく過程は、ヒトの能力の理解には重要じゃないかと思います。

これまでもこれから

松尾 最後に、新学術最終年度、今までどうだったか、どういふところを目指していきたいか教えて下さい。

田中 AIが脳科学に役立つと言われると脳のデータ解析とかツールとしてですが、そうじゃなくてマーの理論で言えば計算論のレベル、何がなされるべきかについて、人間の脳科学は高次脳機能をターゲットにし始めているのでこれから認知科学者、人工知能の発達が脳科学の方向づけに非常に役立つんじゃないかと私は思っています。新学術には最初、私のチーム副リーダー程康さんが入り、発足から半年で急病で亡くなられ交代で入らせていただいた。入って初めて松尾先生にもお会いし、多くを学び、人工知能の研究しておられる方の発想が脳科学の発展に役立つという強い印象を持ちました。DeepMindのHassabisもそう。彼はコンピューター技術と、ヒトのファンクションMRIの神経科学の行ったり来たりをやった稀な存在なんですよ。両方知ってるから、神経科学に対する洞察がずっとやってきた人間と違う味がある。北米神経科学学会の特別講演ではものすごく感激した。こういった、これは面白いなということが非常に楽しかった、これからもきっと楽しむだろう。新学術は大いに期待したけど時間はどんどん経ってしまい限られるけど、もっと勉強したいなと思っています。この近い過去からこれからへの発展についての理解を大いに深め続けたいと思います。

脳科学と人工知能の研究融合を推進する新学術領域「人工知能と脳科学の対照と融合」が立ち上がってから四年が経過し、今年度当該プロジェクトも最終年度を迎えます。融合を推進する企画の一つとしてNeuroscience Research誌において脳科学と人工知能の融合研究に関わる特集号を企画しています。本特集では、当該特集号に本領域から投稿予定の共同研究テーマから三つを取り上げて、どのような内容となるのかを解説していただきました。

特集号に関する情報

論文誌：Neuroscience Research

エディター：坂上 雅道、銅谷 賢治、谷口 忠大

発行予定日：2021年4月

特集タイトル：Brain Inspired Artificial Intelligence

エディターからのメッセージ

AIの新たな発展のために、脳の理解は貢献できるのでしょうか？ 脳機能理解とそのモデル化が、AI研究と接点を持つことにより、新しい脳科学も生まれようとしている。

(坂上 雅道 玉川大学脳科学研究所 教授)

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

谷口 忠大（立命館大学 / パナソニック）、山川 宏（全脳アーキテクチャ・イニシアティブ / 東京大学）

脳の視覚情報処理にヒントを得つつ構築されたディープラーニングは、今日においては画像認識のみならず、音声認識、機械翻訳、強化学習における方策学習といった様々な機能を、データドリブンな形で実現しています。そしてディープラーニングと脳科学の境界分野でも、多数の成果が現れています。例えば、物体認識を実現する視覚系のディープラーニングが学習で獲得した情報表現が人間や猿の視覚野において観察されています。ディープラーニングはその意味で脳に学び人工知能を作り、また人工知能のモデルを通して脳を理解するという循環をある程度生み出していると言えるでしょう。

こうした研究トレンドを踏まえ、2020年以降の新学術領域「人工知能と脳科学の対照と融合」は、次に何をめざすべきでしょうか。新学術領域での研究交流を通して生まれた共同研究のサブグループにおいて領域外部の研究者も巻き込んだ形で議論を重ねてきました。そこで重要と考えるのはディープラーニングや確率モデルの成功を踏まえながら、人の脳のように自律的に自らの末端の情報から学び続けるシステムを生み出すということです。また、そのような自律的なシステムはなにかの認識のためだけのシステムではなく、認識し、行動し、生成し、実世界の中で生存し続けるシステムである必要があります。

そのために機械学習としては人手のラベルデータを前提とするのではなく、原則として自らのマルチモーダルな感覚運動情報を統合することを主としたものである必要があります。なぜなら人手のラベルデータは自律的システムが環境中で行動する中で絶えず供給されるものではないからです。そうした点からみて、現在のディープラーニング研究で不足していると思われる2つの側面について整理してみます。厳密な議論は後ほど出版されるであろう論文に委ね、ここでは素描のみを記述します。

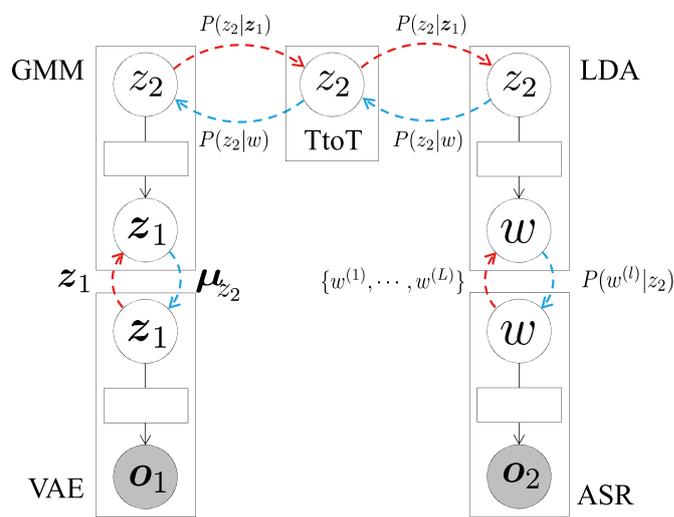
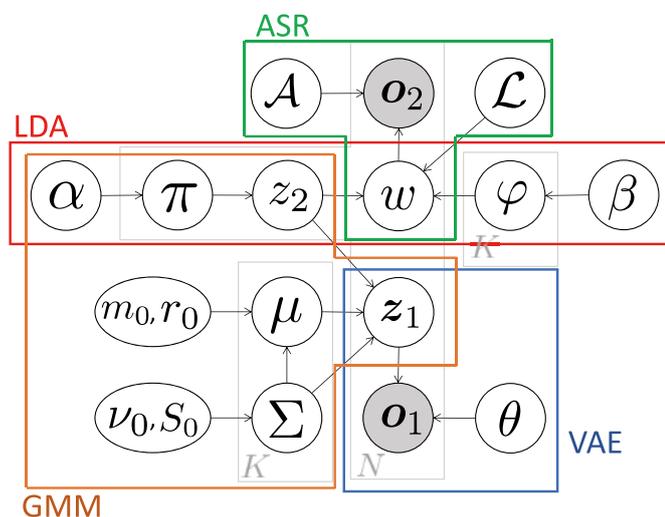
1つ目には、現状で研究開発が進められているディープラーニングのシステムは、人間の脳でいえば個別の認知モジュールに対応する構成単位を「関数」としています。そして、その夫々を最適化することを通して知的なパフォーマンス向上を目指すことを行います。これは、20世紀に主流であった知能を「関数」のような情報処理とみなす、知能観の延長線上にあるものです。オートポイエーシス論などのシステム描像を持ち出すまでもなく、私たち生命は受動的な「関数」ではなく能動的に世界で生存する自律的な「システム」です。その作動を体現しているのが脳であると考えた時に、私たちの研究や議論も、「関数」的な認知モジュールという見方を超えてゆくべきでしょう。

2つ目に、検討すべき点は、現在の多くの研究では、基本的に認知モジュールごと分割して機械学習を行っているということです。これは、実環境で学び続ける人間の知能のモデルを創造しようとする視点からみると、脳内に教師信号を直接注入するようなものであり、相当に不自然です。現在は、知能技術が大きく発展したことを踏まえれば、これまで仮定してきた「全体を一度に考えるのは難しすぎるから個別の認知モジュールに分割して研究しよう」という段階は終わりつつあるように思います。それはつまり、知能のモデルを作るにあたり、各種の感覚運動器の情報を統合し、さらに Lifelong learning な学習をすすめられる脳全体の計算論モデルを構築するということです。このようなシステムの全体を構築するには、その構成要素として、多くの脳器官を実装できる一般性の高い計算モデルを採用する必要があります。さらに、様々な認知モジュールをつなぎ合わせる構造としての認知アーキテクチャが必要になります。認知アーキテクチャの歴史は古く、20世紀には記号処理のための認知アーキテクチャが多く研究されました。しかしながら、神経科学の発展した現在においては、脳を参考としてアーキテクチャを構築することは有力な選択肢です。また、それを支える数学的枠組みも必要でしょう。これを本研究では確率的生成モデルに求めます。そこで、私たちが脳からの学んで目指すべきは、全脳の構造を模倣し、統合的認知アーキテクチャを創造するというアプローチと考えます。

まず統合的認知アーキテクチャを創造するための数理的アプローチについて述べます。すでに述べたように、統合的認知アーキテクチャを構成する認知モジュールは単なる関数を越えた一般性の高い部品である必要があります。そうした技術として、筆者らは確率的生成モデルに注目しています。確率的生成モデルは観測データを予測可能とすることを目的にシステムの内部表現を学習させるという機械学習のアプローチです。

2010年代を通してディープラーニングとの融合であるGANやVAEなども注目を集めました。筆者らは2010年代の記号創発ロボティクスの研究を通してマルチモーダル情報から教師なし学習を通して概念形成や語彙獲得を行う様々な確率的生成モデルを提案してきました^[1,2]。

当該新学術領域の成果として、脳に学びディープラーニングを活用し、スケーラブルに複雑な統合的認知アーキテクチャを構成するための枠組み Neuro-SERKET を提案し論文として発表しました^[3]。これはそれまでに提案してきた SERKET というフレームワークの発展版になります^[4]。Neuro-SERKET では統合的認知アーキテクチャを分割して開発しつつ、これらを学習時および推論時に統合し全体システムとして作動することを可能にする理論的枠組み（フレームワーク）を提案しています。Neuro-SERKET では各認知モジュールが確率的生成モデルで構築されていること、もしくはそれによって近似可能であることを前提とします。左の図は音声データと画像データのマルチモーダル情報から対象のカテゴリを形成する認知システムのグラフィカルモデルです。これが全体として教師なし学習します。これを仔細に見ると音声認識を自動音声認識装置 (ASR: automatic speech recognition) で行い、その認識結果を LDA (latent Dirichlet allocation) でトピック分類する部分と、VAE (variational autoencoder) により画像を表現学習しその潜在空間で内部表現となったベクトルを GMM (Gaussian mixture model) でクラスタリングする部分が結合したモデルになっています。これは異種混合の統合的な確率的生成モデルなのです。このようなシステムの推論アルゴリズムの導出は複雑です。Neuro-SERKET はその推論を各認知モジュールの個別最適化とモジュール間のコミュニケーションに分割します（図右）。



このようなアプローチにより、様々な認知モジュールを脳の結合様式に倣いながらつなぎ合わせ、つなぎ合わせた後でも、全体として教師なし学習が進むことをある程度保証できるのです。

次に脳を参考としてアーキテクチャを構築するアプローチに関してです。先に述べたように、脳を模倣して統合的認知アーキテクチャを構築するため、確率的生成モデルによる認知モジュールを組み上げる足場を脳型の認知アーキテクチャに求めます。すでに、脳型人工知能を構築する足場となる情報を、膨大な神経科学知見から抽出することによる、全脳参照アーキテクチャ (Whole Brain Reference Architecture: WBRA) の構築が進められています。著者の一人である山川が代表を務める全脳アーキテクチャ・イニシアティブ (WBAI) がその開発を主導しています。まず全脳アーキテクチャとは、「脳全体のアーキテクチャに学び人間のような汎用人工知能 (AGI) を創る (工学)」と定義される脳型 AGI の開発アプローチです。WBRA は、その開発を支えるために構築が進められている参照アーキテクチャであり、特定の機械学習モデルを前提とはしていません。WBRA は、主にげっ歯類や人の脳のメゾスコピックレベルの解剖学的構造をベースとし、そこに計算機能についての仮説が付与されています。やや専門的になりますが、WBRA の記述の最小単位は、モジュールを超えて投射する同一タイプのニューロン群です。こうして WBRA を構成するモジュール数は、最終的には数百から数千個程度になることを想定しています。なお、モジュールの計算機能や、モジュール間で伝達される信号の計算論的意味の多くは現状の神経科学で明らかになっていないため、それらについての仮説を構築する研究も、WBAI において促進されています^[5-7]。

なお WBAI は、当該新学術領域が企画した Gatsby Neuroscience Unit との合同ワークショップ (ロンドン、2017) にスピーカーを送り込んだり、共同でハッカソンを運営したりといった連携を行ってきました。今後は、身体を持つ全体の認知システムを動作させる発達ロボティクスの分野で実績のある Neuro-SERKE の枠組みを用いつつ、WBRA を参考として、確率的生成モデルを組み上げることで、統合的認知アーキテクチャ (いわば全脳アーキテクチャ) を創ることにチャレンジしてゆきたいと考えています。

さらにそうした方向性で、Neuroscientific Report で企画される特集号にはより詳細な「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」に関する議論と、脳科学と人工知能の対照と融合に関する道筋、及び到達点をまとめ投稿する予定です。上記に関わる議論にご参画いただくとともに、出版された時には是非ご一読いただければ幸いです。

参考文献

- [1] Tadahiro Taniguchi, Takayuki Nagai, Tomoaki Nakamura, Naoto Iwahashi, Tetsuya Ogata, and Hideki Asoh, Symbol Emergence in Robotics: A Survey, *Advanced Robotics*, 30, (11-12), pp. 706-728, 2016. DOI:10.1080/01691864.2016.1164622
- [2] Tadahiro Taniguchi, Emre Ugur, Matej Hoffmann, Lorenzo Jamone, Takayuki Nagai, Benjamin Rosman, Toshihiko Matsuka, Naoto Iwahashi, Erhan Oztop, Justus Piater, Florentin Wörgötter, Symbol Emergence in Cognitive Developmental Systems: A Survey, *IEEE Transactions on Cognitive and Developmental Systems*, 11(4), pp.494-516, 2019. DOI: 10.1109/TCDS.2018.2867772
- [3] Tomoaki Nakamura, Takayuki Nagai, Tadahiro Taniguchi, SERKET: An Architecture For Connecting Stochastic Models to Realize a Large-Scale Cognitive Model, *Frontiers in Neurorobotics*, 12(25), 2018. DOI: 10.3389/fnbot.2018.00025
- [4] Tadahiro Taniguchi, Tomoaki Nakamura, Masahiro Suzuki, Ryo Kuniyasu, Kaede Hayashi, Akira Taniguchi, Takato Horii, Takayuki Nagai, Neuro-SERKET: Development of Integrative Cognitive System through the Composition of Deep Probabilistic Generative Models, *New Generation Computing*, 84(0), 2019
- [5] Fukawa, A., Aizawa, T., Yamakawa, H., & Yairi, I. E. (2020). Identifying Core Regions for Path Integration on Medial Entorhinal Cortex of Hippocampal Formation. *Brain Sciences*, 10(1). <https://doi.org/10.3390/brainsci10010028>
- [6] Yamakawa, H. (2020). Attentional Reinforcement Learning in the Brain. *New Generation Computing*. <https://doi.org/10.1007/s00354-019-00081-z>
- [7] Yamakawa, H. (2020). Revealing the computational meaning of neocortical interarea signals. *Frontiers in Computational Neuroscience*, 14, 74.

