

連載 「一人称研究」〔第7回〕

# 一人称研究対談： 「孤高の人の思考を科学するとは？」上篇

## How Is Science to be in the Study of One and Only Expert? (1)

下條 信輔  
Shinsuke Shimajo

カリフォルニア工科大学生物・生物工学専攻  
Division of Biology and Biological Engineering, California Institute of Technology.  
sshimajo@caltech.edu, <http://neuro.caltech.edu>

松原 仁  
Hitoshi Matsubara

公立ほこだて未来大学システム情報科学科  
Future University of Hakodate, School of Systems Information Science.  
matsubar@fun.ac.jp

伊藤 毅志  
Takeshi Ito

電気通信大学情報理工学研究科／人工知能先端研究センター  
Graduate School of Informatics and Engineering, The University of Electro-Communications. /  
Artificial Intelligence eXploration Research Center.  
ito@cs.uec.ac.jp, <http://minerva.cs.uec.ac.jp/~itolab-web/wiki.cgi>

諏訪 正樹  
Masaki Suwa

慶應義塾大学環境情報学部  
Faculty of Environment and Information Studies, Keio University.  
suwa@sfc.keio.ac.jp, <http://metacog.jp/>

**Keywords:** first-person's view, Yoshiharu Habu, Shogi, second-person approach, creativity.

「一人称研究」という考え方は、本誌2013年9月号 (Vol. 28, No. 5) の特集「一人称研究の勧め」で初めて提唱された [諏訪 13]。諏訪正樹、堀 浩一が編集し、両名を含めた9名の研究者 (伊藤毅志、松原仁、阿部明典、大武美保子、松尾 豊、藤井晴行、中島秀之) が論文を執筆した。さらに、この特集論文をベースにして、より一般の読者に向けて加筆修正を施した内容の書『一人称研究のすすめ 知能研究の新しい潮流』 (近代科学社) が、2015年4月に出された [諏訪 15]。客観性や普遍性を前提とする従来科学の方法論だけでは、「生きている生身のひと」の知を十全に扱うことができないのではないかという問題意識を共有し、上記の研究者達は、知能にまつわる研究や学問の新しい方法論を模索し始めたのである。

しかし、この新しい研究観はまだ広く受け入れられているとはいえない。執筆した著者達も、一人称研究とは何か、どうあるべきかについて多種多様な理念や方法論をもっている。そこで、上記の著者達が、人工知能・認知科学・心

理学・社会学・哲学などを専門とする研究者を対談相手に選び、一人称研究の考え方について議論する対談を、学会誌連載として寄稿することになった。

対談相手は、必ずしも一人称研究の考え方に賛同しているわけではない。その方の専門分野から見て一人称研究はどう見えるのかを率直に語り、知能の探究において本当に必須な方法論なのかと議論を投げかけてもらうことを依頼した。この対談を通じて、一人称研究とはなんぞや、どうあるべきかについて、問いを深めたい。

上記構想の第4弾として、認知神経科学者であり、知覚心理学者でもあるカリフォルニア工科大学で教鞭をとっておられる下條信輔氏をお招きし、2017年10月9日に、御茶ノ水のデジタルハリウッド大学の会議室において、松原、伊藤、諏訪との四者による対談を行った。下條氏の幅広い知見から、一人称研究に対する従来の三人称研究に対する考え方から、コンピュータ将棋と人間の対戦である電王戦の話題、さらには、脳の潜在的な思考が及ぼす影響に

関する話題など、多岐にわたる重要な視点を提供してくださった。下條氏は2時間ほどほぼノンストップで語り続け、一人称視点研究にとって示唆に富む対談となった。本稿は、その前半部分である。

### 二人称アプローチ

松原：いきなりですけど、下條さんにお聞きしたかったのは、囲碁や将棋のトップレベルの人を研究対象にするという研究は、そもそも例えば下條さんはどうお考えになっていらっしゃるのか、ということ。例えば、ご経験があるとか、そういう点に我々としては興味があって、下條さんにお声掛けしたのです。諏訪：もともと、あの本 (『一人称研究のすすめ 知能研究の新しい潮流』) は私と東京大学の堀 浩一先生が編集をして、共著者の皆さんと一緒に「一人称研究」という新しい概念を世に問うた。僕は松原さんと伊藤さんがやられた羽生さんの研究 [伊藤 13, 松原 15] は二人称研究だと思っていて、一人称研究っていうのは二人称も含むと思うのです。羽

生さん自身が将棋をやるときのことを何か語るとすれば、それは一人称だけど、それを第三者ではない、その将棋のことがよくわかる人間がインタビューするというのが二人称というわけです。

伊藤：1.5 人称ぐらいの感じです。

諏訪：従来の科学としての三人称というのと違って、知能研究においてはこれから三人称ではない研究が必要だろうと思います。今日の大きなテーマとして、一人称研究、二人称研究をどう思うかという論点と、その中の一つの例として、将棋の羽生さんみたいなトップレベルの人のことを調べた研究をどのように捉えるのかという点、この二本立ての話なのかなと想定してします。

下條：二人称アプローチというのが流行っていて、発達心理学の分野で。

諏訪：はい。

下條：ヴァスデヴィ\*1[レディ 17]という人がいてね、発達心理学者です。インド系アメリカ人だと思います。二人称アプローチの本を書いて佐伯 胖先生\*2が高く買っていて、翻訳されています[レディ 15]。

諏訪：存じ上げています。

下條：それで、呼ばれて話をしたことがあるのですよ。佐伯先生もその前に、多分諏訪さんが引用されていたと思うのだけれど、「誰でもいいというわけじゃなくて、誰か詳しく話せる親しい他者がいて、それで二人称の価値みたいなものがある」と言われていて。

ただ、そこには二つの問題があって、一つは  $N=1$  のデータをどう処理するかという問題で、もう一つは一人称のアプローチが何にコントリビュートできるかという問題の二つです。

