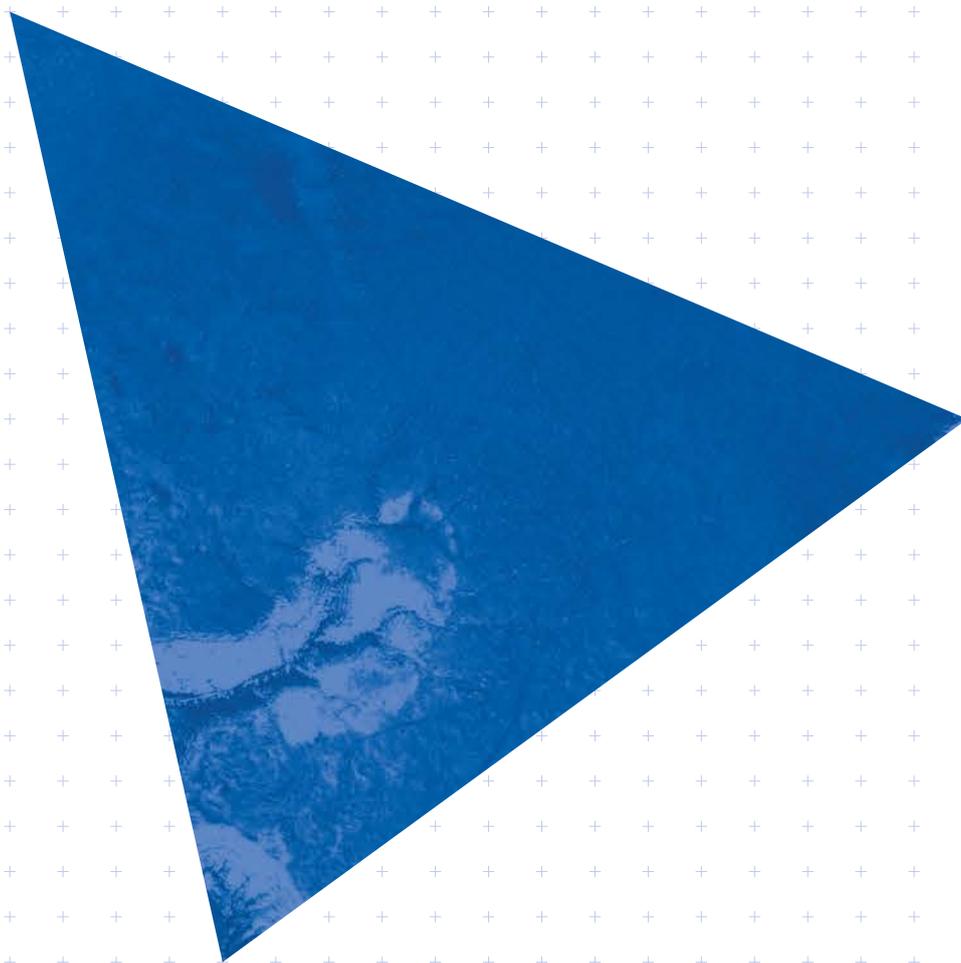


俯瞰セミナーシリーズ報告書

# 機械学習と科学

2020年7月～9月開催



# エグゼクティブサマリー

---

本報告書は、国立研究開発法人科学技術振興機構（JST）研究開発戦略センター（CRDS）が2020年7～9月に開催した俯瞰セミナーシリーズ「機械学習と科学」に関するものである。

CRDSでは機械学習、ロボティクス、ビッグデータなどデジタル技術を活用した新たな価値創造と変革（DX：デジタルトランスフォーメーション）が、科学技術の様々な分野の研究開発にも生じているとの認識から、21世紀の科学技術のあり方の変容についてまとめた報告書「デジタルトランスフォーメーションに伴う科学技術・イノベーションの変容」を2020年4月に発行した。

本セミナーシリーズでは、その先の未来像として、機械学習、スパコン、実験ロボットなどの様々な道具の力による科学研究のあり方の広がり、それにより得られる知のフロンティアの可能性に着目した。セミナーでは、計算機科学・情報科学側と科学技術の分野側の両面から今何が起きているか俯瞰することから始め、可能性と課題を議論した。具体的には

1. 第4の科学（データ駆動科学）までとの違いは何か？ 新たなツールのどのような機能・性能が、人間の性能を超えた（あるいは強化した）ことになるのか？ しきい値のようなものがあるか？
2. 科学による知のフロンティアはどこまで、どのように広がっていくのか？ ツールを使い理解を広げていく方向とは別に、人間が理解できない科学のような方向にも広がるか？
3. 新しいツールは、科学技術における人間の活動のどの部分を代替することになるのか？ 具体的に期待される問題設定・用途や利点は何か？

の3点を中心に幅広い観点で議論し、科学技術と推進方策・戦略の両面からボトルネック課題を探索した。これらの議論を踏まえ、CRDSでは今後国として重点的に推進すべき研究領域と具体的研究開発課題を検討し、研究開発の推進方法も含めた戦略プロポーザルとして関係府省や産業界・学界等に提案する予定である。



## 目次

---

1		開催趣旨	1
2		「機械学習と科学」	5
2.1		科学の変容 —アストロバイオ構成主義— 久木田 水生 (名古屋大学)	5
2.2		データと測定数理制約による科学的モデルの発見 鷲尾 隆 (大阪大学産業科学研究所)	13
2.3		機械学習と物理 橋本 幸士 (大阪大学)	22
2.4		材料における機械学習の現状と将来展望 徐 一斌 (物質・材料研究機構)	33
2.5		データ駆動による因果仮説探索 清水 昌平 (滋賀大学 / 理化学研究所)	45
2.6		Creating the Engine for Scientific Discovery 北野 宏明 (ソニーコンピューターサイエンス研究所)	59
2.7		細胞計測解析を人から解放し、人知を拡張する 太田 禎生 (東京大学 / シンクサイト株式会社)	76
2.8		機械学習、数理科学に基づく健康・疾患層別化と予測 川上 英良 (理化学研究所 / 千葉大学)	83
2.9		因果推論 —決定論的アプローチ— 陳 洛南 (中国科学院系統生物学重点实验室)	96
2.10		記号推論と科学 井上 克巳 (国立情報学研究所)	103
2.11		機械学習によるデータ駆動型研究が拓く創薬と医療 山西 芳裕 (九州工業大学)	113
2.12		プロセスインフォマティクスWS からの内容紹介 眞子 隆志 (JST-CRDS)	127
2.13		モデル駆動型アプローチからみた逆問題の諸相 山本 昌宏 (東京大学)	134

<b>3</b>	<b>総合討論</b> .....	<b>146</b>
3.1	現状 .....	146
3.2	目標・研究開発課題 .....	147
3.3	推進方策 .....	148

# 1 | 開催趣旨

## 俯瞰セミナーシリーズ「機械学習と科学」

嶋田 義皓 (JST-CRDS)

機械学習と科学というタイトルで新しい科学研究の潮流を調査しており、今後、プロポーザルとして提言する予定である。新型コロナウイルスの感染拡大を受け、大人数で集まったワークショップではなく、オンラインのセミナーを複数回実施することとした。広範な分野を扱うが、バラバラにならないよう縦串を通すという意味合いを込め、上田氏（東京大学）、丸山氏（株式会社 Preferred Networks）の2名の方を本セミナーシリーズのアドバイザーとした。全13回の予定で、統計的機械学習だけでなく、因果推論、探索、最適化なども扱う。また、対象分野としても、物理、化学、材料科学、生命科学、創薬、医療などさまざまなものを扱う。

本セミナーの趣旨は「機械学習と科学」に関係する研究開発の動向の俯瞰である。現状・課題の把握、強化するために必要な戦略・方向性の認識、そして、研究開発テーマと推進方策の両面から、ボトルネック課題を探索したい。議論としては「第4の科学までとの違いは何か?」「知のフロンティアはどこまで広がるか?」「新しいツールは、人間活動の何を代替するか?」という点に着目する。

大きな背景として、CRDSで2020年3月に発刊した「デジタルトランスフォーメーションに伴う科学技術・イノベーションの変容」での議論がある。ここでは、ITが科学・研究現場に浸透したことで、科学技術のあり方が大きく変容する様子が描かれている。ビジネスの文脈で取り沙汰されるデジタルトランスフォーメーション（DX）が研究開発の世界にも浸透し、いろいろなことが変わろうとしているのでは、ということである。とくに、データ駆動アプローチと計算理論的アプローチ、計測技術・センサー・ロボット技術とAI技術との融合によるデータ取得・収集、オープンサイエンス・オープンイノベーション型研究プラットフォーム、などが論じられている（図1-1）。我々の興味は、ではこの先に何が待っているのか、ということである。

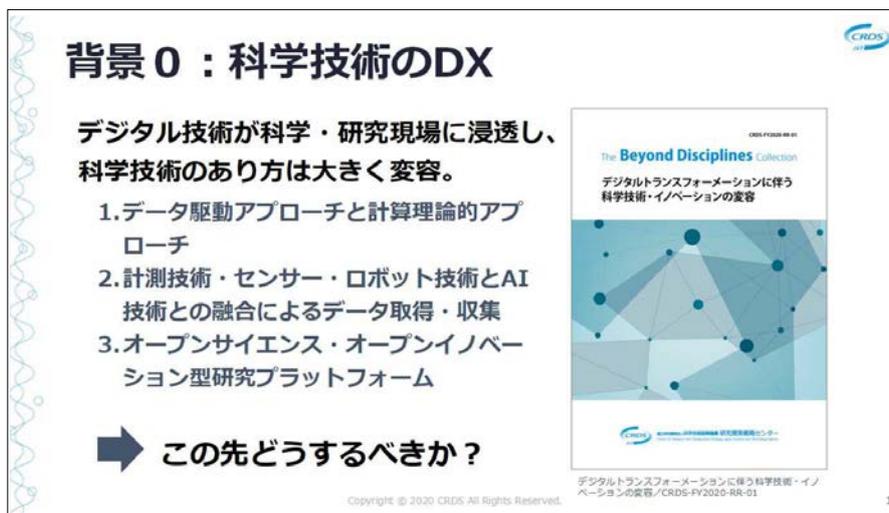


図1-1 科学技術のDX

統計的機械学習が極めて強力なツールになってきていて、科学の方法論をドラスティックに変えるのではないかと、ということで注目している。これまでの科学の方法論としては、理論科学、実験科学、計算科学、データ駆動科学などという形で進められてきた。ではこの先に何が来るのだろうか、というのが我々の興味である。

背景は3つある。1つ目は、やはり機械学習がコモディティツールとして様々な分野の研究の必須のツールになっていることが挙げられる。図1-2に挙げたように、教師あり機械学習、画像認識、スパースモデリングなど、様々なものがツールとして最先端の研究に使われている。2つ目の背景は、要素還元主義の限界である(図1-3)。とくに生命科学では顕著だが、高次元・非線形で複雑な現象を適切に取り扱おうようなモデルが分からないときに、機械学習を用いて近似するというのが登場している。3つ目の背景は、人間の認知限界である。図1-4に示したように、昔から言われてきたことだが、科学というものが人間が進めるものである以上、人間の脳の能力の限界や、認知バイアスの影響を受けている。また、人間のコミュニケーションの方法など、人間の社会的な能力がこれまでの科学を形作ってきたとも言える。このような限界を、ツールによって変えていけるのではないかと、という期待がある。

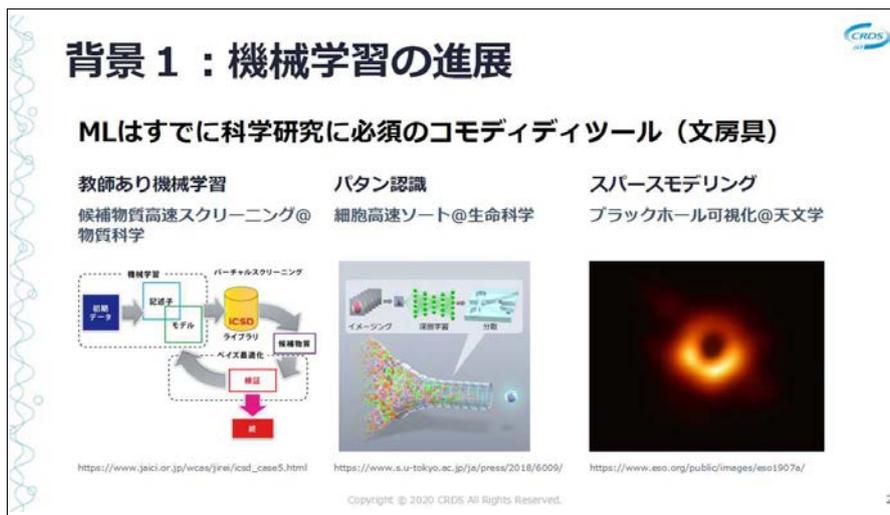


図1-2 機械学習の進展

こういったことの前にある一つの考え方として「科学の自動化」というものがある。仮説生成、定式化・作業仮説化、観測・実験検証、そして理解、といったフェーズに分けて考えると、様々な部分で人手や手計算から計算機やセンサー、ロボットなどへの転換が行われてきた。この先、具体的に何が変わっていくかに興味があるのだが、これだけだと、人間が科学研究をしなくなるという「科学の疎外」ということになってしまう。

一方で、機械学習による科学研究の推進によるご利益というのもわかってきている(図1-5)。現在までに分かっているという直近のご利益ということでは、マクロとミクロの間に落ちてしまって綺麗なモデルで書けないような現象が扱える、パラメーターから直接設計するような逆問題を取り扱うことができる、などが挙げられる。これ以外にもたくさんあるとは考えられるので、このセミナーシリーズでも触れたい。

図1-6にセミナーシリーズで議論したい「3つの問い」をまとめた。1つは、これまでのデータ駆動科学などとの本質的な違いがあるのだとすると、それは何か、という疑問だ。もう一つは、それがどこまで広がっていかれるかということである。図1-6では右側の境界線である。そして、大事なものは、この間に具体的に何があ

るか、ということである。新しいツールがこれまでの科学研究の活動の何を代替するか、あるいは、今まで人間の科学者が行わなかったことを機械によって行うようになるのか、などを探索したい。

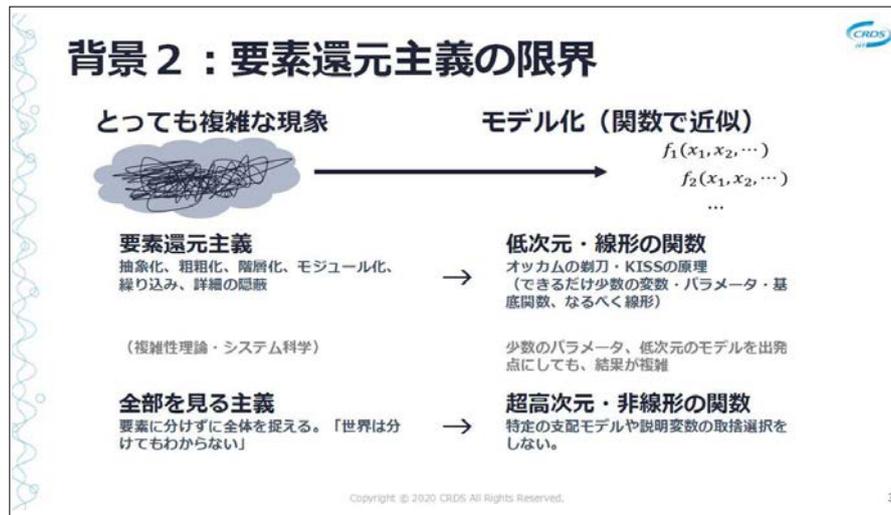


図 1-3 要素還元主義の限界



図 1-4 人間の認知限界

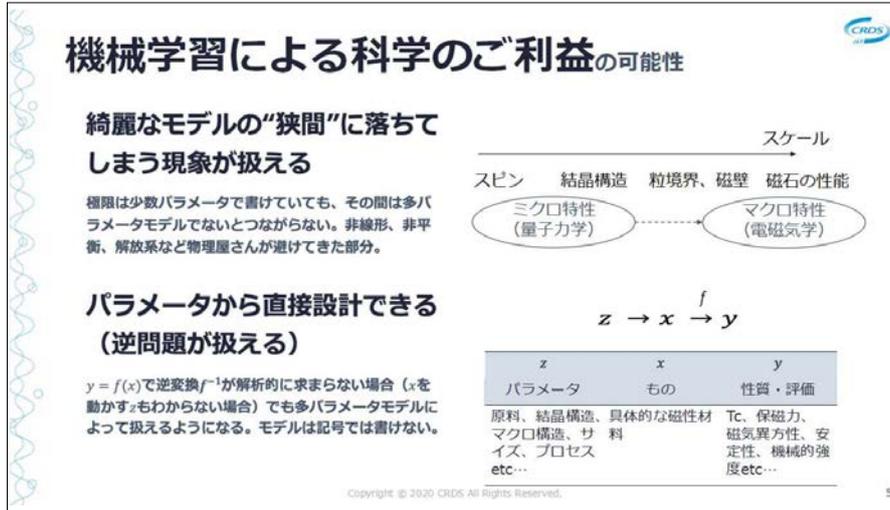


図 1-5 機械学習による科学のご利益

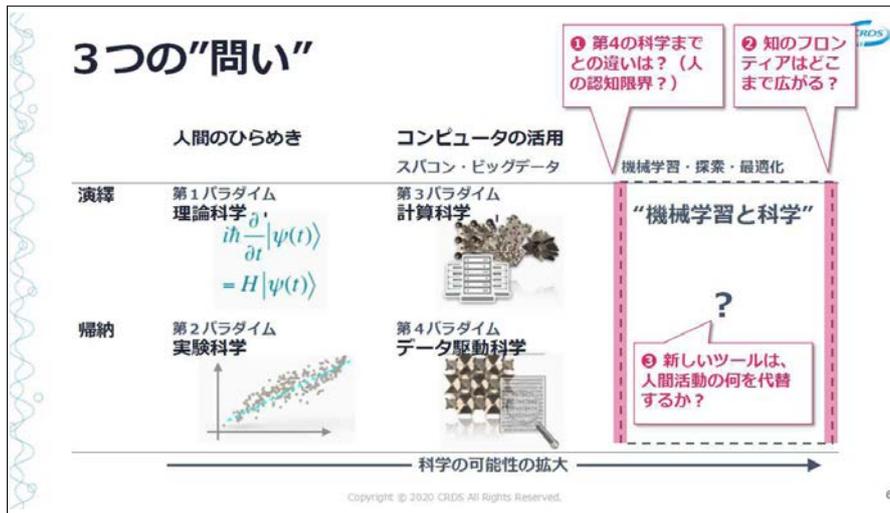


図 1-6 3つの“問い”

## 2 | 「機械学習と科学」

### 2.1 科学の変容 —アストロバイオ構成主義—

久木田 水生 (名古屋大学)

私は科学哲学のど真ん中の専門家ではないが、とくに人工知能がどういふふうにかこれからの科学を変えていくかという問題について考えている。一番の関心の中心は、科学技術と社会との関係である。特に発展が著しい情報技術が、どういふふうにか社会や人間を変えていくのかに興味があり、最近では人工知能やロボットなどの話もよくしている。科学の価値とはそもそもなんだろうかということ問い直すいいきっかけになるかもしれないと捉えている。

本日の話の一つの背景は、機械学習やビッグデータの技術が科学では当たり前のツールになりつつあり、それが科学の実践を変えつつあることである。そうすると、これまで我々が考えていた知識の在り方も、その価値も変わっていくだろうと考えている。

もう一つの背景は、科学における普遍主義と社会構成主義という2つの考え方の対立がある。意外に思われるかもしれないが、この問題と人工知能が科学をどう変容させるかという問題には、密接な関係がある(図2-1-1)。科学は普遍的で、物理学や数学だったら、遠く離れた異星人でも、もし高度な知性をもっているのであれば、地球の物理学や数学と同じようなものを発達させているだろうと考える科学者がいる。一方、社会構成主義では、科学とはそれほど普遍的なものではなく、あくまでも社会に相対的に作られると考えられる。社会構成主義は一般にある現象や概念が社会的な要因によって引き起こされていると考える立場で、ジェンダーや人種、身分、社会階層、精神病、宗教などは社会的に構成されるものだと言われる。科学に関しても社会的に構成されるというのは構成主義の中でもラジカルな立場であり、科学者や科学哲学者にはこの立場に対して批判的な人が多い。

2.1 科学の変容 —アストロバイオ構成主義—

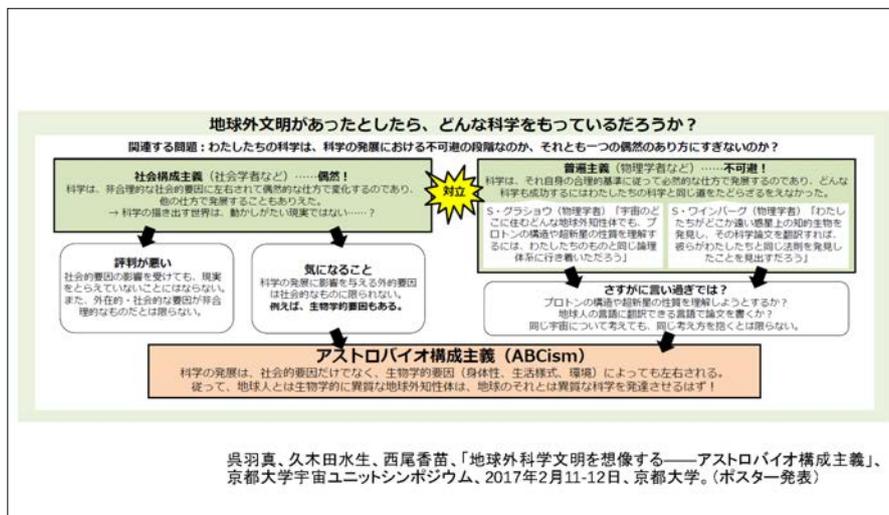


図 2-1-1 アストロバイオ構成主義

このような社会構成主義と普遍主義の対立を、アストロバイオ構成主義という考え方でアウフヘーベンして解消するということを考えた（図2-1-1）。社会構成主義については、社会ばかり考えていて生物学的な要因をあまり見ていないのではないかと。普遍主義の主張に関しては、グラシヨーやワインバーグなどが言っているように宇宙のどこでも同じ様な科学を発達させるだろうというのは、さすがに言いすぎなのではないかと思う。論文とかいった制度はさすがに社会的に構成されていると言ってよいのではないかと思うわけである。

アストロバイオ構成主義は、科学の発展というのは社会的要因だけではなく生物学的な要因によっても左右されるのだとする立場。従って、地球人とは生物学的に異質な地球外生命体は、地球のそれとは異質な科学を発達させるはずだと考える。地球人は地球人の身体性や地球という環境に依存した科学や数学を発達させているのではないかと考える。もし、人間とは全く違うタイプの生き物が科学を発達させたとして、キノコ数学、ヒトデ言語、サメ物理学などを考えてみたが、人間の科学とそこまで変わらないのではないかとも思われる。そもそも人間の科学と全然違うものを想像するのは難しい。もう少しドラスティックに違うような科学の在り方はないのだろうかということを考えていて思い当たったのが、AI科学が発達してロボットが実験を行うイメージである。

このようなことを踏まえて、科学の本質とは何か、価値とは何か、科学が今我々の考えているような科学になったのはどういう理由があるのか、他に科学の有り様があったのではないかと、あるいはこれから違うものになっていくのか、科学は社会的に構成されたものなのか、それとも社会的な要因が果たす役割は些細なものだけしか無いのか、などに関心がある。

科学の自動化に関しては、楽観的な人と悲観的な人がいて、楽観的な人は、自動化によって人間は退屈なルーチンから解放され、人間はもっと創造的な仕事に集中することができ、今までにない新しい科学の在り方が生まれると考える。一方で悲観的な人は、テクノロジー失業と同じように科学者の仕事も奪われてしまうのではないかと、あるいは科学がますますルーチン化してしまうのではないかと考える。データを集めて機械学習のチューニングをひたすらしたり、データにラベル付けをしたりなどが科学者の仕事になっていき、科学から得られる純粋な喜びが失われてしまうのではないかという悲観もある。

**私たちが関心を持っている問題**

- 科学の本質とは何か？
- 科学の価値とは何か？
- 科学が現在のようなものになったのにはどのような理由があるのか？
- 科学の別のあり方にはどのようなものがあるか？
- 科学は社会的に構築／構成されたものなのか？
- 科学は現実の忠実な鏡か？

図2-1-2 関心のある問題

ここで、今我々が持っている「科学」の特徴を振り返って考えてみたい。ユヴァル・ノア・ハラリの「サピエンス全史」の中では、近代科学の特徴は、この世界には我々がまだ知らないが知る価値のある事実があるという無知の自覚をベースにし、数学を使用することだと書かれている。そして、帝国主義や資本主義と協力し合って、科学の発展が新しい力の獲得と結びついているということを指摘している。

一方で、物理学者スティーブン・ワインバーグは『科学の発見』の中で、科学はニュートンが始めたのだと言っている。彼はニュートンが、幅広い様々な現象を精密に支配するシンプルな数学的原理という物理理論の一つの模範を未来に示したとしている。ハラリと共通しているのは、数学を使っているという点に注目しているということである。幅広さや精密さが重要というのは分かるが、なぜ数学を使うことやシンプルであることがそれほど重要なのか。

これは科学に関して、我々が持っている価値観を表しているのではないかと思う。よく、科学と価値は独立だ、科学それ自体に価値はない、などと言う人もいるが、明らかに価値というものがあるように見える。それは、なるべくシンプルな理論、モデル、数式が良く、できるだけ一般的で広範な現象を正確に説明したり予測したり記述できることが望ましい、という価値である。さらにワインバーグは、美的感覚や報酬ということにも言及しているが、これも科学と価値が結びついていることを表している。

文明の発展を促進した3つの「流れ」に注目したい。それは認知的流動性、大きな社会の中での分業と知識の流通、異なる文明間の交流、という3つのレベルでの情報や知識の流れである。

まず認知的流動性であるが、これは認知考古学者のスティーブン・ミズンが提唱した言葉である。人間は博物学的な知能、技術的知能、社会的知能という3つの異なる分野で高い知能を持っているが、これ自体はネアンデルタールでも持っていただろうと言われている。しかし、ホモ・サピエンスに特有なのは、これらの異なる分野間で知識を横断的に持ち運べる認知的な流動性であり、これが決定的に重要であったと彼は指摘している。

社会の中での知識の流通については、人類学者のジョセフ・ヘンリックの本（『文化がヒトを進化させた』）などで強調されている。社会や集団で情報を探求・蓄積・継承し存続させるということが文明の発展にとって非常に重要である。

また、離れた社会の間での価値ある知識やアイデア、産品を交換することが、人類の繁栄の鍵であった、とイギリスのサイエンスライターのマット・リドレーは言っている。同様のことは、ジャレド・ダイヤモンドの『銃・病原菌・鉄』の中にも見て取れる。ダイヤモンドは東西に広がっていることがユーラシア大陸の文明が他の文明よりも早く発展した要因だったと述べる。東西に長いと、同じ気候帯に多くの人が住んでいて、そこでは同じ栽培植物が使えるということで、文明間の交流が促進された、と考えるわけである。これによりアイデアが大陸のあちこちに行き来して、そこでアイデア同士が組み合わさって新しいイノベーションが生まれ、そのことによってユーラシア大陸の文明は他の大陸の文明よりもはるかに速いスピードで進んだのだろう。

このような3つの「流れ」が人類にとって重要だったことに照らすと、今現在我々が持っている科学がいかに便利であるかが理解できる。一般化・抽象化することによって理論やモデルが様々な新しい事例に適用できる。しかも、これがシンプルな数式で表現されている場合には、新しい事例に適用するのが非常に簡単で、認知的コストが少なくて済む。当たり前の話のようだが、このように理解することができる。

ここから考えを進めると、図2-1-3に挙げたような地球人的な条件から、現代科学の特徴の利点が理解できる。地球では環境が非常に複雑で予測が難しく、それと比べて人間の知的資源というのは貧弱である、独立した知能を持つ複数の個体が協力して問題解決に当たる、知識を言葉のような間接的な形で伝達しなければならない、などの条件に制約されている地球人にとっては、今の数学や物理学などが非常に便利だったの

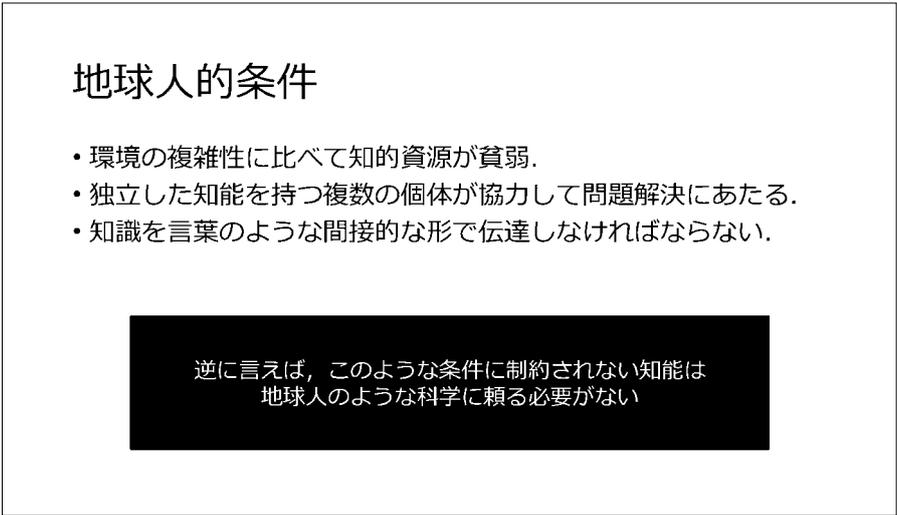


図 2-1-3 地球人的条件

ではないかと思われるわけである。逆に言うと、このような条件に制約されない知能は、地球人のような科学に頼る必要がないだろうとも言える。

エイリアン科学というものについて考えたい。ケヴィン・ケリーが『〈インターネット〉の次に来るもの』の中で、AIという言葉は異質の知性（Alien Intelligence）の略号にもなるだろうと言っている。ケリーは、AIは科学について異星人のように考え、それはあまりに人間の科学者と違う考え方なので、人間も科学に対して違った考え方をするように迫られるだろうと述べている。現在のビッグデータに基づく機械学習システムもブラックボックスだとよく言われる。このようなものが科学でも全面的に使われるときに、それは果たして科学と言えるのだろうか。少なくとも、ワインバーグが考えていたような科学とはぜんぜん異なっている。

地球人の条件に制約されない知能の（架空の）例として、ニック・ポストロムの『スーパーインテリジェンス』の中で描写される「シングルトン」といわれている超知能が挙げられる。これは、環境の複雑さに圧倒されないだけの知的リソースを持ち、言語によらず内部状態をコピーするだけで知識を伝達できる。この超知能がたった1体だけになって世界の隅々までを全部知ってしまうようになる状況をポストロムは想像している。

超知能はシンギュラリティと関連付けて論じられることが多いが、必ずしもシンギュラリティを待たなくても、我々の知能はホモ・サピエンスがから少しずつ蓄積した知識や技術によってどんどん向上している。その意味で私は、シンギュラリティはすでに起こっているというふうに見ている。単体の人工知能が人間より賢くなるのではなく、次世代の人間が前世代の人間よりずっと賢くなっているという意味では、もうすでにシンギュラリティが起こっていると考えてよいのではないか。

コンピューターを用いた数学的定理の証明というのは70年代からあり、例えば4色定理の証明をコンピューターで行ったときには、これは証明と言えるのかどうかについて論争があった。ケプラー予想の証明も、これが証明できたからと言って、我々は数学について理解や深い洞察を得たのだろうかという疑問があった。照井一成先生の『コンピューターは数学者になれるのか?』という本の中では、いずれコンピューターによる検証がスタンダードとして確立する日が来るかもしれないということが書かれている。コンピューターとは関係ないが、数学において「証明」の概念の拡張というのは以前にも起こっている。19世紀末に集合論が創始されたときに無限集合を用いた非構成的な証明が使われるようになり、当時はこれを証明とみなすかどうかについて

の議論があったが、今では数学者に受け入れられるようになった。

科学においても同じように、使えるから別に理解しなくてもいいというようなことは起こってくるのではないかと思う。そこで懸念されるのが「科学の疎外」である。エイリアン科学によって、科学がエイリアネートされる。マルクスなどが使った「労働の疎外」というときの疎外である。

科学にとって、現象を予測して制御する力と物事の在り方を知り理解する喜びのどちらが本質的なのだろうか（図2-1-4）。レオナルド・ダ・ヴィンチの「最も高貴な喜びは理解することの喜びだ」という言葉に表されるように、やはり科学者が科学をやる動機にはそういう喜びが得られることがあると思う。ワインバーグも似たようなことを示唆している。しかし、一般の人はどうかという、結局のところ社会は実用性を求めるだろう。カート・ヴォネガットの小説『ホーカス・ポーカス』の中に「情報というものはエンターテインメントでなければほとんどの人にとって用はない。その事実がおかしくも怖くもなく、自分を金持ちにしてくれるのでもなければ、一体それが何だというのだ」という言葉が出てくるが、これは学問についても言えることだろう。