自然「非」言語への挑戦

三村 喬生 Koki Mimura (量子科学技術研究開発機構 National Institutes for Quantum and Radiological Science and Technology (QST))

「data analysis」という言葉を初めて使ったのはかのJ. W. Tukeyですが、1961年に書かれたその論文の最後にはデータ解析の未来に対する彼の思いが記されています。曰く「データ解析は今後あらゆる分野のサイエンスやテクノロジーに大きく貢献するでしょう。本当に?それは私たち次第です。安易な道ではなく、真の問題を取り扱うという岩だらけの道(the rocky road of real problems)を歩もうとする私たちの意思に掛かっているのです」。本稿では、私たちが取り組んでいる社会性における非言語要素の理解に向けて、現時点で展望している岩だらけの道をまとめてみました。

ある論理体系を発展させるために、そこからいくつかの仮定を除いた公理系から出発し、より一般化された論理体系への展開を試みるという方法があります。例えば脳科学は、知能はヒトに固有(いわばHI)という仮定を取り除き、脳に属する知能(BI)へと拡張することでモデル動物を用いた詳細な神経機能の探究へと発展してきました。同様に、人工知能(AI)という概念の導入は、Natureにのみ属するという仮定を取り除くことで知能研究の裾野を押し広げる試みといえます。この視点に立つとBIとAIの対照と融合は、拡張された知能という概念に正面から取り組むプロセスで発展していくと考えられます。

現在、私たちが取り組んでいる社会性という知能の研究では、AI分野においてヒトの言語にフォーカスした先端的な研究成果が報告されています(自然言語処理)。ヒトの言語は世界に約7千種類ほど知られていますが、その基盤には本質的に同一の構造があり、確率モデルとして解いていくことで文章を要約・分類したり「人間らしい」文章を計算機的に創造していく試みです。一方、BI研究として言語からヒトに固有という仮定を取り除くにはどうすればよいでしょうか。霊長類は種ごとに特異的な音声パターンを持っており自己の内部状態(情動・意図)に基づいて使い分けることが知られています。脳の解剖学的構造が進化的にある程度保存されていることを考慮すると、モデル動物の音声コミュニケーションを調べるのは有望なアプローチの1つに見えます。ただしボキャブラリや統語構造といった点で動物の音声はヒト言語と比べると圧倒的に単純であるため、ただちに自然言語処理と融合した展開は困難かもしれません。

そこで少し視野を広げると、社会性は言語という公理系の中で閉じていないことに気が付きます。「目は口ほどにものを言う」という格言もありますが、言語ほどの複雑で高度な情報伝達手段を持ってしても我々は社会的場面において非言語の要素に多くを依っています。これは例えばモートワークにおけるビデオ会議の重用もそうですし、あるいは私自身は0歳児の親として日々

実感しているところです。乳幼児期の社会的コミュニケーションは自閉症リスクの観点からもよく研究が進んでおり、そのチェック項目は当然のことながら非言語要素によって構成されています。また実際に霊長類の日常生活を観察すると、その振る舞いがとても「人間らしい」ことに気が付きます。彼らは言語に依らないコミュニケーションのエキスパートで、ヒトは言語を発達させることでこうした手段をむしろ単純化してきた可能性すらあります。そうするとこれは話が逆さまで、ヒトの非言語コミュニケーションが「霊長類らしい」と表現するのが妥当です。従ってここに共通の公理系を仮定するアプローチも十分に有望であり、非言語の社会性という視点に立つことでヒトに固有という仮定を取り除きBI研究としての発展を臨むことが可能となります。

このいわば自然「非」言語処理は、AIとの親和性も高いテーマです。例えば映画スター・ウォーズに登場するロボットR2D2はヒトとは異なる外形を持っていますが、エモーショナルな動作がピーブ音と上手く組み合わせ、自然な社会性を感じさせます。ヒトが演じた動きをトレースしたり、それっぽく見える動きのパターンを事前にプログラムするという現在の手法から脱却し、状況に応じて自らの身体構造に合わせた自然な動作を自律的に創造するロボットが作れば、これは非言語的であっても社会性を持ったロボットというに相応しいでしょう。無数にありうる環境情報のライブラリとそれに対応する動作を1つずつ事前に登録していく手法には限界があり、何らかの創発的なモデルが必要です。これは自然言語処理が抱えているのと本質的に同じ課題であり、言語・非言語を跨いだ社会性そのものを理解する上で、まさにreal problemになっていると考えています。

幸いなことに、物体認識技術の躍進によりヒトや動物の身体運動(視線・姿勢・動作など)を詳細にトラッキングすることが可能となってきました。我々も深層学習と物理演算を組み合わせた小型霊長類用のモーショントラック技術を独自に開発し、ボディマーカなどを用いずに顔の向きや頭の位置を三次元的に測定することに成功しています。現在は特定の脳領域(例えばヒトの社会性障害と関連づけられている扁桃体)の神経活動を可逆的に制御する技術を用い、動物同士の社会的場面における姿勢・動作の時系列な文脈を捉えようと試みています。これには自然言語処理から派生した統語構造の解析技術を援用できる可能性があり、ある種の単純な合目的行動の文脈解析において上手くワークしつつあります。こうした私たちの自然「非」言語への取り組みはまだまだBIの範囲に留まりますが、上記のようにN+Iに拡張された知能研究という観点からの長期的な展望を持ち、歩み続けたいと思います。

言語の計算認知神経科学に向けて

- Towards computational cognitive neuroscience of language -

大関 洋平（東京大学 大学院総合文化研究科 言語情報科学専攻 講師）

計算神経科学の父であるMarr (1982)は、人間の脳に代表される情報処理システムを総合的に理解するためには、「計算理論」、「アルゴリズム」、「ハードウェア」という3つのレベルが必要であると主張しました。ここで重要なのは、人間の言語処理も究極的には脳内で実現される情報処理であるという点であり、言語という高次認知機能もMarrの3つのレベルに位置付けることが出来ます。特に、言語の認知科学である言語学は「計算理論」レベル、言語を含む高次認知機能の神経科学である認知神経科学は「ハードウェア」レベルに対応すると考えられ、言語学と認知神経科学の協働により、言語の認知神経科学は飛躍的な進歩を遂げました。

その一方で、様々な問題も指摘されています。例えば、言語学と認知神経科学の間には、研究粒度やオントロジーに関して、簡単に埋めることの出来ない乖離が生じており、言語の認知神経科学は言語処理と神経活動の相関の検証に留まっているという批判があります (Embick & Poeppel 2014)。ただ、Marrの3つのレベルに立ち返って考えてみると、一つの「計算理論」に対して複数の「アルゴリズム」が理論的に存在し、一つの「アルゴリズム」に対して複数の「ハードウェア」が理論的に可能であるという相互依存の関係を考慮すると、中間の「アルゴリズム」レベルを無視して「計算理論」レベルと「ハードウェア」レベルを所謂「力技で」繋げようとする従来の言語の認知神経科学に限界があるのも当然なのかもしれません。従って、今こそ言語という高次認知機能に対する新しいアプローチが求められています。

そこで、この問題を解決するため、言語に対する「計算認知神経科学」(computational cognitive neuroscience)のアプローチを提案したいと思います (図1)。計算認知神経科学とは、Kriegeskorte & Douglas (2018)によって提案された認知科学・人工知能・計算神経科学の複合領域で、Naseralis et al. (2018)によって同名の国際会議も設立され、Marrの3つのレベルを統合するアプローチとして注目されています。ここで、言語の計算認知神経科学を実現するために重要な役割を果たすのが、言語の人工知能研究である「自然言語処理」(Natural Language Processing; NLP)です。NLPは、昨今の機械学習やビッグデータの恩恵を受けて急速に発展していますが、工学的な目的で開発された確率生成モデルや人工ニューラルネットワークなどを高次認知機能の計算モデルとして「アルゴリズム」レベルに介在させることで、「計

算理論」レベルと「ハードウェア」レベルを「自然に」橋渡しすることが可能になります。

言語の計算認知神経科学と一概に言っても様々な可能性がありますが、大別して「NLP→脳科学」と「脳科学→NLP」という2つのベクトルがあります。最近では、NLPで開発された単語の分散表現やエンコーダ・デコーダを神経活動の解析に援用する「脳情報デコーディング」(Huth et al. 2016, Anumanchipalli et al. 2019)など「NLP→脳科学」のベクトルが主流であり、ブレインマシンインターフェイスへの臨床応用など脚光を浴びています。今後は、NLPで開発された計算モデルを人間の神経活動で評価する「神経計算モデリング」(Hale et al. 2018, Toneba & Whebe 2019)など「脳科学→NLP」のベクトルが必要であり、川人他 (2000) が正しく指摘している様に、認知科学からのトップダウン的アプローチと計算神経科学からのボトムアップ的アプローチの融合が求められています。言語の計算認知神経科学の実現に向けて、人工知能・記号創発ロボティクスの谷口忠大氏および計算神経科学の深井朋樹氏 (共に当領域メンバー) と共同研究を実施している真っ最中です。

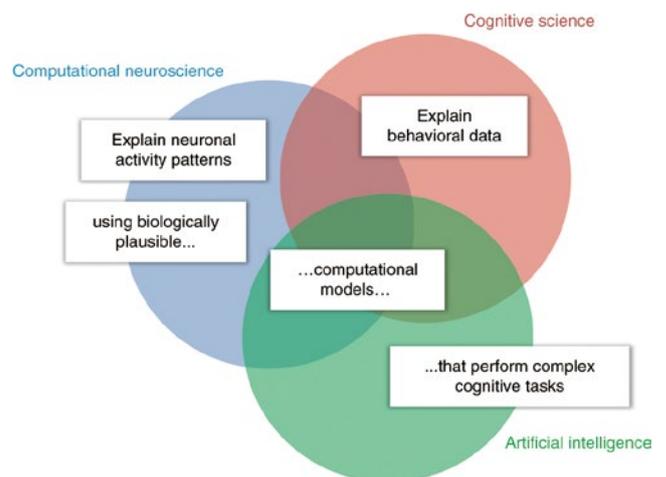


図1: 計算認知神経科学の概要。認知科学・人工知能・計算神経科学の複合領域であり、Marrの3つのレベルを統合するアプローチとして注目されている。図はKriegeskorte, N., Douglas, P.K. Cognitive computational neuroscience. Nat Neurosci 21, 1148-1160 (2018)より引用。

進化的・発達の知能の創成に向けて

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）

論文タイトル: Toward evolutionary and developmental intelligence

著者: Doya K, Taniguchi T

Current Opinion in Behavioral Sciences, 29, 91-96. (2019).