$N=1$  だけど徹底的に三人称ってことはあり得るわけです。実は心理学では昔からそれは扱われていて、簡単に言うと症例研究がそれです。

非常にまれな症例の場合には、一方ではまれなほど価値が高い一面があるけれども、それは統計学的に良いのかということですが、健常者の知覚の研究だっ

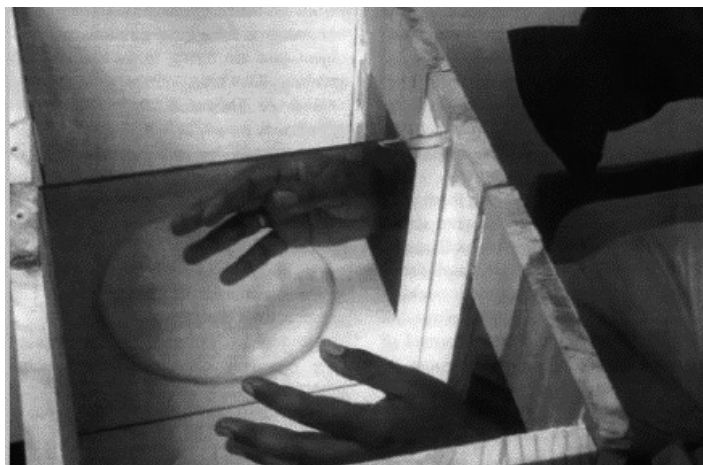


図1 ラマチャンドラン博士が考案した鏡箱（切断によって手を失った人が、現実にはないはずの腕や足に痛みを感じる（幻肢痛と呼ばれる）現象に対して、この鏡箱に手を入れ、健常な手や腕を鏡に写すことで、そこ手や腕があるように錯覚させ、視覚的な感覚フィードバックを利用して幻肢痛を和らげることに成功したという事例）

たら 30 人集めるところを、まれな症例だと一人しかいないとなるわけです。それはその非常にまれな一例をテストしたら、こうなったというので、それは価値があるのか、という問題があるんです。

だからまあ、 $N=1$  の症例研究と一人称研究、その二つを分けて考える必要があるかなっていうこと。

その意味で言うと、一つのエピソードだけドラマチャンドラン\*3 っているでしょ。彼がファントムリム（幻肢痛）の話 [ラマチャンドラン 99] を、15 年前ぐらいにアメリカ神経学会でしたことがあって、この辺に腕を切断された患者には幻肢とあって、存在しない腕を感じる幻覚があるんだけど、体性感覚野で腕の表現は顔の表現と隣接している。そのため、顔の頬とか鼻のわきとかを刺激すると、腕のあたりにかゆみを感じる、という幻覚を、ある症例で報告しています。

すると学会場である人が手をあげて、多分電気生理学者だと思んですけど、その  $N=1$  のデータで、お前は物を言ったつもりなのかと聞いたわけです、専門家ばかりが何千人もいる前で。

そこでラマチャンドランがすかさず何

と答えたかと言うと、「お前の目の前に、今仮に人間の言葉をしゃべる豚が現れたとしよう、お前は一例だからといって、それをなかったことにするのか」って言ったんですよ。すかさず。

だから、それはラマチャンドランがずっとそういう批判を浴びていて、答えを考えていたんだと思うんですけど。それでみんなは「おおー」と言ったという伝説があるのですが……。それは、多分本当にそう言ったらいいです。

ラマチャンドランが、確か Vision 関係の雑誌のエディターに書いた手紙というのを僕は見たことがあって、そこには、「統計に乗るような知見はたいしたことはない」、「お前は私の論文の図 1 を見たのか」、「新しいイリュージョンをお前も見ただろう。それは統計とどっちが偉いの、5%水準とどっちが偉いの？」という手紙が残っていて、皆おもしろがって回覧していた。

そういうことがあるので、でもそれは、 $N=1$  をどう扱うのかという話で、ただラマチャンドランの場合には、彼の強みは知覚イリュージョンが彼の武器だから、エディターが見たら一人称で見えちゃう、つまり一人称研究になってしまうわけで、実際は三人称の研究ではあるのだけれど。

彼の「ネイチャー」の論文の図はたいして読んだ人が見ると錯視が見えるわけです。新しい知覚効果が、例えば網膜上にないものが見えたりするわけです。

\*3 ヴィラヤヌル・S・ラマチャンドラン：心理学・神経科学者。

山下篤子 訳：脳のなかの幽霊、角川書店〈角川 21 世紀叢書〉(1999)、山下篤子 訳：脳のなかの幽霊、ふたたび—見えてきた心のしくみ、角川書店 (2005)

\*1 ヴァスデヴィ・レディ：二人称アプローチを提唱している発達心理学者。

\*2 佐伯 胖：認知科学者、発達心理学者。

だから、そういう意味では重なっているのですが、その二つの側面があって、あの本の羽生さんのインタビューを拝見して感じたのは、心理学の歴史でずっと問題になってきたことが、歴史的に逆向きの順序で問題になっているのだなと思ったわけなのです。

逆向きという意味は、心理学では昔は一人称の内観から入って、その後、内観じゃ科学としてだめだということになって、客観主義になった。心理学が客観主義になったのは物理学のまねをしたからだといわれているのですけどね。

### 客観主義

諏訪：行動主義ですね。

下條：そうですね。行動主義が入ってきて、認知科学といえども行動主義の影響を相当受けていて、インプットとアウトプットは見えているという話だから、その中身はブラックボックスで、そこは推定できるだけという考え方です。そういうことでやってきたのだけれど、やっぱりそれじゃあ、ちょっと何か足りない部分があるんじゃないのというので、アフォーダンスが出てきたりとか、もうちょっと一人称的なアプローチが必要だとか、エンボディードインテリジェンスといったものが出てきたりしたと。もっと主観的で一人称的なものを考えなきゃいけないという流れに舞い戻る感じで、歴史的にそういう順できているのです。

大きく言えば、その最新の部分を取り除くと、そこは認知科学と重なっている近代史だから、もっただけ言うと「主観から入って客観になった」ということで、それでやっとならハードサイエンスのメンタリティー・方法論になれたということで理解されてきた。

