

科学基礎論学会-DLMPST 共催シンポジウム

久木田水生先生講演原稿日本語版

What are explanations worth in science?

「AI」という言葉は非常に多様な技術を指しますが、この講演において私が「AI」というときは、現在最も活用が進んでおり、最も大きな利益を生んでいる AI、すなわちビッグデータに基づいて機械学習を行なうシステムを指すことにします。

本日は科学において AI の活用が進んで行く、その中で科学における説明の価値ということについて考えたいと思います。

科学においては説明は中心的な重要性を持っていると一般には考えられます。しかししばしば指摘されるように現代のビッグデータに基づく機械学習システムはその判断の理由が不透明である、どのようにしてその判断に至ったかのプロセスがブラックボックス化してしまっている。その意味で AI はある種の「説明なしの知識」を私たちに与えます。

科学において AI の活用がこれからますます進んで行く時、果たして説明なしの科学が従来の科学に置き換わっていくのでしょうか。そのような科学は科学と言えるのでしょうか。

こういった問題を科学における説明の価値という観点から考察していきたいと思えます。

物理学者でノーベル賞の受賞者であったスティーブンワインバーグは彼の著書 *To explain the world* の中で 科学の模範は幅広い様々な現象を精密に支配するシンプルなひとまとまりの数学的原理であると述べています。彼はかなりユニークな科学観の持ち主であってこのようなもの以外は科学ではない、したがってそのような形の科学を初めて実践したニュートン以前には科学は存在しなかったとまで言っています。

また彼は 別のところでは地球以外の惑星に知的な生物がいればその生物が発達させる科学の中にも地球の科学と同じような自然の法則が見出せるだろうと言っています。このように ワインバーグ は科学というものを非常に狭い意味で捉えています。

一方で、霊長類学者のロビンダンバーはワインバーグとは対照的な科学観を提示しています。ダンバーによれば科学とは、仮説を立ててそこから導かれた予測を事実によってテストすることを繰り返すことで、この世界についての様々なことを明らかにしていく方法です。それは近代の西洋に限ったものではなく、あらゆる時代と場所の人

間、それどころかあらゆる高等な生物に共通の方法だとダンバーは考えます。このようにダンバーは科学というものを非常に広くとらえています。

科学をワインバーグのように捉えるならば、AIによって生成された「知識」は科学的知識とは言えないでしょう。一方でダンバーのように捉えるなら AI 科学もまた科学的活動の一種であると言える可能性はあります。定義の問題はさておき、科学の現場には AI がどんどん導入されており、そしてこれからも科学において AI の役割は大きくなっていくだろうということは確かだろうと思われます。なぜなら現在、様々な観測機器から得られる膨大で多種多様なデータを AI で処理することなしには手に入れることができない有益な情報が存在するからです。そのような情報を「科学ではない」といって科学者みんなが拒否することはおそらくないでしょう。

ここでワインバーグの見解を振り返ってみましょう。そこでは次のような価値が重視されています。つまり単純さ、正確さ、そして一般性です。正確さが重視されるのは当然に思われます。爆弾の落ちる位置を 1 キロメートルの誤差で予測できるより、1メートルの誤差で予測できた方が役に立つのは理解できます。しかし正確さや一般性についてはそれほど明らかではありません。なぜ科学はこのような性質を模範としているのでしょうか。

この点を考えるために、「情報の流動性」ということが科学（ワインバーグの狭い意味の科学ではなく、ダンバーの広い意味の科学）において持つ重要性に注目したいと思います。ここでは三つの異なる状況での情報の流動性を考えます。それは Cognitive fluidity、Intra-group fluidity、Inter-group fluidity です。