DOI: <https://doi.org/10.1016/j.cobeha.2019.04.006>

この論文は、DeepMindのMatt BovinickとHarvardのSam Gershmanが企画したCurrent Opinion in Behavioral Sciences誌の“Artificial Intelligence”特集号に、当領域の谷口忠大さんと共著で寄稿したものです。

<https://www.sciencedirect.com/journal/current-opinion-in-behavioral-sciences/vol/29/suppl/C>

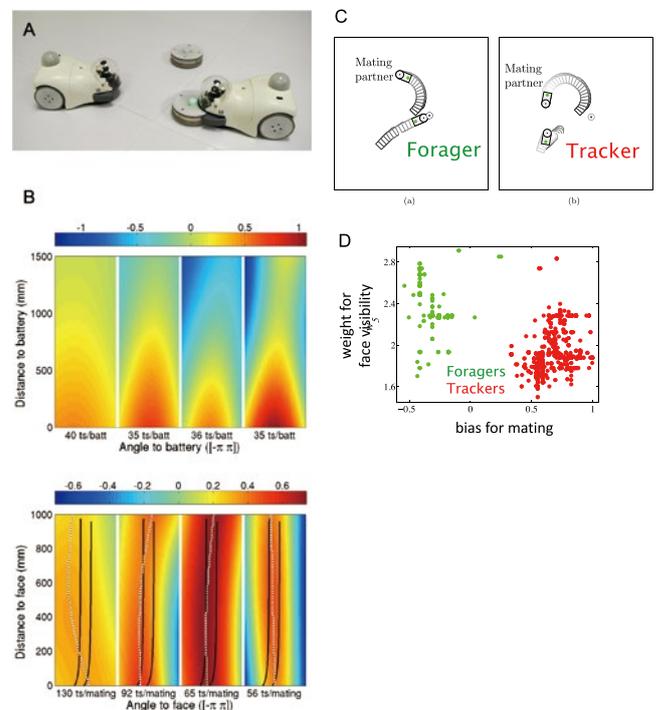
この特集号にはメンタルシミュレーション、メタ学習、認知科学と機械学習の相互作用など、興味ある論文が多数掲載されており、ぜひ眺めてみることをお勧めします。

私たちの論文で強調したことは、Artificial General Intelligenceといったものを作る上で、それが単に個別の専門AIの寄せ集めでは面白くないし限界がある、人や動物が生まれ落ちた物理的社会的環境のもとでそれら内部モデルや方策を獲得していく過程、またそれを可能にする学習アーキテクチャや報酬系が進化して来た過程を理解して活用することが、遠大そうに思えるかもしれないけど結果的には近道になるのでは、ということでした。

まず、発達心理学、認知神経科学の分野での最近の展開や、私たちが取り組んできたメンタルシミュレーションの脳回路(Ferminn et al. 2016)とニューロン機構(Funamizu et al. 2016)を解明しようとする研究について紹介し、脳の発達学習機構を解明しそこから学ぶことが現実的な選択となり得ることを主張しました。また認知発達、進化ロボティクスの側から、環境や人との相互作用のもとでマルチモーダルな確率生成モデルが獲得可能なこと(Taniguchi et al. 2017)や、獲得したモデルを柔軟に組み合わせるアーキテクチャSERKET(Nakamura et al. 2018)について、また強化学習と進化アルゴリズムを組み合わせることにより、ロボットが自らの報酬関数や異なる生存戦略を獲得し得る(Elfwing et al., 2011, 2014)といった例を挙げて、発達し進化する人工知能が夢ではないということを主張しています。

もちろん、生物の進化をそのまま再現するには何十億年かかるかもしれませんが、インターネットに繋がったロボットやAIエージェントでは、テレパシーや脳移植、獲得形質の遺伝や組み合わせなど、本物の生物では不可能な離れ業が可能であり、AIが発達し進化できるような仕組みを用意することの方が、万能AIを人力で組み上げようとするよりも現実的なのではないか、というのがこの論文の主張です。

ご一読いただきコメント批評をいただけたら、あるいはこのような研究の展開に加わっていただけたら幸いです。



分散進化ロボット(A)による報酬関数の獲得(B; Elfwing et al. 2011)と繁殖戦略の多型性(C, D; Elfwing et al. 2014)。

細胞体-樹状突起間の整合性確認による時間的特徴の分割

深井 朋樹 (沖縄科学技術大学院大学 教授)

論文タイトル: Somatodendritic consistency check for temporal feature segmentation

著者: Asabuki T, Fukai T

Nature Communications 11(1):1554 (2020)

DOI: <https://doi.org/10.1038/s41467-020-15367-w>

概要

ニューロンの樹状突起は、入力情報の統合やルーティング、シナプス可塑性の制御において、重要な役割を果たすと考えられています。本研究では単純化した数学的モデルを用いて、細胞体と樹状突起の相互作用に依り入力情報の時間的特徴を検出するための、新しい教師なし学習規則を提案しました。

はじめに

本領域は、「AIを脳の研究に役立てよう」、「脳から学んだ知見をAIに役立てよう」という二つの目標を掲げています。本稿では、後者の目標に関する、私の研究室の最近の試みを紹介します。

なぜ樹状突起なのか

まず、樹状突起が重要だと考える根拠について少し述べたいと思います。大脳皮質などでは、異なる脳領域からのシナプス入力、樹状突起上の異なる部位に投射する傾向がありますが、このことは、全脳レベルの情報の流れの制御に樹状突起が関与していることを示唆します。また樹状突起は、線形加算と非線形演算を連続的に行う線形-非線形フィルターで近似可能なことが知られており、単一細胞に複雑な神経回路と同等の信号処理機能を与えます。さらに樹状突起スパイクや逆行伝搬スパイクなど、樹状突起にはシナプス可塑性を制御する仕組みが備わっています。さらに最近の実験は、細胞体とは異なり、樹状突起が学習する刺激の特徴（例えば視覚刺激の線分の傾き）は、複数あることを明らかにしました。樹状突起の情報表現がもつ冗長性の計算論的役割を理解できれば、脳の優れた情報処理能力の実現に、一歩近づくことが可能です。

今回の研究では、樹状突起計算の数学的原理の定式化を目指して、細胞体と樹状突起を一つだけ有する、2コンパートメントモデルを構築しました。生物学的に詳細な樹状突起のモデルは存在しますが、樹状突起に高度な学習・計算機能を求

める研究はまだあまりありません。単純化された数学的モデルは実験との対応が取りづらくなりますが、工学的応用の面では有利です。

樹状突起と細胞体間に於ける「情報損失の最少化」

一言で言えば、私達が提案した2コンパートメントモデルは、シナプス入力の中に繰り返し出現する時空間パターンを抽出することを、「教師なし」に学習します。このモデルは、ベルン大学の Walter Sennらが提案した「教師有り」学習の2コンパートメントモデルを概念的に発展させて、教師なし学習を実現しています。彼らは、樹状突起はシナプス学習を通じて細胞体の発火活動を予測するという考えを提案しました。細胞体が教師信号を受けて発火している場合、樹状突起は細胞体の活動を真似るように、外部入力のシナプス結合荷重を修正するというわけです。これは非常に面白いアイデアですが、細胞体から樹状突起に向けて発せられる逆行伝搬スパイクの役割が、今ひとつ明らかではありません。

そこで私達は、教師信号は外部から与えられるのではなく、逆行伝搬スパイクに依り、細胞体から樹状突起にもたらされるという考えに基づき、モデルを構築しました。樹状突起は外部からシナプス入力を受け、その総和が細胞体に伝達されて活動電位生成（神経発火）が起きますが、このとき樹状突起活動の確率分布が、細胞体の発火率分布の最適確率モデルになるように、樹状突起のシナプスを学習します。数学的には、この学習は細胞体活動と樹状突起活動の確率分布のクルバック・ライブラー発散（情報損失）を最少化することで実現されます（図1A）。樹状突起の活動分布は、樹状突起も活動電位を生成するかの如く扱うことで計算されますが、樹状突起スパイクを生成するシナプス入力のみ考慮することに対応するのかもしれない。

上述のようにすると、細胞体への陽な教師信号がなくても、ニューロン全体では教師なし学習が実現できます。ただし問題は、このニューロンが学習する入力の特徴は何なのかということです。その答えは、樹状突起が細胞体活動を予測できる

のはどういふ場合か考えれば、明らかになります。樹状突起が単にランダムな外部入力を受けている場合は、細胞体の発火もランダムになり、予測は不可能です。ところが外部入力がある種の繰り返しパターンを含む場合、このパターンを選択的に学習することで、予測が可能になります (図1B)。

実際に数値実験に依り、このニューロンモデルは様々な入力に含まれる繰り返し構造、つまりチャンクを学習可能なことが確認できました。このニューロンモデルを抑制シナプスで結合した競合神経回路は、複数のチャンク構造を学習できますが、特筆すべきなのは、同じ回路モデルを用いて混合音からの音源分離も実行可能なことです (図1C)。従来は問題ごとに個別のネットワークと学習ルールを必要としていた問題が、同一の回路モデルで解けるのは、大きな進歩と言えます。

チャンク学習の重要性

最後に、チャンク学習の重要性について一言述べておきた

と思います。脳は外界の統計モデルを構築して未来に起こる事象を予測すると考えられており、この過程を記述するために、部分観測隠れマルコフモデルなどの数理モデルが提案されています。良い統計モデルを構築するためには、脳は時系列的に入力される情報の中から、外界を特徴づける“変数”を検出し、学習する必要があります。しかし脳が外界を記述する適切な変数を見出すメカニズムは、ほとんど研究されていません。このメカニズムは情報の圧縮を伴うべきものであり、相互情報量最大化のような枠組みでは記述できません。チャンクの検出は、この情報圧縮のために必要なステップであると考えられます。現在、上述した問題に関して、機械学習の研究者との共同研究が始まろうとしています。また日立製作所との間で、提案モデルをニューロモルフィック技術に応用するための検討も始めています。

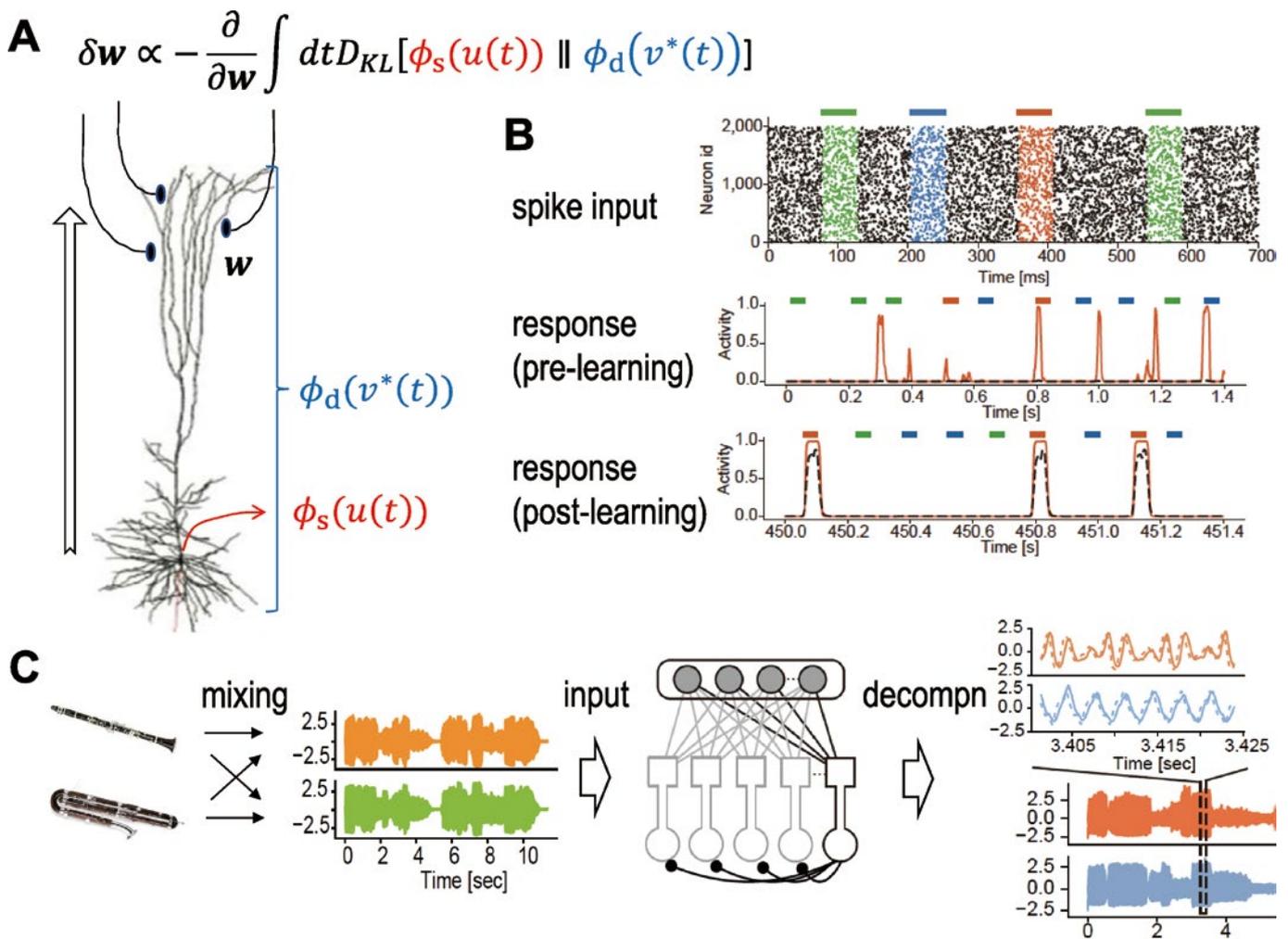


図1: モデルと学習の概要. (A) 細胞体活動 u と樹状突起活動 v で表現される2コンパートメントモデル. (B) 単体ニューロンへのスパイク入力と学習前後の活動. この例では赤色表示された反復入力パターンを学習. (C) クラリネットとファゴットの混合音を、音色を検出することで分離した例.