AIで今このことが問題になっていることが、非常におもしろく思えたのです。歴史的に見てね。なぜそれが問題になったのだらうなというところが、まだ半分しかわかっていないけれども。

あの本\*4がおもしろいし、トリッキー

だと思ったのは、若い人に研究のスタイルを勧めて教えている本なのか、それとも、人間の知の本質はこう（＝一人称）なので、人工知能もそれも学ばなきゃいけないと言っている本なのか。どうも両方、ダブルミーニングを狙っているらしいという気はするのだけど、ちょっと若い人を見ると混乱するんじゃないかなという気もしたんですね。

諏訪：うん、後者はまだ望んではいないです。

下條：あ、まだ望んではいない。

諏訪：というか、まだ身体をもたないコンピュータに後者を望んでも、ちょっと無理でしょって感じがしますね。だからまあ、一応前者のつもりなのですかね。

下條：なるほど。

諏訪：でもコグニティブレボリューションという、認知科学のもとの起こりは、行動主義に対する批判から起こったわけで、「外だけ相手にして何が認知か！ 中身を見ましょう」と言っただけじゃないんだっけ？ というのが、我々の主張なんですよ。だから、科学にするために、いつの間にかまた外からだけ見ていると言っているわけです。

下條：もう一つ言うとね、その心理学における客観主義というのが誤解されているわけです。私の学生のときは、物理学的に存在が証明できるもの、例えばカメラで撮ったら見える行動とかね、こういうものが客観的なデータで、それに基づいてモノを言うべきで、内観に基づいちゃいかんと言われていました。

ところがそこに大きな誤解があって、最近他所で書いたことがあるんだけど[学阪 02]、心理学で言っている客観主義は、実はかなり主観性の濃いことを言っているんですよ。徹頭徹尾。

例えば、実験心理で色知覚の実験をします。今グラフィクス上に二つ黄色が出てくるから、より黄色味の強いほうのボタンを押しなさいと。これ、黄色という言葉に対する共通理解が前提になっているわけです。

一同：うん。うん。

下條：それがないと、全然何をとっているのかわからない。たまに共通理解がなくて、誤って違うもので、例えば明

るさの違いでボタンを押ししたりということが起きる。動物と人間の場合も同じでして、サルで色知覚の実験をするときは、そのおそれが非常に高いわけです。本当に、サルは明るさでもなく、（昔はスライドを使っていたので）スライドのクリック音の違いでもなく、黄色の色味の強弱に応じてボタンを押しているのか、そういう共通理解が成立しているかどうかというところが、動物実験では困難な問題になるわけです。

一同：うん。うん。

下條：ときどきそれでとんでもない失敗をする人がいたりしたわけですけど、でも成功したらそれは共通理解があると。そこを見ても、共通理解が実験成立の大前提になっている。

そういう意味で、最近で言うと、知覚、両眼視野闘争をサルでやって、知覚意識というものがサルにもあるという証明をした論文なんかが有名なのですが[Sheinberg 97]、それはつまり、人間とサルの間に関主観性という意味で、その見える・見えないという主観的な事象に対して、共通理解が生じたからデータが得られたともいえるし、逆に言うと、データを見たら、その両眼に与えられた、例えば縦縞と横縞を与えて、これが相互に抑制して交互に見えるわけですけども、交替頻度をデータとしてとってみたら、人間と質的に同じパターンになったとかね。

また例えば、あるコントロール実験でやったら結果が全く変わったが、その変わり方自体が人間と振舞いが同じわけですよ。そうすると、データから逆にその関主観性が証明されたともいえるわけです。鶏が先か、卵が先か、ですが。

だけど、そこで議論していることは徹頭徹尾、関主観性\*5であって、物理的な客観性ではなかったのです。それでも、僕が学生のときは、そういうのはその地下鉄で二つぐらい先のキャンパスで教わったわけだけど、物理的な客観性だって教えられたからね。

それとフィジオロジー、脳の電気生理

\*4 諏訪正樹、堀浩一 編著：一人称研究のすすめ 知能研究の新しい潮流、近代科学社（2015）

\*5 関主観性：哲学者のフッサールが提唱した用語で、複数の主観で共通に成り立つこと。

学と重ね合わせて、それで理解するのだったという話をしていたから、皆勘違いしちゃった。心の主観的な部分は科学の対象ではない、ということになってしまった。

間主観性というコンテキストでもっていくと、一人称、二人称、三人称の話ってわかりやすいわけですよ。つまり、自分の一人称の経験と、三人称のデータないしは、二人称の相手の経験とが重なり合っているという前提で話をしている。一人称データだって、束ねれば二、三人称だと。

でもよく見たら微妙に違っていて、色覚異常だったとか。その違いも、またその重ね合わせの中から出てくるわけですよ。

諏訪：なのに、三人称のふりをしているということですか。

下條：そう。だから簡単に言うと、物理主義に対するコンプレックスですね。しかも、まずいことに日本の心理学の場合は、文系の理系に対するコンプレックスみたいなのがぶついているので。

松原：人工知能も、ちょっと理系ではヤクザな学問なので、多分物理コンプレックスみたいなのがあって、こうなっちゃったんだよね。

### 一人称の意味

下條：そのおかげでサイドネガティブ、サイドエフェクトというのが歴史的にあって……、何でそんなことを言っているのかというと、歴史的に見ると三人称主義というのも、歴史的なネガティブ、サイドエフェクトの一種というふうに見えています、という話です。

ただね、いろいろなことを一気にしゃべっちゃったので、後でもう 1 回詳録したほうがいいのかもわからないのですが、「一人称」という言葉の使われている意味が相当ルーズですね。皆良い加減に使っていますね。というふうに感じたので、今後リファイン（精練）していただくのなら、永井均\*6 さんという哲学者がいますが、彼式に厳密に一人称の

語用法を突き詰めていく必要があるかもしれないと思います。一人称といっても、今しゃべっているこの私が全部一人称じゃないわけですよ。例えば、諏訪さんが書かれた野球の一人称のレポート、私にとってはもう一人称じゃないわけね。