我々の理解の喜びというものは、食べ物のおいしさというのと同じことではないか。食べるという行為のそもそもの機能は栄養を摂取することである。この意味で、食べる喜びというのは2次的であると言える。2次的だから無くてもいいというわけではないが、結局の所は物質的利益が1次的で、そこから得られる喜びというのは2次的なものではないかと考えることができる。無論、それで科学者の研究の動機が物質的利益であると言っているわけではない。

産業革命のときには、仕事の喜びよりも効率が優先され労働が疎外されるということが起こった。同じようなことは科学でも起こるのではないかと考えられ、そのことには倫理的な懸念もある。科学から喜びが失われるだけでなく、自然や世界に対する畏敬の念や、科学者の責任も失われていく可能性も考えなければならぬ。世の中には「完全栄養食」と謳われる商品がある。しかしそのようなものがあったとしても、我々はそれだけを食べて生きるようにはならず、やはり楽しみのために何かを食べるということは続けていこう。おそらく科学に関しても、人工知能を使って自動化された科学というのは、完全栄養食のようなものとしてあり、一方で楽しみのための科学も残されていこう。ただ、それはかなり縮小されて日曜大工のようなものになってしまうのかもしれないと思うわけである。

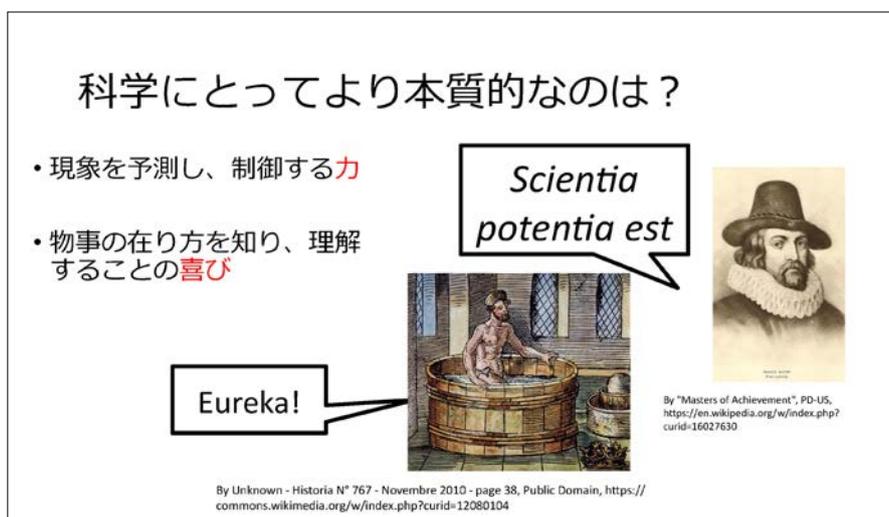


図2-1-4 科学にとってより本質的なのは？

【質疑応答】

Q：科学では合意を出すために様々な手続きがあると思う。例えば統計的仮説検定を行ってP値が5%に入ればいいのかそういうプロトコルがあるが、そのような科学で使われているプロトコルの一覧のようなものはないか？

A：論文に載せるために必要な手続きや満たすべき基準のようなものはおそらく分野ごとにあると思う。多くの場合は、そこまで明示的なものではなく、その分野の専門家の暗黙知のような形で存在しているのではないと思う。数学では、証明であるか否かはかなり厳密に決まっているように思うが、それでも何が証明であるかということも歴史的には変わってきている。私の分野（哲学）では、少なくとも査読する人間の暗黙知にかなり依っているとと思う。

Q：これからの科学というのに理解が必要なのかどうかについてご意見を伺いたい。

A：人間の理解力で科学の範囲というのは実はすごく限定されており、人間の理解という範囲を超えたところにもまだ発掘されていない知識が膨大にある。理解できるところにこだわっていると、逃してしまうものもあると思う。また、理解というのは慣れの問題という気もしており、数学の定理のコンピューターによる証明でも、「ああOK、OK」のような感じで、慣れればそれで理解した気持ちにはなれるのではないか。もちろん危険性もあり、それはまさしく我々が知識から疎外されてしまうことである。

Q：哲学はどうなるのか？

A：私は哲学というのはどちらかというと文学に近いと思っている。哲学に限らず、多くの人文系の人間の仕事は、その時代の人たちが気になっていることなどを何かしっくりする概念や言葉で表現することだろう。これが人工知能にできないかと言われたら、できないことはないと思う。人工知能が短歌や俳句を作ったり、漫画や小説を書いたり、応用の範囲が広がってきている。今のところ、意味の通る長い文章を作れそうな気配は感じないが、分からない。これも慣れかもしれない。

Q：AIが仮説を自動的に生成するとか、それを検証するための実験方法をAIが組むとか、そういうプロトコルが、AIの科学の場合には人間とは全く違った形で取り組むようなことになってくるのか？

A：仮説を立てる作業というのは、おそらく一般的な法則やモデルなどを立てるのだと思う。様々な現象を包摂するような、短く表現できるものの中に圧縮するということだろう。それは帰納論理プログラミングなどで80年代頃に盛んにやられていたが、今はそうでなくなっている。深層学習などでは仮説はおそらくネットワークの構成全体の中にあると考えるべきだろう。もちろん、予測と現実がずれてきたらパラメーターを調整していくという形で仮説検証したりブラッシュアップさせたりということ、ある種の仮説の近似と言えないこともないが、今までの科学での仮説とはかなり違うものだろう。少なくとも、人間が理解して、言葉で伝達したり、何かを書いて伝達したりできるようなものではないと思う。

Q：アストロバイオ構成主義や社会構成主義では、生物ごとや社会ごとに科学ができてくるものだと理解した。その間の共通部分や差分と科学の理解との関係はどのように捉えられているのか？

A：難しい問題。差分をどう理解するかということについて、まず誰が理解するのだという視点が入ってくると思う。この生物から見た世界と、この生物から見た世界との差を考えたときに、どの生物の視点で考えるのか、ということである。私の印象は、この差を考えたときに、何か超越的な視点というものは取れないのではないかということである。共通のことや差について客観的に評価できるような神様みたいな視点というのがあるのかどうか。

Q：普遍主義と社会構成主義で出てくる科学というものにも共通部分があるというような議論は、2つの対立する主義の間であまり議論されてこなかったのか？

- A : もちろん議論はあった。社会構成主義の人たちは「社会や歴史によってこんなに違うだろう。数学でさえこのように変わってきている。こっちの社会の数学とこっちの社会の数学、この時代の数学、これらはこんなに違うのだ」などと言う。それに対して反対する人は「そうは言っても今の数学にはそれほどバラエティはなく、スタンダードな数学に落ち着いてきているではないか。物理学も紆余曲折あったが、やはり今のスタンダードな物理学で多くの物理学者が認めているものがあるではないか。このように収束してゆくのだ」などと言うだろう。これらは両方に理があると思う。科学は現実世界を相手にしているので何でも良いということにはならず、現実世界の我々に与えてくる制約というものに逆らって滅茶苦茶な理論は立てられない。そこで、どこかに折り合いがついていくことになるが、そのときに我々が共通のどこかに収束していけるのは同じ人間だからではないのか、ということ忘れてはいけないだろう。
- C : 今まで行われてきた人間の科学は理解可能なもので、人工知能が予言するものは理解ができないという対立軸のようなものが最初にあるように感じるが、それは問題の立て方にあるのではないかと思う。コンピューターが出てきてモンテカルロで様々な事が予言できたときに非常に画期的だと感じ、使いこなすのにやはり数十年かかっているわけである。深層学習は非線形なイン・アウトの関係の中に非常に興味深い特徴抽出をする能力があり、サイエンスにおける非常に画期的な手法がいま見いだされたところだと言える。それがなぜ興味深い結果を生み出すのかは、これまでの人間の科学による方法で解明していくと数十年かかることは容易に想像でき、この手法がなぜ動くのか、手法の適用範囲がどれぐらい広がるのかということ、これから何十年にも渡って私達は研究していくような対象だという気がする。従って、別々というよりは、知を探索する新しい望遠鏡が得られたというような感覚を持っている。
- A : 研究の現場の方の動きは非常に参考になる。
- Q : 科学の疎外という話は、主に理解というところからの疎外という印象を受けた。その他の科学者のアクティビティーについてはどうか？ 例えば、今は科学者が読みきれないほどの論文が出ているような状況だが、論文という方法で科学の集団的な理解を進めるというような方法はどうなるのか？
- A : 人工知能がどんどん発展してこのままいくとすると、論文という形は多分廃れていくだろうと思う。人間が読むところ、ほとんど趣味の問題になってくるのではないか。人工知能が何かを発見していくが、別にそれを人間に知らせる必要もなく、例えば地震や感染症の予測など、単に予測だけをして人間はそれに従っていけばうまくいくというようなことになれば、もう論文を読む必要はないのかなと思う。そういったことが増えていくと、論文を書いても読む人がいないわけで、人工知能と人工知能のやり取りになってしまえば、論文は要らないということになるだろう。ただし、論文というのは科学者にとっていちばん重要な仕事のひとつで、論文を書くことによって仕事をしたと認められるということもある。従って科学者としての名誉や責任といったものも、変わってくるように思う。ある面ではいいところもあり、今は科学者にあまりに多くの社会的な責任を負わせすぎる風潮がある。新型コロナウイルスに関する専門家会議の中にいた科学者が、自粛する必要は本当にあったのかとか、小学校まで休みにする必要はあったのかとか、そういうことまで科学者の責任にされるというのはよくないと思う。科学者個人としての予測や発言についての責任は、あまり問わないほうがよいのではないかと思う。
- Q : コロナの専門家会議の科学者は、裏にある仮定とモデルをすごくよく理解しての発言だったと思うが、政府の方や我々一般市民の理解のレベルというのには相当な差がある。そういったものが、科学における理解の阻害要因になっているという気がするがどうだろうか？
- A : そうだろう。背景にある細かなことはわかりにくいので、結果的にわかりやすいフレーズだけが伝わる。

ニュアンスなども飛んでしまう。我々科学者じゃない人間にとっては、人間の科学者もすでにブラックボックスなわけである。

Q：その意味では、人工知能による科学というものも、理解はできないとしてもせめて説明できるようなものであってほしいと思う。今般のコロナの話でも、42万人という数字はともかく、人との接触8割減というのがどこから出てくるのかというのが、説明できていないのか、あるいはできないのか、などについてどう考えればよいか？

A：説明と理解はかなり深い関係があるのだろうと思う。何かしらの筋の通った説明は出してほしいが、その説明が説明になっているかどうかというのは受け手の知識レベルなどに依存しており、先述の慣れの問題というのものもある。理解しようという意欲も大きく、そもそも説明しても多くの人は読まない。ここを読んでくれたらいくらかは分かると言っても、やはりあまり読まない。そもそも、我々にはそれほど理解しようという意欲がないのかもしれないという気がしている。自分に都合のいいものはすんなり受け入れ、そうでないものは受け入れたがらないというようなところがある。受け入れる側のプリファレンスや知識レベル、価値観などを踏まえた上で、どういう説明が受け入れられるかを考える必要がある。

C：今回の件では、サイエンスというのは必ず仮説空間の中で生きているので、そのことを全て省略した上で、このままだこうなるという結論を言ったので、混乱が起こったのだと思う。コロナについて分からないことがまだたくさんあるということを最初に断言した上で、最悪のシナリオを考えるとこうなる、それを避けるためにはこうしないといけなく、というような順序での説明であれば、それを受け取った人の反応も違って来たと思う。コロナはこれから何年にも渡って長く共存せざるを得ないことがもう分かり、科学者に問われる課題を突きつけられていると思う。

Q：論文が要らなくなるかどうかは別として、科学的な知見・知識を交換し合う仕組みとして論文というのは「今」のやり方だとは思う。今回のCOVIDでも論文が大量に出ており、十分に皆が追いきれない状況だが、仮に論文を読んで分かったなという気になるのにかかる時間が1秒で済むとしたら、サイエンスの在り方はどのように変わるだろうか？

C：これは人工知能が科学を画期的に発展させるポテンシャルを示しているのではないかと。コンピューターが人間よりも自明に優れているのは、計算能力とメモリーである。これらと人工知能のテクノロジーを使うことによって、ある蓋然性の下での様々な判断がものすごく迅速になると思う。これは個々の研究者ができないことだろう。

A：社会心理学では、科学ということに限定されていないが、集団の中での意思決定や情報共有がどのようになされているかという研究は結構ある。最近よくあるのは、そういったものがインターネットを介したインタラクションになった場合に、どのように変化するかという研究である。対面でのディスカッションと、インターネット越しのディスカッションで、どのような違いが出てくるかというような研究もある。ただし、科学の現場そのものに関して、どのように情報共有がなされたり、新しい発見がなされたりということは、まだそれほど研究がないと思う。どこかの科学のラボに入って観察するという社会学者のような方ももちろんいるが、研究としてはこれからだろう。

## 2.1

### 科学の変容 — アストロバイオ構成主義 —

## 2.2 データと測定数理制約による科学的モデルの発見

鷺尾 隆（大阪大学産業科学研究所）

### 科学的モデル発見の研究：2005年頃までの流れ

データと測定数理制約による科学的モデルの発見について、まず、これまでの研究動向を紹介する。1980年代半ばぐらいから2000年代前半ぐらいまで、実はAI分野で科学的法則式やモデルを発見するためのアルゴリズムの研究が、その研究コミュニティの中で結構活発に取り組まれていた。演繹推論、帰納推論、因果推論、定性推論、発想推論、仮説推論、事例ベース推論など、取り組まれてきた。そういったいろいろな推論メカニズムを駆使しながら、測定データから何らかの物理的関係式を探索する試みが行われていた。その後、ビッグデータが出てきて、データドリブンの機械学習や深層学習のアプローチの研究が盛んになって、この分野の研究は世界的にも、今から15年ぐらい前まででいったん終息してしまった状況にある。

その頃の研究からざっと流れを紹介する。1985年から1990年頃、数個の変量から成る対象のモデルを見つけるような研究があり、Bacon、Fahrenheit、Abacus、Kepler、IDSなど、アルゴリズムに哲学者や科学者の名前をつけたものが多い。ただ、いろいろな式の候補の探索で、やはり組み合わせ爆発が起きる。また、法則式とはどのようなものなのかということをあまり追求しないで、物理法則式に似た形の式を探すというレベルのことだったので、しょせんは実験式を導出していたということである。それから当然、実際のデータはノイズや誤差が入っていて、そういったものに対してあまりロバストでなかった。

一方、単位次元解析という方法があって、それを使うと、物理的な式の形をある程度当たりをつけることができる。そういったものを導入しながら、次元解析の制約に矛盾しない範囲でデータを関係づける式を探索するという研究も、Coper、拡張Abacusなど、若干取り組まれていた。ただ、これは単位次元が不明の非物理系への適用が難しいなど、やはり適用範囲に限られるという問題があった。

それで、より高度な数理制約で効率的に法則を反映する蓋然性の高い関係式やモデルを探索的に発見することが求められた。

### 公理的測定論

私もこの頃、1990年代から研究を始めたが、そのときによりどころとした考え方に公理的測定論がある。これは、物事を測定するとは一体どういうことなのか、徹底して数学的に、その構造や定義を洗い出すことに取り組まれた研究分野で、100年以上歴史がある。

その中にはいろいろな理論があるが、その一つに外延的測定あるいは基本的測定と呼ばれるものがある（図2-2-1）。例えば上皿天秤でおもりの重さを量るということを考えると、天秤が釣り合えば両方の重さが同じだということだが、実際の天秤とおもりという現実の世界（経験的關係システムと呼ばれる）を、数字上の関係（数的關係システムと呼ばれる）に写し取るということをする。現実の世界では、例えばおもりが釣り合っていれば、右側と左側の上皿天秤のおもりを入れ替えても釣り合う。つまり、交換律が成り立つ。また、おもりA×1個とおもりB×2個が釣り合って、おもりB×2個とおもりC×3個が釣り合えば、おもりA×1個とおもりC×3個も釣り合うという推移律が成り立つ。結局、同じであるかどうかは、この交換律と推移律で定義できる。

この現実の世界のおもりの関係、つまり質量の関係について、数字の世界でも同じように交換律や推移律が成り立つように、質量というものに数を割り当てることが測定だと考えられる。言い換えると、測定

## 2.2

データと測定数理制約による科学的モデルの発見

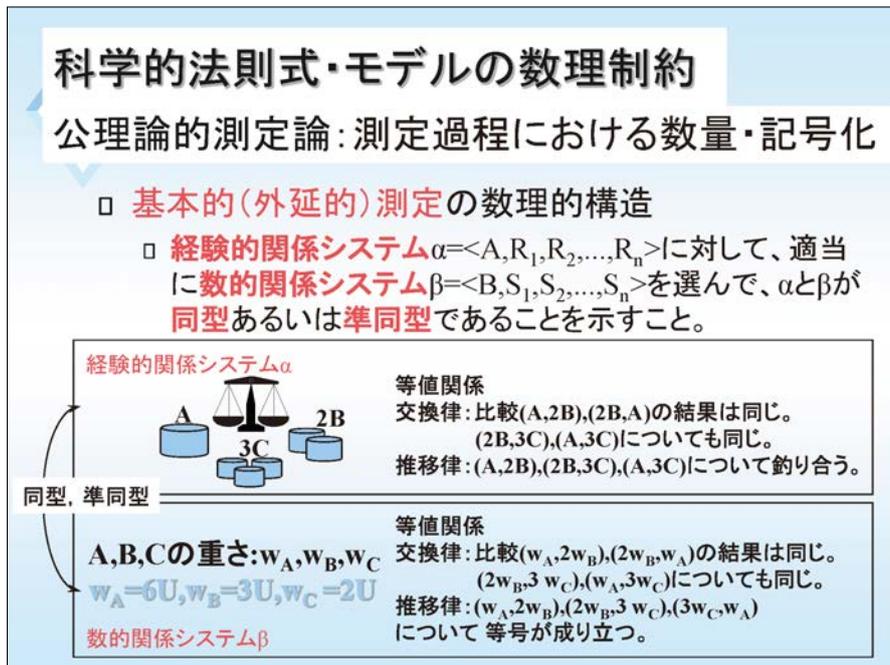


図 2-2-1 公理的測定論：外延的測定

とは、この現実の關係と数字上の關係が同型あるいは準同型になるように写し取ることで、これを公理的測定論では基本的測定あるいは外延的測定と呼んでいる。ここで外延的と言っているのは、おもりやもの長さのように直接測れる物理量を測定することを意味している。なお、温度とかエントロピーというのは直接測れない。他に測った物理量から計算することしかできない。こういったものは、誘導量や内包量と呼ばれる。

それで、公理的測定論には尺度というものが出てくる。尺度とは、経験的關係システムと数的關係システムの対応關係をつけるルールである。つまり、同型・準同型性を保存するルールが尺度である。

### 尺度の種類と許容關係

このような数学的基盤の研究が昔からやられている中に、もう相当古い時代、1946年だが、Stevensによる尺度のタイプの研究がある。図 2-2-2 のように、名義尺度、順序尺度、間隔尺度、比例尺度、絶対尺度などがある。名義尺度は、単純に一致しているか一致していないかだけを定める尺度で、順序尺度は、大小關係だけを定める尺度である。

間隔尺度は、間隔や差を決める尺度で、2つの値の差が同じかどうかということだけを測っている。例えば温度がそうで、摂氏だったら水の氷点から何度離れているかという差で測っている。位置エネルギーもそうで、海拔で測るのか、床から測るのかで、全然値が異なるので、値そのものには意味がない。

それに対して比例尺度は、比の等値性を定める尺度である。相似群と言われ、単位の変換系に従うもので、例えば、お金の値打ち、貨幣、物の伸長、長さ、質量など、これらはゼロとなる絶対原点の意味がある。身長がゼロとか質量がゼロとかお金がゼロというのは、原点に意味があるわけで、原点を勝手に変えられない。先ほど述べた間隔尺度は、水を基準に測るかとか、絶対零度のような何か分子のエネルギーが一番下がったところで測るかとか、人間が適宜原点を選べる。こういう数の割り当て方が違うことから、尺度の違いが出てくるわけである。

基礎理論：尺度のタイプ [Stevens 1946]

尺度のタイプ	基礎的経験操作	数学的許容変換
名義	対象の等値性の決定	置換群 $x' = f(x)$ , ここで $f(x)$ は任意の一対一代入
順序	大小関係の決定	等方群 $x' = f(x)$ , ここで $f(x)$ は任意の単調増加関数
間隔	間隔や差の等値性の決定	一般的線形群 $x' = kx + c$
比例	比の等値性の決定	相似群 $x' = kx$
絶対	絶対的な値の等値性の決定	同一群 $x' = x$

図2-2-2 尺度のタイプ

唯一、絶対尺度は単位がない。物理で言うところの、いわゆる無次元量になる。絶対的な値の等値性を決める。測定するという事は、必ず何かルールを決めてそれで測る。現実の世界を数字の世界に写し取るわけなので、こういったものを定義することが必要になる。

ここで問題は、こういった尺度というものが与えられているときに、尺度を持っている複数の変数の間にはどのような数学的な関係が許されるかということになる。xとyという2つが比例尺度であるとき、xとyの間に  $y = u(x)$  という関係があると仮定する。例えばこの関係がlogだとすると  $y = \log x$  となるが、xは比例尺度なので、ある定数kを掛けて単位を変換することができる。つまり  $y' = \log x' = \log kx = \log k + \log x$  となる。しかし、 $y'$ も比例尺度なので原点は変えられないはずなのに、この式では原点をシフトしないといけないことになって、矛盾が生じる。ということは、2つの比例尺度変数の間にはlogという関係は数学的にあり得ないことが分かる。それに対して、比例尺度xとyの間の関係がべき乗だったら、 $y' = x'^n = (kx)^n = k^n x^n$  となり、スケールが変わるだけで原点のシフトは起きないので許容される。

こういった議論をして、1959年にLuceが2数量間の許容関係、つまり、比例尺度同士、間隔尺度と比例尺度の間、間隔尺度同士といった各ケースでどのような数学的な関係が許されるかを示した。その結果、2変数が比例尺度同士の場合は、変数のべき乗にあるファクターを掛けたものしか関係が許されない。比例尺度と間隔尺度の場合は、間隔尺度変数にlogかべき乗してファクターを掛けて、切片を足したようなものしか許されない。この場合、比例尺度が独立量で間隔尺度が従属量だが、その逆で、間隔尺度が独立量で比例尺度が従属量という場合は、どのような関係も成立し得ないことが証明されている。これは少し考えれば当たり前のことで、間隔尺度は原点の情報を持っておらず、比例尺度は原点とスケールの両方の情報を持っているので、比例尺度の方が情報が多い。ということは、情報の少ない間隔尺度の変量から情報のより多い比例尺度の変量を定義することはできないというわけである。あと、間隔尺度同士の場合は、線形関係のみが許される。

## 尺度に基づく実験操作環境での科学的法則式・モデルの発見

これと並行して、物理学の世界では常識とも言える単位次元解析理論というものがある。古く大正時代から知られている Buckingham  $\Pi$  Theorem や Product Theorem という理論がある。これは比例尺度の変数だけの間の関係だが、複数の物理測定量が与えられて、その間に何か関係があるということが知られているとき、変数をまとめ上げることでいくつかの無次元量に分解できて、無次元量の一般的な関係に還元できるという考え方である。ここで各無次元量は、各比例尺度の量のべき乗の積に何かスケールファクターを掛けたもので必ず表せる。

これは従来、比例尺度の変数だけについて知られていたことだが、私はこれを比例尺度と間隔尺度の間でも成り立つように Buckingham  $\Pi$  Theorem や Product Theorem を拡張した。関係式の中に現れる量に間隔尺度も許される場合を考えて一般化したということである。今からもう 20 年近く前になるが、これを IJCAI 1997 で発表した。

このようなことで、物理的なモデルは、無次元量の間非常に多様な関係と、無次元量を定義する測定量の間関係に、必ず分解できる。それで次に、この定理を使ってこれに当てはまるような測定量の間関係式を列挙して、実際に当てはめたらデータをうまく説明できるような式を探してくるアルゴリズム SDS (Smart Discovery System) を開発した。この探索は、非常に多くの変数の組み合わせや式の組み合わせを調べなくてはならないので、普通にやると組み合わせ爆発を起こしてしまう。そこで、まずは部分的に 2 変数間の関係を調べる。そして、2 変数間の関係でよく合うものが見つかったら、それをさらに 3 変数間の関係と組み上げていって、段階的に全体を説明するような式に組み上げていくようにした。先ほどの Buckingham  $\Pi$  Theorem や Product Theorem を拡張したものを制約に使うと、変数の間のあり得る関係を探していく。最後はいくつかの無次元量になるわけだが、それらの無次元量の間にはさまざまな任意の関係があり得るので、多重線形な関係に絞って探索するようにした。

各変数が間隔尺度や比例尺度であるか、あるいは、無次元量、絶対尺度であるかということだけ与えて、単位の情報も教えずに、あとは実験データを加えることで、対象に関する知識はまったくなくても見つけてくる。実際に計算して評価したところ、理想気体の状態方程式や、運動量保存式、クーロン力、ストークスの式、運動エネルギー保存の式、電子回路の事例など、従来法よりも非常に高速に見つけることができた。しかも、こういった式の制約が強いため、ノイズをたくさん与えても見つけることができた。

## 尺度に基づく所与のデータからの科学的法則式・モデルの発見

ここまで話した研究では、変数の値をいろいろ能動操作して実験できることを前提としている。つまり、実験装置やシミュレーターにおいて、物理的な変数の値を変化させたいろいろな条件で実験を実施することで、さまざまな変数間の関係のデータが用意できるという、ある種、恵まれたケースを考えていた。これに対して、2 年後の IJCAI 1999 で発表したアルゴリズムでは、変数を勝手に動かして実験するという能動操作はできないケースで、所与のデータだけから、先ほどのような制約を満たす式をどこまで導けるかを試みた。

アンケート結果はこういうケースに該当する。例えば、地震のリスクと住宅の購入価格の関係調査、ある地域で地震がどれくらい起こりやすいかの情報があるとき、その地域で住宅の値段がこれくらいだったら、あなたはそれを買うか、買わないか、その程度を答えてもらうというアンケートの結果のデータを考える。地震のリスクは確率・頻度なので比例尺度、住宅の購入価格はお金・価値なのでこれも比例尺度になる。一方、買うか、買わないかの程度は選好度ということで、人間のアンケート結果である。心理学分野でアンケートの尺度構成法というのはよく研究されていて、こういった選好度は間隔尺度や比例尺度で測られている。ここでは、

系列カテゴリー法というアンケート方法を使って、選好度を間隔尺度で測る。そうすると、間隔尺度を目的変数、2つの比例尺度を説明変数とした関係なので、先ほどのような制約式を使って探索できる。

ただ、この場合、人間の嗜好というのは勝手に変えられないので、与えられたデータだけから関係づける式を、先ほどの制約を使ってうまく見つけないといけない。それで、当時何をやったかという、多数のデータを集めておいて、その中で、まず2変数間の関係を調べるのだけれど、そのときに他の変数の値がなるべく似ているデータだけを抽出する、つまり、他の変数は統計的に統制してしまうことをした。そのようにしておいて、2変数間の関係だけを先ほどの制約の式に当てはめることを繰り返す。さらに、3変数についても同様のことを繰り返して、段階的に関係を組み上げていくようにした。

先ほど話したように、比例尺度と間隔尺度の間関係は、log 関係かべき乗関係かの2通りがあり得る。このケースは実際に当てはめてみるしかないのでもやってみると、この社会アンケートについては、log 関係の方がパラメーターの数も少なく、しかも、当てはまりの精度がよかった。これは事後的に当てはまっているかどうかだけの判断だが、当てはまりのよさからいうと、リスクとコストに対する人間の嗜好を表す、ある程度普遍性を持った式が得られたということかもしれない。

### 取り組んできた研究発展の概観

図2-2-3に、この分野で私が取り組んできた研究の発展から主なものを並べた。既に、このリストの中の1番目 (IJCAI97) と3番目 (IJCAI99) について話したことになる。1番目は能動操作ができる実験環境で、3番目は能動操作のできないアンケートのような所与のデータでの科学的なモデル式の発見だった。これらのモデル式は1本の式だが、現実世界はもっと複雑で、連立方程式で表されるモデルが必要になる。それで、連立方程式モデルを発見するようにしたのが、図2-2-3のリストの2番目と4番目の研究である。

2.2 データと測定数理制約による科学的モデルの発見

## 1997～2005年までの研究発展概観

- 実験操作環境での科学的法則式・モデルの発見 [Washio, IJCAI97]
- 実験操作環境での科学的連立方程式モデルの発見 [Washio, AAAI98]
- 所与のデータからの科学的法則式・モデルの発見 [Washio, IJCAI99]
- 所与のデータからの科学的連立方程式モデルの発見 [Washio, ECML01]
- 所与のデータ中で既知法則式やモデルで説明困難な部分の科学的法則式・モデルの発見 [Washio, ICML00]
- 所与のデータからの科学的連立時間微分方程式モデルの発見 [Washio, IJCAI05]

図2-2-3 1997～2005年までの研究発展概観

それから5番目の研究は、ある程度データを支配する法則式やモデル式が分かっているが、一部は分からない、説明困難な部分があるというケースを扱っている。既存のモデル式の部分は分かっているとして、一部説明困難な部分の式を、先ほど述べたような制約を使って、数学的な許容性を満たす範囲内で足りない部分のモデル式や法則式を見つけることを行った。

ここまではすべてスタティックな式のモデリングだったが、6番目の研究は、ダイナミックな時系列データから連立微分方程式、時間微分方程式を同じような原理を使って見つけるというものである。

以下で、2番目の研究、実験操作環境での科学的連立方程式モデルの発見について簡単に紹介する。

### 実験操作環境での科学的連立方程式モデルの発見

既に説明済みの比例尺度と間隔尺度を入れて拡張した Buckingham  $\Pi$  Theorem と Product Theorem は1個の方程式に関するものだったが、これを連立方程式に自然に拡張することができる。元の式が複数本あると、当然、無次元量間の関係式も複数本になる。それぞれの無次元量は、先ほどと同じような比例尺度と間隔尺度の測定量で組み上げられた式に分割できる。このとき、中に含まれる基礎単位（長さ・質量・時間を扱うMKS単位系）の種類数、比例尺度の場合の原点数、基礎原点の種類数などによって、各制約式の中の無次元量の数が決まることが証明できる。それで、連立方程式の場合はやや複雑で、連立方程式系のマトロイドという構造をデータから調べながら、先ほどの拡張した Buckingham  $\Pi$  Theorem や Product Theorem の制約を使って、可能な連立方程式の形を探索する。連立方程式は行や列の代入・消去でいろいろな形になるので、正準系、カノニカルなパターンというものを数学的に定義して、そのカノニカルなモデルを見つけるというアルゴリズムにしている。

性能評価も行い、平行抵抗回路、伝熱過程の関係モデルや、光測定電子回路で方程式の数が14、変数の数も22といった規模のものまで、データを与えてシミュレーションで解くことができた。また、1,000回ぐらいトライアルをやると、ノイズを大量に加えた場合はたまに間違ってしまうことがあったり、方程式の数が多くて変数の数が多くなると誤差が積み上がってうまくいかない割合が徐々に増えたりするが、それでも失敗は18%程度で、多くの場合は成功して正しいモデルに行き着く。ただし、方程式の数や変数の数が多いと探索の組み合わせ爆発が起きて、かなりの計算時間がかかるようになる。

### 課題と展望

私もこういった研究をしていたのは2005年頃までで、世界的にも2000年代の前半くらいまであまり研究されなくなってしまった。その理由として、時期的に少し早過ぎたのかなと今は思う。当時の社会や技術環境として、まず、データが簡単に手に入る時代ではなかった。今のマテリアルズ・インフォマティクスのような実験環境はなかったし、所与のビッグデータもなかった。それから、探索アルゴリズムや機械学習の技術がまだまだ不十分で、組み合わせ計算爆発の問題や、深層学習のように大規模なデータを使って対象を強力にモデル化できる技術がなかったということもある。実際、先ほど示した例ぐらいならば、今ではエンジニアが自分たちで考えてモデルを作れるスケールなわけで、わざわざこういったものを使おうと思わないだろう。社会アンケートのようなところは、若干使い手はあるのかもしれないが。

要するに、科学や産業の現場での具体的なニーズを満たすまで至らなかったということだと思う。複雑で大規模な科学的モデルの発見というところに使えれば、科学者や技術者に必要とされるわけだが、20年前の取り組みではそこまで至らなかったということだ。しかし、近年の計算機環境やビッグデータ化、AI性能向上などを踏まえると、いま述べたような課題を克服して、この分野の新たな発展が期待できると思う。

## 論点を受けて

私からの発表はここまでだが、事前に4つの論点をいただいていたので、それらについて私の考えを述べる。

第1点、第4の科学・データ駆動科学との違いは何か、どのような機能や性能が人間を超えたことになるのかについては、AI自らが重要な答えの「候補」を導出するということだと思う。AI・機械学習によって、物理の関係式が100%決まるわけではない。やはり物理学者・科学者の力が必要で、人間とAIの間のインタラクションによって絞り込むことになる。ただ、これまでのデータ駆動科学と比べて、AIが重要な答えの候補をよりうまく絞り込んで出してくれるようになるだろう。