分布マッチングの観点からみた敵対的不変表現学習の安定化

岩澤 有祐（東京大学 工学系研究科 技術経営戦略学専攻 松尾研究室 特任助教）

論文タイトル: Stabilizing Adversarial Invariance Induction from Divergence Minimization Perspective

著者: Iwasawa Y, Akuzawa K, Matsuo Y

会議名: International Joint Conference on Artificial Intelligence (IJCAI), 2020

はじめに

表現の不変性とは、ある表現が特定の因子に対して独立であることを指す概念です。不変性が重要な例として監視カメラの動画からの不審行動を検知することを考えます。このシステムが背丈などの身体的特徴によらず不審行動を検出するには、身体的特徴に依存しない情報（表現）を活用した判断をする必要があります（ロバスト性）。あるいは、例えば肌が黒い人がいる、という情報を抽出しそのことを判断の基準として利用することは社会通念上問題になる場合があるため、検知システムは肌の色に依存しないことが要請されます（公平性）。しかし、表現学習の手法として近年よく使われる深層ニューラルネットワーク（以降DNNと表記）は、獲得した表現がこのような不変性を持つ保証はありません。

本論文では、上記のような不変性の制約をDNNの学習に明示的に取り入れる新しい手法を提案しました。提案手法は、既存研究である敵対的特徴学習の学習が不安定になるという問題を、ダイバージェンス最小化の観点から着想を得たアイデアにより改良したものです。実験では、既存手法が学習に失敗するようなトイデータで提案法が一貫して最適に近い不変性を達成できること、また実際的な応用例として、分布シフトに対する分類器のロバスト性の改善、ウェアラブルセンシングにおけるプライバシー保護に利用できることを示しました。

敵対的特徴学習の不安定性

本研究の技術的な貢献は、既存手法である敵対的特徴学習（以降AIIと表記）の不安定性について一定の説明を与えた上で、この問題を解決するシンプルな方法を提案したことです。AIIの肝は、「現在の表現がある因子 a の情報をどれくらい持つか」を「ある観測に関する表現 $f_\theta(x)$ から x と紐づく因子 a をどれくらい正確に予測できるか」として捉えなおし、この基準に沿って表現が不変になるように学習することにあります。これは、形式的には次のような敵対的ゲームを最適化することに相当します。

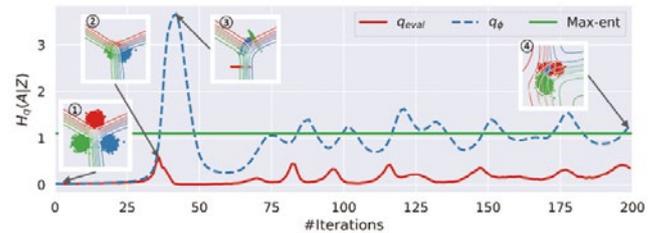


図1 トイデータを用いたAIIの不安定な挙動の可視化。

$$\min_{\theta} \max_{\phi} \mathbb{E}_D [q_{\phi}(a|f_{\theta}(x))].$$

ただし、 D はデータセットであり、 $q_{\phi}(a|f_{\theta}(x))$ は表現 $f_{\theta}(x)$ から属性を予測するカテゴリカル分類器です。直感的には、maxフェーズで表現の不変性を推定し、minフェーズで推定値にもとづき表現が不変になるように学習する、ということになります。また、情報理論的な観点からも、妥当な不変性の尺度である条件付きエントロピー $H(a|z) = \mathbb{E}_D [-p(a|f_{\theta}(x))]$ と密接な関係にあるなど理論的に優れている点が知られており、様々な拡張手法が提案されてきています。

このような利点の一方で、AIIの実践的な挙動は不安定であることが報告されています。わかりやすい例として、図1に3つのガウス分布（図1の①参照）からなるトイデータを、それぞれの分布を属性の違いとしてみて不変になるように（すなわちすべての分布が重なるように）学習した際のAIIの挙動を示します。ここで、緑線は理論上最適な不変性、赤線が実際の不変性、青線がAIIが学習に利用している推定値を表しています。この結果は、推定値が実際の不変性と大きく離れてしまっており、また推定値が乱高下するなど不安定な挙動を示しています。

分布マッチングの観点からみた敵対的特徴学習の安定化

では、このような問題はなぜ発生するのでしょうか?またどのように解消できるのでしょうか?本研究では、この問題について分布マッチングの観点からみた提案を行いました。ここで分布マッチングとは、ある属性に関する特徴表現の分布を、別の属性に関する分布を近づけるような処理のことです。このような分布のマッチングを考える直感的な理由は、「表現がある属性に対して不変であること」は「特徴表現の分布が、属性によらず同じであること」と言い直すことができるためです。より形式的には、すべての属性の実現値のペアに対して特徴表現の分布が同じになる場合のみ、AIIが最大化しようとしている条件付きエントロピーが最大化されることを示すことができます(詳しくは元論文を参照ください)。

残念なことに、このような分布マッチングの観点はAIIが最適化している目的関数(あるいはその勾配)には現れません。図2(中央)に、先ほど述べたトイデータにおいてAIIの学習規則によって得られる勾配ベクトル(属性1=赤点のデータ分布が学習によってどのように移動するか)、およびその勾配ベクトルを計算するために使われる要素を可視化したものを示します。このように、AIIはデータ分布を更新する際に別の属性のデータ分布がどこにあるかという情報を全く利用しない(厳密には、分類境界を介して間接的にしか利用しない)値ということがわかります。そのため、「分布マッチングするかにかかわらず決定境界から離れるほど良い」という学習規則になっています。

本研究では、上記の「分布マッチングの欠如」がAIIの不安定性を引き起こしているという仮説のもと、AIIの枠組みを分布マッチングを考慮するように変更した手法を提案しました。AIIと同様、提案法も敵対的学習の枠組みを利用する、すなわち、特徴表現から因子を予測する分類器を学習中に利用しこの分類器を騙すことで不変性を達成するように学習します。ただし、提案法は分類器を騙す際に、分類器の出力が属性によらず同じになるような方法をとります。因子を予測する分類器はよくDiscriminatorと呼ばれるため、提案法をInvariance Induction by Discriminator Matching (IIDM)と表記します。直感的には、分布を動かす方向がほかの属性に対応したデータ点の情報に依存して決まるようになるため(図2右参照)、AIIより分布マッチングを明示的に制約として取り入れることができます。より厳密な関係については、こちらも元論文を参照ください。

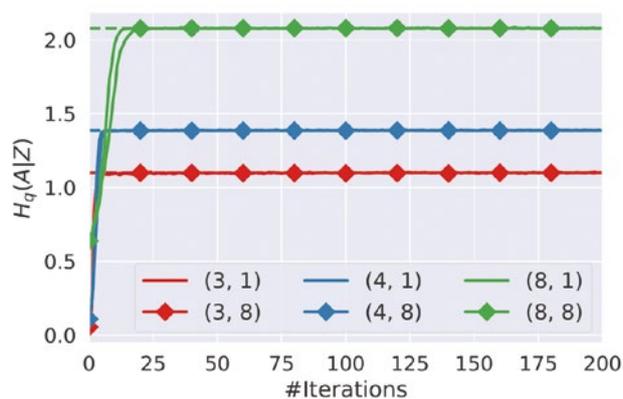
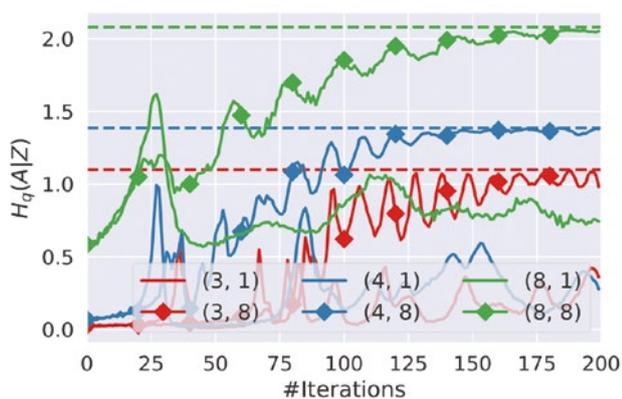


図2 AII(左)と提案法(右)のトイデータでの比較。

実験結果

図3は、図1の実験で利用したトイデータと同様のデータにAII(左)とIIDM(右)を適用した結果を示しています。図中の線の色やマークは問題設定の微妙な違い(ガウス分布の数と、最適化の際に分類器 $q_\phi(a|\cdot)$ をどれだけ学習するか)を表しており、点線が理論的な最適値を表しています。この結果は、(1) IIDMはすべての条件でほぼ最適な値に安定して収束していること、(2) AIIは問題設定を変えても不安定であることを示しています。さらに、画像認識における分布シフトに対する分類器のロバスト性を改善させる、ウェアラブルセンシングにおけるプライバシー保護といったAIIが使われてきた応用例でより良い性能を達成することを示しました。

おわりに

本研究では、DNNがある因子に不変になることをシステムチックに導入する敵対的学習にもとづく手法を提案しました。畳み込みネットが移動不変性を組み込むことで画像領域でよい性能を達成しているように、なんらかの因子に対する不変性を組み込むことは、今回実験として用いた画像認識やウェアラブルセンシングだけでなく様々なドメインで重要になると思われます。一方、今回の手法をはじめとした既存手法は因子がカテゴリカルな場合に限定されているといった課題があり、この点の解決に向けた研究を行う予定です。

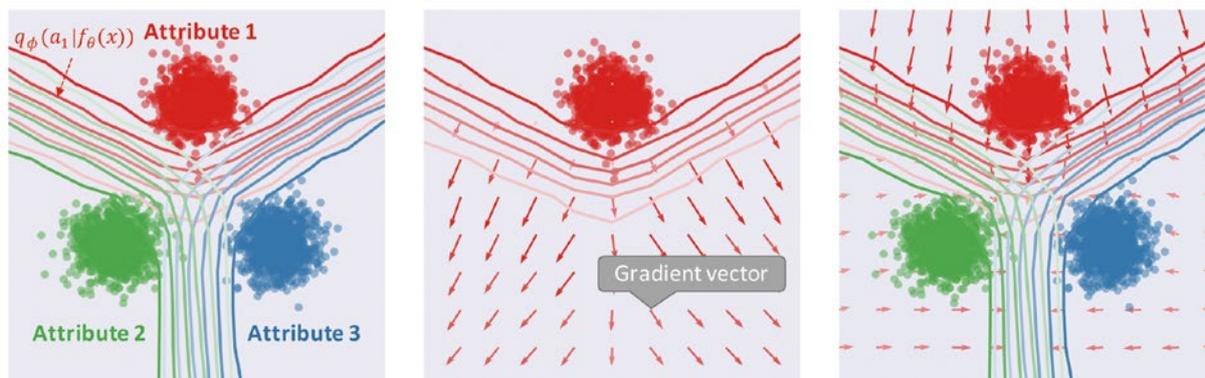


図3 AII(中央)と提案法(右)の敵対的学習における勾配ベクトルの比較。両手法とも特徴空間上での属性分類器(左図等高線)を騙すことで不変性を獲得するが、騙し方が異なる。

全神経活動のイメージングのための神経細胞のアノテーション

豊島 有（東京大学大学院 理学系研究科 助教）・石原 健（九州大学大学院 理学研究院 教授）・
飯野 雄一（東京大学大学院 理学系研究科 教授）

論文タイトル: Neuron ID dataset facilitates neuronal annotation for whole-brain activity imaging of *C. elegans*
著者: Toyoshima Y, et al.
BMC Biology, 18, 30 (2020)
DOI: <https://doi.org/10.1186/s12915-020-0745-2>

概要

神経細胞のネットワークである脳が、いかに情報を処理しているかを生の生物を用いて明らかにすることは、脳科学の主要な目標です。線虫という小さな動物では、302個ある神経細胞のそれぞれに名前がつけられ、お互いのつながりのグラフ構造（コネクトーム）が明らかになっています。私たちはこの生物にカルシウムプローブを導入して、頭部全神経細胞のイメージングによる活動測定を行っています。この活動データを解釈するためには、観測された神経細胞のそれぞれに名前をつける（アノテーションを行う）必要がありますが、実はそれは難しい課題であることが判明しています。そのための基礎知見として、私たちは神経細胞の配置のデータを系統的に収集することを行いました。さらに神経マーキング株を作製し、自動アノテーションアルゴリズムを開発することにより、活動イメージングのデータに神経回路情報を対応づけて解釈できるようになりました。

はじめに

神経細胞のつながりを知った上で活動測定ができれば、脳神経系がどうやって複雑な情報処理を行うかを格段に効率よく解明できるはずですが、しかし現時点でこれができる生物材料は限られています。

線虫 *C. elegans* は体長が約 1mm ほどと小さいため、ニポウディスク方式の共焦点装置とピエゾ駆動型対物レンズを用いることにより、頭部のすべての神経細胞の活動をタイムラプス三次元画像（4D 画像）として取得できます。画像上のそれぞれの細胞が、302 個のうちどの神経細胞なのか、名前をつけることができれば、観察された神経活動の時系列データを回路（グラフ）に対応づけることができるようになり、神経回路が情報を受け取って処理していく様子がわかるはずですが、神経細胞（細胞体）の配置は個体間である程度似ていますが、位置のばらつきも大きく、個々の細胞の位置情報だけ