一同：うん。

下條：で、もう読んでいるから、そこでメタレベルの、あなたも私と同じようにこういう経験をしているらしいというのはもう二人称なわけですね。彼、被験者の彼も、私と同じように色を見ているらしいとか、野球のときに右肘で同じような感覚があるらしいっていうのは、それはもう三人称になる。それが本来、一人称、二人称、三人称の区別なので。

それはヴァスデヴィさんとの二人称のシンポジウムするときにも言ったのですが、あなた二人称、二人称って言っているけど、それ三人称の分析をやっているでしょうって。哲学的に、厳密に言うよ。

諏訪：いやあ、僕もそう思いましたね。

下條：それを放棄しちゃったら、狂人の妄想とね、一人称の主観的な研究でインパクトの大きいもの、どうやって識別するのですかと。そここのところがかなり曖昧です。

純粹に客観的な研究なのだけれど、一人称の視点からに限って、視野の中にこういうメタレベルの記述とか、入替え可能性のようなものが入っていないとか。そのときにサッカーボールはプレイヤーの後ろに隠れて見えないとかね。そういう分析の仕方をするのが一人称だということ、それはちょっと違って、厳密に言うと三人称の中に二通りの分析がある。神の視点とプレイヤーの視点と二通りあるというだけの話だから。でも、その「だけ」の話が大きいと、実は……。という面と、さらにそこから永井均的に言っ、厳密に一人称をどう使い分けるのかという問題とある。

例えばね、僕は「ロボットの意識」について書いたことがあるのですが [下條 15]、ロボットが意識をもてるかって話がよくあって、松原さんもその文脈でご招待して豊橋でシンポジウムやったりしたんだけど、それ相当ナンセンスな

問いなわけですよ。ロボットが厳密に「一人称の意味で」意識がもてますかという問いだと考えると、私自身がロボットでない限りはその問いは意味がないわけですよ。

一同：うん。

下條：だって、そこにロボットがいて、この人が一人称の意味で意識をもっているかどうかってことを推定するのは、それすでに二人称の問題だから。

だから一人称、二人称、三人称って話は相当入り組んだ話で、それを整理していかなきゃいけないよ。

で、整理したうえでなおかつ、そうは言っても一人称でも二人称でも三人称でも、「意識をもつ」という表現を同じように適用して、例えば色知覚が黄色という言葉の意味は共通理解をしているという前提で話が進んでいくわけですよ。そこに前提があって、それで世の中が進んでいるということ自体の中に概念の乖離もあるし、ただその深奥にこそ問題の解決が実はあるかもしれない。そのあたりのところで、議論の混乱があるので、もしこれから先へ進むというのであれば、その辺がどうなのかなっていうのが、ざっと拝見した感想なのですけどね。

あれ、皆さんはどう思われているのですか、AI の分野の中では。

相当新しい提案として受け取られていますか。

諏訪：ヤクザな提案と。

伊藤：あまり受け入れられていないというのが正しいかと。

下條：ちょっとわからなかったのが、僕は AI っていうのは浅い理解しかしてないけれど、古典的にはエキスパートシステムというのがあって、人間のエキスパートの言説とか知識をいろいろコンピュータに入れていって、例えば名医と同等の、またはそれ以上の診察ができるエキスパートシステムとかね。それは古典的な AI のアプローチで、この間松原さんも呼んで、松原さんに喝破されたのだけれども、でも将棋だともう機械どうしが 1 日に 1 万局やって、どんどん強化学習しちゃうというようなことで、ある意味シンギュラリティなんてもう超えていますよという話が

\*6 永井均：哲学者。

改訂版なぜ意識は実在しないのか、岩波書店、岩波現代文庫 (2016)

あった。

そういうことで言うと、その一人称視点が「新しい」という意味がよくわからない。

つまり、羽生さんのまねをして羽生さんの棋譜を読ませて、将棋のプログラムをつくっただけなら、もう弱い、AIとしても古いわけだけど、でもそれが古典的なスタイルだったわけで、それって、見方によっては一人称じゃないのかと、そこがちょっとよくわからなかったのです。

どうなのですか。神の視点のほうが何か、つまり主流だからということですか。

### 状況依存性

諏訪：今、抜けている話が1個あると思うのです。

下條：ええ。

諏訪：それ、状況依存性<sup>\*7</sup>だと思うのですよね。

下條：ええ、ええ。

諏訪：それで羽生さんも、自分がどういう *situated* な行動をしているかは全部、まあ、ぱっとは語れませんよね。

下條：うん。

諏訪：でも何かしら *situated* なことをしているはずで、多分 *situated* などところに重要性を見いだそうとすると、一人称の視点でしかデータは得られないという考え方です。

将棋の世界で、機械どうして何万回も対戦したときに、わっといろいろなことが起こるっていうのは、そもそも、その状況依存的な知の側面がもとから抜けているからです。状況依存的に全く新しいことに着眼する、という話を抜きにすると、もうすでにシンギュラリティは超えているのかもしれない。でも、状況依存的な知に何か価値を見いだそうとすると、やっぱり一人称でしかあり得ない。つまり一人称の視点でしか見ることのできない物事があるのだけれど、それがこれまでの学問の世界にはない。その話が抜けていると思うのですよね。

下條：それは何ですか、フレームの問題とか。

諏訪：そう、フレーム問題の話ですね。

下條：羽生さんもあれでしょ、多分僕もそんなに松原さん達ほど彼の発言を追っかけていないけれど、若い頃は盤面の中だけで考えたいと言っていて、今のこの瞬間の局面で最善手を探す以外に将棋の最適手なんてないと言い切っていたと思うんですよ。最近は少し雰囲気違って、いわゆるフレーム問題のフレームとは意味が違うかもしれないけれど、もっとこの手に長考するとかね。相手はこの手の顔を立てようとしているとかね。その種の要因にもう少し手を、注意を向けるようになっていっているように思うのですけどね。