第2点、科学による知のフロンティアはどこまでどのように広がっていくのか、人間が理解できない科学のような方向にも広がるのかについては、完全に人間が理解できないならば、それは科学として成り立たないと思っている。AIは自ら意思を持つ主体ではないので、人間と意思疎通できない科学は当面発展しないだろう。科学というのはある意味で哲学でもあって、何らかの意思決定がそこに入ってくる。客観性や再現性を重んじるといっても、何を探求するかとか、どこにフォーカスするかというのは、やはり意思決定の問題になる。AIにはそれはできないわけで、そういった点は必ず人間と意思疎通しながらやっていかないと科学としては成り立たない。人間が完全に理解できない科学を、AIが勝手に構築することは、現状あり得ないと思っている。

第3点は、新しいツールは科学技術における人間の活動のどの部分を代替することになるのかについては、先ほど話したように、やはり従来の人間の手に負えない複雑さ、規模、量を持つ対象に関する探究だと思う。複雑度、規模、量において、今の人間の解析能力だけではいくら計算機を使ってももう追えないようなところで、AIがこういう答えがあり得るのではないかと候補を出してきてくれて、それを人間が理解・解釈して選び取っていくといった発展を期待している。

## 【質疑応答】

Q：例えば相対性理論が考え出されたように、現状認められていた理論・常識の枠を超えたモデルを見つけられるようにするには、現状の理論・常識から外れる観測データを与えないといけないことになるか？ その外れるところまで探索範囲を広げるにはどうすればよいか？

A：その通り、外れるデータを与えないといけない。演繹的な思考実験だけで思いついてしまう天才科学者もいるかもしれないが、今の私のアプローチはデータドリブンなので、外れるデータが必要になる。演繹的な推論メカニズムも導入するアプローチも組み合わせると、もっと高度なことができるのかもしれない。

Q：天文学の歴史では、ティコ・ブラーエに至る数百年で天文学のビッグデータが集積され、それらを仔細に分析して、ケプラーが惑星の運動に関する関係式、いわゆるケプラーの法則を見つけた。さらに、その法則を深く考察してニュートンが力学モデルを発見した。これに照らすと、現在のAIが発見可能なのはケプラーの法則の発見までということになるか？

A：その通り、ケプラーの法則までだと思う。いろいろな実験設定をすれば、万有引力の法則もこれで発見できるはずだが、実験設定をすること自体が、何らかの普遍的な法則を意識して行われるので、そこが今のAIにはできない。

Q：機械学習などデータドリブンなアプローチで得られるのは経験則であって、それを科学的法則式に格上げするには、他にいろいろな要件があるという検討を、鷲尾先生はされていたと思う。今日の話はその要件の1つである数学的許容性に絞ったものだったが、他の要件も含めて、経験則を科学的法則式に格上げするための手段は可能性が出てきているか？

A : 科学的法則式と言えるためには、いろいろな条件が検証される必要があり、図2-2-4のような条件があると考えた。今日の話は、このうち数学的許容性に絞ったものだった。しかし、それ以外にも必要な条件がある。その1つは客観性で、これは少なくとも物理的に検証可能であるためには、構成される変数をすべて観測できて、観測データだけから式を計算できなくてはならないというものである。それから普遍性というのがあって、これは非常に広い範囲で成立しなくてはならないというものである。もちろん、同じ条件でいつも同じ結果が得られるという再現性も必要である。あと、健全性というのは、当然、観測データに対して矛盾があってはならないというものである。さらに簡潔性、これはアインシュタインが自然は単純だと言ったが、極力単純な式で、少ない変数で表せている方がいいというものである。このようにいろいろな基準がある。

それで、やはり本当に受け入れられている式・モデルというのは、これら一連の厳しい検証をくぐり抜けてきたものだと思う。今日は数学的許容性の話ばかりしたが、本当はこれら全部をチェックした科学的なモデルを出せるようにするというが目指すところである。ただ、そのためにはさまざまな条件変更や膨大なデータ収集・検証など、まだまだ大きな課題がある。

ただ、こういったものを、候補を出してくれるAIツールとして使う分には、一部の条件に対応したもので役立つ。その候補から先は科学者が検証していけばよい。数学的な許容性だけでもかなり候補を絞り込める。

Q : 機械学習によって得られた経験則について、その外側で図2-2-4のような条件を検証する仕組みを作っていくという流れになっていくか？



## 法則式の条件

式eの対象領域Deにおいて

- 客観性 (式の構成変数を全て観測可能)
- 普遍性 (式が広範囲で成立)
- 再現性 (同一条件で同じ結果)
- 健全性 (観測に関し無矛盾)
- 簡潔性 (極力少数の変数より成る)
- **数学的許容性 (式がS及びAに従うこと)**

図2-2-4 法則式の条件

- A : それもあると思う。また、科学や産業の現場で求められるシナリオとして、ゼロから完全スクラッチでデータを表すモデル式を作るというケースだけでなく、ある程度は既に分かっているのだが、分からない部分の一部残っていて、そこをAIで何とかしたいというケースが考えられる。図2-2-3のリストの5番目の研究は、この後者のケースを支援しようというもので、現場のニーズとしてはむしろこちらが多いのではないかと。非常に重要なテーマになるだろうと思っている。
- Q : 法則式の発見にAIが関わっていくと、図2-2-4の簡潔性の扱いは変わっていくか？  
例えば、非常に高次元なモデル、変数が非常に多くて膨大な式を人間は理解・解釈できないが、AIは発見できるのではないかと？
- A : いわゆるブラックボックス問題のようなことは起きてくると思う。しかし、そうすると、先ほど述べたような人間との意思疎通の問題が生じる。それで、ブラックボックスのままでは科学として止まってしまうと思っているので、ブラックボックスを人間が解釈できるように支援することが必要になる。説明可能AI (XAI) の研究が進んでいるが、こちらの分野にも入れてさらに発展させていくことも必要になってくると思う。
- Q : 量子力学でいう確率振幅みたいなもの、測れないのだが、測定結果は個点のデータとして確率、統計分布として得られるようなタイプの変数はどのように扱われるか？
- A : これまで話してきたのはデータミニスティックなモデルなので、量子力学のようなモデルはスコープに入っていない。機械学習も確率密度関数をデータから学習するようなものも出てきている。このとき、確率の間でBuckingham  $\Pi$  Theoremのような数学的許容性は一体どのようなものになるのか。しかも、純粋に統計・確率論での話ではなく、測定量として確率密度関数が満たすべき制約のようなことまで遡って考えると、まったく新しい発展がありそうで、世界中でまだ誰もやっていない、非常に面白いテーマだと思う。
- Q : 関係式を推定するとき、実際にはデータにノイズが含まれるが、線形より非線形関係はノイズに敏感だと思う。非線形でノイズを含んでも、うまく関係式を推定できるか？
- A : 深層学習のような最新の機械学習は、測定ノイズやプロセスノイズをうまく扱う理論やテクニックが非常に発展しているので、カオスのような極端な例を除けば、ある程度対処できているように思う。
- Q : 強化学習を考えると、AIは環境から学ぶことができる。それで、ある条件下で生き残れという適切な指示を与えると、その与えられた環境の中で生き残るための戦略をある確率で獲得することになる。ある意味、意思を持っているように見えるかもしれない。
- A : ただし、その強化学習のための効用関数自体は、今のところ人間が与えなくてはならない。AIが意思で選び取るわけではない。

## 2.2

データと測定数理制約による科学的モデルの発見

## 2.3 機械学習と物理

橋本 幸士 (大阪大学)

私は素粒子物理学の研究者で、素粒子物理学はある意味で要素還元主義に徹するところである。今日はそこから見た科学と機械学習の関係についていくつか話題提供したい。

「ディープラーニングと物理学」というタイトルで、私は数年前から研究会を開催しており、特に素粒子物理学に限らず物理の広い分野の若手の研究者が大変盛り上がっているというのを目の当たりにしている。今はオンラインで隔週のセミナーを開催しており、700名ほどが登録しており、毎回100名以上が参加している。新分野を作ろうとか、一旗揚げてやろうとかいう雰囲気を感じられる。このような動きは日本だけのことではなく、米国の友人が立ち上げた Physics and Machine Learning というオンラインシリーズもあり、これはヨーロッパとアメリカの1000人以上の研究者が登録しており、盛り上がりを見せている。機械学習の分野が非常に発展しているというのを聞きつけて、物理学者もなんとかそれをやってやろうというような雰囲気は全世界にある。日本物理学会で「機械学習と物理」というシンポジウムを開催したところ、これも大盛況であった。また、京都大学基礎物理学研究所で開催した国際会議では、甘利先生を囲んで統計やニューラルネットワークの話など議論に花が咲いた。2017年頃から、世界各地で同様のものが行われているようで、機械学習が元になっているグループもあれば、物性物理学が元になっているものもあるし、私のように素粒子物理学や弦理論などから生まれているグループもある。それらが今、ようやく横のつながりができつつある状況だろう。

1-1

「機械学習と物理」のトピック

量子多体系の基底状態の波動関数  
[Carleo Troyer Science355 (2017)602]

NNの時空解釈と物理系への応用  
[Tanaka Tomiya Sugishita KH Phys.Rev.D98(2018)046019]

自己学習モンテカルロ、有効理論構築  
[Liu Qi Meng Fu Phys.Rev.B95(2017)]

分子動力学法  
[Behler J.Chem.Phys.145(2016)170901]

相図の予想、相転移の検出 [Ohtsuki Ohtsuki JPSJ85 (2016)123706]  
[Tanaka Tomiya JPSJ86(2017)963001]

物理量の関係間の写像  
[Fujimoto Fukushima Murase Phys. Rev.D98(2018)]

解の探索と分類 [He Phys.Lett.B774(2017)564]  
[Carifio Halverson Krioukov Nelson JHEP09(2017)157]

図 2-3-1 「機械学習と物理」のトピック

2.3

機械学習と物理

22

CRDS 国立研究開発法人科学技術振興機構 研究開発戦略センター

CRDS-FY2020-WR-13

手法としての機械学習という観点から、物理のどういう分野にどういう関係があるのかという議論、すなわち、超弦理論に重きをおいたもの、数学との関係を議論するもの、機械学習に物理の考え方を導入すると機械学習がどう効率化されるか、など様々な観点のものが開催され、研究者のつながりができつつある。個々の研究者は、個々の物理学の特定の問題に興味があるので、その問題に「ひょっとしたらこれが役に立つのではないか」というようなものを見聞する場としてこれらの研究会は機能しており、ツールとしての国際学会というようなものに発展する可能性は十分にある。

物理学における機械学習といったとき、画像認識やデータからのシグナル検出などでは、深層学習のブームが到来するずっと以前から物理学に取り入れられてきている。例えば、最近のものでは、すばる望遠鏡での遠方銀河や超新星の観察がある。画像認識をつかって、大量の超新星が発見されるという研究があった。また、素粒子実験では得られた多量のデータの中から、それをジェットと認定したり、ジェットの中からどういう素粒子が出てきているのかの同定など、標準的に画像認識の機械学習が使われている。

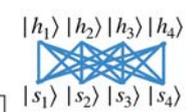
今回私が取り上げたいのは実験ではなく、理論のほうでも機械学習との関係でこの3年ほどの間で起こった発展があるので、いくつか例示したい(図2-3-1)。まず、ハミルトニアンが与えられたときに波動関数を求めるという問題が物性物理の基礎的問題としてあり、そこにニューラルネットワーク型の波動関数を仮定することで非常に大きな発展が見られている。

1-4

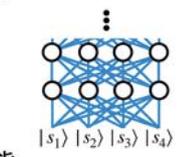
ニューラルネットワーク状態の登場

- ボルツマンマシン状態  
[Carleo Troyer `17], [Nomura, Darmawan, Yamaji, Imada `17], ..

$$\psi(s_1, \dots, s_N) = \sum_{h_A} \exp \left[ \sum_a a_a s_a + \sum_A b_A h_A + \sum_{a,A} J_{aA} s_a h_A \right]$$


- 深層ボルツマンマシン状態  
[Carleo, Nomura, Imada `18], ..

$$|\psi\rangle = \lim_{\tau \rightarrow \infty} e^{-\tau H} |any\rangle = e^{-\Delta\tau H} e^{-\Delta\tau H} \dots |any\rangle$$


- フィードフォワードニューラルネットワーク状態  
[Saito `18], ..

$$\psi(s_1, \dots, s_N) = \sum_i f_i \sigma \left( \sum_j W_{ij} s_j + b_i \right)$$

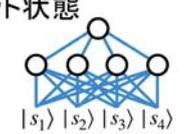


図2-3-2 ニューラルネットワーク状態の登場

ニューラルネットワークは脳のニューロンの相互作用を模倣するモデルだが、記憶のメカニズムとして知られるホップフィールドモデルは、物理系でいうとスピンモデルである。従って、ニューラルネットワークは物理系が主体にあって、それをニューラルネットワークに移したという歴史的経緯がある。スピンというのは空間の中に並んでいて時間発展するものなので、そのようなニューラルネットワークは元来物理学との親和性が高い。そこで、ニューラルネットワークの時空解釈という研究が進んでおり、それを物理系に応用する研究がある。

量子多体計算を行うときに、様々なサンプルをモンテカルロで生成し、そのサンプルを使って期待値などを計算する手法がスタンダードだが、この生成効率は従来非常に悪い。それをうまく調整する方法として様々なモンテカルロが考案されてきたわけだが、途中に自己学習型のニューラルネットワークを入れ込むことで、よいハミルトニアンを提案してもらい、生成効率を良くする手法が近年進んでいる。このようなことが上手く行き始めると、物理としていちばん重要な部分となる、複雑な理論の有効モデルを少数の自由度で構築することが可能となる。この少数自由度とはなにかというのを選ぶところまでニューラルネットワークにやらせるということもあり得る。

分子動力学計算（MD）は、原子の電子状態などを規定せずに、勝手な有効ポテンシャルを導入して計算するが、これは職人芸である。これを用いて多体計算をするのだが、こういった部分を機械学習に置き換えるということがある。

波動関数は画像データでもあるのでそこからどんな相にあるのかをニューラルネットワークに判別させたり、相転移をニューラルネットワークの重みのヒートマップから検出するというような研究もある。また、高密度核物質の状態方程式が与えられたら原理的には中性子星の半径や質量を予言できるわけだが、その途中の計算はとても大変であるため、そこをニューラルネットワークに置き換えてしまうことで、実際の観測で中性子星の半径が分かったときに、もとの状態方程式がこういうものだったというように逆解きする研究もある。

リストのはじめの2つについて簡単に紹介したい。1つめの量子多体系の基底状態の問題は次のとおりだ（図2-3-2）。物性物理学の基礎的な問題は、基底状態を求めるという問題である。ハミルトニアンと呼ばれる、その物質の性質の情報が入っているシュレディンガー方程式を解くことができれば、そのハミルトニアンで規定される物質の基底状態が分かる。N個の電子の波動関数というのは、スピンのN個並んでいるものと考えたとして、スピンのそれぞれ上か下かの状態を数え上げれば2のN乗通りの可能性があり、そのそれぞれの係数を求める必要がある。この非常に大きい種類の係数を全部決めるのは困難なので、仮定関数が試されてきた。例えば、行列積状態というのは、並進不変性のようなものを課して、隣同士のみがつながっているという条件を課するという、「物理系としてはこういうものになってないといけない」というある種の帰納バイアスを入れるわけである。

2017年のカーレオとトロイヤの論文では、これにニューラルネットワークによる関数を使っている。ニューラルネットワークには万能近似定理があり、十分大きなニューラルネットワークを使うと、どんな関数でも近似できる。しかも、機械学習はある種の最適化なので、エネルギーを最小化するための誤差関数を設定すると、その向きにこのネットワークを調整してくれることになる。このボルツマンマシン状態というのを試行状態として使えば、非常に良い波動関数を求めることができるのではないか、という提案であった。論文で書かれている例では、2次元の反強磁性ハイゼンベルグ模型について、隠れ層の数を多くしていくと、従来の仮定関数よりもエネルギーが下がるような波動関数が発見された。非常に簡単なテクニックを使うことで、従来型の仮定関数の限界を突破できるということで、大きな注目を浴びた。この方法は、派生的なものが数多く出ており、すでに引用数が700を超えている。

素粒子物理の最後の宿題は量子重力理論である。これは、重力理論が、素粒子理論の数学的な体系である場の量子論の重要な要素であるくりこみ理論というものの適用対象外になってしまうということが宿題となっている。重力理論を場の量子論的に扱えるようにするというのが、素粒子物理の最後の課題である。

機械学習においては、ニューラルネットワークに突っ込んだら何でもできるというわけではなく、帰納バイアスを従前に考える必要がある。ここで帰納バイアスというのは、例えば学習のゴールを何に設定するか、どのようなアルゴリズムを使うか、どのようなニューラルネットワークのアーキテクチャを用いるか、などである。

これを人間がやっけてあげないと機械学習は何もうまくいかない。

素粒子物理学における帰納バイアスは、明らかに場の量子論である。場の量子論はこの3～40年非常にうまくいっている。その場の量子論の中に重力というが入っていないから、これはどうするんだとなるわけである。従って、素粒子物理学で、場の量子論を超えるということを考えなければならないとき、機械学習で様々な帰納バイアスが試されているという面を採り入れて重力理論の可能性を考える、というのは重要な意味があると個人的に思っている。そのためには、まずは従来の場の量子論が機械学習の中にどのようにフィットするかというのを考えないといけない。重力理論では、時空をネットワークとして考えるということが古くから行われてきたので、そことニューラルネットワークには親和性がある（図2-3-3）。

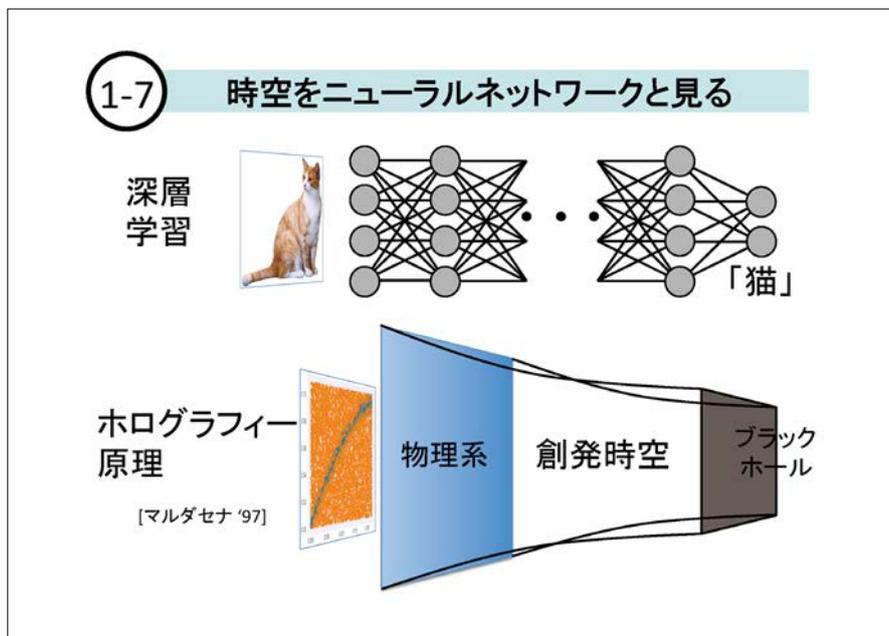


図2-3-3 時空をニューラルネットワークと見る

私は、この20年ぐらい量子重力理論の分野を牽引してきたホログラフィー原理と深層学習とを対応付けるという研究をした。ホログラフィー原理というのは、量子重力理論を場の量子論の帰納バイアスに落とし込むというような位置づけのもので、それを超えて量子重力が存在するかどうかというのが、今業界の一番の重要な問題になっている。図2-3-4のようにニューラルネットワークと時空の関係をつけることができ、そのようなフレームワークで学習を行うと、ホログラフィー原理に適切な重力模型を機械学習によって見つけることができると分かった。分野は違うが、ホログラフィー原理とボルツマンマシンによる波動関数の話は、実はほとんど同じものになっていて互いに関係している。これがもう少し幅広く関係してくるような状況になると、大変面白くなる。

物理学の革新と機械学習の関係をもう少し議論したい。まず物理学の問題には図2-3-5のような種類があり、例えば関数 $y=f(x)$ というのがあったとき、 $y$ が分からないという順問題に加えて、 $x$ が分からない問題と $f$ そのものが分からないという2つの逆問題がある。この $f$ そのものが分からない、つまりシステムが分からないような問題を解くことが、これまで物理学に革新を起こしてきた。このシステム決定問題は、 $x$ と $y$ が与えられ

1-7 時空をニューラルネットワークと見る



ホログラフィー	深層学習
創発空間方向	層の深さ
重力場	ネットワークの重み
量子系の実験データ	入力データ
ブラックホール条件	出力データ
重力相互作用	活性化関数

[田中、富谷、杉下、橋本 '18] 28

図 2-3-4 ホログラフィー原理と深層学習の対応関係

2-1 物理学の「問題」には種類がある

関数  $y = f(x)$  についての問題とは？

	$f$	$x$	$y$
順問題	知	知	未知
逆問題(1) 初期値問題	知	未知	知
逆問題(2) システム決定問題	未知	知	知

逆問題の一般的性質

- ・ 直接測れない
- ・ 結果から原因を知る
- ・ 物理定数を決定する
- ・ 物理法則の同定

逆問題(1)  
逆問題(2)

32

図 2-3-5 物理学の「問題」には種類がある

て、多数のデータの間関係性を決める問題であり、機械学習ではそれを形式的に解くことになる。これは、第一義的には機械学習が物理に役に立つのではないかと思わせるところである。

科学の流れというのを図 2-3-6 のように整理してみたい。まず実験とモデルがある。モデルというのは理論のことである。実験と理論は車の両輪のように組み合わせられて科学が進んでいくとよく言われているが、その

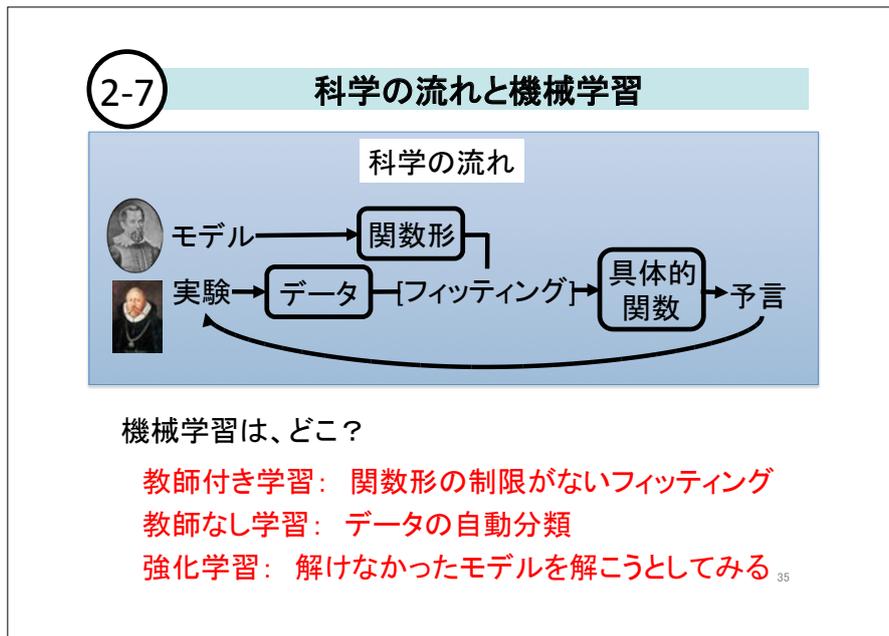


図 2-3-6 科学の流れと機械学習

状況は今でも変わっていない。実験からデータが取られ、そのデータを見てモデルを組む理論屋がいる。このモデルは少ないパラメーターで表された関数のことである。この関数を用いて大量のデータをフィットするわけである。フィットによって決まったパラメーターを使った関数に、今度は予言をさせる。その予言が実際に正しいかどうかというのを、実験や観測をしてデータを取ることで確かめる。この繰り返しで科学は流れている、という考え方が自然であろう。ティコ・ブラーエとケプラーの話はこういう位置づけである。では、機械学習はどこかという、教師あり学習はフィッティング関数のところ、教師なし学習は実験データを人間が整理しやすいように分類するところだと考えられる。また、強化学習はモデルを提案する前の段階で、人間が従来解けなかったモデルの解き方を示唆してくれるもの、と考えることができる。

これらの3つはどれも人間が理論家として従来やってきたことではない。それを手助けすることではあっても、理論家がすることではない。したがって、ここは分けて考えたほうがよく、機械学習が物理学にどういう意義があるのかを考えるときには、モデルを作るというのは科学においてどういう意味があるのかということをよく考えたほうが良いと思う。

物理学において機械学習とは何か、と聞かれた時に、私は「第3の実験」ではないかと思っている。まず、第2の実験というのは、仮想世界での実験（シミュレーション）であり、スパコンの台頭など20年くらい前からの大きな流れとなっている。実験と理論が両輪となって科学を動かしていく中で機械学習はどこにあたるのかと考えたときにいろいろな考え方ができるが、第3の実験として位置づけてはどうかというのが私の作業仮説である。

この第3の実験としての機械学習が可能にすることは、先程紹介した7つのトピックにつながるが、それらの中での機械学習の使われ方を総合するに、(1) 職人芸の置換、(2) 解や理論への示唆、(3) 情報科学との融合・数理的な分野、という3つの観点があり得る。(1)については、波動関数やMDなどそれに非常に長けた人が直感でというのがこれまで大きな割合を占めてきたが、そこがボトルネックになっている場合もあるため、今後職人芸が置換されていくかもしれない。(2)というのは、「こういう解もありますよ」という

ところで手助けしてくれるようなことである。その解釈までは与えてくれないが、人間の側では「そんな解はない。なぜなら場の量子論でそれは許されないから。いや待てよ、それを許すためにはこういうふうな変更をしないといけないぞ」というような協働があり得るだろう。

(3)に関連して、物理学といってもいろいろあり、実際は中では解くべき問題がエネルギースケールによって分かれている。例えば、素粒子論は還元主義的なところに重きをおいており、新理論形式の探索や経路積分の具体的な実行などが至上命題になっている。一旦その素粒子物理学を認めたら、その上には原子核物理学というのが乗っている。そこでは核種や陽子・中性子などを基本とする経験的な理論体系が積み上がっているのである。そこで解くべきものは、単なる経路積分ではなく、量子の少数多体系を解くことが至上命題になる。こういったものをつなぐ縦串のようなものとして、近年ノーベル賞も出たトポロジーが挙げられる。素粒子や原子核ではトポロジーは古くから使われており、そういった考え方の他分野への応用もされてきた。現在ではトポジカル物質の観点を素粒子物理学に輸入するというも行われている。量子情報のエンタングルメントというキーワードもこの縦串だろう。このような縦串として、機械学習というのが入り得ると思う。

機械学習は目的に応じて多彩な方法が開発されてきているが、それぞれがどう物理学的の方法に対応するのかということをはっきりマッピングすることが、コミュニティを作る上でまず最初に必要ではないかと考えている。図2-3-7は多彩な問題と手法をツールという観点で線でつないでみたものである。ニューラルネットワークがもともと物理から来ているというような経緯や、線形回帰自体も、ガウスやルジャンドル、ラプラスといった人たちが天体の観測データのフィッティング手法を開発し、フィッシャーの最尤法などの開発から今の機械学習に至っていることを踏まえ、時代を振り返って見て、こういったマッピングをしていくのが大事なのではないかと思う。

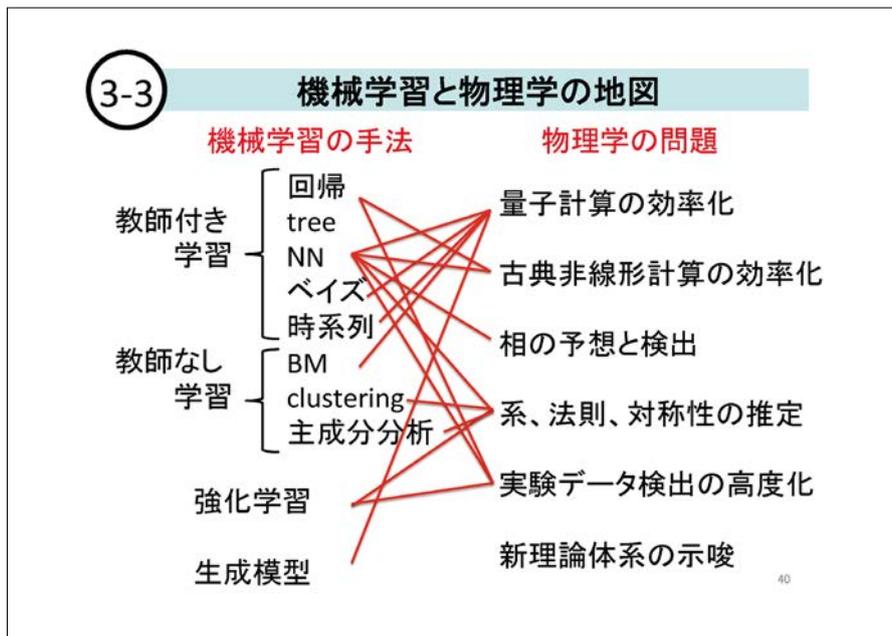


図2-3-7 機械学習と物理学の地図

このような実験的なことに機械学習を適用するだけで、理論では無理なのかということに対して、私の個人的な意見は、少なくとも素粒子物理に関しては帰納バイアスがもう40年くらい場の量子論というので設定されてきていて、それを機械学習によって超えるものをというのは、あまり可能性はないかもしれない。もちろん問題の設定の仕方に依存するので、大規模な波がこれから起こるかもしれないが、人間がモデルを理解して新しい物理概念に到達するということまで、現在の機械学習が役に立つというような感じはしない。ただし、素粒子物理ではオッカムの剃刀をひっくり返されたら困るというような観点も入っている。例えば神経科学などの分野ではオッカムの剃刀を捨てるような方向を真剣に発展させるのは面白いとは思う。しかし、それが物理にどう影響があるかは判断できない。

人間が理解できない科学のような方向性には私は否定的である。人間が理解するからこそ科学であって、人間が理解しない科学というものは科学ではなく「自然」だと考える。私が「第3の実験」と呼んでいるのは、それは人間がこれから理解すべき対象すなわち「自然」の一部であるという意味で呼んでいる。新しいツールとしての機械学習ということでは、職人芸の部分と、人間が解ききれなかったモデルの解法の示唆の2つは今の所大きい。それ以外の部分はまだ手探りで、機械学習の様々なツールと物理の問題との関係性をもう少し詰めていき、個々の問題について研究する必要があると思う。

### 【質疑応答】

- Q：波動関数を求めるのにニューラルネットワークを使うという話は、教師データはふんだんに得られるものなのか？
- A：教師あり学習の流れで説明してしまったので誤解を生んだ。まず、入力としてスピン配位のようなものを入れたときに、波動関数の値みたいなものを出力するニューラルネットワークを用意する。そしてその出力を使ったエネルギーの値を誤差関数として選び、それをどんどん小さくするようにニューラルネットワークのパラメーターをチューンする。従って、教師データを生成しておく必要はなく、自分が与えたハミルトニアンから、エネルギーをどんどん下げなさいという最適化をするということになる。
- Q：第3の実験としての機械学習では、モデルがこうなるかというような示唆を与える役割があると言っていたが、帰納バイアスについても教えてくれるという意味か？
- A：帰納バイアスがここなのではないか、ということまで行ったような話はまだ聞いていない。むしろ、超弦理論ではある方程式を解くというときに、強化学習による探索で解いたときに、学習で作られたアルゴリズムを見てみると、実際に超対称性みたいなものを課して試してみなさいとかというようなアルゴリズムが自然に学習されている。今まで人間が仮定をおいて解くというように経験的に積み上げてきた帰納バイアスが、確かにある範囲ではリーズナブルだったというような学習結果は出ている。
- Q：「こういうものを予測したいときに、こういう説明変数を入れるとより精度は高いけれど、この変数はあまり効いていない」というようなことを物理学者に示唆するなど、物理学者がいろいろ中身がどうなっているか知りたいときに機械学習を使ってみると何かアイデアが湧いてくるというような感じで使われるイメージか？
- A：例えば、系の対称性などを知りたいというときに、その時系列のデータだけがあって、簡単な古典力学を考えて、データからハミルトニアンを推定する問題で既知のハミルトニアンが再現できるかどうか、のような研究は最近たくさん出てきた。しかし、そこでの結論は、今のところ「この帰納バイアスは実は外しておいたほうがよかった」というような最先端の物理学のところまでくるにはまだ距離がある。
- Q：素粒子物理では、トリプルブラインドなど、科学者の持っているバイアスをできるだけ外そうという仕組み

みがあると聞いたが、そうなのか？

A：素粒子実験ではスタンダードに行われている。ブラックホールの撮像でも、4つほどのチームが別々の手法でブラックホールの絵を出してみても、それで比べてみるということをしている。ヒッグス粒子の発見でも、少なくとも2つのチームが独立にデータを分析して同じ結果を出すかどうか最後に突き合わせるということが行われていた。一方で、理論の研究ではそこまで真剣にすることは少ない。例えば、私も日々の研究でゼロになるはずの計算でゼロにならなかったら何かおかしいとは思いますが、ゼロになったらもうそれでいいとする。慎重にするときには、共同研究者に独立に計算してもらい確認することはあるが共同研究者にも同じバイアスがあるから、絶対に回避できない問題だと思う。