で名前がつけられるかはわかりませんでした。

1. 神経細胞の配置のデータの収集

そこで私たちは、さまざまな細胞特異的プロモーターを用いて、GFP などの蛍光タンパク質を特定のサブセットの神経で発現させた線虫株を 35 種類作成し、計 311 個体の頭部の三次元蛍光画像を共焦点顕微鏡で取得しました。蛍光タンパク質を発現した細胞を目印にして、他の神経細胞にも可能な限り名前をつけ、個々の細胞体の位置を調べました。その結果、ひとつの細胞が存在する範囲は、隣の細胞の存在範囲と大きく重なっている場合が多く、こうしたばらつきのため、細胞の位置だけを頼りに名前をつけることは難しいことがわかりました（図 1）。卵からの孵化直後は神経細胞の配置は一定していますが、成長の過程で個体ごとに細胞体の配置が変化するようにです。

2. マーカー株の作製

細胞の配置がばらついていても、目印となる細胞があれば、特定の細胞に名前をつけることができます。目印細胞の数を増やしていくと、手がかりは増えるものの、目印細胞同士の区別が次第に難しくなります。そこで私たちは、発現パターンを複数組み合わせることで細胞を塗り分けることにしました。色の異なる蛍光タンパク質を 2 つ用いれば、組み合わせで 4 色に塗り分けることができます。35 種類の細胞特異的プロモーターのパターンの組み合わせから最適な組み合わせを選び出し、神経細胞が 4 色に塗り分けられた線虫株を作成しました。この線虫株では、ひとつの個体内で名前のつく細胞の数は約 160 個（従来比 3.6 倍）に増え、線虫頭部のほとんどすべての神経細胞に名前がつけられるようになりました（図 2）。

次にこの株に、蛍光の強さにより神経活動を測れる蛍光カルシウムプローブタンパク質（カメレオン）を導入することで、神

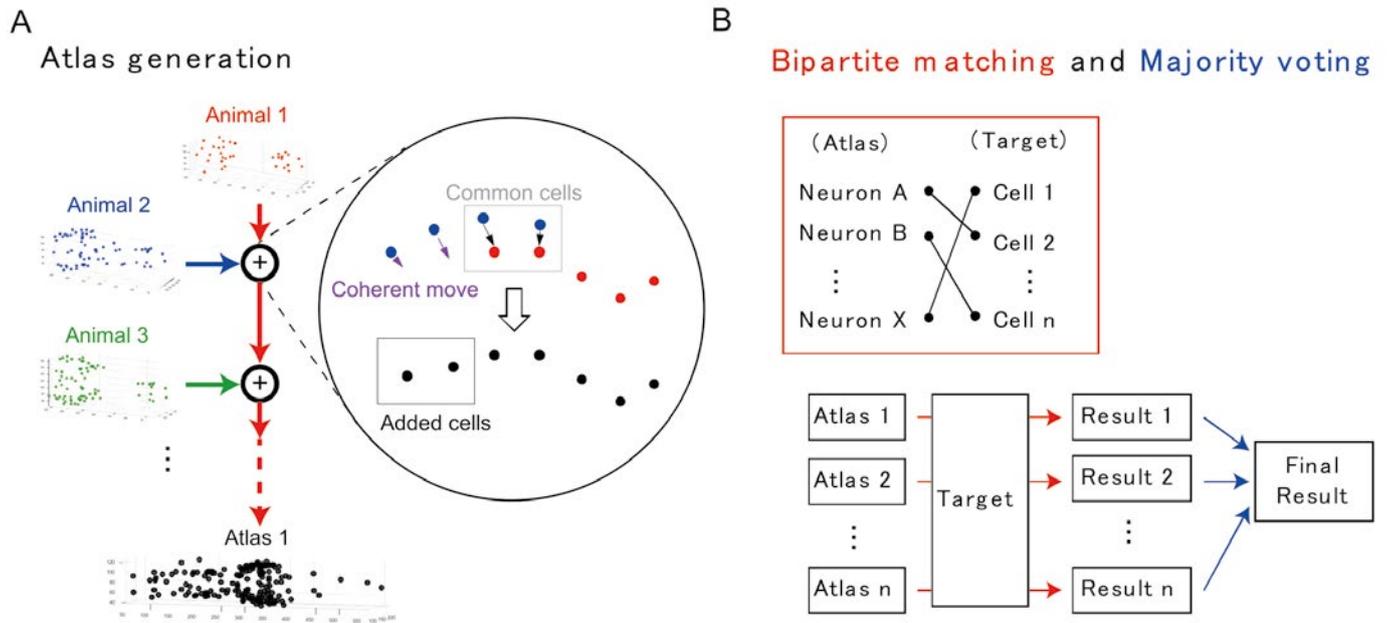


図3:自動アノテーションのためのアルゴリズム。

A) 311 個体のそれぞれは名前付けに成功した神経が異なる。共通に名前がつけられている神経を手掛かりに細胞位置をアラインメントしていき、最後にほぼ全神経を含む細胞位置アトラスが出来上がる。

B) 複数のアトラスを用いたアンサンブル学習により最適解を得る。

運動中の伸張反射応答は視覚的な身体情報を介して調整される

伊藤 翔 (NTTコミュニケーション科学基礎研究所 研究主任)

論文タイトル: Visually-updated hand state estimates modulate the proprioceptive reflex independently of motor task requirements

著者: Ito S, Gomi H

eLife: 9:e52380 (2020)

DOI: <https://doi.org/10.7554/eLife.52380>

概要

人間の身体動作は意識に上らない様々な感覚-運動制御機構によって支えられています。伸張反射は、そのような無意識の運動応答の一つであり、姿勢の安定化やすばやい運動修正に役立っていると考えられています。しかし、その応答ゲインを調整するための基盤となる身体状態推定の詳細は明らかになっていません。本研究では、運動中の伸張反射応答を計測する実験を行い、身体状態を表す視覚情報に操作を加えることで反射応答が小さくなることを示しました。この結果は、伸張反射の応答ゲイン調整が体性感覚情報だけでなく、視覚情報も統合した身体表象に基づき行われていることを示唆します。

はじめに

人間の運動は、多重の感覚-運動制御ループによる階層的な情報処理によって安定に実行されていると考えられます。そのうちの 하나가伸張反射（筋紡錘からの感覚入力に対して筋収縮を起こす運動応答）であり、体性感覚入力によって駆動される応答時間が比較的短いフィードバックループを形成しています。伸張反射応答のうち、特に経皮質ループによって生成される長潜時成分については、課題や環境の変化に依存して応答の調整が見られることが種々の先行研究で示されています。脳はこのような反射系の調整を通して、状況に応じた適切なフィードバックゲインを設定していると考えられます。一方で、反射系の調整がどのような身体表現を介して行われているかは、詳細には明らかになっていません。一例として、伸張反射の調整が単一の感覚情報（体性感覚）のみに基づくのか、あるいは複数モダリティ間の感覚統合（視覚+体性感覚）を含む状態推定を介して行われているのかについてはこれまで未解明でした。本研究では、この点を検証するため、運動実行中の自己位置の視覚フィードバックを操作し、伸張反射に影響がみられるかを調べました。

視覚的な身体情報の操作に伴う伸張反射の調整

実験において、手首の屈曲運動による視覚目標への到達課題を行いました。この運動の途中で、手首周りに摂動トルクを作用させることで伸張反射を誘発し、その大きさを計測・評価しました。ここで実験条件として、手の位置を表すカーソルの移動方向に回転変換を加えることで、視覚による身体情報と実際の運動の間に不一致を生じさせました（図1）。ただし、視覚的な目標位置や運動開始位置も同じ回転変換により移動させることで、行われる運動が常に同一であるようにしました。この条件下で伸張反射応答の大きさを比較することで、視覚情報の伸張反射への影響を調べました。結果、視覚フィードバックの回転変換が大きくなるほど伸張反射の長潜時成分が小さくなることがわかりました。このことは、伸張反射応答の調整に視覚情報を含む多感覚統合による身体表象が関与している、という仮説を支持する結果といえます。反射系の調整の基盤となる状態推定も、随意運動と同様、複数の感覚情報を統合することによって推定の精度を高めていると考えられます。

状態推定の不確かさが伸張反射に影響を与える

第一の実験では視覚フィードバックの回転変換に応じて、伸張反射応答が小さくなることを示しました。この現象の説明として、回転変換によって運動中の身体状態推定が不確かになり、そのことが伸張反射応答の減衰につながった可能性が考えられます。この仮説を検証するための実験として、運動中に自分の手の位置を示すカーソルを消去する条件を与えることで身体情報の不確かさを操作する実験を行いました（図2）。カーソルの提示時間が短いほど運動到達位置のばらつきは大きくなり、運動中の身体状態推定がより不確かになっていると考えられます。一方、伸張反射を計測したところ、カーソルの提示時間が短いほど、長潜時の反射応答が小さくなること

わかりました。よって、この実験の結果から、運動中の状態推定の不確かさに応じて伸張反射が調整されていることが示唆されます。身体状態が曖昧な状況では誤った方向に大きな運動修正を行ってしまう危険性があるため、反射応答を小さくすることでリスクを回避する調整を行っていると考えられます。

おわりに

本研究の結果は、伸張反射系の調整が 1. 視覚情報を統合した状態推定を介して行われていること、2. 状態推定の不確かさに応じて行われていること、を示唆しています。反射系の制御を確度の高い状態推定に基づき行うことで、運動中、刻一刻変化する身体状態に応じた適切な応答の調整が可能になり、随意運動を含む上位の制御系と協調して柔軟な運動制御を行うことに寄与していると考えられます。

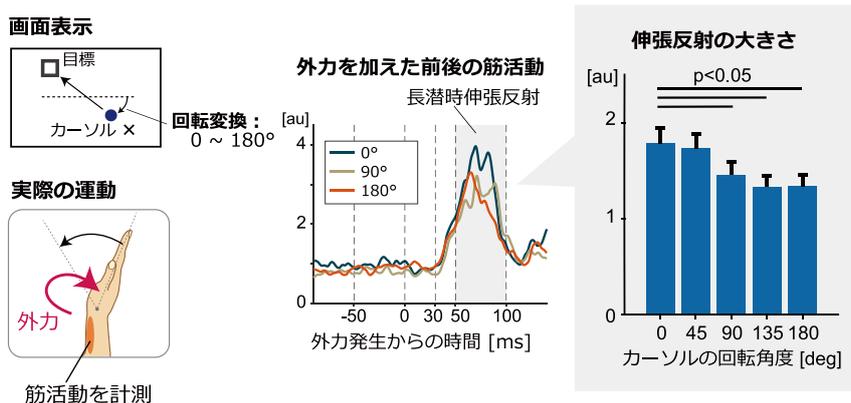


図1. 視覚フィードバックの回転変換が伸張反射に及ぼす影響

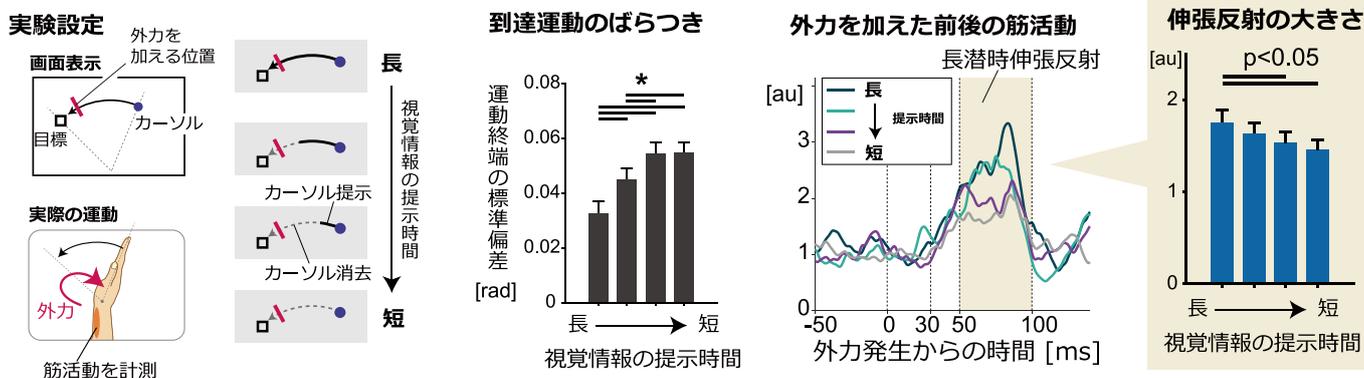


図2. 視覚フィードバックの提示時間が伸張反射に及ぼす影響

社交不安・対人恐怖をひき起こす脳のしくみ

～共感性に着目したアプローチ～

高橋 英彦（東京医科歯科大学大学院 医歯学総合研究科 教授）

論文タイトル: Brain and behavioral alterations in subjects with social anxiety dominated by empathic embarrassment

著者: Tei S, Kauppi J-P, Jankowski K F, Fujino J, Monti R P, Tohka J, Abe N, Murai T, Takahashi H, Hari R
Proceedings of the National Academy of Sciences of the United States of America 117(8), pp. 4385-4391, 2020.
DOI: <https://doi.org/10.1073/pnas.1918081117>

はじめに

人前でスピーチをしたり注目をあびる場面では、多くの方が緊張したり不安を感じます。しかし、これが過度になると、対人交流の場面から著しく身をひくことになり、社会生活に深刻な支障をきたします。事実、仕事や学校に行かず、また家族以外の人とほとんど交流ができなくなっている人の中に、人前でのふるまいで、「相手にどう思われるか」、「恥をかくのでは」という様なことを感じやすい心理的気質が関わるケースが数多く報告されています。しかし、このような社交不安・対人恐怖をもたらす脳のしくみは、まだ十分にわかっていません。対人交流の不安を低減させるには、対人場面で不安を引き起こす脳神経メカニズムを解明することが大きな手掛かりになると考えられます。本研究では、対人交流において重要な役割を担っている「共感」機能に着目し、社交不安・対人恐怖の気質が関わる脳のしくみを調べました。