どうですかね、それ。

伊藤：よくわからないけれども、そこは、その印象はあんまりないですね。割ともう盤上の最善手を考えるというアプローチだと思います。

下條：今、藤井君<sup>\*8</sup>がそういう感じですよ。

伊藤：はい、はい。

下條：今の瞬間の中で考えている。

伊藤：そういう点では、羽生さんはあまり変わってないような印象を受けるのですけど。

松原：でも多くのプロ棋士は年を取ると、下條さんが言ったような経過をたどるのは確かです。若い頃は盤面だけ見ていて、人生経験で、相手の顔色がこういいうときにはこう思っているんじゃないとか。多分人間は相手が今の局面をどれくらい良いと思っているか、悪く思っているかというのは、かなり重要な情報で、それは多分経験によって得られることが多いので、あと勝負にもかなり実際は効くので、多くのプロ棋士はそれを使っていると思うのですが。

下條：そういう意味で言うとね、松原さん、いつも将棋・囲碁は完全情報二人ゲームだって言うけれど、今言ったその盤面の外側の話をに入れていくと、人狼に近くなってくるわけね、大げさに言うと。

松原：はい、はい。

下條：こいつはぼやいているけれども、これは三味線じゃないかとか、そういう話まで出てくるとね。

松原：だから人間は完全情報とかいうのは、あれは情報处理的なというか、数学的な言い方で、人間はそういう意味では、将棋は不完全情報のゲームなのですよね、人間の営みである以上は、明らかにリアルタイムでやっているし。

### 電王戦

下條：で、おもしろいなと思ったのは、松原さんご存じのとおり、僕は電王戦<sup>\*9</sup>ウォッチャーでね、Web ロンザ (ASAHI WebRONZA : <http://Webronza.asahi.com/>) に 5 ~ 6 本書いたのです。事件が起きるたびに、塚田九段が泣いたときとか、いろいろ書いていますけど。

非常におもしろいなと思っているのは、プロどうしの——話があちこち飛びますけれど——プロどうしの常識が覆された場面がたくさんあってね。プロが一目無理解という手をポナンザが指して、押し切っちゃったとかね。だったら、それまでのプロの「一目無理解」という共通理解は何だったろうかと。

あるいは今の盤の外側の話で言うと、時間攻めにしてもコンピュータには効かないみたいなことをプロが言ったりとかね。

普通は時間攻めで予想外の手を連発していくと、終盤だったら相手が間違えてくれるってことがあるのだけど、それも全く通用しないとかね。

そういうことで考えると、人間どうしのシチュエーションとか、エンボディ<sup>\*10</sup>のものとか、盤の外側のインタラクションみたいなものがなくなったからこそ、そういう限界も露呈したと最初に思ったのですが、そうでもなくて、シチュエーションがなくなったわけじゃなく、それとは違うすごく変なものに置き換わった

<sup>\*9</sup> 電王戦：ドワンゴが主催をしたプロ棋士とコンピュータの将棋の対戦。2011 ~ 17 年までいろいろな形で行われて最終的にはコンピュータが名人に勝利して終わった。

<sup>\*10</sup> エンボディメント：身体をもっていること。身体性。

<sup>\*7</sup> 状況依存性：物事が状況に依存しているという性質のこと。

<sup>\*8</sup> 藤井聡太：将棋棋士。中学生でプロ棋士になり、いきなり 29 連勝したことで有名。2018 年 7 月時点で七段。

んですね。こんなピカピカなアーム\*11で指しますとか。これが対局開始のときは一応お辞儀するのだけれど、あとは全くコンピュータの計算上の優先順位に基づいた評価でもって指してきて、それが恐らくは前の手の顔を立てるとかいうことを考えずに、現在の時点で評価して指してくるということが、ある意味、それ全体がシチュエートに立っていて、それに対してプロ棋士が全く用意ができていなかったと。

そういうことで盤の中ではなしに、盤の外でもプロ棋士は鍛えられてきたんだけれど、そういうことに対して鍛えられてきていなかったから、その人間どうしのシチュエーションとか、人間どうしの盤外文脈みたいなものになじんできたのが、全く異質の文脈に放り込まれてお手上げだったっていうのが、あの電王戦だったんじゃないかと思っているんですよ。

諏訪：それも多分、人間は学びますよね。

下條：学びます。だから、先崎学八段が最後のときに観戦記で書いていたのは、プロ側がようやく機械と戦う勘所を見つけ始めた。これから人間側の反撃が見られるのでは、と。

で、その反撃というのがどういふのかというと、阿久津主税八段のハメ手\*12とかね、ルール違反を誘発\*13したりとか、そういうほとんど盤外、プロレスというリング外での乱闘で勝ったみたいな、棋譜が二つぐらい出てきたりします。

\*11 ピカピカなアーム：電王戦ではAIの指し手を「電王手くん」と呼ばれるピカピカなロボットアームに伝え、代わりに駒をもって指した。

\*12 阿久津主税八段のハメ手：相手のコンピュータ将棋がある局面で必ず悪い手を指すという情報を知ったうえでその局面に誘導して悪い手を指させて勝ったことを指す。対局のルールとして対局の数か月前にプログラムを固定して変更してはいけないということになっていたので、事前にプログラムの「バグ」を見つけて利用することができた。

\*13 ルール違反を誘発：電王戦で永瀬拓矢六段がSeleneというコンピュータ将棋を相手に、角の不成(普通は成る)に対応していない「バグ」があることを事前対局で見つけて本番で角を不成と指して反則を誘発して勝利したことを指す。

松原：出てきましたね。

下條：非常におもしろかったのは、片上大輔六段という東京大学出身の棋士がいて(当時、将棋連盟の担当理事でしたけれど)「人と機械が戦うとは、つまりこういうことです」と言って、すまして自分の席に帰って来ちゃった、大変な事件を起こして帰っちゃったのだけれど、でもそのとおりでね、人間と機械が戦うと、お互いの裏をかくというか、出し抜くとかいうか、そういうことになる。というのは、もうルールの枠が相当ギリギリのところ、外に投げ捨てたりとか、いろいろなことやっている、そういう意味で非常にダイナミックだったし、おもしろかったと思っていますよ。