Q：物理というのは人間が理解しないといけないということを言っていたが、この人間とは個々の物理学者のことを指すのか？ それとも物理学のコミュニティ全体のようなものを指しているのか？

A：コミュニティのことである。相対論のときもアインシュタインが発表したときには世界で10人ほどしか理解していなかったと言われている。今は、相対論はおそらく数十万人の規模で世界に理解者がいる。理解者は全人類である必要はなく、コミュニティであればいい。コミュニティの中で共通の理解に達した物理の基盤として使われていくことが必要である。

Q：そのコミュニティというのは人の集合でもあるし、私達が歴史的に持っている知識の積み上げを含めたコミュニティでもある。そうすると、例えば数学とか、観測のための機械とか、いろいろなものがあって、その総体としての集合知性なので、今後その中に機械学習というのが入ってくるかもしれないという理解で良いか？

A：そのコミュニティの中に、「機械学習くん」のようなものがたくさん入ってきて、それで一緒に共同作業するという将来は容易に頭の中に浮かぶ。それは人間から見ると、自然の一部のように個人的には感じる。連続的につながっているのかどうかは分からないが、一緒にやっていく相手のような感じだろう。

Q：システム決定問題とそこから逆問題として法則式を発見するようなところというのは、どのぐらい可能性があるか？

A：物理のシステムは何で決定されるかと言うと、ハミルトニアンである。あるいはラグランジアン。それが1個決まれば、それを量子多体的に使うか、素粒子的に使うか、いろいろなやり方があるが、ハミルトニアンが1個決まれば、それはシステムが決定されたことを意味する。そういう狭い意味でのシステム決定問題だとすると、今の機械学習のやり方で、有効ハミルトニアンを推定するなど、システムを決定することができつつある。しかし、難しいのは「本当にハミルトニアンから出発するだけでいいんですか」というような物理の問題を考えた時である。例えば、場の量子論をさらに拡大しないと重力の問題が取り込めないと思うのか、それは場の量子論の範疇で絶対できるはずだと思うのか、という信条の違いでもある。ハミルトニアンによる定式化を超えたところで物理をやらないといけないと思う段になると、今知られているような機械学習を物理に応用する方法だけでは、その上のシステムを決定させてくれるということにはなっていない。違う問題になっているように思う。

Q：理解というのはどういう意味で使われているのか？ 前の回答のように、システムの上のレベルまで見つけてくることを理解と呼んでいるのか？

A：波動関数を誰かにもらったときに、「これは波動関数ですよ」と言われてそれで私が理解できるというのはどういう意味かということ、その係数や重みなどの波動関数のデータを見たときに、それがもともとの物理のハミルトニアンならそのハミルトニアンの性質を違うものについてまで自分が拡大して予想できる、というようなときに理解という。例えば並進対称だと見て取れた場合には、「もともとの系も並進対称性

があったに違いない。そうすると、それに付随する保存量があるはずで、その保存量を変えたときの波動関数はどうなるのかな」というのが理解である。従って、波動関数がニューラルネットワーク型の仮定関数でできたというだけでは、理解したということにはならない。「なぜニューラルネットワーク型の波動関数を使うと基底状態のエネルギーがさらに下がったのか」というと、従来知られていた厳密な求め方とニューラルネットワークのアーキテクチャーの形にこういう類似性があるからだ」といったことが理解には必要となる。そのようなことをやっている研究もあり、経路積分で基底状態を表現する方法はまさにディープボルツマンマシンの形をしている。これらが似ているとか似ていないとかいう議論をすることで、解釈可能とか理解できるとかいう話になるのだと思う。

Q：ニューラルネットワークとボルツマンマシンの類似性のようなところは、人間でないといけない部分ということか？

A：そうである。そもそも、ニューラルネットワークのアーキテクチャーを人間が決めているものなので、そこに人間の手が必ず入っている。

Q：素粒子物理の加速器実験をはたからみていると、従来の統計的な検定を使っているような気がしている。別の分野の計測関係の問題に取り組んでいるのだが、そこでは、実験装置の癖を学習することで、統計検定だけに頼らずにその癖を使って少数のデータから精度よく実験結果が良いかどうかを検証できるということが分かってきている。そこまで考えると、今の実験科学の方法論自体に、根本的に何か新しい考え方を導入できるのではないかと思うが、どのように考えるか？

A：それはすばらしい話である。私が知っている素粒子実験に機械学習を用いるという話は、画像認識の考えからデータの抽出を行う話や、従来の統計で5シグマないと発見と呼ばないとかそういう話であり、最先端のところは私も知らない。加速器のコミュニティーでの機械学習の勉強会というのが欧米で行われており、そこで議論はされているのかもしれない。

Q：オッカムの剃刀もある種の帰納バイアスだろう。素粒子理論では、これは帰納バイアスだと明示的に分かっているものと、そうではなく職人芸のように暗黙知的に皆がやっているというものがあるように思うが、コミュニティではどれぐらい認識されているのか？

A：素粒子理論の中ではかなりはっきりしている。コミュニティのほとんどの人間が信用している帰納バイアスというのは、場の量子論の枠組みである。これは摂動論でも経路積分でも、非常にはっきりとその無矛盾性が見えているので、信用に値する。実際に素粒子の標準模型のラグランジアンはその範疇で書かれている。今までの素粒子理論の歴史を見ると、例えば1950年代には100種類ぐらいあったハドロンを6種類のクォークに落とし込めたことがある。このようにオッカムの剃刀が実際に働いて科学を進展させたということをずっと信用しているので、ここから先に行くにもこの考えを基本にして、例えば大統一理論だったらパラメーターの数をこれだけ減らせるとか、こういう新粒子を予言できるとか、だからこれは信用に値する理論である、というような観点になっている。より少ないパラメーターで検出されるはずだというオッカムの剃刀が帰納バイアスになっている。もう一つは、例えば今の物質粒子にはフェルミオンしかないので、ヒッグス粒子はボゾンだが、それすらフェルミオンではないといけないうような考え方もあり得るので、それに従った学派がある。それを使うと、こういう粒子が予言され、こういう超対称性があるはずだと予言がされる。つまり、今までの帰納バイアスに基づいているいろいろな流派があり、それらが予言する多数の粒子を実験で一個一個確認していくというように、素粒子物理学は進んできている。従って、大きな帰納バイアスとしては場の量子論があり、そこからはみ出る可能性を考えているのは重力をやっている人ぐらいで、それ以外の普通の素粒子物理学は、その中のさらに細かい帰納バ

イアスから出てくる様々な理論の信条のようなものを戦わせていくという構図になっている。これらは全てオッカムの剃刀には基づいているわけである。

Q：モデルと言ったときには、そのバイアス込みで、少数のパラメーターの関数で書かれているということか？

A：そうである。原子の時代から現在の素粒子物理学までの百数十年は多分それずっと来ている。その前の、量子よりも前の時代の話になると、例えばニュートンが決定論的な力学の体系を世界に広めてしまった後で、熱力学が生まれ、それはニュートンの決定論的な世界観と矛盾するので、2つの相容れない科学観が衝突したというストーリーはよく知られている。最終的には統計力学が定式化されてその衝突を中和することになり、その名残がまだ今でも量子的にどうなんだという流れはあるにせよ、一応解決を見たことになっている。その後は、量子力学を含めた決定論的な世界観でオッカムの剃刀を使うという流れがずっとあるので、ハミルトニアンさえ与えればもう安心だということになるが、そのハミルトニアンをどうやって与えるかというところに、帰納バイアスを使うわけである。今我々が使っている帰納バイアスには、ハミルトニアンにあるたくさんの項のうち、これとこれは無しにしていよいよというような対称性の考え方や、場の量子論としてコンシステントになるようにハミルトニアンの数を制限するようなくりこみ可能性の考え方がある。対称性なら代数幾何、くりこみならくりこみ理論など、数学的な知見がその背後にあり、その数学の考え方を使って拡大された中でラグランジアンを書いたらどういうまずことが起こるのか、というようなことをしきりに我々はチェックしている。上手く行ったところだけが太い道として、今残っているようなものである。

Q：機械学習は緩和数理だという表現があったが、これはどういう意味か？

A：例えばトポロジーは素粒子や物性に役に立った事例というのがいくつもある。そこでは位相幾何を隔々まで使ったというよりは、非常に特殊な位相幾何の概念をいくつかつまみ食いして、そのつまみ食い部分が物理に幅広く適用されるということで縦串がついたという印象である。素粒子物理では原子核の種になる陽子や中性子をトポロジカルソリトンというもので作るが、それはトポロジーの分野では、どういう整数の質量を持つかということで分類できる。その分類に使われたK理論というのは今度は物性物理学で幅広く活用されている。トポロジーの全部が役に立つということよりは、トポロジーの中で培われたある一部に物理に非常にフィットする部分があるというイメージである。機械学習でも同様に一部分が非常に役に立つのではないかと考える。ここではある誤差関数のようなものを最適化する方法論は物理に非常に役に立つ場合があるという例を紹介したが、そのようなゆるい意味で緩和数理という呼び方をした。おそらく、ここが接点となり、そこから先の機械学習の様々な研究分野のところを探っていけるのではないかと考えている。

C：ある誤差関数を最小にするようなパラメーターセットが分かりますという意味で、機械学習が役に立つということと言っていると理解する。その意味では、いわゆる機械学習の範疇にはないかもしれないが、ブラックボックス最適化やベイズ最適化のような技術が近年進化している。例えばPFN社のOptunaは元々深層学習のハイパーパラメーターのチューニングのために作られたものだが、一般の関数の形が分かっていない世界でのパラメーター最適化がうまくいく。そういう意味では、橋本先生がご覧になっている機械学習というのは、いわゆる機械学習よりもう少し広い、別の領域も入っているのかもしれない。

## 2.4 材料における機械学習の現状と将来展望

徐一斌（物質・材料研究機構）

本講演では、材料科学におけるマテリアルズ・インフォマティクス (MI) について、自分の研究における実例と其中で遭遇した問題点を紹介したのち、将来の発展についての個人的な見解を述べる。

材料研究のすべての課題は、プロセス (Process, P)、構造 (Structure, S)、物性 (Property)、いわゆる PSP の関係の中で生じる。すなわち、プロセスによって材料の構造を制御し、材料の構造制御によって望む物性を実現することが材料研究の中心であるといえる。

これまで、この中心を形成した研究手法は実験であった。近年では、計算科学が著しく進展し、状態図の計算や、第一原理による物性の予測の例が大変増えてきてはいるが、全体的にみると、材料研究は相変わらず trial-and-error の実験が中心となっている。そのため、機械学習の手法によって、この材料研究の伝統的な方法を徹底的に変えて、材料の開発の時間とコストを大幅に短縮するという期待は多くの研究者が持っている。

材料研究の難しさは、それを記述するパラメーターが非常に高次元かつ広大であることに起因している。空間的には原子のサイズから、ナノスケール、ミクロスケール、バルクまで、時間的にも、電子が関与するフェムト秒レベルから、腐食やクリープのような現象の数十年、数百年レベルまで幅広い、空間・時間のスパンで現象が生じている。

我々は、このパラメーター空間が何次元あるのかを、実質的に把握できていない。それぞれの材料分野に対して、独自のサブスペースを定義して対処しているのである。例えば、金属を例にとると、構造材料の視点からは、チタンとアルミニウムとバナジウムの合金 (Ti-6Al-4V) と認識しているが、結晶構造の視点からは、チタンの結晶としてどういう構造を取っているのか (六方細密構造の  $\alpha$  相か、体心立方構造の  $\beta$  相か) を問題にしている。そのため、同じ記述子で記述された材料でも実は違うものであったり、違う表記をしているもの

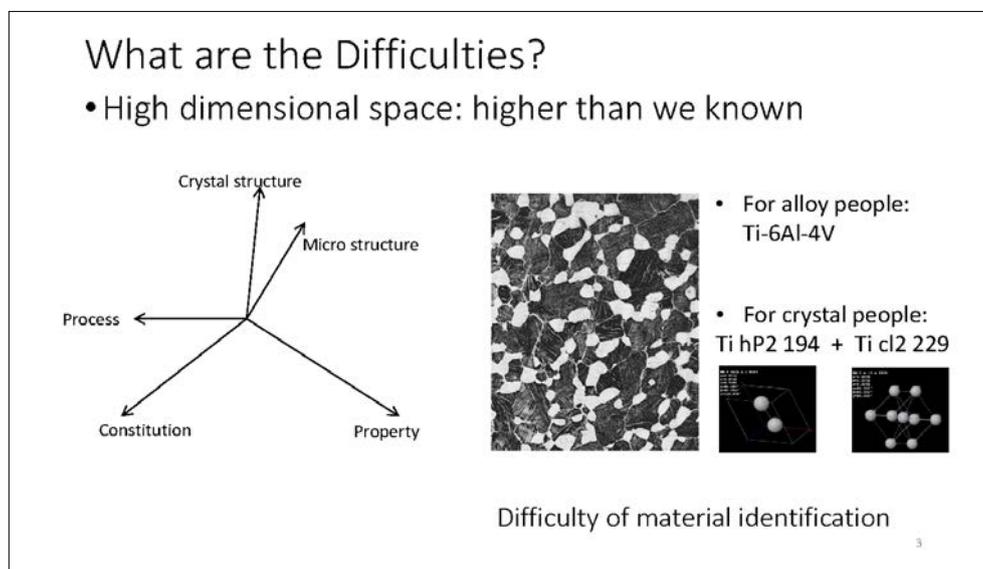


図 2-4-1 材料研究の難しさ

が実は同じものであったりということが起こりうる。こういうことが起こると、将来期待されている「分野を超えたデータのシェア」は非常に困難になる。研究分野によって、それぞれ一部の特徴しか特定していないということが問題になりつつある。

もう一つの問題点は、この材料の空間は大きいだけでなく、パーティションされていることにある。例えば合金の状態図を思い浮かべると、金属の組み合わせによって、2つの相が共存する場合もあれば、単一の相が存在する場合もあり、相としても結晶構造の違いや、液体・固体といった集合状態の違いもある。機械学習においては、最終的にパターン認識によって、パラメーター領域の中の規則性を探すことになるが、それはこの相のパーティションのところで不連続となり、規則性は崩れる。MIについて、よく言われている「内挿はできるが外挿はできない」という問題は、この相境界を跨ぐ場合、さらに厳しい問題となる。

次に、材料データ(=機械学習の基になっている材料データ)の現状がどうなっているのか見てみる。

巨大な材料空間の中で、すでに人類が「知識を持っている」、「探索したことがある」、「作ったことがある」空間は、図2-4-2左側に黄色で示したような局所的なものであり、その中でさらに、現在データを入手できる部分は本当にごく一部の黒い点ようになる。なぜそういうことになるかというと、一つは材料の歴史が非常に長い(土器・陶器の製造から始めるとほぼ1万年の歴史があり、冶金以降でも数千年の歴史がある)ことによる。そのような太古のデータは、現在資料として残されていないため、今入手できるデータは、ほぼ論文などに発表されているものに限られ、多くても直近200年分くらいしかない。このように、「入手できるデータが少ない」、「しかも局所的なものである」ということは、材料探索したい区域は、ほとんど既知領域の外にあることを意味する。材料に対する機械学習は、解決しなければならない問題がほとんど外挿問題となり、内挿出来る問題は本当に少ないのである。これらが材料への機械学習の適用を難しくしているもうひとつの問題点である。

ここで、材料データは材料空間をどの程度カバーしているかについて、1つの例を示す。

結晶構造データは、材料科学におけるデータのうち、最も網羅的に収集されているデータだと思われる。Springer、Elsevier他の、大きなプロジェクトで1880年代から、この結晶構造のデータ収集は始まっている。

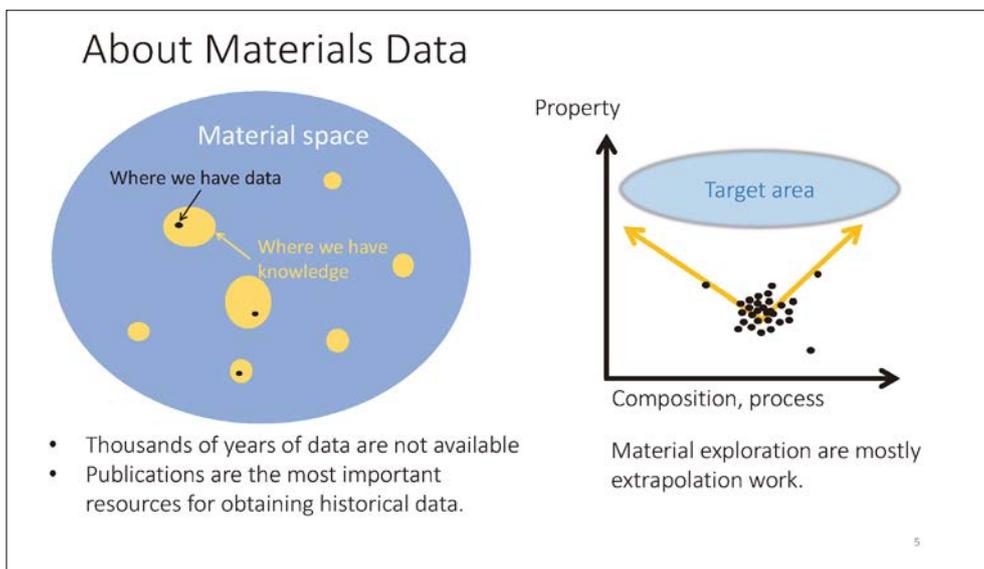


図2-4-2 材料データ量と存在範囲

これまでに得られたデータは、1元系に関しては、周期表にある100種類の元素に関して、ほとんど全て(～100%)網羅されている。2つの元素から構成される2元系になると、考えられる5000種類程度のうち72%と、これも概ね網羅しているといえる。しかし、これが、3元系になると、3つの元素からなる化合物の全ての可能性に対して、今、収録されているのが16%、4元系に至っては0.6%、さらに5元系以上はほとんどデータがないという状況になる。今の材料開発が、徐々に多元系の材料を目指していることを考えると、その周辺のデータがないということは大きな問題である。

この問題に対するため、2012年にオバマ大統領下の米国スタートしたMGI(Materials Genome Initiative)プロジェクトを契機として、世界中で、様々なMIプロジェクトが立ち上がった。これらのプロジェクトには、大きく分けると2つの種類がある。1つは新しい物質探索、分子や新しい結晶構造を含めての物質を探索するタイプのプロジェクトで、代表的にはアメリカのMGI、ドイツのNOMAD、スイスのMARVEL、日本はMI<sup>21</sup>などがある。この種類のプロジェクトは、目指している材料が化合物、あるいは結晶構造で、利用しているデータも、この結晶構造のデータベースに加えて第一原理計算で得られた電子構造など、比較的網羅的に材料データを入手しやすい分野となる。

一方、近年は特に中国を中心とした新しいMIのプロジェクトなどのように、実用材料を目指すタイプのプロジェクトも多い。合金、セラミック、複合材料のような実用材料を探索する際には、利用されるデータは、シミュレーションではなく、ほとんどが実験データに基づいたもので、日本のその種のプロジェクトとしては、SIPプロジェクトと、後で詳しく説明するMI<sup>21</sup>がある。材料の実用性の観点から見ると、予測すべき物性は実用上のデータである必要があるが、その場合の問題点は、そうしたデータがない、入手できないということになる。これについても後で詳しく説明する。

ここで、私が直接参加しているMI<sup>21</sup>というプロジェクトについて、少し詳しく説明する。

MI<sup>21</sup>プロジェクトは、JSTとNIMS共同のプロジェクトで、2015年から2019年の5年間行われた。プロジェクトリーダーは、前半が寺倉先生、後半は伊藤聡さんである。

プロジェクト全体のグループ構成は前半と後半で多少の違いはあるが、大きく3層構造となっている点は変わっていない。一番下の層がデータプラットフォームで、もともとNIMSが保有していたマテリアルデータベースに基づいて新しい機械学習を用いた使いやすいプラットフォームの構築を行う。この部分は私がリーダーになっている。真ん中の層は、データ科学やシミュレーションなどの計算科学のツール、要素技術の研究を目的としている。この層に含まれている研究者は材料研究者だけではなく、例えば数学系の人や、統数研のデータサイエンス専門家、あるいはバイオインフォマティクスの専門家も多く含まれている。一番上の層は、材料の出口のレイヤーであり、ここでは3つの材料系、蓄電池材料、磁性材料、伝熱制御・熱電材料をターゲットとしている。ここでも伝熱制御材料グループは私がリーダーとなっている。

これらのグループには、トータル200人ぐらいのオールジャパンの研究者がメンバーとして加わっていてコンソーシアムも形成されている。また、主に企業向けのコンソーシアムもあり、5年プロジェクト終了時点で、90社以上が参加し400人規模の活動員を供出していた。これらから日本の企業、産業界もこのプロジェクトにかなり注目していたことが分かる。

プロジェクトは5年間に9回の、協会セミナー、フォーラム、MI<sup>21</sup>フォーラム、社会向けの公開フォーラムを行い、毎回、企業のメンバーを中心に大体200から300人ぐらいの参加者を集めていた。フォーラムで行っていたアンケート調査によれば、MIに関する印象は、3回目(プロジェクト2年目)の時点で、半数以上の意見が「まだまだ時間がかかりそう」であったのに対し、第8回時点では、半分以上の回答では、「今すぐにも使えそう」と変化している。まだ、様子見の人という回答ももちろんあるが、その割合はかなり少なくなった。

この点からも、このプロジェクトは機械学習の材料科学への応用をかなり大きく促進したと考えている。

ここからは、2つの材料開発に関して、機械学習がどのように使われ、どのような成果が得られたのについて説明する。

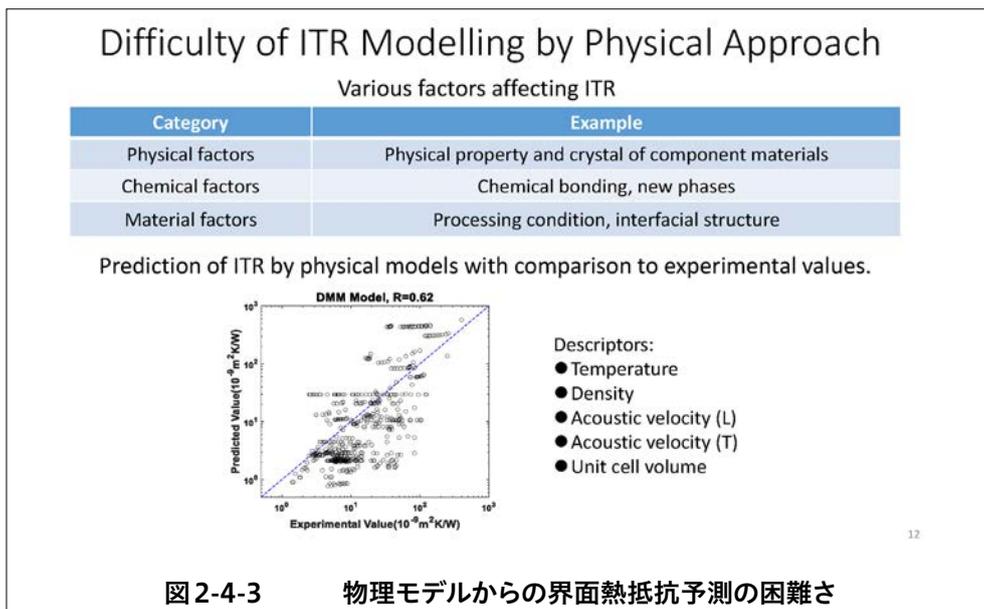
一番目の例は、無機断熱材料の開発である。断熱のコンセプトとしては界面の熱抵抗を利用する。2種類の異なる材料を接合した界面では、熱のキャリアに当たるフォノンや電子が散乱・反射され、それにより熱抵抗が生じる。この熱抵抗をうまく利用できるようなナノ材料を作り、多くの界面を含むことによって、低熱伝導の材料を作ろうというコンセプトである。この技術は、どのような材料の組合せで界面熱抵抗が一番大きいかという問題を解かなければならないが、現在我々が知っている約15万の物質の間の全ての組合せを考えると膨大な数となる。

その膨大な可能性の中から、どうやって界面熱抵抗の一番高い材料を探すかが、一番の課題となる。界面熱抵抗の物理というのはかなり研究されていて、その予測や制御が大変難しいということもよく知られている。この難しさの理由は、界面の熱抵抗に影響を及ぼす要因が多数存在することにある。それらは、(1) 物理的な要因「その材料の結晶構造およびそれ自身の本来の物性」、(2) 化学的要因「材料接合界面の化学反応」「原子間の結合」「異相の発生」(3) 材料学的要因「プロセス条件」「界面構造」などが複雑に絡み合っている。

物理的な界面熱抵抗予測モデルも存在しており、中でも一番よく用いられるのが、ディフューズ・ミスマッチ・モデルである。このモデルの詳細は説明しないが、単純には、2種の材料の物性のミスマッチから界面熱抵抗が生成すると考えて、その両側材料の物性の差から界面熱抵抗を計算する、というアプローチである。

図2-4-3下側に、このモデルを用いて予測した界面熱抵抗値(縦軸)と、実測値(横軸)の関係を示す。見て分かるように、大体の傾向は取られているが、全く精度の高い予測にはなっていない。特に中央あたりに、理論的に同じ値が予測された1列の多くの点があるが、実測値は2桁にわたって分布していることが分かる。

これは、記述子が足りない、すなわち、両側の材料の物性値だけを考えていては界面のことは決められないということを示している。そこで、これを機械学習の手法で何とかできないかと考えた。



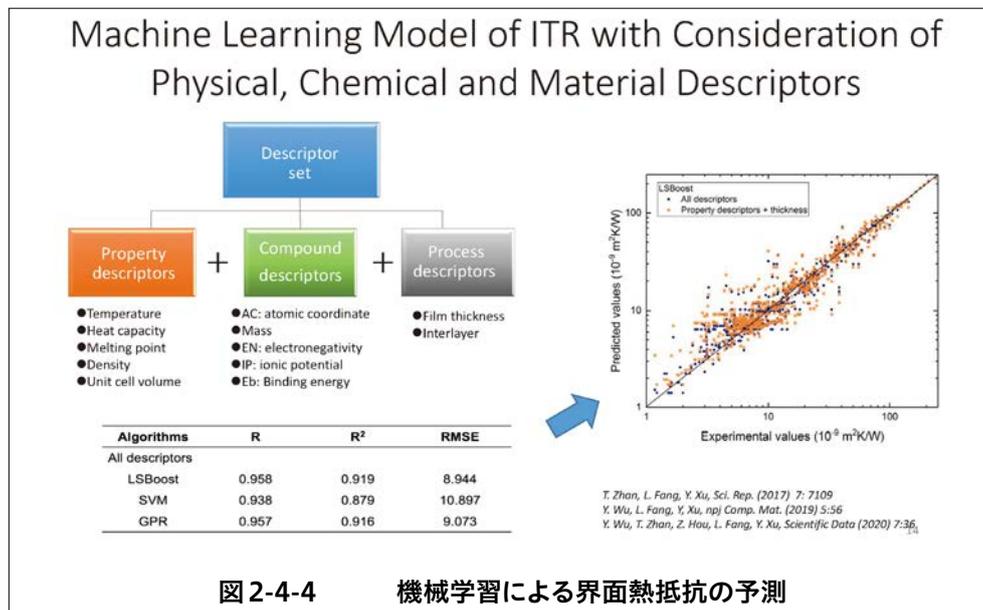


図2-4-4 機械学習による界面熱抵抗の予測

まず、機械学習にかけるデータを収集するために、論文から界面熱抵抗の実験データを集めた。これを行うだけで、3-4ヶ月かかったが、これにより、1,300点程度のデータがあつまり、約450の界面サンプルの点を収集できた。

次の難題は、記述子の選定である。先ほど述べたように、界面熱抵抗に影響を与える物理的、化学的、材料学的要因があるため、その材料に対して、それらの値がどのようなプロパティで表すことができるか、そもそもそのプロパティデータを入手できるか、データの信頼性がどうか等いろいろ検討しなければならない。いろいろ試してはみたが、結論から言うと最後は専門家が決めることになった。我々は界面の熱抵抗に関し、10年以上経験があるため、我々の経験で比較的入手しやすくデータの信頼性が高いと思われるデータを抽出し、記述子を12個程度選んだ。

こうして得られた、データと記述子を利用して機械学習を行うことになるが、次に問題となるのは、多くの手法の中からどの機械学習手法を使うかということである。事前には、この目的のためにどの手法が良いかは分からなかったため、一般に良く使われている5、6種類の手法をすべて試してみたところ、その中で比較的精度の高いものがいくつか見つかった。

図2-4-4の一番下の表で表したLSBoost (LSB)とSupport Vector Machine (SVM)とGaussian Process Regression (GPR)の手法は比較的精度がよく、それらの方法を最初に利用することにした。

この学習結果の表だけを見ると、3つの方法が統計的にはほぼ同じ精度を表しているが、この3つの方法を使って、8万種類の界面の組合せに対して、「どの界面の熱抵抗が高いか」という予測を行わせたところ、それらはあまり同じ結果を返しては来なかった。それぞれの手法が、異なる結果を出してくるため、最終的にそれらの結果を専門家が精査して、「作りやすさ」、「環境に対する配慮」、「界面の安定性」等を考慮した上で、最終的に選んだのがビスマス/シリコン (Bi/Si)という界面になる。

このBi/Siの界面は、専門家には興味深いものである。というのも、この組み合わせは、普通は、材料の熱伝導を低くしたいという用途では思いつかない(=作らない)材料系だからである。

そこで、まずは、ビスマスシリコンの界面熱抵抗を、実際に界面を作って測定してみた結果が表2-4-1である。

実験値と先ほどの3種の手法の予測値の間には若干の差もあるが、比較的良く一致していることが分かる。どちらかというと、実験値は、このLSBの値に最も近く一致しているとは言える。

表 2-4-1 予測値と実験値

	ITR( $10^{-9}m^2KW^{-1}$ )
SVM	61.13
GPR	60.32
LSB	50.68
実験値	51.80 ± 4.5

次に、ビスマスとシリコンの界面をたくさん含む複合材材料を構築、コンビナトリアル手法を用いて、組織構造の制御を含めた様々な材料を作り、この熱抵抗を測定した結果を図 2-4-5 に示す。図からも明らかのように、かなり低い熱伝導が得られている。報告がある無機ナノ材料の熱伝導と比較しても、少なくとも2018年の当時は最も低い熱伝導が得られている。無機材料であるにも関わらず、ほとんど高分子と同じぐらいの低熱伝導を実現することができた。

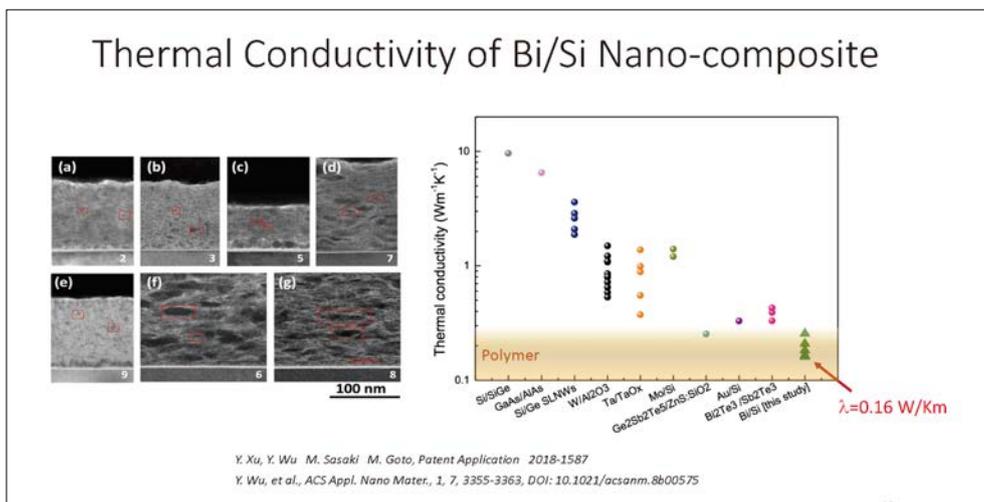


図 2-4-5 Bi/Si ナノコンポジットの熱伝導率

二番目の例は、逆に高熱伝導の高分子を開発するというものである。高分子は、普通は大体0.1~0.2(W/m K)という低い熱伝導率を示すが、それを高くしたいということである。コンセプトとしては、高分子の高次構造は考えず、分子の1次元鎖の分子構造をフィンガープリントにして、それと熱伝導率との相関性を機械学習で探すというものである。

高熱伝導の高分子を設計していくステップで、利用したデータはNIMSのポリインフォというデータベースであるが、これは元々1995年にJSTで作られ始めたデータベースで、それが2003年にNIMSに移管され、その後、今日までずっと20年以上開発を続けられたものである。論文からデータの収集が行われている。