研究手法

本研究では、まず被験者に、普段どの位の頻度で対人恐怖・社交不安を感じるかを、日常的な具体例が書かれた質問紙により調べました。そして、機能的MRIにより、被験者がカラオケ大会のビデオ映像を視聴している最中の脳活動を測定し（図1）、対人恐怖のレベルと深く関わる脳領域を調べました。ビデオ映像では、歌が苦手な人たちが、うまく歌えていないことを自覚して恥ずかしそうに歌っている場面と、音の外れに気づかず得意げに歌っている場面を提示しました。これにより、このビデオを視聴した被験者たちが、恥ずかしさを共感する際に、(1)「他者の感じる気持ちを自分の感情として感じる」という感情的側面、(2)「相手の立場・状況を推測して文脈的な理解をする」という認知的側面の脳活動を別々に測定しました。

さらに本研究では、上記のビデオ視聴中の脳活動を使用

し、各被験者の神経ネットワークの働きも調べました。ここでは、まず、被験者たちの脳活動を、同期解析（Inter-Subject Correlation: ISC）を用いて解析しました（図2）。これにより、映像刺激によって引き起こされた脳の時系列活動の中で、どの脳領域が同調（共有）しているかを特定しました。また、これらの領域が全て含まれる全脳のマスク（ひな形）を作り、このマスク内での神経ネットワークの強さを計算しました（functional connectivity analysis）。なお、これらの解析をする際に、近年、脳画像解析で指摘されている個人差にも着目し、緻密な解析を行っています。具体的には、個人差を解析で加味する目的で、本研究のために新たに解析手法を開発しました（ISC-based functional segmentation Maskingとfunctional networks analysis by mixed neighborhood selection method）。この手法により、被験者群の脳活動強度を平均化して算定する過程で、外れ値として見落とされがちであった脳の前頭部分への感度が高まり、脳活動をより詳しく評価することが可能となりました。

結果

本研究の主たる結果は、対人恐怖・社交不安を感じる人ほど、(1) 共感に関わる脳の活動がアンバランスであった、(2) 共感に関わる神経ネットワークの働きが低下していた、という2点です。今回の解析により、対人恐怖を強く感じる人には、共感性を支える脳活動に特徴がある事がわかりました。

具体的には、日常生活において対人恐怖を強く感じる人ほど、共感の中でも、「他者の感じる気持ちを自分の感情として感じる」感情的側面に関わる脳活動（扁桃体）が高く、反対に、「相手の立場・状況を推測して文脈的な理解をする」認知的側面に関わる脳活動（側頭頭頂接合部）は低いという、対人恐怖と共感機能のアンバランスな関係が明らかになりました（図3）。さらに、対人恐怖を強く感じる人ほど、共感に関わる

脳全体の神経ネットワークのつながりが低下していることがわかり、情報処理・認知機能の変調との関連が示唆されました。

おわりに

今回、対人恐怖・社交不安に関わる脳のメカニズムを明らかにしたことにより、社交不安症の病態理解を深められる事が期待されます。また、不安を低減する認知行動療法、社交やコミュニケーションに困難さを抱える発達障害などの精神疾患

に対する支援法開発にも貢献できる可能性があります。なお、今回観測された共感に関わる脳活動の変調と社交不安の因果関係については未解明で、共感性の変調が社交不安に結びつくのか、あるいは社交不安の増強で共感性に変調をきたすのかは、まだわかっておりません。今後の研究では、こうした問題にもアプローチして、実際の介入に役立つ知見を得たいと考えています。

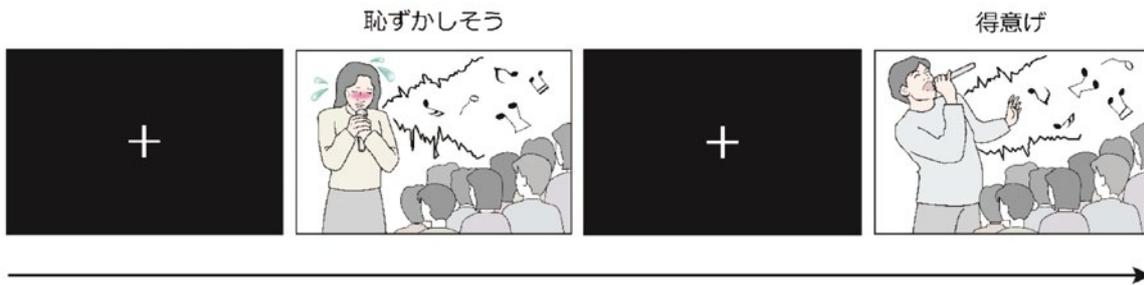


図1 機能的MRIの撮像中に被験者に視聴してもらった、歌が苦手な人たちがカラオケ大会で歌うビデオ映像。

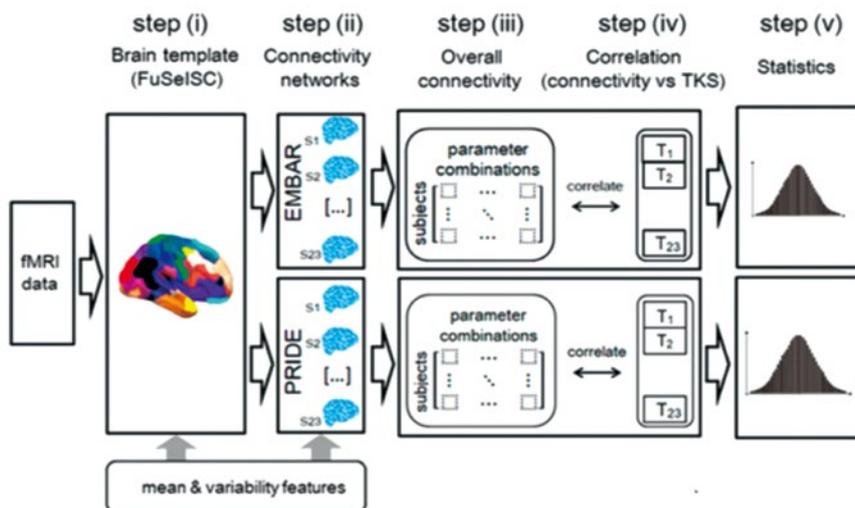


図2 同期解析 (Inter-Subject Correlation: ISC) の流れ。

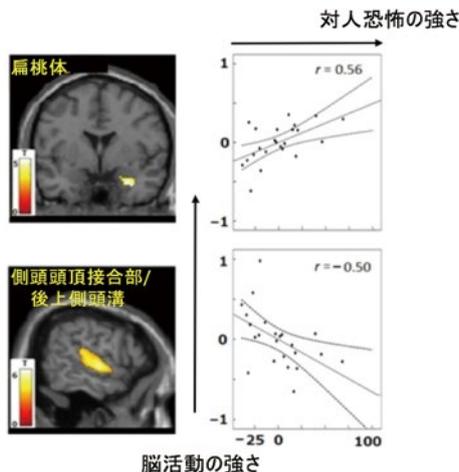


図3 対人恐怖の強さは扁桃体の脳活動と正の相関、側頭頭頂接合部の脳活動と負の相関を示した。

脳の結合性によって明らかにされた統合失調症と自閉症の重なりつつも非対称な関係性

高橋 英彦（東京医科歯科大学大学院 医歯学総合研究科 教授）

論文タイトル: Overlapping but Asymmetrical Relationships Between Schizophrenia and Autism Revealed by Brain Connectivity

著者: Yoshihara Y, Lisi G, Yahata N, Fujino J, Matsumoto Y, Miyata J, Sugihara GI, Urayama SI, Kubota M, Yamashita M, Hashimoto R, Ichikawa N, Cahn W, van Haren NEM, Mori S, Okamoto Y, Kasai K, Kato N, Imamizu H, Kahn RS, Sawa A, Kawato M, Murai T, Morimoto J, Takahashi H.

Schizophr Bull. 2020 Apr 17:sbaa021.

DOI: <https://doi.org/10.1093/schbul/sbaa021>

はじめに

幻覚や妄想、思考障害などの精神症状を主とする統合失調症スペクトラム障害（Schizophrenia Spectrum Disorder: SSD）、社会性やコミュニケーションの障害などの精神症状を主とする自閉症スペクトラム障害（Autism Spectrum Disorder: ASD）という2つの精神疾患の関係性は、この半世紀の間、議論されてきました。ASDの幼少期からの発達歴とSSDの思春期以降の発症時期は異なりますが、最近の生物学的な研究では、2つの疾患の重複性や共通点が示されています。脳MRI画像の研究では、共通した脳部位の灰白質の体積異常、脳の活動性の異常が指摘されています。このように未だにSSDとASDの関係性については十分に解明されたとは言えない状況です。

精神疾患の関係性が明確でない根本的な理由は、精神疾患の信頼できる生物学的な指標がないこと、また、米国の精神疾患の診断基準のように、精神疾患の診断に用いられる指標の多くが患者の症状と行動から規定されることにあります。そのために、生物学的なエビデンスと診断との間にギャップが存在します。SSDとASDの関係性を明らかにするためには、生物学的な手法と従来の疾患分類的な診断手法の両方が必要となってきます。そこで、人工知能技術を用いて、SSDの信頼性の高い判別法の開発を目指しました。さらに、以前に開発したASDの信頼性の高い判別法も利用することにより、一個人でSSD、ASDの診断の確実性であるSSD度、ASD度を算出することを目指した。このことにより、SSD度とASD度の二次元の座標として、個々の診断の確実性を表示し、同時に集団としてのSSD群、ASD群の関係性を定量的に評価することが可能となる。

研究手法

本邦で募集された成人の研究参加者:日本人データ 170人 (SSD患者68人・健常者102人)の安静時のMRI上の脳機能的結合(9,730個)から、(株)国際電気通信基礎技術研究所(ATR)・脳情報通信総合研究所で開発された革新的な人口人工知能技術を適用することで、SSDの診断に関わる16個の機能的結合を特定しました。16個の機能的結合の重み付け和でSSDの診断の確実性(SSD度)を定量化し、SSDもしくは健常者いずれかの判別法としました。本判別法が独立データにも汎化可能か確認するため、海外3施設のデータを入手し、適用しました(図1)

結果

判別率の精度を示すROC曲線のAUCは、SSD判別法を慢性期のSSDを含む京都データに適用した場合、0.83と高い判別率を示しました。開発されたSSD判別法は、海外の米国の慢性期のSSDデータに対してはAUCが0.75、オランダの慢性期のSSDデータに対しては、AUCが0.66の判別の精度が得られました。一方、米国の初発エピソードのSSD患者データに対してはAUCが0.42の低い判別率を示しました(図2)。

SSD判別法により定量化された個々のSSD度を横軸として、また、以前に研究グループが開発したASD判別法により定量化されたASD度を縦軸として、二次元の座標上に、京都データのSSDと健常者、本邦のASDと健常者を示しました(図3)。SSD-ASDの座標面では、SSD群の中心※8がASD軸上で健常者群の中心よりもASD度が高くなりSSD群とASD群が重なり合う要因となっています。一方でASD群の中心はSSD軸上で健常者群の中心と差はありませんでし

た。また、ASD 群の中で、SSD 度が高いほど、ASD 度が高くなる傾向が有意に認められ、一方で SSD 群の中では SSD 度と ASD 度の関係性は認められませんでした。このように SSD 群と ASD 群は、重なり合う関係性と、非対称な関係性の両方の特性があることを、SSD-ASD 座標上で確認された。

おわりに

本研究では、生物学的な安静時の脳機能結合と、従来の症状や行動から規定される精神疾患の診断を基に、高度な機械学習アルゴリズムを用いて SSD の判別法を新たに開発しました。SSD 判別法は、本邦のデータだけでなく、MRI 機種や国籍を超えて海外データにも汎化されるものです。さらに、こ

の SSD 判別法と研究グループが以前に開発した ASD 判別法を応用し、SSD と ASD の診断の確実性を定量的に視覚化することで、2 つの疾患の関係性を世界に先駆け明らかにしました。つまり、両者には重複する部分も多いが、SSD には ASD 傾向があるのに対して、ASD には SSD の傾向がないことがわかりました。本研究で、2 つの精神疾患の関係性を定量的に視覚化することが可能となり、今後は、他の精神疾患についてもそれぞれの判別法を開発し、また、生物学的な診断の確実性を計量することで、様々な精神疾患の関係性を広く検討することができます。今後、精神疾患の個別化医療への応用が期待されます。