諏訪：似たようなことはサッカーでもありますよね。

伊藤：そうですね。

諏訪：選手が初めて海外に行って試合をすると、もう何か場外の雰囲気とか、アジアの笛とか、何かいろいろなことで試合が左右されることを経験する。

伊藤：はい、はい。

下條：あとは反則の使い方とか。

諏訪：そう、反則の使い方とか、それにいきり立ってしまう若い選手が、ああ、しまったと、いきり立ってはいけなかったか。

下條：サッカーってものの自体のルールは一緒なのに、そこで何か根本的に違うもので。

諏訪：だからこう、人間の欲と欲のぶつかり合いのような状況に放り込まれるみたいなね。

伊藤：今の話だと、羽生さんも実は、羽生さんが出てきた当初は、ちょっと機械的だと相手に思われた、と言われていました。

諏訪：そういうふうに使われていましたね。

伊藤：佐藤康光さんだったと思うのですが、「何か機械と戦っているような気がする」、ということをやったっていう話があって、だから羽生さんはそういう人間的なところをかなり排除することを目指していたんじゃないかなっていうのが。

諏訪：ああ、意図的に。

伊藤：意図的に。

下條：で、あんまり変わってないって言われたのは。

伊藤：例えば羽生マジック\*14みたいなものというの、今までの顔を立てるような手をいっさい考えずに、いきなりそこで最善手を指されたことが、多分そういうふう感じられたのではないかという気はしますね。

松原：だから逆に言うと、羽生さんが昔から機械的だったと。時代を超越して、もしかしたら藤井君とかの他のプロ棋士が追いついてきた。

伊藤：追いついてきて、羽生さんに対する見方が変わったっていう気もします。

松原：逆に言うと、相対的にちょっと今弱くなっているように感じたのが、羽生さんが年を取ったというよりは、他のプロ棋士がそれぞれ機械的なやり方を学び、身につけちゃったので、結構良いレベルまで来て。

下條：あ、そうそう、だからさっき、話の前半で終わってしまったけれど、人間と違うシチュエーションに放り込まれた。最初は棋士も右往左往して手も足も出なかった、常識が通用しないと。ところが今の若い人は、人間を相手にしていなくて機械から学んでくるから、その新しい状況に割と慣れてきているわけですよ。

そうすると、その人間が羽生さんと戦っても、そんなに違和感はないというか、その点では同じ土俵というかね。そうすると、やっぱり新しい人のほうが知識も吸収は早いし……ということが起きているのかなと思ったんですけど。

そうすると、僕は将棋弱いだけれど、お二方はアマ強豪でいらっしゃるから、最終的にどうですか。藤井君みたいな人は、藤井君は人間なのだけれど、機械に鍛えられた人間なわけじゃないですか。最終的にコンピュータが人間より強いのは当たり前でね、あれ誰だったか、ちょっと忘れちゃったのですけれど、電王戦のあおり映像だったかな、「機械が人間に勝つなんて当たり前です」と。「これからは1世紀ぐらいをかけて、人間は没々

\*14 羽生マジック：羽生氏が中終盤に見せる観戦者を含め多くの予想を上回る妙手のこと。

それを認めていくだけの話なので、それの先駆けになるでしょう」みたいなこと、憎々しげに言ってね、悪役を買って出たと思う。それがそのとおりになっちゃったのだけだね、実際には。

だけど、そういうふうな世の中に思ったより早くなりつつあると。

実際に機械が人間の天才といわれた者を負かす場面を我々は目撃することになった、AlphaGo もそうだしね。

そのときに、最終的に一つはエンタテインメントの側面はどうなのだろうという心配もあるし、それから片上さんが言った、人間と機械が真剣に戦うようになるのだと言った意味はね、僕の解釈ではいろいろあって、ルールの外側のシチュエーションとか、そこら辺の部分のところで仕掛けるというもあるし、あとはメタ認知の齟齬そごみたいなものも起こると。

例えば、機械側は相手が不成り＝成らなかつたときの場合については、原則的に駒が成らないってことは普通あり得ないので、計算時間を省いていたと。でも、人間側がそれに気がついて不成りをやったら、相手側にそのメタ認知がなかったから、つまり相手側に対するセオリーオブマインド（心の理論）がなかったから完全に出し抜かれたと、あの状況ではそういう解釈ができるわけだけだ。

そういうことがどんどん、例えば証券業界で起きるのか、とかね。まあ、将棋というものの自体は前にも松原さんと話したみたいに、AIのテクノロジー自体としては比較的、後進的な方法を使っているわけじゃないですか。まだディープラーニングも使っていないし。

それ使ってどんどん行ったときに、最終的にやっぱり人間と機械の掛け合いはどうなるのかと。それを認知心理学側から見るとね、ドメインスペシフィック（領域特異的）には人間の常識が覆されたところがおもしろくて、もしかすると人間の認知のバイアスが明らかになるかもしれない。プロ棋士は全員人間だから、人間の本性としての認知のバイアスみたいなものをもって、それはプロ棋士どうしてやっているときは間主観性だからばれない。「シェアド(共有)

リアリティ」というやつです。

そこへ機械ってものが入ってきたためにばれてしまうと。こんな攻め筋は一目無理ですなんて言っていたのに、そのまま攻めつけられて、良いところなく負けちゃったということが起こったわけです。

だからドメインスペシフィックにはそういう興味が、認知心理学の観点から見るとあるけれども、もうちょっとAIウォッチャーとして見ると、やっぱりその最終的にどうなるの。最終的にどうなるかっていうのは、人間はやっぱり機械をコントロールできるのですか、とかね、でも、人間はどこまで行ってもAIにまねできない何かをもっていますよね、とか。