このポリインフォの中のホモポリマーは、全部で1万4,000種類程あるが、その分子構造を利用して今回の研究を行った。しかし、問題点は、分子構造は多数あるものの、高分子の熱伝導のデータは非常に少ないということである。1万4,000種類の高分子に対して熱伝導率の値が収録されているものは、数十種類しかなかったため、このデータだけではとても機械学習はできない。そこで利用したのは特性間の相関性で、例えば、熱伝導率と高分子のガラス転移温度をプロットしてみると両者の間には線形的な相関性が見える。また、熱伝導率とガラス転移温度だけでなく、ガラス転移温度と融点の間にもほとんどこの線形的な関係がある。

そこで、高熱伝導のポリマーを見つけるために、データ数の少ない熱伝導率そのものが高い材料を探すのではなく、データ数の多いガラス転移温度と融点が高い材料を探すことで、高熱伝導率の材料が見つかるアプローチで分子設計を行った。

ガラス転移温度と高融点の高分子を探索すると、そのような材料は大量にあり、安定に存在する可能性を考えなければ何万件もの構造が、すぐに設計できてしまう。そのため、次のステップとして、ポストスクリーニングが必要で、ある程度の構造を絞ってから条件探索を行うことにした。今回の場合、図2-4-6に示す液晶高分子に集中することとして、液晶特有の分子構造を持つかどうかで候補材料をフィルタリングし、高分子が生成できるかどうかも考慮したうえで熱伝導を計算し、最初に設計された1,000個の高分子の中から24の候補に絞った。

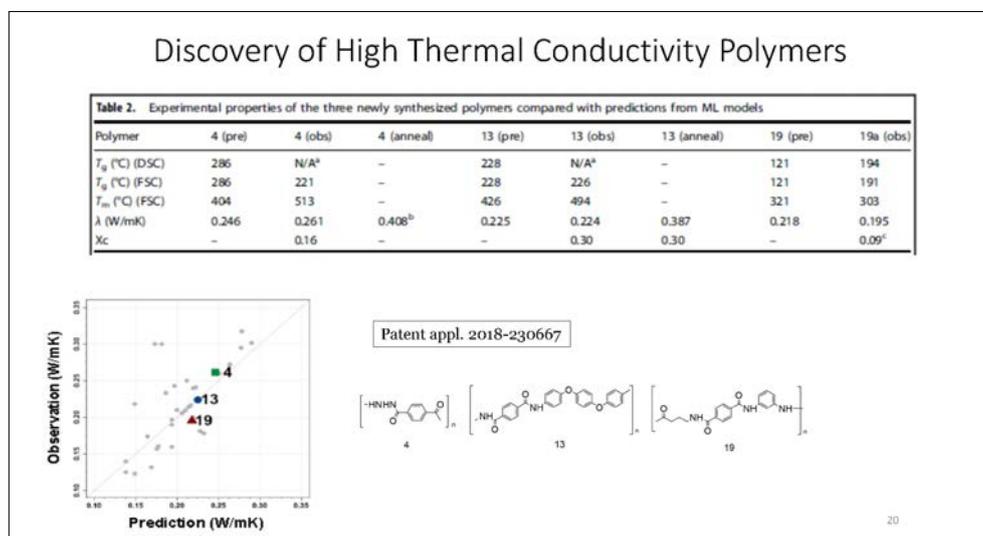


図2-4-6 設計したポリマーの熱伝導率

絞った24個の中から、4種を実際に合成したところ、1つは合成できなかったが残り3種は合成でき熱伝導率を測定することができた。この3種の材料の測定値を図2-4-6に示す。合成直後で0.26と0.2以上あり、アニールを行うと0.4に近い値となっている。先ほど述べた高分子の熱伝導率の平均が大体0.1から0.2の間であることから、高分子の中では比較的高い材料が得られたことが分かる。

一般的な感想としては「やはり機械学習は材料研究に確かに使える」ということである。その利点は、現象自体が複雑なケースに特に現れ、たとえば初めの例の界面熱抵抗のように、あまりにも複雑で物理モデルが簡単に立てられないケースに対しては、このデータモデルは非常に有効だといえる。また、新しい分子の設計

や、新しい物質の設計にも有用で、第二の例の高分子の探索の例や、網羅的に実験条件を探索するときに、ベイジアン最適化が使えば、すべてを網羅的にやらなくても済むため、早く最適化のベストポジションにたどり着くことができる。結果、研究が効率的になり、実験と計算の量はかなり減らすことができる。

ここまでは、材料研究に機械学習を応用する場合の良い面だけを述べてきたが、これからは、課題・問題点を述べる。まず、第一には、なんとといっても、データ量がボトルネックとなる。世の中では、情報の問題と言えばビッグデータ、つまり、データ量が多過ぎてコンピューターが処理し切れないことが問題とされるが、我々の場合は、ビッグデータがあればむしろ問題がなく、実際にはスモールデータが問題である。著しい場合には、ノーデータとなる場合もあり、問題はかなり深刻になる。

材料のシミュレーションから得られるデータは、まだ現実を完全に反映しておらず、ある程度の参考にはなっても、まだ実用材料への適用には距離があると言わざるを得ない。一方で、実験データについては、そのコストが「高い」という問題を抱えている。「高い」の意味は、材料を合成する際のコストだけでなく、物性を測定する際のコストも含んでいる。近年では、コンビナトリアル実験によって、自動化、ロボット化することで網羅的に材料データを収集するアプローチもあるが、個人的な見解では、自動化によって人件費は確かに減らせるが、実際に材料を作るときに使用している原材料費やエネルギーなどに関しては特に削減されていない。したがって、それほど多くのタスクを網羅的に行うことは現実にはできない。

また、そもそもMIの導入目的としては、「最小限の実験回数で済ませたい」、「最適な条件のピンポイントで実験したい」というものがあるので、機械学習のデータ収集のために網羅的に実験を行うことは本末転倒であり、全般的に見ると、あまり根本的な解決手法にならないと考えている。

最後の問題は、機械学習が作ったモデルが正しいかどうかの判断の手法がないということである。機械学習自体は、専門知識全く持っていない赤ちゃんみたいなもので、この赤ちゃんは与えられたデータに対しての最善のモデルを作り、必ず答えを出してくれる。この答えが、「良い」ときもあれば、「悪い」ときもあるというのが、大きな問題なのである。今、ここで言う「良い」「悪い」というのは、「材料の視点から見て正しいか

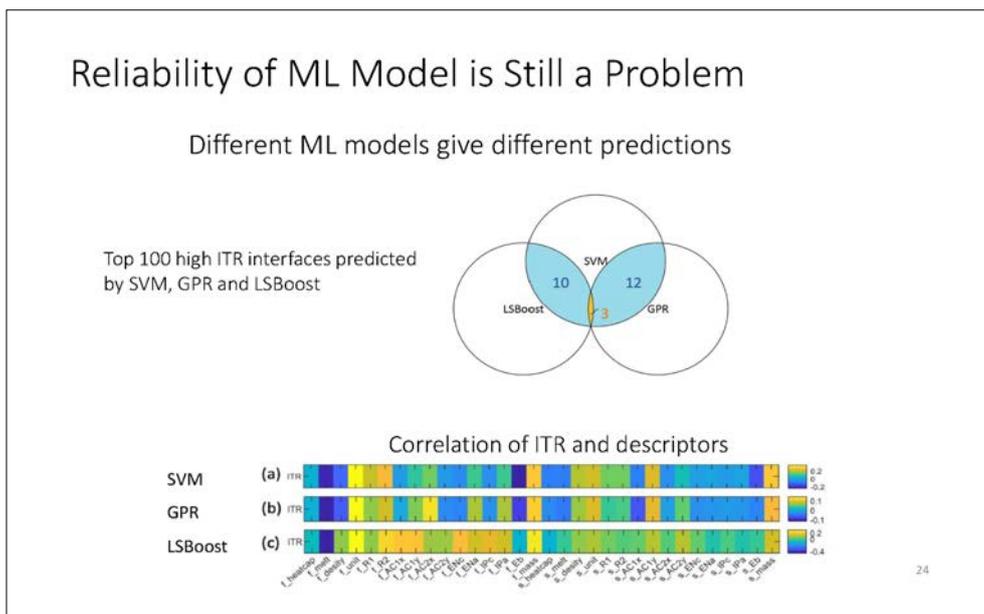


図 2-4-7 機械学習モデルの妥当性

どうか」、「このように材料を作って本当に良いものになるかどうか」、それが唯一の判断基準で、別に原理を理解しなくても、材料科学としては全く構わないといえる。

しかし、全く理解することなく、モデルが正しいかどうかを判断することができないという理由において、最終的にある程度理解する必要があるが出てくる。つまり、「理解することが目的ではない」にも関わらず、「理解しないとモデルの適用可能範囲が分からない」という理由で、「我々は理解する必要がある」という皮肉な構造が生じている。

熱抵抗の例で、もう少し詳しく説明する。図2-4-7に示すように、我々が用いた3つの手法に、トップ100の候補化合物を出力させると、全てに共通の材料は3種しかない。それぞれの手法が予測する材料は異なっており、2つ以上の手法で共通する材料も合わせて25種しかない。

これから、手法によって機械学習の予測結果に大きな差があることが分かる。そこで、我々は、最終的に得られた界面熱抵抗と、最初に設定した記述子との相関性を分析してみた。図2-4-7下に、各手法がどの記述子と結果の相関性を高いと見ているか(どの記述子を重要と見做しているか)を色分けで示しているが、同じ色のパターンであれば、同じ記述子を重要と見ているということになる。図から明らかなように、SVMとGPRは、ある程度近い記述子を見ているが、LSBは全然違うものを見ている。つまり、この3つのモデルにとって、重要な記述子が異なるということになる。実際には、「本当は何が重要か」という答えはあるはずで、最終的にこの例に対して我々は、共通しているもの取るという手法を取った。

もし、今回行ったように、機械学習の内部(記述子との相関)を調べることをしなかったら、どのモデルの予測値を信頼すればいいか判断できず、実験をミスリードしてしまったり、間違った結果が出たりすることも十分考えられる。

ここまですがMIの現状であり、次はMIの将来の発展に関する私見を述べる。まずは、「ノーデータ」というデータが少ない問題を解決しなければならない。これをブレイクスルーしない限りMIをブレイクスルーはできない。

しかし、先ほども述べたように、実験データの量を増やすことに関しては、明らかに限度があるため、可能性が一番高いのはやはりシミュレーションデータを増やしていくということになる。シミュレーションから得られるデータは、近年、品質も向上し、量もかなり増えて来ているが、私がさらに期待しているのは、量子コンピューティングである。量子コンピューティングが10年以内に実用化されれば、「もう少し現実に近い材料のシミュレーションができるのではないか」、「本当にサイバー空間の中で自動、全部材料を設計できて、その結果に基づいて実験できるのではないか」、「実験の際には、マシンラーニングを応用して最小限の実験回数で結果を得られるのではないか」といったことが、私が期待していることである。

次に、AIも人間のような知識の蓄積法を行う必要があると思っている。今のAIは数値データからしか学べないが、材料分野は、すべてが数値データとして保管されて、昔のデータ通りに今も実験がなされているわけでもない。昔の実験データは、すべて知識や理論として残されているため、AIも人間が勉強するように物理の本を読んで材料の理論を勉強する必要があるのではないか。つまり、すべてを数値から学ぶのではない手法を構築する必要があると考える。もしそれができれば、AIも人間のように歴史から材料科学を勉強でき、本当に巨人の肩に立つことができる。さらに、人間のように定年の制限とか寿命の制限がなく、さらに無限に発展する可能性があると感じている。

私は、AIには人間にはできない、AIにしかできないこともあると考えている。それは、「現実を見る」ということである。今でも、科学者として一番重要なのは現実を正確に捉えることであるが、人間にはどうしても限界があり、よく心理学で言われているように、「人間に見える世界は現実より自分が見たいものである」と

いうバイアスがある。AIにはこれがなく、人間が錯覚するだまし絵(例えば、人間には、見方を変えることで、若い女性にも老人にも見えるような絵)をAIが見た場合には、本当に新しい生物とか新しい法則とかを見つけ出すことができるのではないかという期待である。

逆に、AIに多分できないと思っていることもあり、その中でも、一番重要なものは、「発明に対するモチベーションを持つ」ということである。人間がサイエンスを始めるモチベーションは好奇心であり、人間の弱みであり、欲望であり、それらから全てのサイエンスが始まると思うが、コンピューターやAIが、「次にどんな材料を作りたい」といった希望を持つことは、少なくとも今のAIにはできないのではないかと考えている。そして、もし、将来AIにも人間の感情を与える、人間の欲望も与えることができたとしたらどうなるだろうかとも想像している。その場合、すなわち、そのAIが、人間のアドバンテージを全て持ったとしたら、多分、人間の弱みも持つのではないかと思う。

現在、我々は、「AIには何ができる」といった「できるかどうか」だけを議論しているが、「私たちは本当にどんなAIが欲しいのか」とか、「どういうAIを作りたいのか」も考えるべきではないかと考えている。

#### 【質疑応答】

Q：Bi/Si界面熱抵抗に関して、実験値に近い値を与えたLSBは、そのほかの界面に対しても良い値を与えるのか？

A：実験と比較したのはこれだけである。テストデータに関する予測は、どれもほぼ同じ程度に正しい値を返す。その場合SVMが一番近い値になることもある。今回はLSBが一番良かったとしか言えない。

Q：3つの方式が違った値を返していて、SVMとGPRが比較的近い値を返していたら、こちらが正しいのだという判断もあるが、それはどうか。

A：その点こそが、機械学習の信頼性、何を信じたらよいかの問題になる。今回は記述子との相関を調べることで色々なことが分かった。結局、モデルを理解しないと信頼できないと考えている。

Q：材料の世界では外挿することが多いということだが、Bi/Siの場合は、界面熱抵抗の大きいところを探しているという意味で外挿ということではないか。もしそうだとすると、内挿を行った場合の精度については評価したことはあるか。

A：材料の場合、内挿、外挿という区別がつけにくいという点がある。計算から得られたMIの結果を実際に合成した例が他にあまりないので、精度については分からない。

Q：シミュレーションの結果が、なかなか実際とは合わないというのは、計算量を増やしていけば精度を上げられるという性質のものか。

A：例えば第一原理計算では、せいぜい1,000原子ぐらいまでしか計算できない。実際の材料はもちろんそうではなく、もっと多数の原子が複雑に存在している。今の計算はまだ全然そこまでの領域まで追いついていない。もし量子コンピューティングができればもうちょっと複雑なものを予測できると期待している。

Q：ここで使われている記述子は、人間が実験するときに操作するパラメーターのようなものなのか。それとも機械学習のために何か最適化したような特別なものを設定しているのか。

A：今回選んだ記述子は、人間が選んだもので、材料の選択、実験条件の制御など、コントロールできるパラメーターである。コントロールできないものはそもそも分からないものが多い。記述子の選択を自動で決めさせようとしたが、そうすると結構重要なものを落としたり、信頼性の低いデータが選ばれたりするので、専門家が選ぶ方式を取った。

- C : AIはインベンションのモチベーションがないという話は示唆に富む。モチベーションとは、いろんな選択肢の中で優先順位をつけるということだと。「自分はこれを大切に思う」というのは、重みづけ、いろんな特徴の中、選択肢の中で重みづけを行うというのがまさに重みづけということ。
- A : ここでのインベンションは選択肢の中で選ぶというより、何もないところから、例えば月に行きたいとか、宇宙の外で何があるとか、それは特に人間のもともとの好奇心は選択されたものじゃなくて、そのようなもうちょっと広いところの、もうちょっと基礎的なインベンションという意味はしている。
- C : そうすると、全ての選択肢を完全にフラットに考えて、そこから選ぶということ。人間の脳の大部分はまだ未解明だが、少なくとも完全に同じ重みでサーチをすると数学的には何の特徴も、特徴抽出ができないということが証明できるので、そういう意味で人間が何かインベンションするというのは何か脳の構造とか、環境とかいろんな影響によった、我々が認知できない作用を及ぼした結果なのかもしれない。それをAIに教えるというのは何かとても大変な気がする。
- C : 少なくとも今我々が持っている機械学習などの人工知能の技術の中では、やっぱりそのモチベーションをどう与えるとか、目的関数を、例えば最適化問題として定式化できたとしても、どういうふうに目的関数を与えるかとか、それから、そもそも機械学習をするときにその帰納バイアスがあるのを、何らかの形で我々は仮定しないといけないとか、そういうような問題があって、だからその辺はまだ人工知能を50年やっても解けてないというのが実情である。
- A : どうやって実現するか全く想像つかないが、人工知能の研究者達が実現してくれるかもしれないと考えた。問題は、もし実現できたら人間みたいにAIには自分が欲しいものが出てきて、それが私たちと同じものになっているかも分からないし、それで、また客観性も失うという人間の弱点も持つのではないかということを書いたかった。
- Q : 責任者をされているNIMSのデータプラットフォームの説明をして欲しい。
- A : 私は、JSTが構築したデータベースが全てNIMSに移管され、NIMSでデータベースステーションが立ち上がり、材料データベースのインターネットサービスが始まる2002年時点から参加している。NIMSには、インターネットで全て無料で公開していた有名なMatNaviというデータベースシステムがあったが、そのデータページの開発とメンテナンスをずっとやってきた。以前のMatNaviは、ウェブアプリケーションで、ユーザーがGUIを経由してデータを一件一件見ることしかできず、ダウンロードできないということが弱点であった。今回、MI<sup>2</sup>Iのプラットフォームでは、機械学習に使いやすいようにAPIを開発し、計算機能にシミュレーションとデータと機械学習の機能を全部統合した。60台ぐらいのPCクラスターでプラットフォームを構成しているが、それも開発はプロジェクトで行っている。
- Q : このデータは、海外の研究機関と共同のデータベースになっているのか。
- A : 無機材料データベース (AtomWorkとAtomWork-Adv) は、スイスのMPDS社と著作権共有している。データの作成も、スイスを中心にヨーロッパで作成している。データ自体は論文から収集されたもので、世界中の論文から収集して専門家、定年になった研究者が論文一本一本読んで作られたデータベースである。最初からインターネット上で無料公開しており、最近は色々な公開データベースが出てきたが、少なくとも5年、10年前には、世界最大のデータベースと言われていた。高分子、無機材料など様々な材料のデータベースがある。

- Q：シミュレーションで、「こういう結晶構造はこういう性質を持つ」というようなデータも、データベースに入るか。
- A：結晶構造をデータベース持っているので、そのデータをもとに第一原理計算で電子構造を出して、物性計算も試してみた。第一原理計算でできる特性も、難しい特性もあるが、計算能力の課題等もあるため、まだ公開していない。世界的には計算物性データベースも存在する。
- Q：企業が持っているデータはどのくらいあるのか。
- A：実は分からない。企業のデータ自体は外へ持ち出す制限が強く、専用性・秘匿性の高いシステムを作って自社データとNIMSのデータと併せて機械学習するという環境を用意しても、結果としては利用者ゼロであった。
- C：物質関係、プロセスに関わるところというのは全て知財と非常に密接な関係があるために、多分、共有化が最も難しい領域ではないかと考える。例えばJSTなどで安心して企業がそういうのを提供できるような良いスキームを作れば、総合知としてはすごくいいものができると思う。
- Q：よく機械学習で1億件のデータを学習したとか言われるが、手作業の実験でそのデータを集めるというのは、ちょっと不可能ではないかと思うがどうか。
- A：それは不可能である。データによって収集しやすいデータと収集しにくいデータがあるが、論文からのデータ収集なら、年間1万～2万の結晶構造は収集できる。実験に関しては、コンビナトリアル実験でサンプルを作るのは非常に早い（シリコンウエハー1枚の中に組成傾斜を作れば数百個のサンプルが一度に取れる）ものの測定には時間がかかる。200サンプルのTEM測定を行えば、一年はかかる。
- Q：測定を機械でやるということも考えているか。データを自動的に作るアプローチだが。
- A：考えているが、まだ測定できる特性が多くない。光特性なら比較的簡単だが、熱伝導の自動測定は、幾つかの会社が開発しているが、まだ出てきてない。全体的に、コンビナトリアルで行う実験は、まだまだ技術の問題が大きいと考える。
- Q：データに基づいて出てきた結果のうち、専門家から見て当然でないものがあるのか。
- A：さきほどの高熱伝導高分子については、驚くほどではなかったが、Bi/Siの界面熱抵抗については正直言うとちょっとびっくりした。何故なら、Bi/Siというシンプルな材料組み合わせが、何故か誰にもやられていない。これは、研究者が研究している領域は限られていて、熱伝導率が決して低くはない金属材料であるビスマスをあえて使うことが少なかったということだろう。私が低熱伝導の複合材料を設計しようとすれば、まず金属は使わない。
- C：機械学習も我々が思いもよらないような予言を実際に行っているんだと思う。その予言をしたとき我々はどう解釈するかというと、それはアノマリーだから無視しようとする。ここが、将棋と決定的に違う点で、将棋の場合はルールが分かっているので、機械が予言したことは正しい。棋譜に載っていようが、載ってまいが正しい。最終的に、この信頼性とか機械学習の中身の理解に掛かってくる。結局、思いもよらない予言と、その予言をした機械の信頼性の取引関係だと思う。ここが、ゲームなどとは決定的に違う領域。信頼性と、意外性の取引関係から我々は逃れることができない。そこを決めるのが、最終的に今のところやはり専門家以外にないということだと思う。

## 2.5 データ駆動による因果仮説探索

清水 昌平 (滋賀大学/理化学研究所)

因果推論とはデータから因果関係を調べるものであって、例えば「相関があるからといって因果関係があるとは限らない」事象の因果関係の有無を示す方法論である。近年この因果を扱うための方法論が発展し、いわゆる機械学習の国際会議で論文が多数見られるようになり、AIに組み込まれたりして市民権を得てきた。

例えば、国民一人あたりのチョコレート消費量とノーベル賞受賞者数との間には相関がある (図2-5-1)。しかし、仮に国策としてチョコレート消費を増加させても、当然のことながらその国のノーベル賞受賞者は増えるとは限らない。

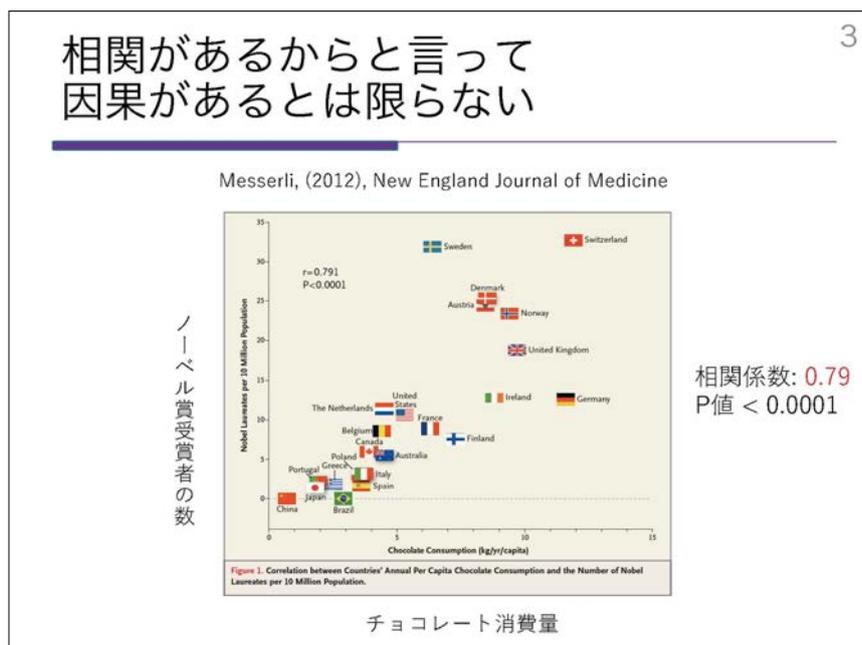


図2-5-1 各国のチョコレート消費量とノーベル賞受賞者数

機械学習の一番単純な予測モデルならば、チョコレート消費量とノーベル賞受賞者数との間の関数を推定して、チョコレート消費量からノーベル賞受賞者数予測できればよい。しかし、因果理論という観点、つまりノーベル賞を増やすためにどうしたら良いかというような仮説を得たい場合には、このような予測だけでは不十分である。

また、仮にチョコレート消費量とノーベル賞受賞者数の間に因果関係があったとしても、原因と結果がどちらであろうと同じ図となる。さらに、チョコレート消費量とノーベル賞受賞者数の間には直接の因果関係がなくても、GDPを高めるとチョコレート消費が増加し、同様にGDPを高めると科学予算が増えてノーベル賞も取るようになる、という場合でも同じ散布図となる。

これまでは、分析者が因果関係の構造を推定した図で表した「因果グラフ」が用いられてきた。因果推論

という、何が因果で何が結果かを探るような語感があるが、因果の大きさを推定するものが主流であった。図2-5-2の例では、GDPは「潜在共通原因」であり、基本的な因果の大きさを推定するために「傾向スコア」、「バックドア」といったツールが開発されてきた（図2-5-3）。観測できない情報 $z$ を潜在変数として用意することによって、 $y$ から $x$ への因果構造が明らかになる場合もある。

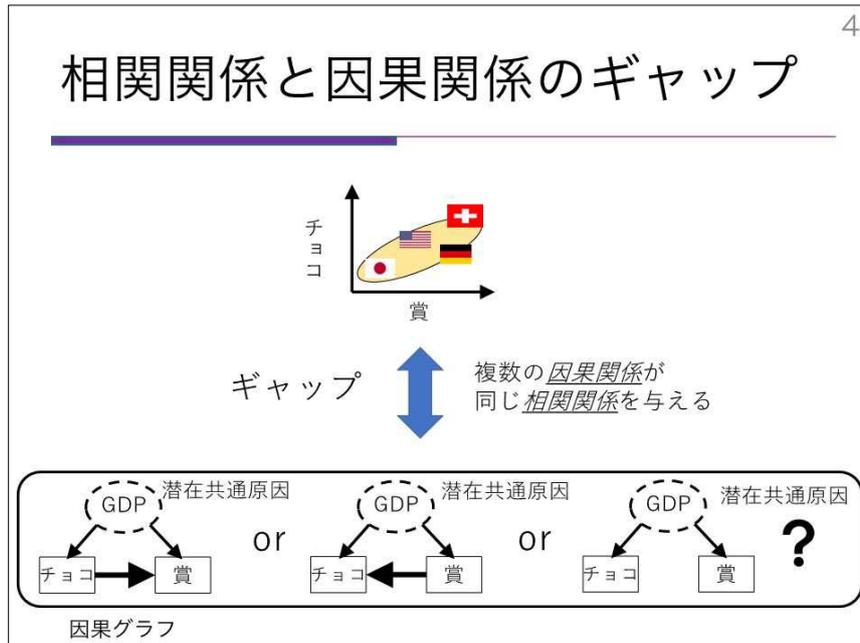


図2-5-2 相関関係と因果関係のギャップ

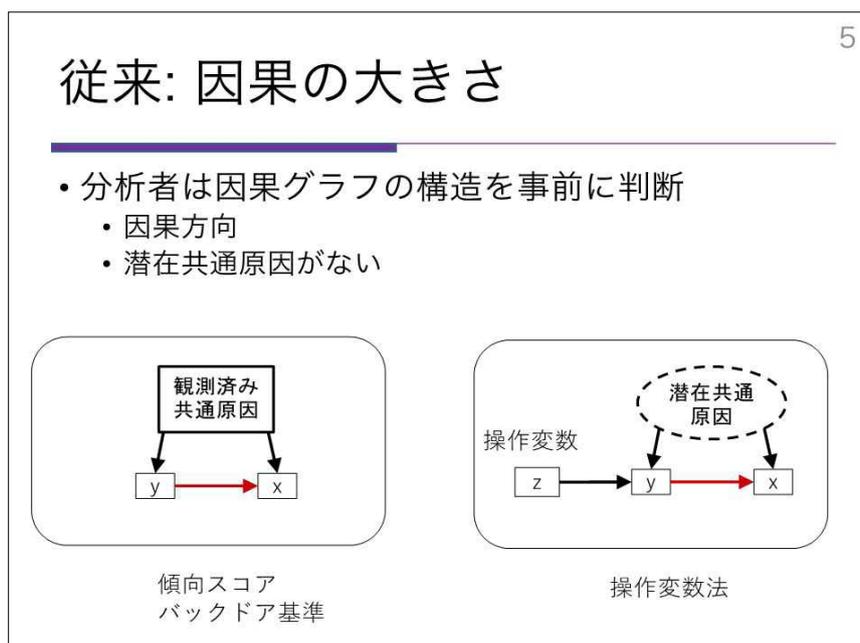


図2-5-3 因果の大きさ

このように因果グラフは、各変数の関係に基づいて研究者が自分で作成するものであった。領域知識がない場合にはデータから推定する。また、「データ」に時系列情報が含まれる場合には活用し、そうでない場合ももちろん取り扱う。

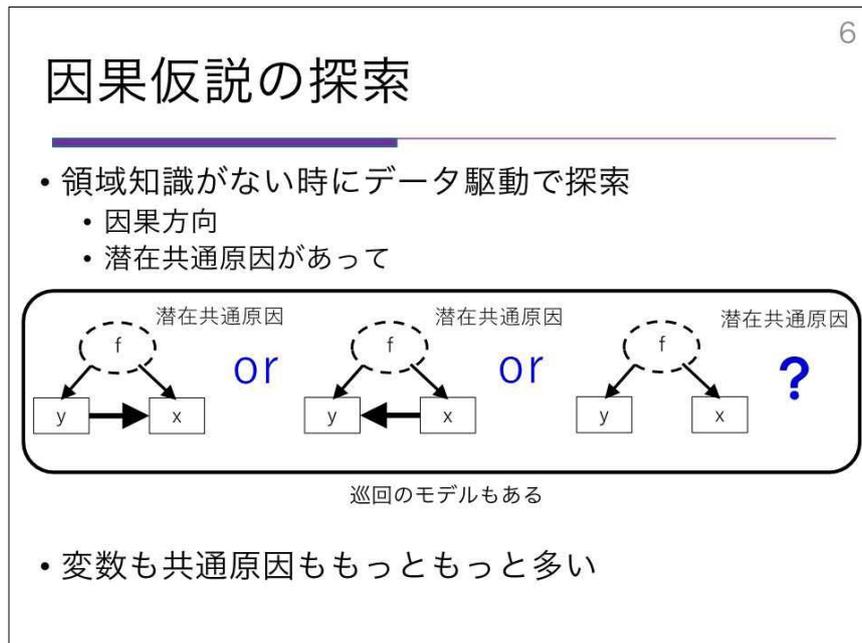


図2-5-4 因果グラフ

このような分析法は、実験が行いにくい分野で需要があつて、実験を行う前に有望な仮説を推定する場合もある。特に疫学や経済学の分野で活用されており、膨大な変数を扱うモデルもある。

例えば睡眠障害と鬱の因果関係を検討する場合、変数自体は少なくとも、共通原因の候補は無数にあり、非常に多数の変数を扱う必要がある。

経済学では、雇用と売上げとR&D、データ取得が1年に1回の場合などは、データがとれない期間に生じる様々な要因による変化をモデル化する。脳の領域間の因果関係なども関心の対象になっている。化学分野では、分子シミュレーションで様々なデータを作成し、分子間の距離等がどのように吸収スペクトルに影響を及ぼし、どこを変えると吸収スペクトル増やせるかを探るといった需要がある。

図2-5-6は、京都大学のWPIのプロジェクトで、SOX17とかBLIMPという遺伝子がどのように影響し合っているかは、基本的には各ステップで1個1個実験して因果構造のグラフを書く。私はここで、データから遺伝子ネットワーク因果関係の推定を行っている。今は4変数、4つの遺伝子しかやっていないが、今の技術でも、200から2,000変数で分析を進められる。

また、「因果モデル」と「予測モデル」の組み合わせも行っている。Yを予測主体として、予測値として $\hat{Y}$ を推定している。

例えば健康診断の項目から将来の罹患を予測するモデルは多いが、その予防のための提案として仮に体重を減らすと、どのくらい病気のリスクが下がるか示すことができれば良い。この場合、今の予測モデルでは、体重だけ変えてそれ以外の計測値は同じ値を入れた場合の予測値を返すため、体重を減らした結果として減

少したコレステロール値などが反映されない。多くの変数の間の因果関係、因果構造も考慮した予測モデルの開発が応用例として考えられる。

もちろん、予測モデルが既に活用されている分野では、それをやめて因果モデルだけに変更するのではなく、その予測モデルは生かしつつ、因果を考えたモデルも使う、という活用の仕方もある。先の健康診断