SSDの判別法の開発

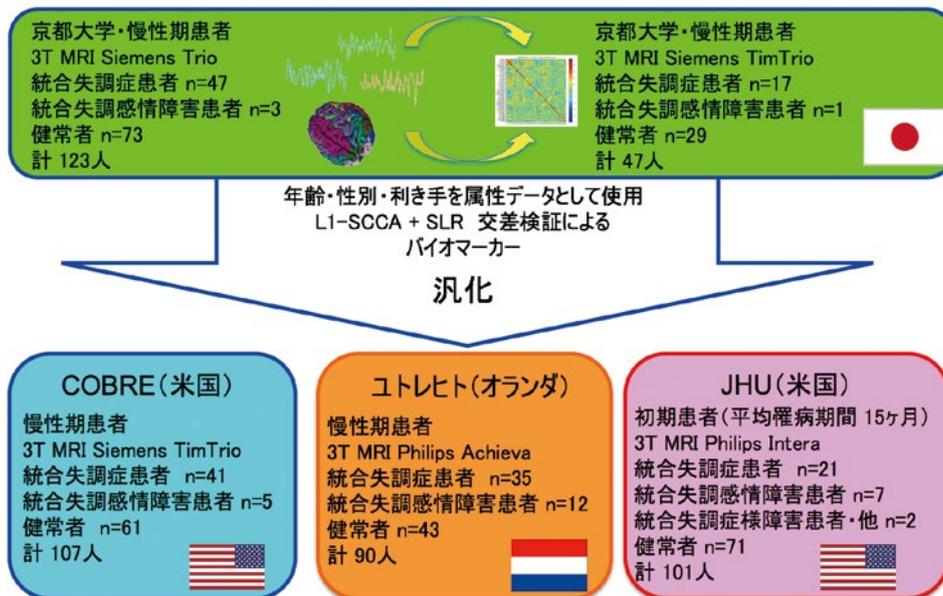


図1 SSDの判別法の開発。本邦のデータを基にして、判別法を開発し、海外データへの汎化性を検証した。

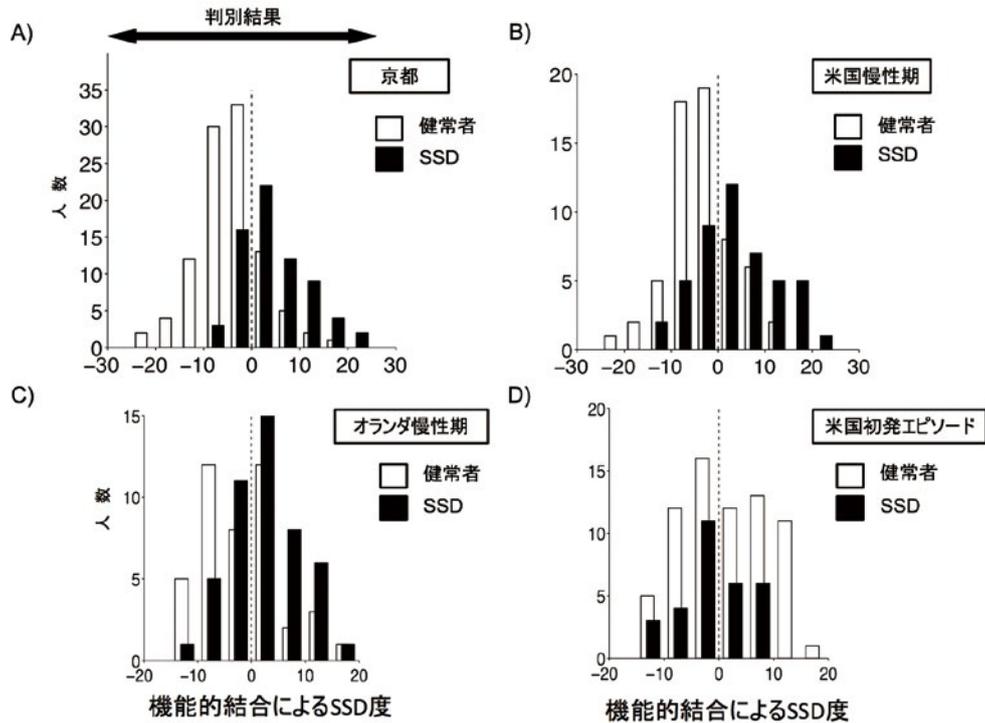


図2 本研究で開発されたSSDの判別法による各データの結果。A) 京都データ、B) 米国の慢性期患者のデータ、C) オランダの慢性期患者のデータ、D) 米国の初発SSD患者のデータ

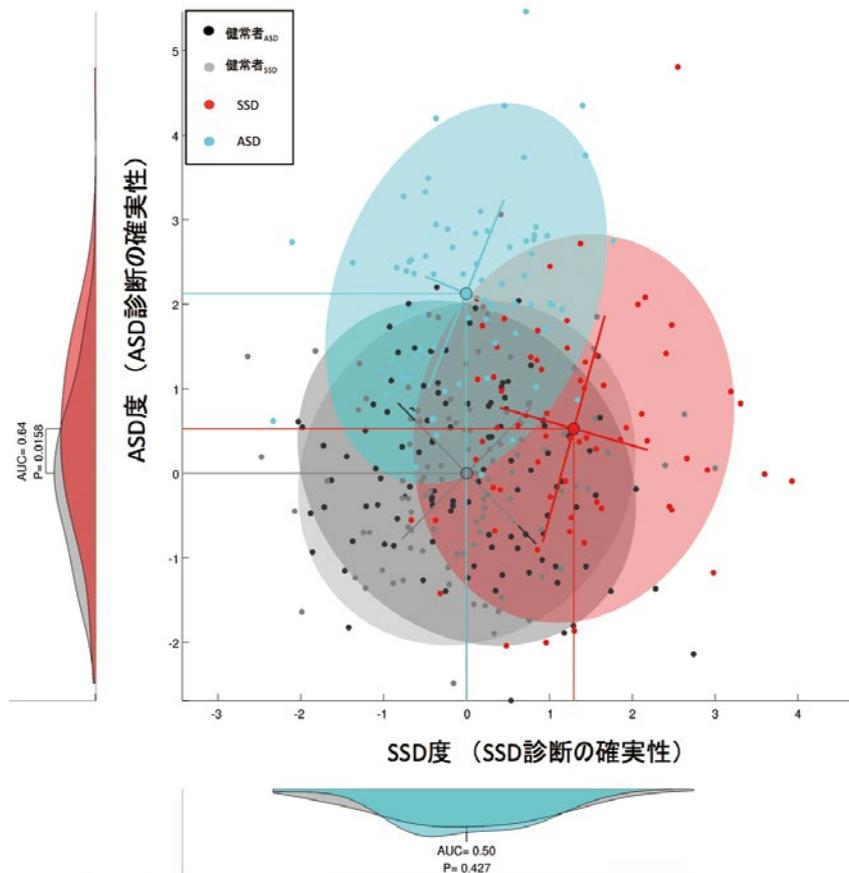


図3 SSD度 (SSD 診断の確実性、縦軸) とASD度 (ASD 診断の確実性、横軸) によるSSD (赤色点)、ASD (青色点)、健常者 (黒もしくは灰色) の定量化と視覚化。楕円は、それぞれ、赤色はSSD群、青色はASD群、灰色は健常者群。図中のAUC値、P値について、SSD度はASD群と健常者群との比較、ASD度はASD群と健常者群との比較を示す。

心を知るための人工知能

認知科学としての記号創発ロボティクス

谷口 忠大 (立命館大学 情報理工学部 教授)

書籍タイトル：心を知るための人工知能 認知科学としての記号創発ロボティクス

著者名：谷口 忠大

出版社名：共立出版

ISBN：978-4320094659

人工知能の研究は一つには有用なツールを作ることにあります。もう一つはやはり「心を知る」ことにあります。構成論的アプローチという言葉がありますが人工知能と脳科学の研究の対照と融合とはまさに構成論的アプローチ——作ることにによって理解するという科学的取り組みなしには語れない新学術領域だと言えるでしょう。本書では記号創発ロボティクスと呼ばれる筆者らが取り組む人工知能研究のアプローチを人の心を理解するための科学——認知科学の文脈からその位置付けを描きました。大きな背景は以上のようなものですが、野心的な記述も含めて2020年代に向けての視点からいくつかの重要な記述をおこなっています。

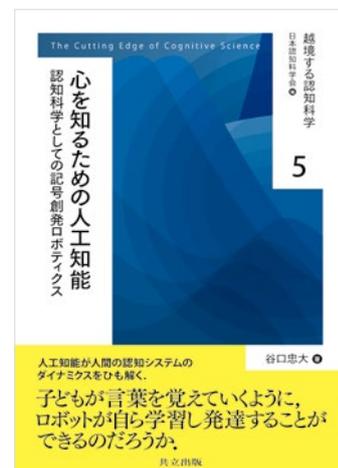
一つにはロボットを使った構成論的研究の位置付けです。ディープラーニングの潮流は私達の人工知能技術を進化させて実世界情報（パターン）の知的処理を大いに発展させました。人間の知能が実世界適応の中で進化的および発達的に形成されてきたことを踏まえれば実世界で活動できる知能に知能の本質を見出そうとする研究はさらに進展していくでしょう。その先にロボットという身体を用いた構成論的研究があります。これを理解するためには「モデルとは何か？」をあらためて問う必要があります。ロボットを用いた構成論的研究に取り組む研究者、また、ロボットを用いた研究を「人間の知能の研究なのだ」と言い切る工学者の主張に疑問を抱く研究者に読んでいただきたいと思います。

二つ目は「記号」概念の再整理と「記号接地問題」に対する反省です。私達の人工知能研究は歴史的な背景から、記号論理学との接点を強く持ってきました。その上で「記号」という言葉に関する混乱を孕んできたのです。ディープラーニングにより刺激応答的なパターン処理が可能になり高次認知へと至ろうとする昨今であるからこそ、そこをあらためて整理し、記号接地問題とは実際には何だったのかに関する筆者の

理解を示します。その上でそれを乗り越える記号描像として記号創発システムと、新たな問題「記号創発問題」を示します。

三つ目は認知のモデルとしての確率的生成モデルと記号創発ロボティクスの研究成果です。確率的生成モデルはこれまで認知科学におけるカテゴリ化や概念形成のモデルとして扱われてきた数理モデルと人工知能やロボティクスにおけるモデルに架け橋を与えます。これまでその関係に関してボヤッとした認識のみを持っておられた方々に読んで、考える機会にさせていただければと思います。

さて、そもそも認知学会の企画として依頼を受けて書いた本書ですが、2020年以降の研究を行う上で、時機を得た一冊になったと思います。執筆の経緯などは「あとがき」に書いております（ちょっと遊びすぎだと言われましたが）。人工知能と脳科学の対照と融合を議論する上でも参考になるかと思えます。もしよろしければ一度手に取って頂いて覗いていただければ幸いです。



Report for ICLR 2020, The International Conference on Learning Representation

Dongqi Han, Cognitive Neurorobotics Research Unit, Okinawa Institute of Science and Technology



The International Conference on Learning Representation (ICLR, pronounced as “ai-clear”) is a relatively young international conference on artificial intelligence and machine learning, but has rapidly developed into one of the most impactful symposiums in its field. This year, 2020, it was planned to be held on the continent of Africa, in the city of Addis Ababa, Ethiopia, which made the participants excited about the unexplored place. As one of the attendees, I was also looking forward to presenting my work to the researcher all over the world during this interesting travel.

Unfortunately, due to the pandemic of the notorious COVID-19, we have to accept the sad fact that we could not meet physically in the southern hemisphere. However, though the exciting trip to Africa was cancelled, the organizers of ICLR2020 conference decided to change its form to a virtual online conference, which is also another novel experience for me. Obviously, it is not easy to organize an online meeting with hundreds of presentations and thousands of participants. Therefore, hybrid forms of interactions between the authors and audiences were used to ensure success of the event. First, the authors were asked to pre-record videos of both their

presentation and the corresponding slides, which were 15-minutes long for featured orals and 5 minutes for other accepted papers. Also, the authors needed to be present in two 2-hour poster sessions which allow other participants to do live interactions with the authors using ZOOM. Furthermore, asynchronous chat was also available in a virtual channel during the conference.

Our paper, titled “Variational recurrent models for solving partially observable control tasks” was accepted as a poster presentation. However, this time, it was not presented in the form of a poster, but as a 5-minute video of slide presentation (as the screenshot shows). In the live-interaction sessions, people came and asked questions to me, just like in a physical poster presentation. Although the feeling was a bit weird, I had several joyful chats with researchers from different places. Asynchronous chatting rooms were also used for those who were not in a convenient time zone for the live-interaction sessions, which is a kind of advantage compared to physical meeting.

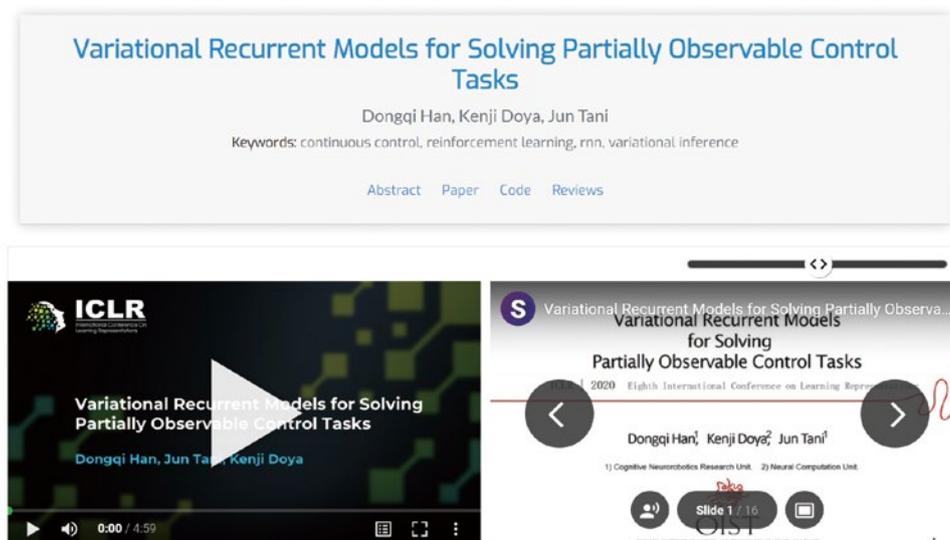
ICLR has been known as a top-tier conference for reinforcement learning and unsupervised learning. Besides the presentation by myself, there were also

numerous other studies appeared interesting. A work related to my research scope is so-called “Dynamics-Aware Unsupervised Discovery of Skills”, which proposes an algorithm to learn motor skills represented in a latent space without any supervision or reward signal. The algorithm is based on mutual-information maximization and it can employ either a discrete or continuous latent skill space. The learned skills can be further used for motor planning or hierarchical control to solve difficult robotic tasks.