### 人間と機械

下條：この間、ノーベル賞財団と学術振興会がノーベル賞ダイアログっていうのをやって。

松原：ああ、はい。

下條：僕も呼ばれて行って、なぜか、何かAIウォッチャーとして認知されているらしくて、そのAIの将来のための……。

松原：堀 浩一さんも出ていたんじゃないですか。

下條：出ておられましたよ。ノーベル賞受賞者も利根川先生はじめ、たくさん出ていて、それで最終的にAIは人間に類する心と意識をもち得るか、というセッションに入れられたわけ。で、國吉康夫さんがコーディネータで。でもね、集まってきていた人は、一人はAIの人でアメリカ人でしたけど、フランスの人文学者の女性とかは絶対そんなことあり得ないって言うわけですよ。人間並みの心をもつなんてことはね。よくある話だけでも、人間には蓄積があるし、人間の社会どうし、人間どうしのインタラクションも歴史も機械では起こり得ないだろうと、そういう論理なのだけだ。

僕はね、あんまりそういうふうには思っなくて、30年も40年前からずっとそう言ってきたのに、そのつど覆されているじゃないかと思っているわけですよ。意思決定なんて昔はできるわけないって言っていたのに、会社のいろいろなファ

イナンシャルな意思決定でも簡単な部分はもうとっくに置き換わっているし、もっと大昔のことを言うと、16世紀ぐらいのヨーロッパに行くと、商人にね、計算、お金の計算は機械ができると思うかって聞いたら、そんなの無理だって言うに決まっているでしょ。あんな高級な認知的な仕事を機械がやれるわけないだろうっていうのが常識ですよ、あの頃で言えば、それだって簡単に覆されちゃったわけだからね。

最近で言えば、セオリーオブマインドとかね、今は人狼が最前線かもしれないけれど、どんどん出て、10年か15年ごとに人間はこれだけは絶対に譲れないみたいなものもどんどん覆されてきて、それでディープブルーの敗戦があったり、電王戦があったりしたと思うので。そんなに簡単に安心しないほうがいいんじゃないかって言ったんですけど、すごく評判悪いですよ。

伊藤：そうなのですか。

下條：うん、日本の大会社の偉い人とノーベル賞受賞者がそうだけど、言っちゃいけないと思っているのかな。それとも信じてないのか、どっちなのかわからない。

諏訪：僕はあんまり信じてないほうですけどね。

下條：あ、そうですか。

諏訪：やっぱりフレーム問題だと思いますよ。別の言葉で言うと、着眼と解釈というか。

昔からクリエイティブな知を、ずっと研究テーマとして追いかけてきたんですが、クリエイティブって何かっていうことを考えていくと、何かに着眼するという点と、着眼したことに自分なりの解釈を施せるという、この二つに凝縮できるのです。でも現状のAIは、とてもじゃないけど、まだその両者をできるレベルにない【諏訪 18】。

下條：そのとおりで、そこにはサンプリングバイアスが入っていて、1万人、10万人いる中の一人の天才を取り上げて、結果においてポストディテクティブに素晴らしい着眼点、クリエイティブな世の中に影響を与えた、と。これは機械にまねできないとおっしゃっているわけですよ。

だけどそんなの、1万個の機械をつくって、乱数入れて、この範囲ってもちろん可能な社会の変数があるわけだから、その中で結果において、どれが一番成功したか、後でポストディテクティブに評価すれば、この機械は天才だったと、後づけでいくらでもものは言える。

それをね、クリエイティビティを実現したと呼ぶか、呼ばないかは際どいところだと思います。実際、歴史上の天才評価はだいたい後付けですから。

伊藤：うん、何をもってクリエイティビティと言うか。

下條：だけど、そういう攻め方ができるわけです。例えばですが。

松原：僕の「小説をつくるプロジェクト」で100万冊つくって1個名作ができたのを、後からAIはこんな名作をつくって天才だと言うか。100万個作るばかりだとか、チンパンジーがキーボードを押しただけではないかと言うか……。

下條：小説家志望の若い人、数十万人いるわけだから。たいていは駄作をつくっているわけだから、良いじゃないかという……。

伊藤：いや、だからそういうプロセスも含めて、そういうことできないってことをおっしゃっているわけですね。

諏訪：そんな気がする。

伊藤：だからそういうプロセスを含めて、そのランダムにやって何か答えを一つ見つけたっていうのを、それをクリエイティブと呼ぶかということですよ。

諏訪：いや、「範囲」って言っている時点ですでに、変数を有限個に決定しているじゃないですか。そこだと思わずね。

下條：多くの凡才はその範囲を誤って定義したりとか、見落としているわけですね。

天才だけがもう一つ別な次元があるじゃないと気がついて、それもサーチの範囲に入れるとか、あるいは皆が察してすぐに捨てるとかね、羽生さんがいつも言っているみたいに。

松原：下條さんと諏訪さん、ちょっとごめんなさい。

1万とかおっしゃるけれど、1万をそもそも選ぶってこと自体は、人間

は1万人いるから勝手に1万通りできるけれど、AIにそもそもある条件で1万通りってやらせること自体も含めてできないってということではないですか。

諏訪：そうそう。

下條：うん、そこもあるね。それもあるし、もうちょっとそう、だからその、だから変数1万個って誰が言った。

主流が人間である限りはAIは無限に進化するって言い方ができますよね。だけど、主流が機械になっちゃったらどうなるのかっていうことが。

一同：そう、そう。

諏訪：だから、機械プラス人間であるよね。

下條：それは認めます。

だけど、主流が人間である間は、金融業界であれ、それが将棋であれ、主流が人間である間は、それに対してどんどんどんどん、良いところを吸収して学んでいくということができる。

だけど、その1万個あるパラメータを全部機械が決めることになったとき、どうなるかというのはちょっとこれ、想像を絶するところがあるかもしれない。

#### まとめ

下條氏は、まず、ラマチャンドランの例をあげて、客観主義と一人称研究について歴史的視点に我々をいざなってくれました。一人称の視点の重要性については理解しながらも、それをどのように客観性を保っていくのかについて深い議論がなされました。一方で、我々がこれまで信奉してきた三人称視点の研究も共通認識という危うい前提に基づいて議論されているという視点にも気付かせてくれました。

その後、「状況依存性」というキーワードから話が広がり、将棋電王戦で見られた人工知能と人間の対戦にも話が及びました。一見、脱線したかと思いきや、人間の知の本質は、人工知能が行っているような膨大なデータからの学習ではないのではないかという視点を提供して、一人称研究の話に回帰しています。