Cognitive fluidity は認知考古学者の Steven Mithen によって提唱された考えです。ミズンはこれを、Homo sapiens とその他の hominids を分ける重要な知的特徴と考えました。これはある domain において得られた知識やアイデアを別な domain に輸出して応用する能力です。Homo sapiens は social intelligence、natural history intelligence、technical intelligence といった、特定の domain に特化した知能を持っています。しかしこういった知能は Homo sapiens に近い hominids、例えばネアンデルタール人なども持っていたらと思われれます。Homo sapiens だけが、個々の領域で得られた知識やアイデアを抽象化やメタファーを使うことで他の領域に応用することができた、そしてそれが Homo sapiens と他の hominids の大きな違いを作った、というのが Mithen の説です。

もう一つの fluidity は集団内での情報の共有、継承です。これを強調するのが、人類学者の Joseph Henrich です。彼は、人間と他の霊長類の認知能力を比較したとき、人間の最も顕著な優位は social learning にあるということを示す研究に注目します。人間は他者の振る舞いを見て、模倣し、そのことによって知識や技能を獲得するのが非常に得意です。「to ape」という言葉は「人の真似をする」という意味で使われますが、実は人間こそが aping をします。こうして世代を経て集団の中に有用な知識や技能が蓄積されていきます。

いかに知的に優れた人間でも一人では到達できない知識や技能が社会には溢れています。Henrich のお気に入りの例は、キャッサバの処理方法です。キャッサバは栽培が容易で栄養が豊富な食料源ですが、毒が含まれているのでそのままでは食べられません。かなり複雑な工程を経てようやく毒が除去されます。このような知識や技能は社会的学習を通じて集団の中で獲得、継承されていきます。それゆえ Henrich は共同体を our collective brains と表現します。

Alex Pentland もまた集団内の情報の流れの重要性に着目している研究者です。彼は様々な規模の集団において、social network interactions と idea flow が creative outputs と productivity の major drivers であることを実証的に示しています。例

例えば Pentland らの研究によって、都市における social network に沿ったアイデアの流れは、一平方マイル当たりの GDP といった経済的な統計的指標をかなり正確に説明することが分かりました。

集団間の情報の流れについては、Jared Diamond の研究を参照しましょう。

Diamond によれば、ユーラシア大陸で最も早く科学技術が発達した究極の要因は、ユーラシア大陸の東西に広がる地形でした。東西に広がっているということは同一の気候帯に広い土地が広がっているということで、これによってある地域で発見された栽培種が他の地域に伝播することが容易になります。こうしてユーラシア大陸では早い時期から離れた地域間での交流が行なわれていました。そのことがアイデアや製品の交換を促し、科学技術の発展につながった、と Diamond は考えます。

以上の話から得られる教訓は、新しい発見や発明のための秘訣は、いかに情報やアイデアが一つのところから別の場所に滞りなく流れるようにするか、ということだ、ということです。そのことを考えると、なぜ科学において単純性と一般性が重視されるのか、ということが理解できます。ある人が観察や実験に基づいて、それを説明する何らかの理論、モデル、数式を立てます。この理論その他は、他の新しい事例に適

用されることで始めて意味を持ちます。そしてこの目的のために単純で一般的な数学的理論、数式、モデルが理想的に適しているのです。この点は Dowek 教授の講演にあった、説明のアルゴリズムの広さや速さに関連しています。

ここで論理学者、Solomon Feferman の言葉を参照しましょう。彼は代数やトポロジーにおける公理は、自明な真理を表現したものではなく、ある種の数学的構造に共通する特徴を抜き出して、より一般的な構造を定義するものであると述べます。

彼はそのような公理を structural axioms と呼び、それらの価値は数学的仕事を組織化することにあると述べます。数学者がそういった公理を考えだすのは、「a continuing need to package and communicate our knowledge in digestible ways」の故にであると述べます。

論理学の観点から図式化するということ。introduction の後に elimination が続く証明は、「非正規的」と呼ばれますが、しかし集団的な知的活動としての数学には不可欠な要素です。なおかつこのような証明は、新しい抽象的な概念を生み出し