There are also a number of interesting workshops along with the conference. For example, “Bridging AI and Cognitive Science” and “ML-IRL: Machine Learning in Real Life”. Because of the difference

of time zones of participants, it was hard to follow the full live workshop. However, one can watch the videos of invited talks and see the discussions in the chatting room at any time. This allows a participant to attend all the workshops, if he/she wishes.

As a short summary, the virtually held conference was a brand-new experience for me. Although remote conference should not be the best option for most people, it provides a possibility of attending the presentations and discussions for those people who cannot go to the conference physically due to e.g., health conditions. It is likely that in the future there will be academic conferences held in a hybrid form of on-site and on-line meeting.



UCL-NTT Collaboration Meeting 報告

高椋 慎也、五味 裕章（NTTコミュニケーション科学基礎研究所）

NTTコミュニケーション科学基礎研究所ではUniversity College Londonの認知脳科学研究所(ICN)との間で、2016年よりDeep Brain Communication Science Project IIというタイトルで約4年間の共同研究プロジェクトを立ち上げ、感覚運動情報処理、発話情報処理、聴覚情報処理、対話コミュニケーションなどの研究を進めてきました。そのプロジェクトの一環として、2018年10月にOISTで当新学術領域と連携して開催した会議（JNNS大会のサテライトとして実施）に続き、2019年12月にUCLにおいてプロジェクト会議が開催されました。

初日午前のセッションでは、UCLからは、Institute of NeurologyのJames Kilner博士とSven Bestmann博士がそれぞれ、心拍リズムと運動タイミングの関係、運動計画の時間配分に関するご研究を、ICNのMariana Rebelo博士（Patrick Haggard博士の共同研究者）が運動行為と空間位置記憶の関係について発表されました。NTTからは当領域にも参加する安部川と伊藤がそれぞれ、眼球運動のアンカー効果、伸張反射の多感覚性について発表。運動研究者が集まったこともあり、運動の評価指標など、細かい側面に至る密な議論が行われました。

午後1つめのセッションでは基礎研究の社会応用の試みについての発表がありました。UCLからは障がい者向け技術の革新に取り組むGlobal Disability Innovation Hub (GDI Hub)のCatherine Holloway博士(GDI Hubの創設者の一人)とYoungjun Cho博士がそれぞれ、発展途上で利用する障がい者向け用具の開発や低コストな感情認識技術についての研究を紹介、NTTからは新学術領域メンバーでもある五味が、触錯覚を使ったモバイル力覚提示デバイスについての研究を、デモを交えて発表しました。

初日最後のセッションでは、UCLの外部連携に関わられているJoseph Devlin博士(Faculty of Brain Science)とBjorn Christansson博士(UCL Consultants)を招き、”Tech transfer, translation, and development pathways.”というタイトルで企業とアカデミックの共同研究について様々な議論が行われました。

二日目午前のセッション、UCL-ICNからはPatrick Haggard博士、新学術若手サポートを受けたIrena Arslanova研究員、Sophie Scott博士がそれぞれ、触運

動情報の統合知覚過程、ヒト脳における音声情報処理について、Ear InstituteのMaria Chait博士が聴覚におけるサリエンシーについて、University of LondonのAntonio Cataldo博士が眼球運動とラバーハンド錯視の関係について発表されました。触覚研究で著名なVincent Hayward博士も聴講参加され、感覚情報処理の時間スケールの多様性についてなど、興味深い議論がありました。

二日目午後のセッション、UCLからはInstitute of Cognitive NeuroscienceのBahador Baharami博士(LMUと兼任)とAntonia Hamilton博士が各々、アドバイザーが複数名いる場合の助言戦略や、自然状況におけるヒト間インタラクションの研究について発表され、NTTからは熊野と高椋がそれぞれ、他者のために賭け事をするときのヒト間インタラクションや、自閉症成人の腕運動の特異性に関する研究を発表しました。

基礎的な神経科学の研究から社会応用まで、多岐に渡るテーマを扱った今回の会議でしたが、夜の意見交換の場でHayward博士から面白いお話を聞きました。PCのインタフェースがCUIからGUIに切り替わったとき、多くの視覚障がいの方が仕事を失ったそうです。そのことを聞いて不憫に思われた博士は、触覚フィードバックでウィンドウの形状などを提示するシステムを開発されました。この装置は画面に水平な方向にのみ力を提示する装置でしたが、不思議なことにユーザーの中には凹凸の感覚を報告する方がいたそうです。博士が2001年にNatureに発表されたlateral forceによる触覚提示技術（現在のモバイルデバイスの主要な触覚提示手段）の着想は、このときの報告に基づくものだそうです。コラボレーションにはWin-winな関係が重要としばしば指摘されますが、イノベティブな発見にはもともとずっと意外な関係性が必要なのかも知れません。



会議で触錯覚デモを披露する様子

第8回領域会議報告

鮫島 和行 (玉川大学脳科学研究所)

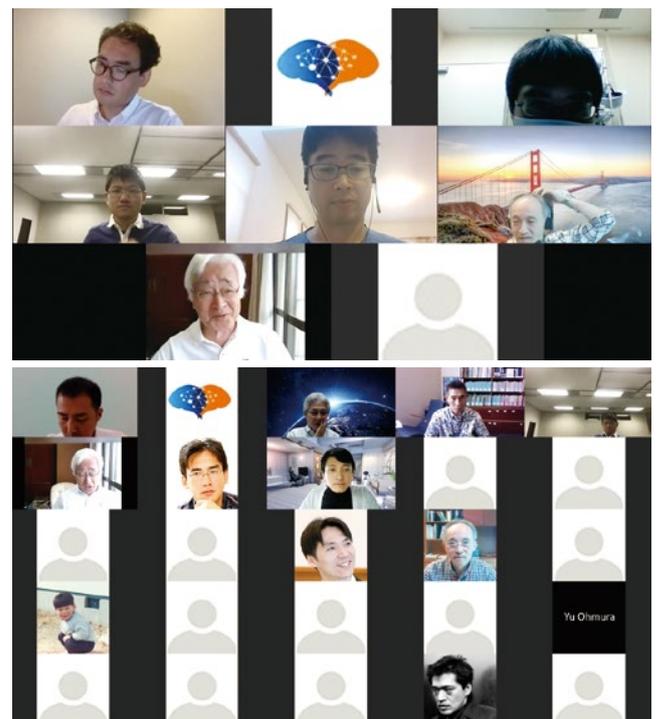
2020年6月15、17、19日の三日間にわたって、新学術領域「人工知能と脳科学」の第8回領域会議がオンライン上で行われました。まず、銅谷領域代表からのご挨拶からはじまり、本年度が最終年度となる旨、中間報告での指摘点、および最終年度での成果とりまとめスケジュールや、国際シンポジウムについてのお話がありました。私にとっては、後半から公募研究として第6回の領域会議に参加させていただいてからあっという間に1年が過ぎ、もうとりまとめの時期なのかと驚いております。

初日(15日)には、8件の公募研究班の発表がありました。飯野先生の実験系に強化学習を導入し、対象動物に適切な刺激を与え、系全体での解析をしようという試みは、刺激になりました。井上先生からは最新の遺伝学的方法やウイルスベクターを使った多シナプス伝搬回路同定を大脳皮質—大脳基底核ループについて解析しており、線条体の情報表現と合わせて考えれば、前頭極や内側前頭前野と皮質下との情報処理に新しい仮説が立てられるのではないかと、思いを巡らせることができました。これらは私の興味からの例にすぎませんが、多くの研究の関連性に気付くことができるのも、領域会議のメリットの1つです。その意味で、発表が一通り終わると、オフラインにいきなりなった瞬間に自宅に引き戻される(自宅からリモート参加でした)のは、最初は戸惑いました。いつもの領域会議の様に、コーヒープレイクや食事をともにしながら、「今日話していた、あれは、これと関係あるのかなあ」や、「質問しそびれたのだけれども、これはこういう解析?」などなどの四方山話が領域会議の楽しみの1つになっていたことを実感する瞬間でもありました。

1日開けて17日、19日に合わせて12件の発表がありました。池上先生のオルタとよばれるロボットと人のジェスチャーを通じたインタラクション実験とその解析や、心理評定の話は、私が以前参加していた非言語情報によるコミュニケーションの研究領域に近く、今のAIでは限定的である自律性とは何か?のヒントとなる非常に興味深いものでした。他にも大村先生のセロトニンニューロン操作と強化子価値低減パラダイムの関係性や、小松先生のマーモセットの全脳ECogにおける側頭葉と前頭葉の大域ネットワークのダイナミクスなど、最新の実験的知見だけでなく、理論の側からも、深井先生の樹状突起での計算を取り込んだリカレントネットの回路学習、細谷先生のモジュールVAE(変分自己符号化器)による小サンプル学習とIT皮質の対比など、理論と実験の両面での進展を一度に聞けるのもこの領域の魅力の1つです。

3日目は、残り2件の発表の後、ブレイクアウトセッションとして、4つに分かれて、この領域での問題となっている事柄についての議論が行われました。私は1つめの表現学習グループに、田中啓治先生、深井先生、庄野先生、細谷先生とともに参加させていただきました。この領域に求められている現在のAIにないものを如何にして脳科学から提言できるのか?という少し大きな話題が中心となりました。工学的な問題定式化、すなわち自然画像からのパターン認識などの明確な目標が必要なのではないか、という提言や、現在の「AI」の「強い」ところではなく「攻めあぐねている」問題とは何かを考えることが重要なのではないか、という話もありました。私としては今のDeep NNは外界の情報を取り込んだ、うまい入出力変換であって、強化学習における探索の問題などの、内在する情報を生み出すメカニズムが必要とされるのではないかという話をしました。それは、なぜ脳に自発発火があるのか、そのメカニズムにヒントがあるのではないか、という議論に発展し、どこかの細胞やどこかの局所回路が作りだしているというよりも、領野間をまたぐグローバルなネットワークが複雑な環境に適応する時に大事なのではないか、という問題提起が行われました。

この様に、リモート参加とはいえ充実した情報交換と議論が行われ、大変刺激的な3日間でした。今後ともこのコミュニティを通じた議論で新しい研究が展開できることを嬉しく思うと同時にあと1年なのかという思いも残りつつ、リモート画面がプチッと切れて一人考えこみました…。



イベント情報

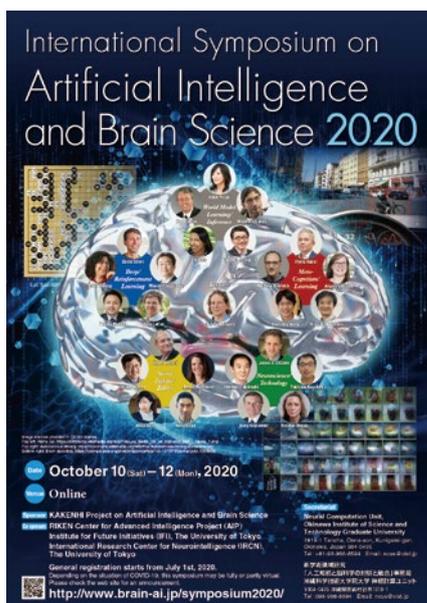
主催イベント

International Symposium on Artificial Intelligence and Brain Science 2020

日程：2020.10.10-12

場所：オンライン開催

<http://www.brain-ai.jp/symposium2020/>



第9回領域会議

日程：2020.10.13

場所：オンライン開催

脳と心のメカニズム第21回冬のワークショップ

日程：2021.1.6-8

場所：北海道、ルスツリゾートホテル

<http://brainmind.umin.jp/wt21.html>

共催、協賛、後援、関連イベント

第28回脳の世紀シンポジウム「依存症と脳」

日程：2020.9.16

WEB ライブ配信 10:10-16:00

特別講演：おおたわ史絵

「依存症家族の経験から語る」

「脳を創る」講演：高橋英彦

「脳イメージングでみる依存症ーギャンブル依存を中心にー」

[https://www.braincentury.org/index.php?](https://www.braincentury.org/index.php?page=brainsympo_index)

[page=brainsympo_index](https://www.braincentury.org/index.php?page=brainsympo_index)

第8回神経法学研究会

日程：2020.11.3 10:00-13:00

場所：オンライン開催

講演：銅谷賢治



発行／編集 新学術領域研究「人工知能と脳科学の対照と融合」
お問い合わせ 新学術領域研究「人工知能と脳科学の対照と融合」事務局
Mail ncus@oist.jp
2020年10月発行

www.brain-ai.jp