後半では、再び一人称視点の意味を問い直しながら、意識の潜在性にも話が及びます。後半もご期待ください。

#### ◇ 参考文献 ◇

- [伊藤 13] 伊藤毅志, 松原 仁: 羽生善治氏の研究, 人工知能学会誌, Vol. 28, No. 5, pp. 702-712 (2013)
- [伊藤 15] 伊藤毅志, 松原 仁: 羽生善治氏の研究, 諏訪正樹, 堀浩一編著: 一人称研究のすすめ, 近代科学社, pp.45-84 (2015)
- [レディ 15] ヴァスデヴィ・レディ 著, 佐伯 胖 訳: 驚くべき乳幼児の心の世界, ミネルヴァ書房 (2015)
- [レディ 17] ヴァスデヴィ・レディ, 松沢 哲郎, 下條信輔: 発達心理学の新しいパラダイム—人間科学の「二人称アプローチ」, 中山書店 (2017)
- [荻阪 02] 荻阪直行, 下條信輔, 佐々木正人, 信原幸弘, 山中康裕: 意識の科学は可能か, 新曜社 (2002)
- [ラマチャンドラン 99] ヴィラヤヌル・S・ラマチャンドラン 著, 山下篤子 訳: 脳のなかの幽霊, 角川書店 (角川 21 世紀叢書) (1999)
- [Sheinberg 97] Sheinberg, D. L. and Logothetis, N. K.: The role of temporal cortical areas in perceptual organization, *Proc. Nat. Acad. Sci.*, Vol. 94, pp. 3408-3413, <https://doi.org/10.1073/pnas.94.7.3408> (1997)
- [下條 15] 下條信輔: ロボットは意識を持ち得るか?, *ASAHI WEBRONZA*, <http://webronza.asahi.com/science/articles/2015073000004.html> (2015)
- [諏訪 13] 諏訪正樹, 堀浩一 編: 特集「一人称研究の勧め」にあたって, 人工知能学会誌, Vol. 28, No. 5, p. 688 (2013)
- [諏訪 15] 諏訪正樹, 堀浩一 編著, 伊藤毅志, 松原 仁, 阿部明典, 大武美保子, 松尾 豊, 藤井晴行, 中島秀之: 一人称研究のすすめ—知能研究の新しい潮流—, 近代科学社 (2015)
- [諏訪 18] 諏訪正樹: 身体が生み出すクリエイティブ, 筑摩書房 (2018)

2018年7月27日 受理



## 著者紹介



### 下條 信輔

1978年東京大学文学部心理学科卒業。1985年マサチューセッツ工科大学(MIT)大学院修了, Ph.D. 1997年カリフォルニア工科大学生物学部/計算神経系准教授, 1998年から同大学教授(現職)。2004~10年科学技術振興機構 ERATO 下條潜在脳機能プロジェクトリーダーなどを歴任。専門は知覚心理学, 視覚科学, 認知神経科学。現在は特に, 情動と意思決定, 感覚統合と脳の可塑性, 知覚と意識などに関心をもつ。1999年サントリー学芸賞, 2004年日本神経科学会時実記念賞, 2008年日本認知科学会独創賞, 2009年中山賞大賞受賞。著書に「サブリミナル・マインド」(中公新書, 1996), 「(意識)とは何だろうか」(講談社新書, 1999), 「ブラックボックス化する現代」(日本評論社, 2017)ほか。朝日 WEBRONZA (科学・環境欄) レギュラー執筆者。



### 松原 仁(正会員)

1981年東京大学理学部情報科学科卒業。1986年同大学院工学系研究科情報工学専攻博士課程修了(工学博士)。同年, 通産省工技院電子技術総合研究所(現 産業技術総合研究所)入所。2000年公立はこだて未来大学教授。2016年同大学副理事長。人工知能, 認知科学, ゲーム情報学, 観光情報学などの研究に従事。著書に「鉄腕アトムは実現できるか」(河出書房新社, 1999), 「人工知能とは」(近代科学社, 2016), 「AIに心は宿るのか」(集英社インターナショナル, 2018)など。元 本学会会長。日本認知科学会, 情報処理学会, ロボット学会, 観光情報学会会員。



### 伊藤 毅志(正会員)

1988年北海道大学文学部行動科学科卒業。1994年名古屋大学大学院工学研究科情報工学専攻博士課程修了, 博士(工学)。同年, 電気通信大学電気通信学部助手。2005年よりデジタルハリウッド大学客員教授。2010年電気通信大学大学院情報理工学研究科助教。2017年より電気通信大学人工知能先端研究センター兼務。将棋や囲碁などの思考ゲームをプレイする人間の高度な認知過程の研究に従事。著書に「先を読む頭脳」(新潮社, 2006; 新潮文庫, 2009, 2018), 「ゲーム情報学概論—ゲームを切り拓く人工知能」(コロナ社, 2018)など。日本認知科学会, 情報処理学会, ICGA, IEEE 各会員。



### 諏訪 正樹(正会員)

1984年東京大学工学部卒業。1989年同大学院工学系研究科博士課程修了(工学博士)。同年, (株)日立製作所基礎研究所入社。推論学習の研究に従事。1994~96年スタンフォード大学 CSLI 研究所にて客員研究員。1997年シドニー大学建築デザイン学科主任研究員(Senior Researcher)。2000年より中京大学情報科学部助教授, 2004年より同学部教授。2008年4月より慶應義塾大学環境情報学部教授。身体知の学び, 感性を育む方法論, コミュニケーションのデザインの研究に従事。共編著に「一人称研究のすすめ 知能研究の新しい潮流」, 「知のデザイン 自分ごととして考えよう」(ともに近代科学社, 2015), 単著に「「こつ」と「スランブ」の研究 身体知の認知科学」(講談社, 2016), 「身体が生み出すクリエイティブ」(筑摩書房, 2018)。日本認知科学会, 日本デザイン学会各会員。