て、数学の世界（あるいはより広い科学の世界）を拡張するという意味でも価値のあるものです。

数学、論理学、あるいは計算論の創造的側面については、最近出版された Masao Morita の『計算する生命』が良い案内を提供してくれます。計算とは単に決まったルールに従ってトークンや記号を動かして結果を出力するだけのものではなく、記号に意味を見いださないではられない人間という生きた主体が、世界に新しい意味を作り出していく過程であり、自らの精神を理解し拡張する過程だ、と著者は論じます。

メタファー、シンボル、言語、数学は、一人一人は乏しい知性しか持たない人間が、極めて複雑な世界がどのように動いているかを把握し、伝達し、それによってそこで生き延びていくために考案した道具です。その一方で、現代の人工知能は人間では手に負えない世界の複雑さに、人間には理解しえない仕方に対処できるようになっています。人工知能は、私たちには理解できないけれど、でもそれに従えば効率よく利益が得られ、リスクを回避できるような判断を提供してくれます。

しかしそのような利益やリスクに対処することが科学のすべてではないでしょう。再びワインバーグの言葉を参照しましょう。To Explain the World の最後でワインバー

グは、科学のもたらす喜びに言及しています。何世紀もの間に、人間はどのような理解を、どのようにすれば獲得できるかを次第に学んでいきます。そうして私たちは次第に理解、説明そのものに対して満足や喜びを見出すような美的感覚を発達させたのだ、とワインバーグは言います。

ワインバーグは近代科学のことだけを言っていますが、理解することに喜びを見いだす感性はもっともっと古くからあったに違いありません。実際、ダヴィンチは、
` `The noblest pleasure is the joy of understanding"と述べたと言われていまし、
アルキメデスは、風呂に入っている時に複雑な形状の物体の体積を計る方法を発見して、喜びのあまりに裸で町中を走ったと言われます。AI 科学のウェイトが増すことはこのような喜びを得る機会が減少することを意味するのでしょうか。

コンピュータを使った数学の定理証明は、AI 科学について考えるヒントを提供するでしょう。ケプラー予想のコンピューターによる証明について、Szpiro は次のように述べています。` `After examining the proof of the Kepler conjecture, one cannot help asking: `What have I learned from this proof?, `Did I get any deeper insight about mathematics?", or `Am I any smarter because I learned the proof?' The answer is, unfortunately, `No'."

コンピューターを用いた定理証明のもう一つの有名な例である四色定理の証明について、哲学者の Thomas Tymoczko は、それが伝統的な意味での証明といえるものではなく、それを証明として受け入れるならば私たちは数学における定理や証明についての考えを変えなければならない、と主張します。

実際に数学の定理についての考えは歴史的に変わってきたし、これからも変わっていくのかもしれませんが。実は数学における「証明」概念の拡張は以前にも起こっています。19 世紀末、集合論が創始され、無限集合を用いた非構成的な証明が用いられるようになった当初は反対もありましたが、その生産性（定理が証明できて論文がかけられる）ゆえに非構成的方法は数学者に受け入れられるようになりました。数学者の Kazushige Terui は、いずれはコンピューターによる証明が数学的厳密さの標準として受け入れられるかもしれないと言っています。

Masao Morita は最近の著作『計算する生命』のなかで、人類の知的歴史を「計算」という観点から描写しました。彼は、計算とは単に決まったルールに従ってトークンや記号を動かして結果を出力するだけのものではなく、記号に意味を見いださないではいられない人間という生きた主体が、世界に新しい意味を作り出し、その意味によ

って世界と自分自身を豊かにする過程であると論じました。コンピューターに依存しすぎることはこのような新しい意味の創造によって世界と自身を豊かにする機会が失われることを意味するのでしょうか。

おそらく AI 科学はその生産性ゆえに普及して、科学の中でのウェイトを増していくでしょう。それに伴って科学が人間にとって持っている価値や意味も変わっていくかもしれません。それは産業革命以来、労働の在り方が大きく変わり、労働の価値や意味が変わっていったのと同様です。おそらく科学から喜びが失われるというだけでなく、自然世界に対する畏敬や、人々がそれに対して持つ責任も失われるかもしれません。