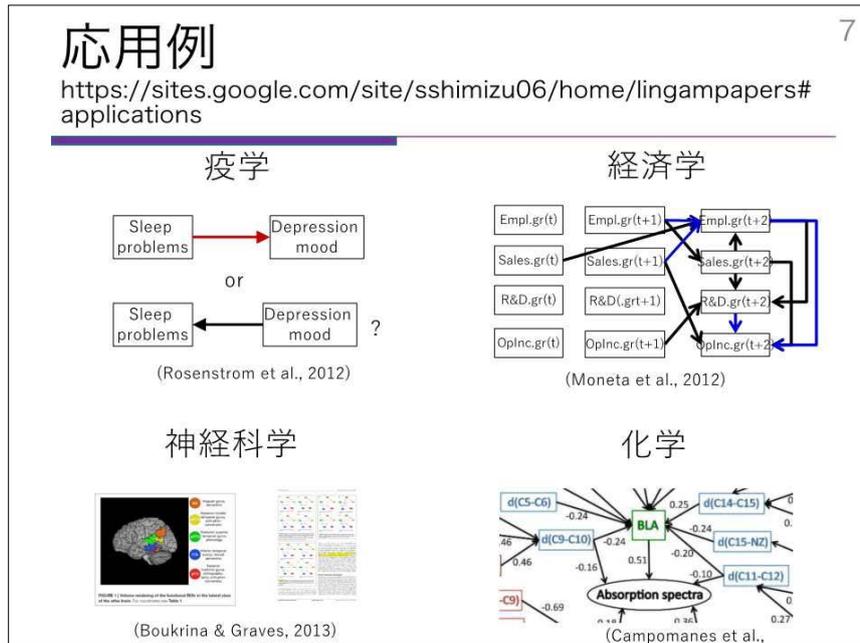


図 2-5-5 様々な分野での活用例

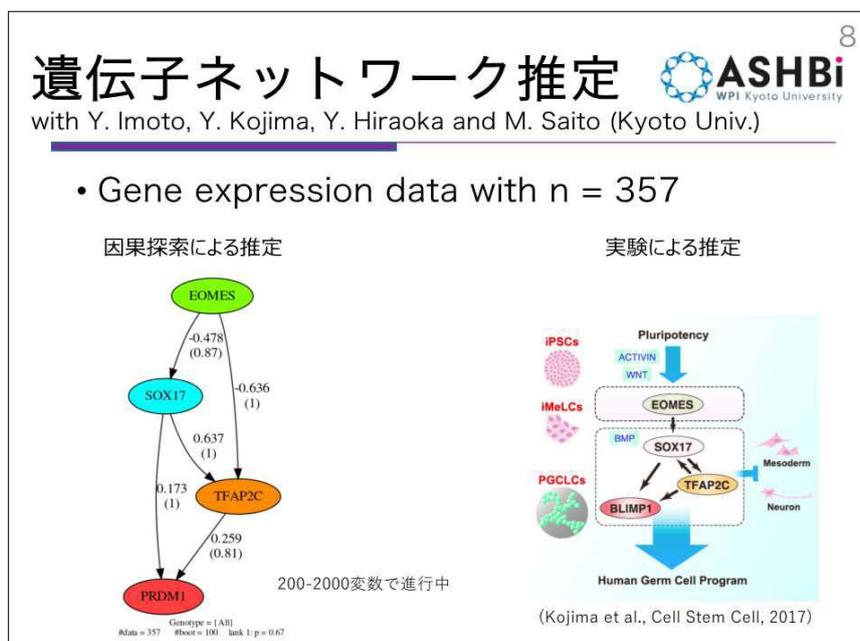


図 2-5-6 遺伝子ネットワーク推定

の例では、体重を変えた場合の他の変数の値の変化については、因果モデルを使用し、スナップショットから病気のリスクを予測するということは、最先端の機械学習モデルを使ってもよい。

ここで一旦まとめると、因果推論では、どんな状況でどの変数がどの変数にどのぐらいの因果の大きさを持つのかというのをサジェスチョンしたい。これまでの説明では、1個の変数に対して1個の変数が介入しているモデルを使ったが、複数の変数の介入もモデルとしては可能である。

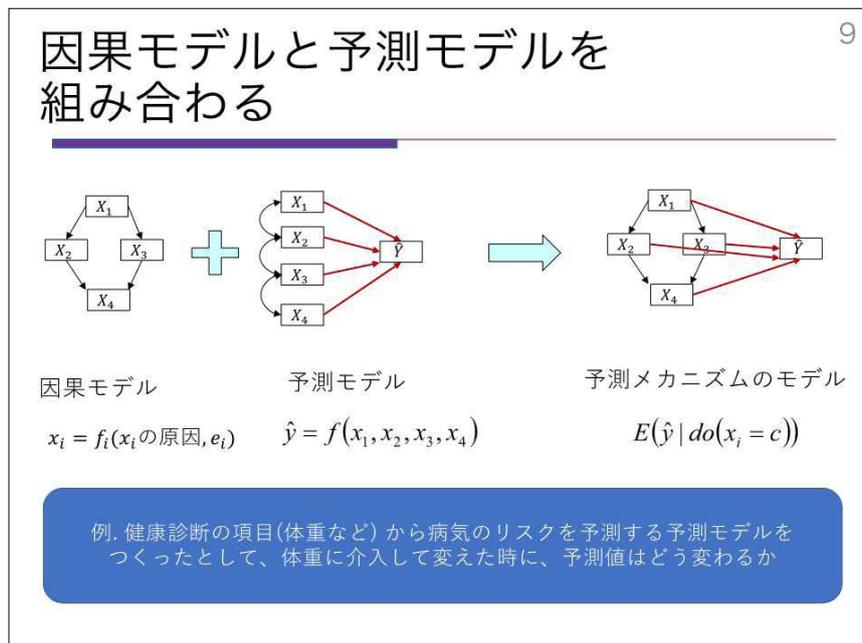


図2-5-7 因果モデルと予測モデルの組み合わせ

製造業や農業分野においては、製造・栽培条件と原料から、最終的な特性として不良率や収量の推定が必要となる。その中間にある多くの変数を明らかにするために因果関数を使うというニーズがある。

こういう仮説を出して、実験する場合にデータから仮説を出したり、調査を行うことにより、効率化が進む。

### 因果推論

因果推論には、基本的に3つのトピックがある。

1つは、「因果を式で表す」という試みで、大体今2つぐらいのフレームワークに収められており、さらに大体1つのところに収束している。次はそのフレームワークの中で、因果の大きさを推定するもので、3つめは因果探索で、これはまだ発展途上。

これは機械学習分野でも非常に注目されている。例えば因果情報を追加情報にして予測するものである。この場合、集団が変わっても因果構造自体は変わらないと考えられるため、これを利用すればよりよい予測ができる。もう1つは、説明可能性と因果。あと、逆に機械学習を使って因果構造の大きさを調べるもの。複雑な関数関係でも、機械学習を使えばうまく推定できる。機械学習と因果推論という組み合わせは近年注目されている。

結局、因果の大きさを調べようとすると、因果グラフが必要となる。因果グラフが解ると、結構いろいろな

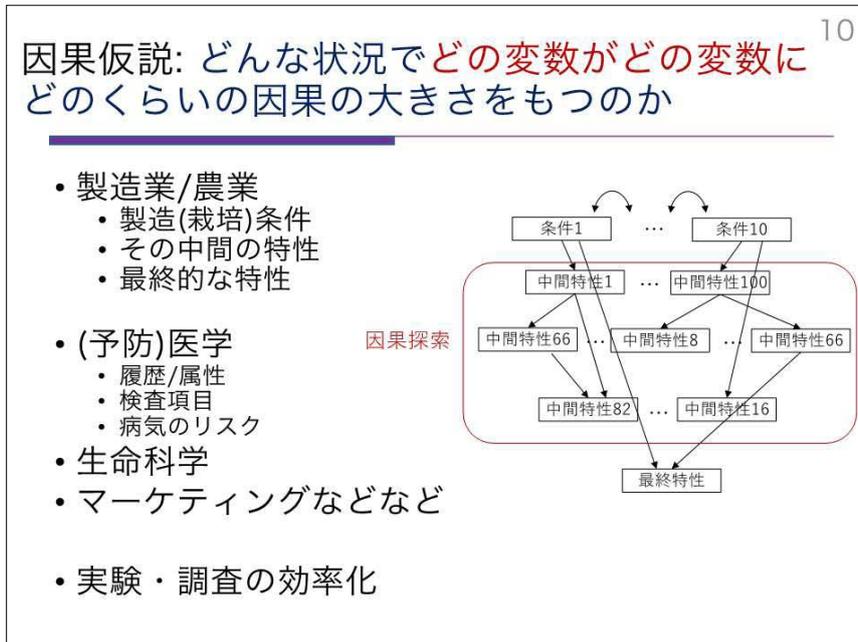


図2-5-8 因果仮説

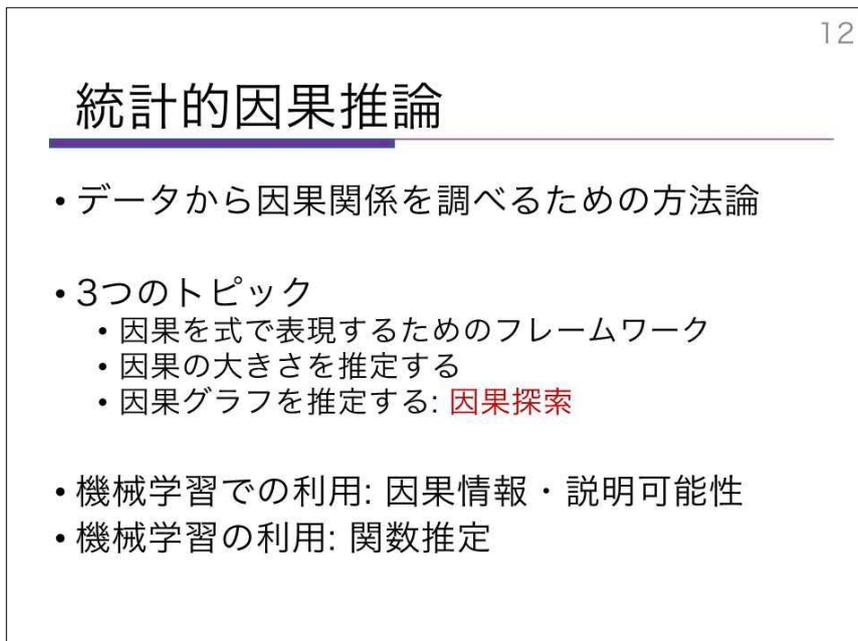


図2-5-9 統計的因果推論

ことができるが、残念ながら、現状では領域知識を用いてヒトが推定している。この場合、ユーザーは必ずしもこの因果グラフに確信が持てないので、これを「因果探索」で支援したい。

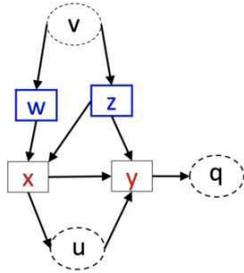
16

## 因果グラフが必要

---

- 因果の向き
- 共通原因の有無

} 因果グラフで表現



- xからyへの因果の大きさを推定するなら、wとzを観測して層別(調整)
- 現状: 領域知識を用いて推定
  - ユーザーは必ずしも確信なし

図2-5-10 因果グラフが必要

### 因果探索

データを使って因果グラフを推定する場合、現状では線形であることが多いが、それを非線形にもしたい。線形から非線形にしようという試み自体は非常に多く行われているが、いずれにせよ最初に説明した未観測共通原因に対処しなければならない。

また、因果グラフを推定する場合、必ずしも因果グラフの候補は1つに絞れるとは限らない。理論的にこの

21

## 因果探索技術の話題と課題

---

- モデルの拡張
  - 線形・非線形の軸だけではない
  - 未観測共通原因
  - 離散と連続変数の混在
  - 領域知識の取り入れ・領域に合わせたモデル
- 推定の精度向上・高速化
  - 高次元
    - 因果グラフ構造の候補が莫大
  - (小)サンプルサイズ
- 選択バイアス、欠損値、外れ値など一般的な話題も

図2-5-11 因果探索技術の話題と課題

変数間の向きはデータから一意的に決まるものなのかというのを検討しておく必要があり、これは予測とは異なる。推定されたものと同等に良いものが他にないのか調べる必要がある。

また、機械学習と因果探索の技術を融合する上で、現状では因果探索の方法が、離散のみか連続変数のみかでアルゴリズムが分かれているので、両方が混在した状況をうまく扱えるようにする必要がある。ある程度一般論が構築されても、モデルを実際に使うときには領域知識の取り入れ、領域に合わせる必要がある。

因果グラフの候補は膨大なので、それを探するためにはある程度のサンプルサイズ必要。小さいサンプルサイズでも可能にすることが重要。推定の精度も高めていかなければならない。

#### 第4の科学との関連

因果が分かるということがまず科学の興味としてある。また、何かアクションを起こすときの予測を考えてもやはり、因果を考える必要がある。因果を適切に議論するためのフレームワークや方法や標準ツールを誰もが知っており、使えるという状況に持って行って、それをもって科学を進めるというようなことが大切である。チューリング賞受賞者のパールは、これを「Causal Science (因果の科学)」と呼んでいる。

22

(1) 第4の科学(データ駆動科学)までとの違いは何か? 新たなツールのどのような機能・性能が、人間の性能を超えた(あるいは強化した)ことになるのか? しきい値のようなものがあるか?

---

- **Causal Science (因果の科学)**
  - Pearl (7<sup>th</sup> July 2020, <http://causality.cs.ucla.edu/blog/>)
  - 従来から科学の興味は因果
  - 相関ではなく因果を適切に議論するためのフレームワークや方法が標準ツール
- **領域知識+データ**
  - 人間が思いつかなかった「因果」仮説
  - 人間が迷っていた「因果」仮説のデータによる順位付け
  - 実験や調査の効率化
    - 仮説生成と実験による検証のサイクルの高速化

図2-5-12 第4の科学

人間が思いつかない因果仮説をデータから見つけるとか、領域知識だけでは解らないが、データを加味することによって、追加の情報を人間に与えることができる。それから、実験や調査を効率化するというような役割はあると思う。

例えば、ターゲットを人間が決め、それに効く介入の候補変数を推定するといった使い方がある。コンピューターを使って候補を探すということは現在行われているが、基本的に相関のある変数を探すので、それに関連しているという意味で使ったとしても、介入しているその変数の値を変えたからといって、ターゲットの値が変わるとは限らないため、目的のために正しい方法を使うというのが標準になる。

23

評価指標?

---

- ターゲットを人間が決めたとして、次の2つを比較
  - それに効く介入の候補変数(複数でも良い)を推定
  - 人間が選ぶ
- 十分な効果を生む介入を見つけるまでの実験回数
- コンピュータにより探すことは今でも行われているが、相関のある変数を探しているのが、介入してもターゲットの値が変わるとは限らない

図2-5-13 評価指標

24

(2) 科学による知のフロンティアはどこまで、どのように広がっていくのか? ツールを使い理解を広げていく方向とは別に、人間が理解できない科学のような方向にも広がるか?

---

- 変数が高次元になると人間には理解が難しくなるだろう
  - 因果グラフの全体像
  - 因果効果が大きい状況
  - 複数の変数への介入
- 「信じる」ために
  - 仮定の透明性
  - 理論保証
  - 統計的信頼性評価
- 領域知識のみで因果仮説を生成するのは限界がある

図2-5-14 知のフロンティアの広がり

変数が高次元になると、理解は難しくなり、因果グラフの変数が増え、因果グラフを人間が理解するのは困難になり、結局「因果効果によるランキング」のような理解になると思う。人間が何らかの方法で信じられるよう、パールは仮定を透明化することが必要、仮定を表すために因果グラフを使う、と言っていると思う。因果グラフが一意にデータから決まるものなのかどうか調べておく必要があると言ったが、それは理論保証

で、反対向きの因果グラフも同じぐらい良いということであれば、最初に出てきたグラフを、その向きが正しい向きと信じて、何かアクション起こすというのは間違いにつながるので、そういうことを防ぐために理論的に調べておく必要はある。

領域知識のみで因果仮説を生成するのは限界があるので、因果探索の技術が発展してきたら、それを使って人間の判断を補うということは必要だろうと思う。

この分野の科学技術への貢献は、仮説生成機能、因果に関する仮説で、どういう条件、状況で、どういう変数がどの変数にどのぐらい効くというような仮説を生成する役割、それは人間が文献とか論文を読んで当たりをつけて、実際に実験することで生命科学は進んできているが、その辺りを補助できると思う。

25

(3) 新しいツールは、科学技術における人間の活動のどの部分を代替することになるのか？ 具体的に期待される問題設定・用途や利点は何か？

---

- 仮説生成機能
  - 従来: 人間が文献を読み考える
  - その時点での領域知識とデータによる「自動」因果仮説生成
- 実験の効率化
  - 因果に関する仮説を因果推論のフレームワークで生成
  - 相関解析で代替はできない
  - Rightなツールで因果に関する仮説を生成
- 生命科学
  - 仕組みの理解と利用
  - 実験可能
- 制御(ターゲットの値を変えたい)
  - (予防)医学
  - 製造業
  - 農業

図2-5-15 科学技術への貢献

**【質疑応答】**

Q：2つのフレームワークとはどのような考えか？

A：いわゆる統計系のルビン流とパール流という2種類のモデルのことである。ある集団に薬を飲ませたときに、効果があるかどうか調べるためには、この集団全員に薬を飲ませたときと、集団全員に薬を飲ませなかったときの治癒の割合を比較すればよい。しかし、全員に薬を飲ませると、薬を飲まなかった場合のデータが取れない。そこで、もし薬を飲んでいたら（あるいは飲まなかったら）どうなったか、という実際には起きなかったことを想像する必要が生じるが、それを反実仮想と言う。ルビンはこの2つの状況に2つの確率変数、つまり薬を飲んで治癒するか、薬を飲まずに治癒するかという確率変数を新たに用意した。

一方パールは、Xが薬でYが治癒したかどうかを示した図で、介入前はこういうふうデータが生成するという方程式を最初に考えておいて、薬を全員に飲ませるということは、薬を飲むかどうかを表す変数の値を1に設定することだと考え、薬を飲んだときに治癒するかどうかを、Yの分布、治癒するかどうか

14

## 因果推論の数学的フレームワーク

- Rubin流とPearl流
- 因果を式で表現
  - 「同じ状況だが、薬の有無だけが異なる場合」を表現
  - 異なる介入

**反実仮想モデル**

- もし集団全員に薬を飲ませた場合
- もし集団全員に薬を飲まなくてもらった場合

**doという介入を表す記号を導入**

- 介入do(x=1)をした集団と
- 介入do(x=0)をした集団を比較

If  $p(y|do(x=1)) \neq p(y|do(x=0))$ ,  $x$  causes  $y$

図2-5-16 2つのフレームワーク

28

## 介入をした集団とは

- 介入前のデータ生成過程を方程式で表す

$$x = g_x(z, e_x)$$

$$y = g_y(x, z, e_y)$$

構造方程式

因果グラフ

- 介入do(x=1)すると新しい集団ができる

$$x = 1$$

$$y = g_y(x, z, e_y)$$

構造方程式

$p(y | do(x=1))$

因果グラフ

図2-5-17 フレームワークの表現

かの分布と表現した。

ルビンは方程式を考えずに、2つ記号を用意して状況を表現し、パールは、新しいデータの生成プロセスを作ったという違いがある。しかし、表現している内容は同じであって、どこまで遡るかが違う。現状では、両方を必要に応じて使い分けている。

- Q：機械学習と今の統計的因果推論の関係は、機械学習と関係ない相関を見つけて、その中から何かフィルタリングするような考えではなく、注目したい因果、この部分に因果関係があるということを最初に人間が想定して、それを先ほどのような数学的な記述で確認しつつ、その結果からグラフを組み立てていくということか？
- A：機械学習との関係から言うと、まず因果関係があって、その結果として相関が生まれてきていると考えている。機械学習は因果を調べないので、相関のレイヤーだけ見て、予測モデルを考える。因果推論は因果関係のモデルを使って、候補を出し、元の候補の中でデータに矛盾しない候補を絞る。
- Q：因果推論のために機械学習を活用する、あるいは、機械学習の結果に対して因果推論で絞り込むといった機械学習の使い方についてはどうか？
- A：機械学習でできることは、例えば病気のリスクを単に示すだけでなく、因果推論を組み合わせると、検査項目の1カ所が変わると、そのことが他の項目にどのように影響を及ぼし、結果としてリスクはこのぐらい下がるという追加情報を与えることができる。また、関数関係の推定は機械学習で技術が進んでいるので、因果モデルの構築にも貢献できる。
- Q：図2-5-7では、最終的には $X_1$ だけ分かれば良いように見えるが、 $Y$ を予測するには $X_2$ も $X_3$ も $X_4$ も必要か？
- A：機械学習との関係から言うと予測精度を上げるという意味では、 $X_2$ も $X_3$ も $X_4$ もあったほうが良いが、1個変えるならどこを変えるとリスクの予測値が変わるかという議論をするときは、この図ならば $X_1$ を変えると良いとは言える。しかし、 $X_1$ は必ずしも変えやすいものとは限らないので、そのときは変えやすいものの中で一番効果の大きなものを探すという事はあり得る。したがって、 $X_1$ だけあればいいということを導くためではなくて、この場合だと、我々が知りたいのはどうすると病気のリスクが下がると予想されるかということなので、そのときにどこを変えるとよいかを提案するのが因果推論の役割である。
- Q：そうすると、因果モデルの $X_3$ 、 $X_4$ というのは、 $X_1$ 以外に、 $X_2$ も従属することなのか？
- A：例えばこの $X_1$ から $X_2$ の矢印は、 $X_2$ が完全に $X_1$ だけで決まるという意味ではなく、 $X_2$ 固有の何かがあり得る。したがって、固有の何かも含めたほうが予測精度は当然よくなる。この図では、完全に決まるのではなく、今取り上げている変数の中での因果関係を扱っている。
- Q：因果関係を推定する上で相関の強さというのはどれぐらい加味するものなのか？
- A：原理的には相関の強さは関係ない。相関が強いから因果が強いとも言えないし、因果が強いから相関が強いとも言えない。相関と因果に何らかの一定の関係があるわけではない。従って、因果を知るのが目的ならば、必ずしも相関が大きいから因果効果も大きいという推測は理論的にはできない。
- Q：推定した因果の正確性はどのように実証するのか？
- A：推定した因果関係の正しさは、実験しなければ示すことができない。予測であれば、新しいアルゴリズムを作り、良い予測ができれば、「性能は上がった」と言えるが、因果の場合は実際に介入して実験しなければ示せない。検査値と疾患の例で例えると、因果関係を示すには、実際に介入群、非介入群にわけて、数年後に健康状態を測定する必要がある。しかし、そのような実験は実現できない場合が多いので、代わりに出てきた因果グラフの中でこの部分は領域知識に合っているから、ほかの部分も信じるといった評価をして代替することが多い。因果推論の評価ということであれば、実験するしかない。

Q：因果グラフの候補が複数あった場合、どこを実験すべきか分かるか？ またその信頼度、例えばこの相関関係のうち何割が因果でしか説明できないということは解るか？

A：どういう順番で実験をしていくと一番効率よく因果グラフを明らかにできるかという研究はある。生命科学やマーケティングの分野が一番多い。また、因果的には正しいけれども、相関は十分に説明できていない、あるいは取り上げている変数ではうまく説明できてないということは起こり得る。

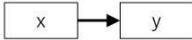
19

## 基本アイデア

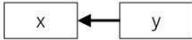
1. 因果グラフに仮定をおく
  - 非巡回
  - 未観測の共通原因なし(すべて観測されている)
2. 仮定を満たす構造の中で、データと(最も)つじつまの合うグラフを選ぶ

3つの候補

(a)



(b)



(c)



「データでxとyが独立」なら、一番右の(c)を選ぶ  
(a)と(b)の区別がつかない(一意に決まらない): **同値類**

図2-5-18 基本アイデア

因果探索の使い方をユーザーにすぐ納得してもらうことは簡単ではない。最初に仮定を置き、候補を絞る。2変数しかない場合、XからYへ矢印、YからXへ矢印、YとXの間に矢印がないというのが候補になる。このような領域知識と仮定から候補を決めてデータと突き合わせる。図2-5-18の(a)と(b)では、XとYの間に矢印があるので、XとYは相関しているはずである、つまりXとYは相関しているはずだということが理論的に導かれ、実際にデータを見て相関の有無を確認する。(c)が正しかった場合には、XとYは相関がない。だから、因果モデルを考えて、この候補だったらこういう相関関係が現れるはずだというのが理論的に導けて、その理論的に導けたものが実際に合っているのかどうかをデータで照合する。データで相関があれば(c)は捨てて、(a)と(b)が残る。もし相関がなければ、(c)だけが残る、という方法で因果探索では候補を絞る。だから、相関を説明できれば良い、という意味ではなく、相関をたくさん説明するということに対応するような基準というのは、つじつまが合っているかどうかを探ることが対応する。したがって、どのくらい信じるかというのは、最初の仮定が、どの程度納得できるかということにかかってくる。

何が原因かを決定するプロセスは、それぞれの因果モデルから導かれる相関関係と矛盾しているかチェックして、合わないものを捨てていく。実際にデータの分布を見てどちらのほうに矛盾していないかというのを調べて、矛盾していないほうを採用する。アクセスできるのはデータだけなので、データ

2.5

データ駆動による  
因果仮説探索

が生み出された背景についてモデルを考え、一番良いモデルを探す。因果の向きを調べるためには、相関だけでなく、分布全体を調べる必要がある。領域知識とデータから得られる情報の両方を用いて、より良い因果仮説を探していくことが重要である。

Q：未観測共通原因について。因果仮説の探索は、未観測であっても、実験的な探索で見つかるかもしれないといった仮説生成に因果グラフは使えるか？

A：潜在共通原因の存在を、場所とともにフィードバックすることはできる。しかしそれが何かは類推するしかない。

## 2.6 Creating the Engine for Scientific Discovery

北野 宏明 (ソニーコンピューターサイエンス研究所)

### ノーベルチューリングチャレンジ

科学的発見のためのエンジンを作ること (Creating the Engine for Scientific Discovery) を考えている。機械学習は道具として使うが、本当にやろうとしていることは、科学の自動化 (Automation of Science) で、自律的なAI科学者 (Autonomous AI Scientist) を作ろうと思っている。ターゲットとしては、2050年までに生理学・医学分野でノーベル賞級の科学的発見をできるAIシステムを作るというのを掲げている。ノーベル賞級という言い方をするかは、少し品がない表現のようにも思って迷ったが、大きな科学的発見というのでは何かが分かりにくい。AIの分野で科学的発見の研究は昔からやられていたし、ケプラーの法則をもう一回発見したという話では、実用的な技術としては全然駄目で、やはりノーベル賞級くらいのインパクトのある発見をAIがほぼ自動的にするというぐらいの踏み込んだ目標にせざるを得ないと考えた。このグランドチャレンジ目標については、2016年にAI Magazine誌<sup>1</sup>、人工知能学会誌<sup>2</sup>に論文を書いたし、日経エレクトロニクス誌<sup>3</sup>などにも寄稿した。もう4年も経ったので、現在、フォローアップの論文を書いているところだ。

ところで、AIがノーベル賞を獲ると言い始めた頃、ノーベルの遺言には「the person」と書かれているので人間じゃないと駄目だと丁寧に教えてくれた人がいた。もちろんそれは分かってはいたのだが、そのあたりが気になる人もいると分かった。それでノーベルチューリングチャレンジ (Nobel Turing Challenge) というふうに言い方を変えた。要するに、ノーベル賞委員会が人間だと思って賞を与えたら、実はAIだったということである。そこまでいなくても、ノーベル賞委員会が少し気をつけようというふうになったら事実上の目標達成だろうと思う。

ここで、人かどうかは簡単に分かってしまうかという、実はそうでもない。数学の分野で有名な「ブルバキ数学原論」という本は、当初、ブルバキという人が本当にいると思われていた。ところが、実際はブルバキという人はいなくて、20人ぐらいの若手数学者がこの名前で論文や本を書き続けていたということが後で分かった。

それは昔の話じゃないかと思うかもしれないが、今、サトシ・ナカモトはどうだろう。もしブロックチェーンやビットコインでノーベル経済学賞を授与しようとしたら、サトシ・ナカモトしかあり得ないと思う。でも、サトシ・ナカモトは誰なのか分からない。AIではないと思われるのは、今まだAIはそこまでできないという共通認識を皆が持っているからであって、AIが自分で発見できるようになり、サトシ・ナカモトのようにブログか何かに論文を載せて、それがすごかったらどうなるか。生身の人間かどうか、会ったことがあるかというのは、もはや大きな問題ではないし、ステイホームの時代になってしまったので、生身の人間に会うこと自体が少なくなってしまったから、なおさらである。

- 1 Kitano, H., "Artificial Intelligence to Win the Nobel Prize and Beyond Creating the Engine for Scientific Discovery", AI Magazine, Spring 2016.
- 2 北野宏明、「人工知能がノーベル賞を獲る日、そして人類の未来一究極のグランドチャレンジがもたらすもの」、人工知能学会誌、2016年3月号。
- 3 「ノーベル賞級の発見をするAI、人の限界を超えた科学研究へ」、日経エレクトロニクス誌、2016年7月号。

**Grand Challenge:**  
By 2050, develop AI systems that can make major scientific discoveries that worth Nobel Prize in Physiology and Medicine.

**Two sub-goals:**  
First, we will need to develop an AI and robotics system that can perform biomedical and biotechnology research fully autonomously that leads to major discovery.  
Second, the machine must be able to make strategic choice of the topic of research, communicate in form of publications and other means to explain the value, methods, reasonings behind the discovery, and their applications and social implications.

図2-6-1 グランドチャレンジと2つのサブゴール

### 2つのサブゴール

このグランドチャレンジ目標に2つのサブゴールを掲げた(図2-6-1)。その1つは、AIとロボティクスで非常に大きな科学的発見をするというもので、もう1つは、それが自律的な判断をして人間のようなコミュニケーションができるというものである。

実は今これを見直し中で、今思っているのはAutomation(自動化)とAutonomy(自律性)である。たぶん最初のゴールは科学的発見の自動化、つまり、科学的発見を自動で行うエンジンを作ることだと思う。もう1つは、科学的発見の自律化で、自律的に何を研究すべきかを決められて、どうしてそれに価値があるかを説明できるというものである。人間の価値観に沿って価値のあるものを研究テーマとして戦略的に選んでいけるということである。

皆あまり気にせず大発見かどうかを言っているが、実はここには人間社会における価値基準が非常に明確に入っている。例えば、線虫の病気を治すという発見ではできるかもしれないし、実際マウスで病気はいろいろ治っているけれど、多くの場合、それ自体で大きな発見とは言わない。人間のがんが治るとか、コロナが治るとかいうならば、それは大きな発見だと言われる。マウスでの発見もそれが、人間に当てはまる可能性が高いとなると重要な発見とみなされ始める。また、基礎的な発見でもそれは、実用的なものではないにしても、純粋な学術上の価値というものも含めて人間の価値観のどこかに結びついていることが大きく評価される。だから、1つ1つの科学的発見の価値の大小を評価するのは非常に難しいことで、まずは、なるべく信頼度の高い発見をたくさんする馬力が必要になると思う。そこで見つかったものが大きな発見かという重要性の評価はたぶん別の問題で、そこには人間の価値観が入る。ここで、自律的に大きな発見をするというものは優秀な科学者の資質なので、科学的発見をするAIにもそれが求められることになる。なので、ノーベルチューリングチャレンジで人間かどうか区別できないという先ほどの話は、当初は半分しゃれたかったけれど、よく考えると、実はそこに科学的発見の本質があるのではないかと最近思うようになった。

したがって、図2-6-1にはまだそこまで書いていないけれど、科学的発見AIではAutomationとAutonomyが非常に重要になると思っている。

### 生物学における科学的発見の現状

では、なぜこういうことが必要になってくるかについて話していきたい。

私の最初の生物学の研究は、細胞老化（Cellular Senescence）に関するもので、1990年代に現在ワシントン大学セントルイスで教授をされている今井眞一郎さんと一緒にやった<sup>4</sup>。細胞老化に関わり、細胞分裂の回数が上限に達するというヘイフリック限界（Hayflick Limit）がある。当時はオミクス的な手法はまだできなくて、個別的に遺伝子発現を見るようなことがされていたが、シミュレーションで解明しようとした。今になって考えてみると、生物学的な仮説というのは、結局、分子機構（Molecular Mechanism）とインタラクションから決まってくるものだ。つまり、図2-6-2に示したように、生物学的な現象の説明は、基本的には、どのような分子機構が働いて、そのどのような組み合わせが可能で、そこにどういう分子が関わっていて、分子機構間の相互作用や細胞間の相互作用があるからだということになる。そこに、実験的な制約や生化学的な制約やシステム論的な制約がかかってくる。

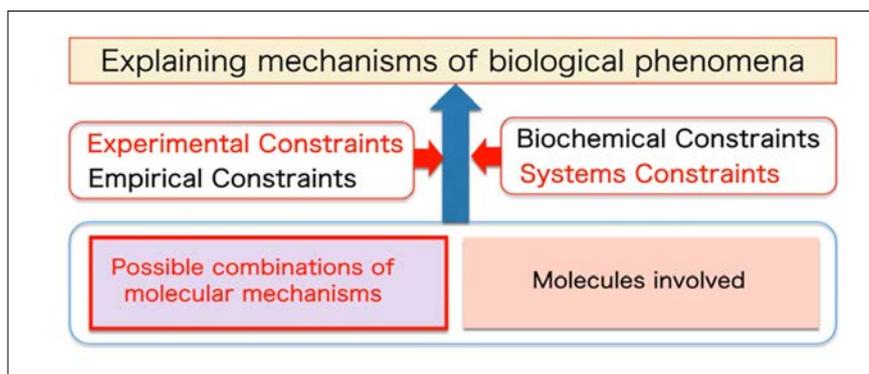


図2-6-2 生物学的発見の構造

我々は、こういう組み合わせに対して、実際に膨大な数のシミュレーションを実行して、どういう条件だったら細胞老化のデータに全部合うのかをひたすら調べた。Unbiased（バイアスのない）と言っていいようなシミュレーションを半年くらい行った。その結果、割と正確なものが出て、数学的に説明できるモデルが1つだけ存在することが分かった。それを今度は生物学の面からその制約を満たす物質は一体何なのかを検討して、Sir2蛋白がそれを満たすだろうということが分かった。それで、今井さんが実際に実験して確認するとNAD依存性ヒストン脱アセチル化酵素であることが分かって、これが抗老化遺伝子Sirtuinsにつながった。

このようにしてSirtuinsファミリー遺伝子群が見つかって、その後、今井さんはマウス実験などを進め、2016年から人間での影響を見る臨床試験に至った。要するに20年かかった（図2-6-3）。ところが、この間の中で非常に体系的・網羅的に取り組んだのはComputational Modelingのところだけだった。それ以外

4 Kitano, H. and Imai, S., "The two-process model of cellular aging," *Exp. Gerontol*, Aug; 33(5):393-419, 1998.  
 Imai, S. and Kitano, H., "Heterochromatin islands and their dynamic reorganization: a hypothesis for three distinctive features of cellular aging," *Exp. Gerontol*, Sep; 33(6):555-70, 1998.  
 Imai, S., et al., "Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase," *Nature* 403(6771): 795-800, 2000.

のところはたぶんこうだろうといった直観による旧来科学のやり方をしている。今考えれば、我々は運がよかったということになる。途中の仮説の組み方やマウス実験などで違うことをやっていたらここに至っていない。

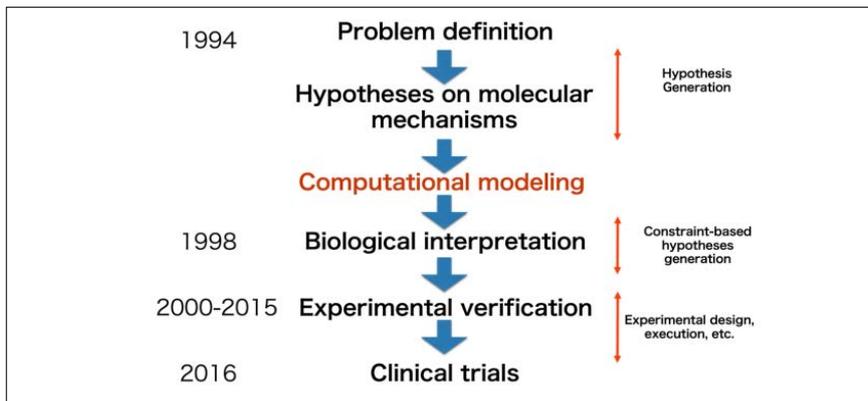


図 2-6-3 生物学的発見の経緯

そういうことを考えると、科学的発見というのは、まだ産業革命以前のレベルだと思う。オミクスのシーケンサーやマイクロアレイなどいろいろな解析ツールが使えるようになって、データ取得のスループットが向上してたくさんデータが取れるようになった。しかし、結局、仮説生成と検証のプロセスは人間が、ああでもない、こうでもないとやっている。それが自動化されない限り科学のスピードは向上しないだろう。逆にデータが多い分、皆が混乱してしまっているというところがある。

科学的発見のプロセスについては、セレンディピティー (Serendipity) や、偶然性 (By Accident) や、科学的直観 (Scientific Intuition) といふことが言われる。これは結構運任せで、その背後の原理はあまりよく分かっていないということだ。本当にやるべきなのは、セレンディピティー、偶発性、科学的直観というが、その本質はいったい何なのかを考えること、科学の科学 (Science of Science) ということだ。

また、システムバイオロジーを提案してから20年やってきて分かったことは、システムバイオロジーは人間がやる学問ではないということだ。昨年 OIST (沖縄科学技術大学院大学) でシステムバイオロジーの学会 (ICSB 2019) を開催したとき、これを言ったら皆のけぞったのだけれど (笑)。なぜそう思うかというと、やはりシステムの複雑性が非常に高いし、データ量が膨大になってしまって、それを人間の頭で考えようとしてもとても追いつかない。

それで、システムバイオロジーの分野ではどうしているかというと、1つは、モデルの大きさはある程度限定したところで、精密に定量的にやっていたいモデルを作ることで生物学的発見をするというもの。限定された範囲で、質の高い研究が進められている。もう1つは、大規模なデータに基づいてある程度統計的な解析と機械学習を用いてこうだろうと機能推定などをするもの。どちらかというトラディショナルなゲノミクスやバイオインフォマティクスに近くて、あまりにも大規模複雑でメカニスティックなシステム論ができていない。1個1個はそれほど大変ではなくても、ものすごい数の分子インタラクションに対して、ちゃんと実験してパターン同定をするのは非常に大変になる。そこに研究者の人生をかけさせるなんていうのは現実には不可能だと思う。また、モデルも複雑になりすぎて、しっかりとしたもの構築できない。

### 科学的発見における認知的問題とAI活用

その背景には人間の認知能力の限界がある。2016年の人工知能学会誌の論文にも書いたが、科学的発見において図2-6-4のような認知的問題が起きる。

1. 情報の地平線問題 (Information Horizon Problem)
2. 情報ギャップ問題 (Information Gap Problem)
3. 表現形記述の不正確さ問題 (Phenotyping Inaccuracy Problem)
4. 認知バイアス問題 (Cognitive Bias Problem)
5. マイノリティーレポート問題 (Minority Report Problem)

図2-6-4 科学的発見における認知的問題

生命科学分野の論文数は年間150万本を超えていて、これは1日当たり4100本ということになる。人間が読めるペースを大幅に超えている。

また、例えば出芽酵母シグナル伝達系 (Yeast Signaling) のマップ作成は、数千本の論文を読んで情報を抽出・整理するのだが、最初に作ったときは論文になるけれど、そのアップデートが進められない。そのアップデートも論文にできないことはないかもしれないが、その論文のインパクトは高くないので、ポストドクに1年かけて取り組んでもらうような話にはならない。いま我々はこれをAI技術で自動アップデートするシステムを作っていて、何とかかなりそうだと思っているが、10~20年前には人力でやるしかなかった。

他にもあるが、図2-6-4にあげたような問題は、要するに人間の認知限界によるもので、これを解決するにはAIの力を使うのがよい。

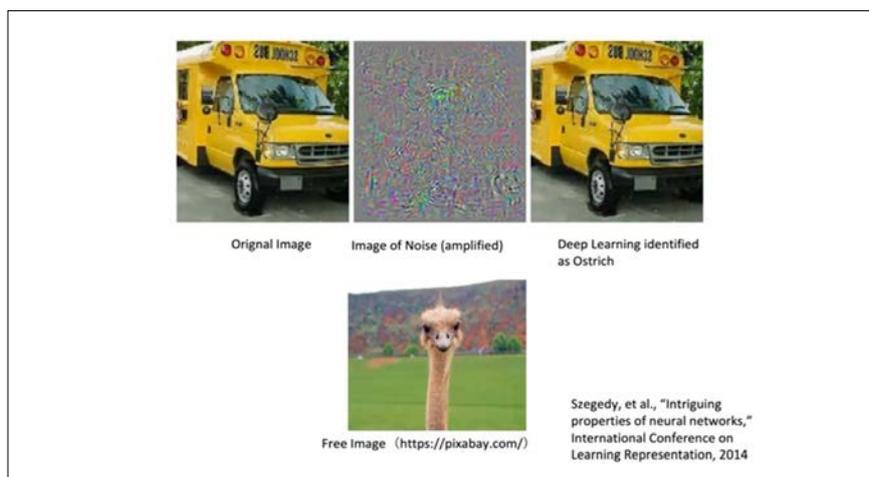


図2-6-5 Adversarial Examples

図2-6-5は深層学習のAdversarial Examples（敵対的サンプル）で有名な例<sup>5</sup>で、左の元画像は人間にも深層学習にも黄色いスクールバスだと認識できるものだが、真ん中のノイズ画像を重ね合わせた右側の画像は、人間には同様にスクールバスと見えるが、深層学習はダチョウだと認識してしまう。実は真ん中のノイズ画像がシステムティックに作られていて、よく見ると、下のようなダチョウの顔の特徴が出ているようにも見える。この例は、深層学習をだますことができる例として示されることが多いのだが、別な見方をすると、右側の画像について、人間にはダチョウの特徴が出ていることを見抜く力はないが、深層学習はそれが見抜けたということになるかもしれない。このような実験の場合、ノイズパターンに意味のある画像のようなものが見えたりは限らない。なので、この特徴は、アーティファクトとみなす部分と、人間に見えないパターンを見出しているのかは、さらに深く研究する必要がある。

生物学の分野では、マウスの行動から遺伝子異常のレベルを推定するという研究がある<sup>6</sup>。ハンチントン病は、遺伝子異常のCAGリピート（ハンチンチン遺伝子中のシトシン・アデニン・グアニンの繰り返し配列）の数が多くなると行動異常が生じるということが分かっていたが、逆に行動異常から遺伝子異常を推定することは皆が試みていたが、できていなかった。この研究では、マウスの行動を24時間ずっと計測するとともに、ゲノム配列解析を行いCAGリピートのデータを取って深層学習にかけると、行動異常から非常に高い精度で遺伝子異常のレベルが推定できたというものだ。生物学分野でフェノタイピング（ある生物のもつ遺伝子型が形質として表現されたフェノタイプを計測すること）に深層学習を使うというのは非常に革命的なものだった。

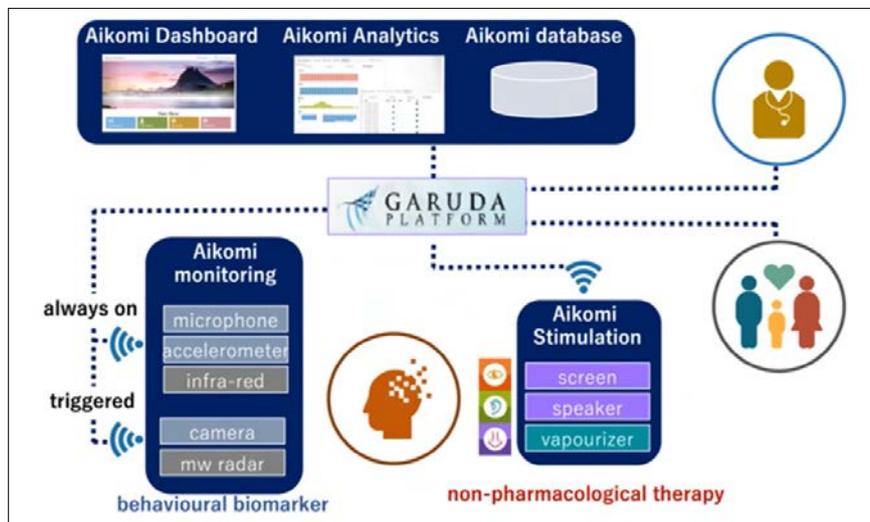


図2-6-6 Aikomi：認知症の非薬物療法

5 Szegedy, C., et al., “Intriguing properties of neural networks,” International Conference on Learning Representation, 2014.  
 6 Ruderfer, D., Dudley, J., “Deep phenotyping predicts Huntington's genotype,” Nat Biotechnol, 34:823–824, 2016. Alexandrov, V., Brunner, D., Menalled, L. et al., “Large-scale phenome analysis defines a behavioral signature for Huntington's disease genotype in mice,” Nat Biotechnol, 34:838–844, 2016.

我々はこの考え方を使ってAikomi<sup>7</sup>というベンチャーを作った。この会社は武田薬品のスピンアウトだが、我々（SBX<sup>8</sup>）も出資していると同時に、技術的サポートもしており、Garuda Platform上にサービス基盤が構築されている。認知症に対して非薬物療法を行う。そのため、行動療法を用いるが、症状が改善したかを正確・客観的に評価するために深層学習を使っている。従来のアンケート方式では粗すぎて、中枢系・認知症の客観的評価に限界があった。

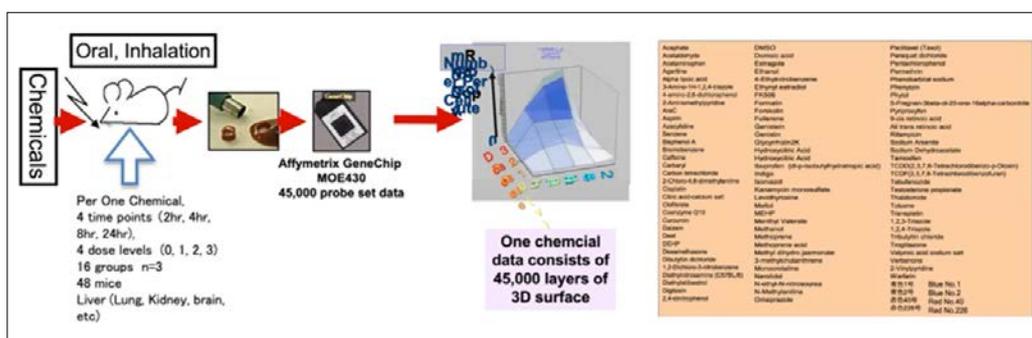


図2-6-7 Percellomeプロジェクト

また、国立医薬品食品衛生研究所の毒性部とPercellomeプロジェクトを行っている<sup>9</sup>。マウスにいろいろな化学物質を使って、臓器ごとに単回投与や複数回投与でさまざまな遺伝子の応答を全部見ることで、どの化学物質のときに何がアップして何がダウンしたか、何が不安定化したかを知りたい。一個一個見ることは簡単だが、組み合わせが膨大で、何千万というケースを調べないといけないため、実現できなかった。これに深層学習を適用した結果、ほぼ自動的に膨大な数のアップ/ダウン、不安定化した/ほぼ変化なしといった分類ができるようになった。

### Search and Optimization

科学的発見AIの原点としては1960年代のDENDRAL<sup>10</sup>が挙げられる。最近では、Ross KingらのAutomation of Science<sup>11</sup>がある。彼らのAdamというシステムは、出芽酵母の遺伝子とその機能について簡単な仮説を立て、それを検証する実験を設計して実行できる。

私がやろうとしているのは科学的発見の再定義で、科学的発見というのは基本的に、仮説空間のMassive Search and Verification (大規模な探索と検証) に基づいて進めるのがよいと思っている。ノーベル賞を獲

7 <https://www.aikomi.co.jp/>  
 8 <https://www.sbx-corp.com/japanese> 代表取締役社長：北野宏明。創業・生命科学の研究開発、そのためのソフトウェアプラットフォーム（Garuda）の提供等を行っている。  
 9 Kanno, J., et al. "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays," BMC genomics 7(64), 29 Mar. 2006.  
 10 Lindsay, R. K., et al., "Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project," McGraw-Hill, 1980.  
 11 King, R. D., et al., "The Automation of Science," Science 324(5923):85-89, 03 Apr 2009.

た科学的発見においても、Search and Optimization（探索と最適化）のプロセスで理解できるものもある。

図2-6-8の左、山中先生のiPS細胞の発見もSearch and Optimizationだった。まず、リプログラミング(細胞の初期化)というターゲットを立てて、FANTOM DBから24の遺伝子をスクリーニングしたところで、24遺伝子を全部入れて試した。そこから1つ抜いて実験して(Leave-one-out)、「山中ファクター」と呼ばれる4つを特定した。全部入れるというのは普通やらないと思うので面白いし、これが鍵だったと思う。24を見つけたところがSearchで、Leave-one-out実験のところがOptimizationで、この発見プロセスは典型的なSearch and Optimizationだと思う。こういった類のステップは、割とストレートに自動化できそうに思える。

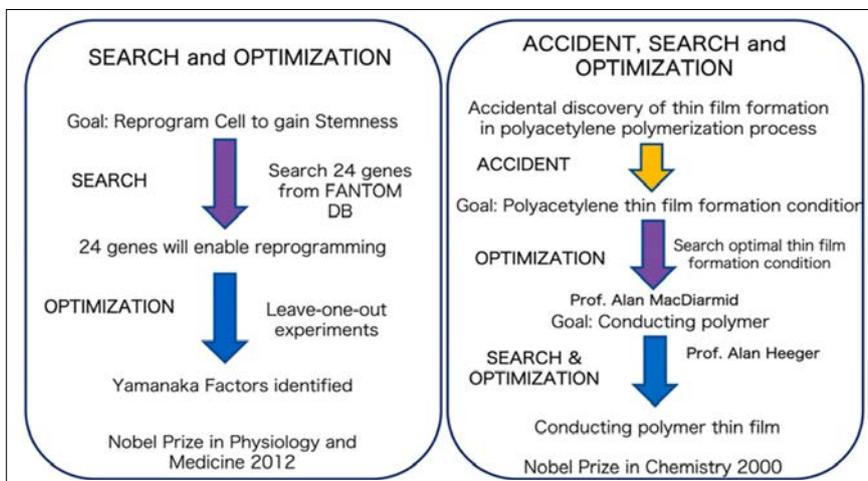


図2-6-8 科学的発見の事例におけるサーチ&最適化

もう一つは図2-6-8の右、白川先生のConducting Polymer（導電性高分子）の事例で、こちらはSearch and Optimizationのプロセスを2段階やっている。まず、ある留学生在が新フィルムで薄膜を作る経験をしたというやって何かすごく変なのが出来てしまった。実はその留学生在が溶液の濃度を1,000倍ぐらい間違えたというアクシデントだったわけで、要するにもともと全く考えてなかった遠いSearchポイントだった。その結果、今度はできることが分かったからOptimizationして非常に安定的に薄膜ができ、誘電性があるものができるようになった。ここまでが1段目で、今度はConducting Polymerをやろうという、非常に明確な次の目標が定まって、どういう条件かをSearch and Optimizationしたというのが2段目になる。その結果、Conducting Polymerという薄膜の伝導膜にたどり着いた。

いろいろなケースを調べたが、かなりのノーベル賞級の発見でSearch and Optimizationが行われている。ただ、このSearch and Optimizationでチャレンジングなのは、Searchポイントが非常に遠い、非常に離れたものをつなぐということだ。まだノーベル賞を受賞していないが、CRISPR-Cas<sup>12</sup>が古細菌と化膿連鎖球菌のコネクションから生まれたというのは、その典型例だ。

12 このセミナー講演は2020年8月に行ったが、その後、10月にゲノム編集技術CRISPR-Casの業績で、エマニュエル・シャルパンティエとジェニファー・ダウドナがノーベル賞を受賞した。

**現在の取り組み**

ノーベルチューリングチャレンジで目指しているのは、全ての結果がノーベル賞級という意味ではなく、AIを使って実行した結果の中から非常に大きなものはノーベル賞級になればよい。科学的発見AIにもやはり得意/不得意がある。

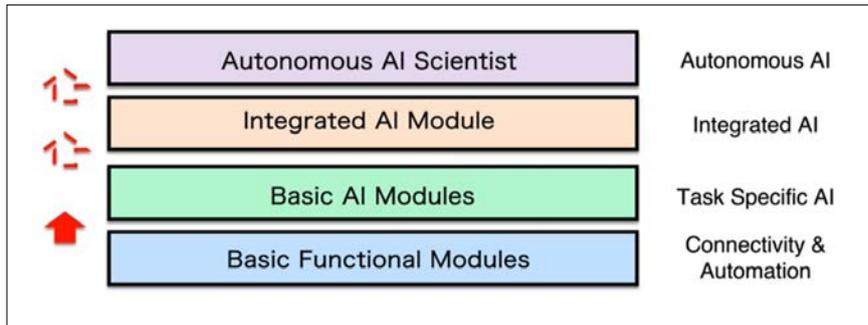


図2-6-9 科学的発見AIプラットフォームの技術スタック

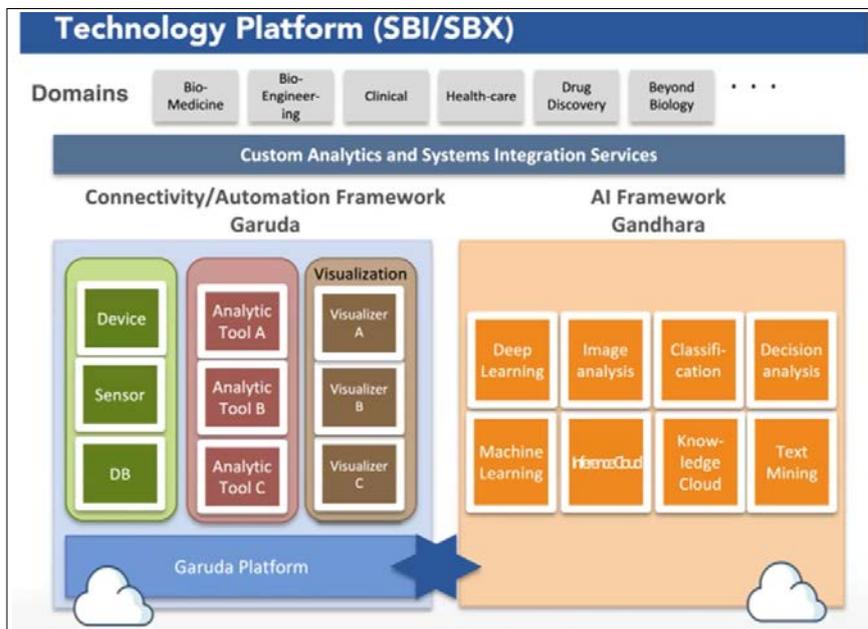


図2-6-10 Garuda PlatformとGandhara AI Framework

それで、我々はこういう取り組みを進めている（図2-6-9）。まずBasic Functional Modulesで、この全てに機材のコネクティビティーやソフトウェア解析のパイプラインがないといけない。まずコネクティットラボを作るということが非常に重要だ。その上にBasic AI Modulesを乗せて、AIがツールとして使えるようにする。その上のIntegrated AIは、Basic AI Modulesを半自律的に組み合わせて新しい実験計画を立てる。最後に、Autonomous AIを研究して、人間の価値観に合わせて自律的に重要なトピックを選ぶことができる

ようにする。このようなプラットフォームを作ろうと、SBI<sup>13</sup>ではコネクティビティーと自動化のプラットフォーム Garuda や AI フレームワーク Gandhara を開発している。

いくつかの機器メーカーとコラボレーションが進んでいるが、製品発表までした事例として島津製作所のマルチオミクス解析パッケージ<sup>14</sup>がある。これにSBIからはGarudaプラットフォームや高度な解析ソフトウェアを提供している。他の事例として、自動走行車の学習データ作成等への応用や、薬剤誘発性不整脈の解析への適用等もある。

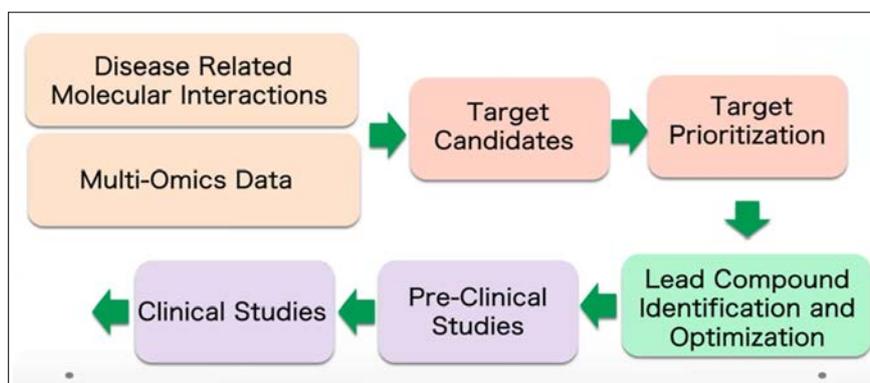


図 2-6-11 Very Simplified Process of System-Driven Drug Discovery

では、そのようなプラットフォームの上でAI技術がどう使えるかだが、ERATO河岡プロジェクト<sup>15</sup>では、Garuda上に解析パイプライン（図2-6-11）を構築して、各ステップにAI技術を組み込んだ。まず、疾患関連分子間相互作用（Disease Related Molecular Interaction）のステップは、ウイルスが侵入して複製して、そこを出るサイクルを表すネットワーク関係図を作るのを、以前は文献をもとに全部人手で作っていたが、今は自動的に作れるようになった。自然言語処理による知識抽出技術で、およそ95～98%ぐらいまで自動抽出できるようになった。今、Covid-19関連のプロジェクトでこれを拡張しながら使っているので、あのときこれをやっておいてよかった。人手で作った部分と、自動抽出した部分があるが、人手の部分にはやはり人間の癖がある。そこには3つの可能性がある。1つ目は、文献に書いてあるが人間は見落として、AIは発見したというケース。2つ目は間違いだったというケース。3つ目が面白くて、文献には書いてないが、その後で発見されたというケースで、実はこれが想像以上に多かった。我々の知識抽出システムは、文献に明に書かれていない関係を推論して仮説として出すので、3つ目のケースのようなことができる。例えば、2000年までの文献から我々の知識抽出システムが出した結果と、2000年および2010年までの文献に出ている結果を比較すると、2000年までの文献に出ている結果と、我々のシステムが推論して新たに出した結果について、その後の10年間で実際に確認されたものがたくさんあるということだ。要するに、我々のシステムで、そういう仮説生成がで

13 特定非営利活動法人システム・バイオロジー研究機構（The Systems Biology Institute）、会長：北野宏明  
<http://sbi.jp/index.html>

14 <https://www.an.shimadzu.co.jp/lcms/tq-option/multiomics.htm>

15 ERATO河岡感染宿主応答ネットワークプロジェクト [https://www.jst.go.jp/erato/kawaoka/index\\_j.html](https://www.jst.go.jp/erato/kawaoka/index_j.html)

きるようになってきたのだと思っている。

こうして分子間ネットワークが再構成されると、制御理論的にどこをターゲットにするとよいか分かる。つまり、ウイルス複製サイクルを止めたいわけなので、そのためには分子間ネットワークのどこをターゲットにするとよいか、ネットワーク構造から分かる。その一方で、実験から取得されたオミクスデータ（図2-6-11のMulti-Omics Data）からも、どこをターゲットにするのがよいか分かる。つまり、ウイルス複製サイクルの生物学的プロセスとウイルス関連分子の関係からも、ウイルス複製を止めるポイントが見つけれられる。それで、制御理論的に求めたターゲット候補と、オミクスデータから求めたターゲット候補で、オーバーラップしているところが非常に有望だということになる。

これがターゲット候補のステップ（図2-6-11のTarget Candidates）で、次のステップであるターゲット優先付け（Target Prioritization）では、OISTのチームがDocking Simulationを大規模に行った。OISTではいろいろあるDocking Simulationツールを買って試したが、商用のもアカデミックフリーのもあまり精度が高くなかった（精度50%以下）。しかし、それぞれに得意／不得意があるので、得意なところだけうまくピックアップできると非常に精度が高くなり得ることが分かった。そこで、機械学習を使って、どういう条件だったら、どのツールの精度が高いかを予測できるようにしたら、精度が82%まで上がって、ほぼ実用レベルになった。それを使って、先ほどのターゲット候補とケミカルライブラリーとで、全対全のDocking Simulationを行って、その結果を詳しく評価して上位候補をターゲットとして絞り込んだ。その後さらに、河岡先生がパイプラインの最終検証を行っているところである。

SBIとツムラで、インフルエンザによく効くという麻黄湯の解析も行って、いろいろ面白い結果が得られている<sup>16</sup>。

これらは、まだ自律的な発見というところまで至ってはいないが、その前段階として、全体プロセスをつないで、その途中のそれぞれのステップを機械学習とか制御理論とかで武装しておく必要があるということである。

### これからの取り組み

今OISTでやろうとしているのは、人間とロボットが協調するラボだ。高度に自動化したCovid-19の検査ファシリティーを作って、PCR検査、抗体検査、T細胞エピトープ検査をやるようにする計画である。そのために単腕・双腕のロボットをいろいろ導入する予定である。

図2-6-12は、東京医科歯科大学の田中先生との共同研究<sup>17</sup>で、タンパク質間相互作用ネットワークに基づいて創薬ターゲット分布を調べたものだが、面白いことが分かる。左側のグラフの横軸は相互作用の数Kで、縦軸はそのKのバンドの遺伝子群がどれくらいの確率で創薬ターゲットになっているかを表している。それで、Kがだいたい5~20あたりのところが高い確率を示しているが、Kが20以上になると薬になる確率がすごく減っている。Kが大きい遺伝子は、一般的に重要だと考えられるのだけれど、副作用が多過ぎて駄目なのだと思う。

16 Nishi, A., et al. Deconstructing the traditional Japanese medicine “Kampo”: compounds, metabolites and pharmacological profile of maoto, a remedy for flu-like symptoms. *npj Syst Biol Appl* 3, 32 (2017).

17 Hase, T., Tanaka, H., Suzuki, Y., Nakagawa, S., Kitano, H., “Structure of Protein Interaction Networks and Their Implications on Drug Design,” *PLoS Comput Biol* 5(10): e1000550, 2009.

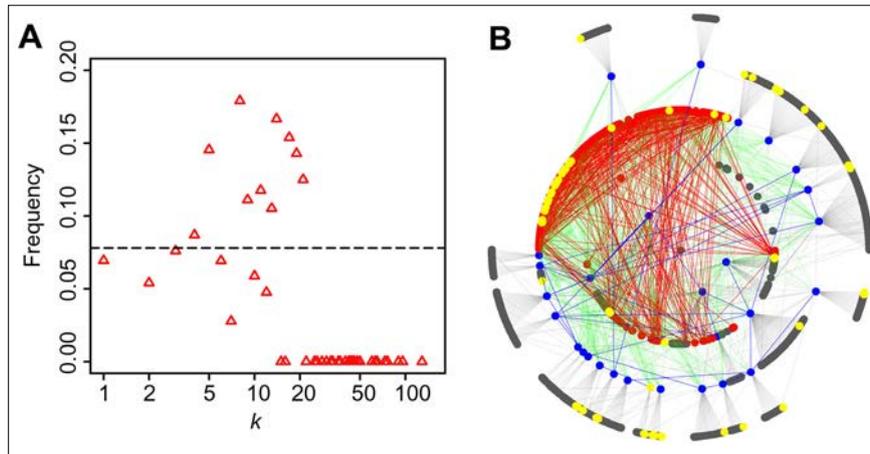


図2-6-12 創薬ターゲットの分布

逆に、Kが1とか2とか相互作用がすごく少ない遺伝子は、意外に創薬ターゲットになりやすいというのが分かってきた。そのKが1や2の遺伝子そのものの数はものすごく多い。つまり、一般的には重要だと思われていなかったロングテールのところに、たくさん創薬ターゲットがありそうだと分かってきた。しかし、数はたくさんあっても、1個1個の確率は高くないので、そこを調べるためにはコンピューターで自動化しないと無理だということになる。それで、我々のGarudaプラットフォームでAutomationとUnbiased Searchをするというオープンコラボレーションをやっている。Garuda上にパイプラインを構築して、それで動くレシピを用意している。レシピの通りに解析フローや実験フローが自動実行される。そこでできたものの検証するプラットフォームも作っている。

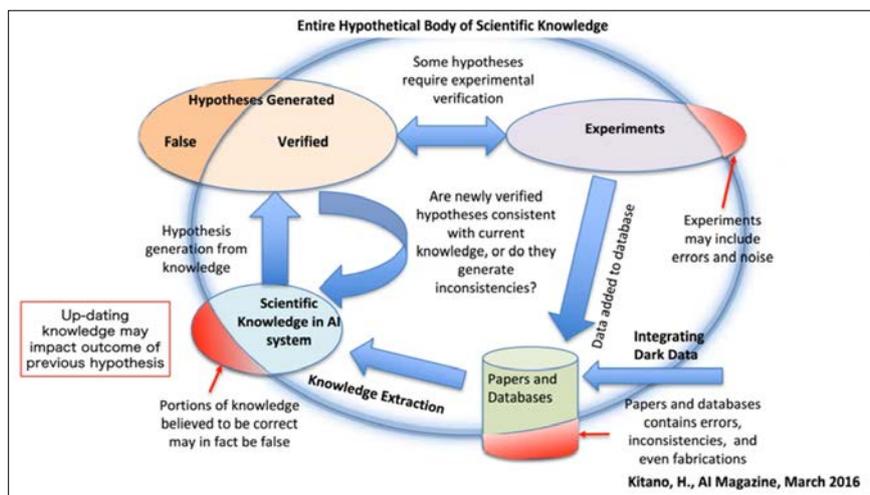


図2-6-13 仮説生成とその反証・検証プロセス

最終的に重要なのは全体の流れだ。既に存在する論文や知識ベースをもとに仮説を作って、それを検証して、実験して、修正するというサイクルを回す (図2-6-13)。もとにする知識が全て正しいと想定できるなら

ば楽だが、実際には、間違っただけのものも含まれているし、記述が全部書かれていないものもある。そういう非常に不安定な知識の集大成から推論しながら精度を上げていくというサイクルを回さないといけない。これをトワイライトゾーン推論と呼んでいる。トワイライトゾーンのように正しいものと間違っているものが混在していて、どれか分からないようなところだが、全体としては正しい方が多かったり、そういう中での中央値というのは意外と正解だったりとかいう感じがしている。こういうサイクルを回すシステムを作っていくといけないし、そのための理論も必要だと思う。

### 科学的発見 AI の可能性の広がり

AlphaGoは深層学習と強化学習を使っている。深層学習によって、過去の棋譜を全部見て、現在の局面で次の手を非常に高い精度で推測する。さらに、AlphaGo同士で戦わせて、いろいろな局面でモンテカルロサーチを実行して、すごい数の状況評価をした結果、どういう盤面で何が最善手かという学習を対局前にしてしまっている。

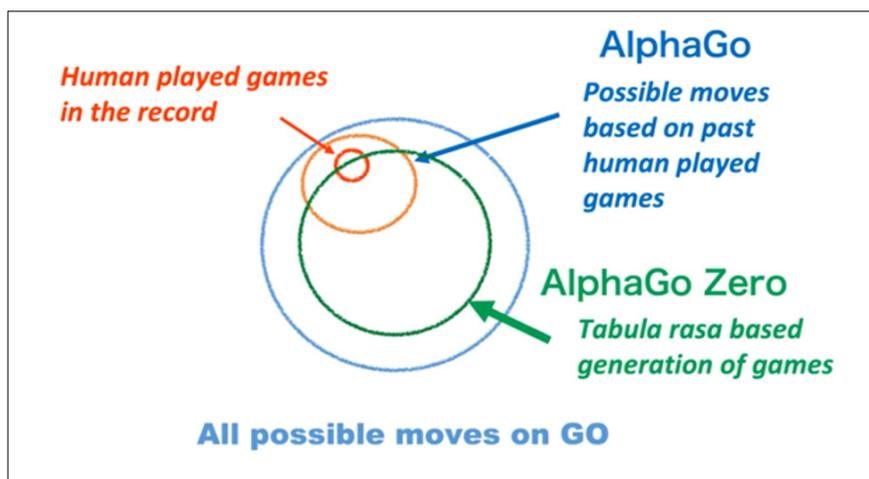


図 2-6-14 AlphaGoとAlphaGo Zeroの探索空間

図 2-6-14 のベン図で示すと、一番外の青い円は碁で可能な全ての動き、膨大な盤面状態を表している。棋譜にあるのはそのごく一部、小さい赤い円のところで、AlphaGoはこれを学習して、さらにモンテカルロサーチをやるので、その周辺（オレンジ色の円）まで全部学習していることになる。イ・セドルらプロ棋士にAlphaGoが勝てるのは、彼らが見てない世界を全部知った上で対局しているから。第4局でイ・セドルが勝ったのは、AlphaGoが事前に見ていなかったようなものすごく意外な手をイ・セドルが打ったのだと思う。続いて開発されたAlphaGo Zeroは、棋譜を学習せずに、最初からランダムサーチと強化学習で探索しているので、さらに広い可能性（緑色の円）を見ている。

ここには非常に面白いインプリケーションがある。これまで人間が打っていた碁は、碁というゲームの最善手にはなっていない、人間が打ちやすい手を打っていたにすぎないのかもしれない。AlphaGo Zeroは、ある程度Unbiased Searchを行ったことで、もっと強い碁の打ち方が見えてきたということだ。

科学的発見（Scientific Discovery）の現状もたぶん同じで、我々が今知っているものは人間が発見でき

る範囲であって、人間には発見できていない法則性というものもまだまだたくさんあるのに、我々はそれに全くアプローチできていないということなのだと思う。それで、Unbiased Searchをすることによって、人間にはなかなか見つけることができない科学的発見が可能になるのではないか。

ただ、AlphaGoの場合は盤面状態の評価がかなり明確だが、生命科学・自然科学での科学的発見はそうはいかない。先ほど話したようなアプローチでは、AI・機械学習、モデリング/シミュレーション、高精度実験（High Precision Experiments）が必要になるが、この高精度実験のところの問題になる。仮説に対する評価をどうやってスピードアップするか。最後は実験を実施しないといけなくて、そこが律速になるが、そこに行く前に、どこまで探索空間を絞り込めるか。そのときに、既存データや新たなデータ取得だけでなく、仮想的なデータ生成（Virtual Data Generation）が必要になって、その有効性確保も重要だと考えている。

ここで改めて科学的発見について話すと、Asking Right Questions、正しい問いを問うことが重要だと言われる。これは現状そうだと思うが、こう言われる背景として、科学者のキャリアにおいてアクティブに研究できるのが30年くらいというように、人間には限りがある中で重要な問題にフォーカスすべきということだ。そのような限界がないのであれば、全ての問いを問えばよい。そうしたら、その中に重要な問いと重要な答えが入っている。ただし、その中で何が重要かについては、科学的に重要だということと、社会的に重要だということがあるので、ソーシャルバリューが関わってくる。それで、前にも述べた通り、仮説空間におけるMassive Search and Verificationだということになる。

ただし、ソーシャルバリューを入れるというところの議論は非常に重要で、OISTで開催した国際会議ICSB 2019のときもすごい議論になった。例えば、現時点では重要でないけれど、もしかしたら将来重要になるかもしれない仮説があって、それが現時点では重要でないために削られるような評価関数になってしまうとか、小さな規模でいろいろな人がいろいろな研究をするのが重要だと主張するのならば、各々の発見の価値判断は事前にはできないはずだとか。

そうすると、Unbiased Searchをするという結論になる。解きたい問題が明確なときは、そこからサブゴールを設定して解いていけばよいけれど、それとは異なり、あらかじめ想定してないような目標をボトムアップに浮かび上がらせようというときは、Unbiased Searchをせざるを得ない。Unbiased Searchを高速に効率的に実行するにはAutomationをしないとイケない。人間がやると、どうしても成功させようという意図が働

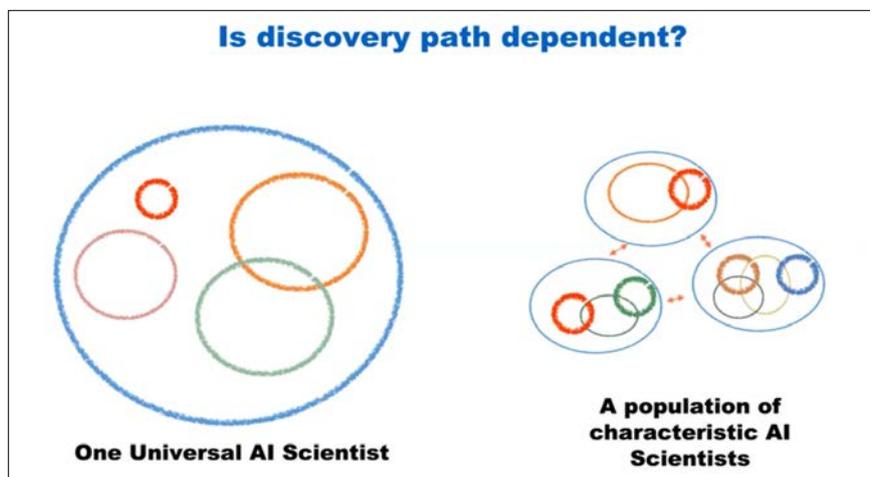


図2-6-15 仮説空間の探索戦略

くため、Unbiasedにはならない。その意図はいいのだけれど、現実には、正しくない選択をしてしまうことが多い。

また、仮説空間のUnbiased Searchの戦略にもいくつかある。まず考えられる戦略は、1つのユニバーサルなAIシステムが全空間をランダムに探索するというもの（図2-6-15の左）だが、別な戦略として、キャラクターを持った複数のAIシステムがインタラクションしながら探索するというもの（図2-6-15の右）が考えられる。この大きな探索空間をどのようにして探索するのが効率的なのかという問題である。もしキャラクターを持ったAIがよい成果を出し始めたら、人間の成功した研究者に似た論文の出し方をするかもしれない。そうすると、いよいよもってAIか人間かの区別がつかないだろうと思う。

科学的発見のプロセスについて、セレンディピティー、偶然性、科学的直観ということが言われると最初の方でも話したが、今日話したような取り組みから、その本質が分かってくるだろうと考えている。その取り組みの中で機械学習はツールとして使うが、私がやろうとしていることの本質は、科学の科学（Science of Science）であり、科学的発見の新しい形態を作りたいということだ。

#### 【質疑応答】

Q：科学的発見にソーシャルバリューを入れるかどうかは極めて興味深く、特に生物医学で重要だと思うが、自然科学ではそうでない場合の大発見の方が多いのではないか。例えばティコ・ブラーエはそれ以前の数百年に及ぶ天文学データを整理し、そのデータに基づいてケプラーはケプラーの法則を発見した。それに基づいてニュートンはニュートンの法則を導き、その上に立ってアインシュタインが一般相対理論を発見した。このいずれの場合もソーシャルバリューはほとんどなかったと思う。こういうことを考えると、分野によってどういう価値観に基づいて科学を評価するかは分かれてくるように思う。

A：それはその通りで、数学、物理の素粒子論などもソーシャルバリューはあまり関係ないように思う。生物医学もすごい基礎のところはそれでいいかもしれない。医学になるとソーシャルバリューが入ってくる。そして、ソーシャルバリューの側からやり過ぎると問題解決は非常に効率的になるが、科学的な大発見はできなくなる気がする。全然関係ないところの発見がないと最後の重要なピースが埋まらなくて、Unbiased Searchをやった方がいいということになるかもしれない。科学的発見のシステムができて、ソーシャルバリューからゴールドリブンを動かした場合とボトムアップに動かした場合との比較実験ができれば非常に面白い。

Q：そのバリューをどんどん人間に近づけていくと、人間的なバイアスがかかってしまうという懸念も出てくるのではないか。

A：その通りで、バイアスがかかってつまらない話しか出てこないというのもあり得る。意外なものをつなげるUnbiased Searchがやはり重要で、山中さんのケースはリプログラミングを可能にする遺伝子24個から4つに絞ってという、わりとストレートフォワードなケースだが、CRISPR-CasのケースはUnbiased Searchを考える典型例になると思う。古細菌と化膿連鎖球菌の研究者が会ってCRISPR-Casが生まれるという、重点化した領域からは生まれないコネクションだ。幸いにジェニファー・ダウドナは、その経緯を本に書いているので、CRISPR-Casのケースはとてもよいベンチマークになると思っている。

Q：最近話題のOpen AIのGPT-3は、大量の事例データを取り込んで、そこから関連する事例を引き出すアプローチかと思っていたら、実は溜めた中にいろいろなつながりができていて、ある種の推論のようなものができてしまっているようにも思う。すると、今回の科学的発見にもつながるような気がする

が、どう見ているか。

A : 私も GPT-3 の中に何か内部構造・フレームのようなものが作られていると思う。それを我々が解釈できるようにするメカニズムが欲しい。我々が今まで生み出した知識の構造から、我々がまだ見えていない知識の構造を推定・推論して生成することができるようなフェーズに、AI のフロンティアは入りつつあるということだ。GPT-3 と我々のシステムを比べてみるとか、我々の持っているデータや論文情報等を GPT-3 の仕組みに入れてみるとか、面白いと思う。

Q : 北野さんがやられている生命科学・システムバイオロジーと比べて、マイリアルズ・インフォマティクス (MI) の分野はどう見られているか。MI ではデータがなくて困っているようだが、科学的発見のアプローチが変わってくるか。

A : MI は生命科学と比べてプリンシプルのところは明確な気がする。ただ、生命科学の分野も、文献情報や画像情報は結構あるけれど、それ以外はデータがあまりないという問題は MI と変わらない。例えば、製薬会社の前臨床試験や基礎探索段階のデータは少なくして深層学習は使えない。化学に近いところやオミクス的なところはデータがあるが、それ以外は無理で、特に生体に絡んだところは、動物実験をコントロールすべきという話もあって、やはりデータは集まらない。なので、生体を使わずに生体に近いところがやれる技術がいろいろできてこないとスケラブルにはならない。製薬会社の中では、マウスと人で同じ結果が出ない薬は怖くて使えないからやらないと言っているところもあり、それも一つの考え方だと思う。データドリブンだけで進めるのには限界がある。

Q : 北野さんがノーベルチューリングチャレンジを提唱してから、国際的にも取り組みが立ち上がってきていると思う。最新の注目動向はどうか。

A : 2020 年 2 月にロンドンで AI Scientist Grand Challenge Workshop が開催された。英国のアラン・チューリング研究所 (The Alan Turing Institute) の主催で、英国・米国・日本からの参加が中心で、私も参加してきた。具体的なプログラム<sup>18</sup>がスタートするので、私もアラン・チューリング研究所を兼任するように頼まれた。生物医学・環境科学・材料科学が中心に議論されているが、他分野もあると思う。米国・英国のファンディング機関も興味を示している<sup>19</sup>。

Q : 海外で戦略的な研究投資が始まっており、CRDS からも戦略提言を出すべきと考えている。日本の作戦や国でどうファンドしていくかについてアドバイスをお願いしたい。

A : これは本当にしっかりやるべき。科学的発見のエンジンを作った国が科学においてすごい競争力を持つから、米国はナショナルセキュリティのための高優先度の手段だと認識して、真剣に取り組もうとしていると理解している。日本も技術立国として復活を目指すのなら、科学的発見の効率化・自動化は必須だ。これから人口が減って研究者も減るといふときに、新しい科学の武器を作らなければ勝ち目がないと思う。どうやるかについては、ゴールを明確にすることと、科学的発見のサイクルを自動化するインフラを構築することが重要だ。個別研究で画像認識や機械学習のツールを作っておしまいではいけない。Automation と Autonomy の両方をやって、仮説を生成して検証するサイクル全体を回す仕

18 The Alan Turing Institute: The Turing AI scientist grand challenge  
<https://www.turing.ac.uk/research/research-projects/turing-ai-scientist-grand-challenge>

19 "Voices from DARPA" Podcast, Episode 31: Science 2.0 でもノーベルチューリングチャレンジや AI Scientist の話題が取り上げられている。<https://youtu.be/FdDsNqbGpsw>

組みまで構築しなくてはならない。

- Q : Unbiased Search をする際、何の空間でサーチするか自体は、現状、人間が選ぶしかないと思う。そこには人間のバイアスがかかってしまうという理解でよいか。
- A : Automation のレベルではそう。このエリアで Unbiased Search をやるという、ある種のツールができて、それだけでも大いに使える。ただ、その先に Autonomy、何をやるべきかを AI が決めるというレベルがある。そこにはたぶんいろいろな考え方が出てきて、哲学的な議論にもなり、非常に面白いチャレンジになると思う。
- C : やはり科学的発見というものは一律に言えないところがあると感じた。物理学や数学では、根本原理を発見するときに物事を単純化するとか、要素に分解するというのがかなり必須の作業になる。他方、生物学では、要素を分解すると生命体を失う可能性があるわけで、複雑な対象物のまま扱わないといけない。そこには人間が全てを認識することができない面があり、機械学習が非常に力を発揮することになる。

## 2.7 細胞計測解析を人から解放し、人知を拡張する

太田 禎生（東京大学／シンクサイト株式会社）

今日はいただいたテーマを受けて、細胞計測にAIや機械学習を取り込むことにより、どうやって人の限界から解放して、どうやったら人知を拡張するところまで進められるか、という我々の取組に関してお話しさせていただく。

### 自己紹介

私は東大で修士中退した後、アメリカで機械工学の博士号を取った。東大ではマイクロ工学、アメリカでは光物理を扱っていた。アメリカでPh.D.を取って帰ってきた後、JSTのさきがけで採択していただいて、そこで機械学習駆動型・融合型の細胞解析技術を開発した。これを基にシンクサイトという会社を起業して、現在そちらでCSOをやりながら、2年前に東大の先端研で研究室を構えて、さらなる新技術をつくっているところである。

研究領域は生命計測で、多岐にわたる専門技術領域を持っている。例えば、マイクロ流体・液滴の技術、メカや電気を組み合わせて並列化や自動化を行なう工学技術、あとはマテリアル合成、ゲノム工学などを組み合わせて研究を行なっている。光イメージングや物理量計測の技術を中心に、そこに他の技術を融合させてきたが、ただくっつけるだけでなく、つなぐことで価値をつくるということを大事にしている。

### 従来のサイトメトリーの問題点

まず、JST・さきがけのポスドクの時に開発した、光イメージングと機械学習とマイクロ流体、工学を組み合わせた新しいサイトメトリー技術を紹介する。サイトメトリー（Cytometry）は、細胞を意味するサイトと測定を意味するメトリーを組み合わせた用語で、数多くの細胞を一つ一つ計測・解析した結果を基に、分類・分取して細胞を研究するためのツールである。例えば、細胞の表面にある分子を既知のバイオマーカーと呼ばれる別の分子で細胞を染め、輝度や大きさといった情報を測定・カウントし、それを解析する。得られた定量的な結果に対して、ある仮説に基づいた基準に見合ったような細胞を取ってくるといったことを行なう。従来のサイトメトリー技術は、人の仮説と検証に基づいて使われてきており、たくさんの細胞の中から人が決めた基準に当てはまる細胞だけを探し出すという作業になっている。

サイトメトリーはわらの山から針を探す作業に例えられる。ただ、人間がどんな針があるか知らないし、針を見分けるためのマーカーの存在も知らないことから、実際にはわら扱いされた細胞が本当は針である可能性がある。開発当初に考えたのは、私たちのバイアスを超えてサーチする機械学習のエンジンが作れば、私たちが考えているよりも私たちが欲しいものにマッチしたアルゴリズムを作れるのではないかと。それにより、人による仮説と検証に縛られていた細胞計測の考え方を改めて、細胞から計測されるデータの価値を本質的に生かせるのではないかと考えた。

### ゴーストサイトメトリー開発のきっかけ

ゴーストサイトメトリーは、機械学習を駆使することでイメージフリーで画像解析を行ない、その結果に基づいて細胞の解析を行う技術であり、AIを融合することで精度と速度で人を超える計測マシンを開発した。

開発のモチベーションは、細胞の画像情報に基づいて、たくさんの細胞を見て選択的により分けていきたいというニーズに発している。例えば、血中にがんの組織から循環がん細胞が漏れ出すと転移を引き起こす可能性があると考えられている。循環がん細胞は血中の細胞の中ではレアであるが、これを見た目で探し出し、分離して、DNAシーケンスなどに回せると、早期診断や高精度な診断、治療方針の決定というところに役立つと考えた。別な例として、最近になって細胞治療、細胞製造分野が立ち上がってきているが、体に入れる実際の細胞を選ぶため、細胞を染色せずに、見た目で見えない細胞か、がん化し得る細胞かというのをチェックし、安全な細胞だけを取ってくる必要がある。

従来のサイトメトリー技術の代表格として顕微鏡があるが、非常に情報リッチな反面遅い。ソートまで行なうと非常に遅く、なかなか実用になかない。一方、フローサイトメトリー技術は、細かい管を高速で流れる細胞の定量計測を一個一個行なって、その計測結果に基づいて細胞を選別可能な技術だが、こちらは高速でソートできる反面、細胞1個で平均化された情報しか取得できず、細胞当たりの情報量は少ない。情報の質とスピード、処理可能な量を両立させた技術というのが存在せず、先に述べたニーズに応えられなかった。私たちは、従来のフローサイトメトリーに対して高速イメージングと機械学習を有機的に結合することでニーズを満たそうと考えた(図2-7-1)。具体的には、1万分の1秒で画像を計測、10万分の1秒以下で選別を行うというゴールを立てた。

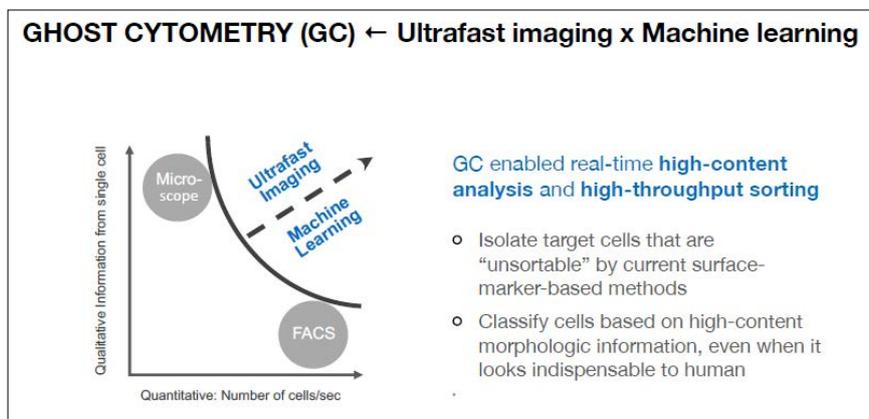


図2-7-1 ゴーストサイトメトリーの目指すところ

### ゴーストモーションイメージング

一番最初にゴーストイメージングをベースにゴーストモーションイメージングという技術を作った。一般的なカメラは、奥にピクセルが並んだアレイ型の素子がついていて、外の空間情報を中の素子上の空間パターンとして結像するのに対し、ゴーストイメージングは1画素の素子でこの空間を捉えるイメージング技術である。今までのゴーストイメージングのやり方だと、サンプルに対してランダムな照明光パターンを時系列方向に多数回与え、それぞれの光と物体の相互作用をシングルピクセルの信号情報として繰り返し撮影する。照明パターンがわかっているならば、時系列の強度情報から2次元画像の再構築が可能である。弱点は、数十ヘルツ程度のプロジェクターで少なくとも1,000回は投影しないと1枚の画像を再構成できない、非常に時間がかかることである。その代わりに我々は、サンプルの動きを利用して撮像することを考案した。構造照明を当てた流路中に蛍光標識した細胞を流すと、流れに伴い別々の場所が励起されていく。高速で流れているため普通の

2.7  
細胞計測解析を人から  
解放し、人知を拡張する

カメラだとぼけてしまうが、シングルピクセルのPMT（光電子増倍管）で時系列に撮ると、波形の強度データが得られる。つまり、細胞内の蛍光体分布パターンと励起光の強弱パターンの重ね合わせが、細胞の動きとともに蛍光シグナルの時間方向の強弱パターンとなって圧縮して変化する（図2-7-2）。もともとどういう光を投影していたかというのを知っていると、元の画像（=蛍光体分布パターン）を再構成できる。細胞を2色に染め分けて、2色別々に2つのPMTで撮れば、マルチカラーのイメージングもできる。

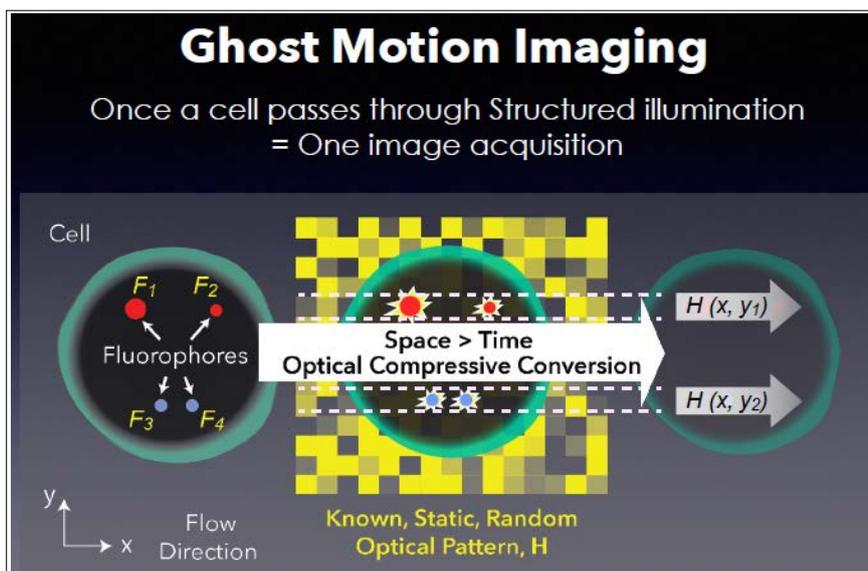


図2-7-2 ゴーストモーションイメージング

### イメージフリーでの情報処理方法

ゴーストサイトメトリー技術のコアはイメージング技術ではなく、イメージフリーのイメージングサイトメトリーを可能とした情報処理のやり方にある。従来のイメージ解析手法は、アナログのデバイスから得られる信号を計算器内で2Dや3Dのイメージに再構成し、そこにある基準を当てはめて解析する。この情報処理において時間とお金がかかるのは、画像を作るプロセスである。一方、ゴーストモーションイメージングにおいて、圧縮センシングで撮られた波形の中に画像を再構築するための情報はもう入っていることから、わざわざ2Dと3Dの画像形式に変換しなくても、波形データに直接機械学習を充ててあげれば、画像再構築用にかかっていた時間もコストもかからず、データサイズも小さいため、いいんじゃないかということ考えた。

教師あり学習のアプローチに関して説明する（図2-7-3）。波形データを取得する際に、ground truthとなる識別マーカーをつけて、同時に測定する。それにより、この波形を発している細胞はマーカーがポジティブかネガティブか、というラベルづけを行なうことができる。このペアになったデータセットを大量に集め、そのデータを基にモデルを作ると、波形からマーカーを直接予測することが可能であると考えた。例えば、同じ細胞に対しミトコンドリアを染めたものと細胞質を染めたもので輝度がマッチする場合、従来のフローサイトメトリーのように蛍光の総量を測ろうとすると見分けがつかないが、ゴーストサイトメトリーで波形を学習・分類させると分けられる。また、測定は1万分の1秒以下なのに加え、判別も10万分の1秒以下に達した。

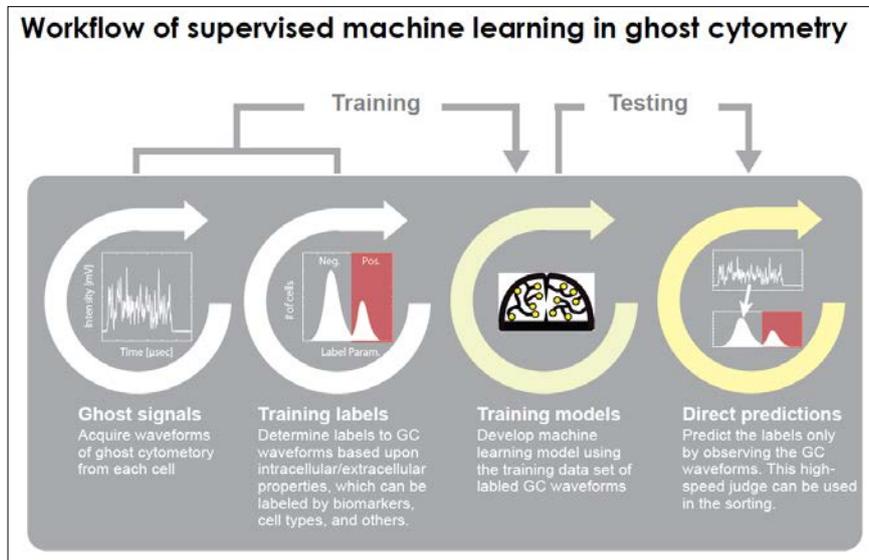


図2-7-3 ゴーストサイトメトリーにおける教師あり学習モデル適用のプロセス

ここにソーティング技術を組み合わせることで、AI 駆動型のイメージング情報識別セルソーターを実現した。マイクロ流体のデバイス内で細胞が流体力学的にフォーカスされ、構造照明を通して出てきた波形をFPGAに組み込んだ機械学習モデルが直接分類し、その結果に基づいてピエゾデバイスが目的のポジティブな細胞をソートする、という流れである。一番最初は手作り感満載であったが、会社をつくって技術移転してプロのエンジニアに任せ、さらに日立と一緒に組むことで、より安定した箱ものとしても上市するという状態まできている。

### ラベルフリー・ゴーストサイトメトリー

海外発のCAR-Tという細胞治療法が日本で上市されたが、人の体から取ってきた免疫細胞を改変して再び体に戻すというプロセスのため、体に戻す上で抗体付加や染色はできれば避けたい。また、1人当たり数千万円かかるということでニュースになったが、安価にしてたくさんの人に届けるためには、今後AIの活用などによる生産の自動化が重要になると考える。CAR-Tを始めとした細胞治療や診断の分野への展開も考え、ラベルフリーのゴーストサイトメトリーも開発した。学習のワークフローは蛍光の場合とほぼ一緒だが、ラベルフリーでは、構造照明による光の散乱や回折を検出し、細胞内の空間に分布する光学的特性を計測している点異なる。

蛍光とラベルフリーのゴーストサイトメトリーの2つを開発したが、それぞれに見合ったニーズがあると考えている。蛍光や発光が生命科学での計測においてスタンダードとして君臨しているのは、非常に高い特異性とS/Nを持っているためである。興味のある分子や細胞内構造が分かっているときは、その情報だけを性能よく取り出したいことから、手段として蛍光が非常に合っている。ゴーストサイトメトリーにおいては、例えば細胞のあるオルガネラや構造だけを標識するツールを使ってトレーニングデータセットにすると、例えば、核の波形や細胞質の波形をあらかじめ学習させることにより、機械に予測を行わせることができる。一方、ラベルフリーが得意としているのは、分子標識を用いたくないケース、または、特徴が画像として我々の目で視認しにくいケース、例えばどうやって基準を決めたらよいか分からないとか、ラマンのようにラベルフリーの情報の使

2.7  
細胞計測解析を人から  
解放し、人知を拡張する

い方自体がよく分からないといったケースである。そういったケースでは、ラベルフリーの用途は広いと考えているところである。

## ネットワーク化計測

ゴーストサイトメトリーの開発中から気づいたのは、現存で発表されているサイトメトリーやイメージフローサイトも含め、本質的に人知を超えるのは容易ではないということである。というのも、トレーニングは我々の知識をベースにしており、その基準も我々が決めているからである。3年ぐらい前から、本当に知識を超える、人知を超えるにはどうしたらいいのかということを考えてきた。我々にとって細胞はいまだにブラックボックスであり、だから細胞研究を続けるのであるが、細胞計測において我々はそのアノテーションを正しくできない中で、ブラックボックスの見分け方をどうやってAIに教えたらいいのかという点を課題だと考えた。個人的には、この点は生命計測全般にあるAI適用の課題だと思っている。

今行われているAIの手法は大きく、Supervisedな（教師ありの）機械学習と、Unsupervisedな（教師なしの）機械学習の2つに分けられる。教師ありの機械学習のモデルは、例えばイメージデータを見たときに、これは腫瘍細胞Aですよ、これは健常細胞ですよ、というように人間の知識や経験、能力に基づいてラベルをつけるものであり、学習の結果出来上がるのは「非常に優秀な人間」である。一方、教師なしの機械学習では、人間の介在を減らしてクラスタリングなどを行なうことができるが、我々人間の解釈はつけにくい。今までのバイオロジーや科学の言語で読み解くことができないため、これで知識が広がっているとはなかなか言い切れないと考えている。

この2つの間を補うためには、細胞を多面的に見ることが必要だと考える。フローサイトでそれを実現するには、今までの装置設計や計測技術の設計の角度を変えなくてはいけないと考えている。例えば、100万の細胞からあるモダリティデータAとBをそれぞれ計測したいと考えた時、A、B単独の計測がハイスループットで行なえたとしても、計測装置は別々のメーカーから作られていることが多いため、多数の細胞群を対象とした一細胞計測を別々の装置を使って高速に行なうというのは難しい。これを克服するための手段としてネットワーク化計測というものを考えている。

## 【質疑応答】

Q：ゴーストサイトメトリーにおける分類に関して、入力の変数の数はどのぐらいのオーダーなのか。また、還元主義的なモデルではなく、全体を見て判断をしているという理解でよいか。

A：構造照明の数、細胞の流れに沿った構造照明の幅が規定しており、画像再構成のためによく使うのは1,300～1,400ぐらい。一方、画像情報判別を最終的な目的とした場合、画像再構成に必要なデータ数の半分とか4分の1とかでも精度が保たれることもあり、画像を使わないメリットの1つである。ゴーストサイトメトリーで生成されるモデルは還元主義的ではないと思う反面、正直根本的なボトルネックを乗り越えたという感覚ではなく、従来の機械学習でできる予測を超えているとは言い難いため、完全に還元主義から外れているわけではないと認識している。

Q：現時点でゴーストサイトメトリーの律速になっているのは何なのか。

A：ハードウェア上、サイトメトリーの速度におけるボトルネックは、ソーティングのスピード。解析スピードは10万分の1秒程度で、ワークフローの上は律速になっていない。

Q：道具によって細胞についての科学が進歩してきた歴史の中で、今回のフローサイトの開発は、顕微鏡の発見やシーケンサーの発明といったものに匹敵するのではと感じる。これら以外で、細胞学を変え

たと思われる科学の道具には何があるのか。

- A : フローサイト自体は50年ぐらい前に発明され、例えば今の免疫学の発展はフローサイトなしでは成し得なかったことから、ノーベル賞に値する発明だと個人的には思う。細胞内のコンポーネントを読み出していく上では、オミックス技術群はどれも非常に重要な技術。その分析装置としての質量分析器やシーケンサーも大きな価値がある。一方、顕微鏡やフローサイトは、細胞を総体として扱う技術だが、非破壊で見るための有力な手法の1つである。我々が開発を進める新しいフローサイト群が本当にどこまで大事かは、我々が頑張って価値を発見できるかにかかっていると思う。
- Q : 「人知を超える」ための試みについて、そもそも「人知を超える」というのはどういう意味で使っているのか。細胞の構造が非常に複雑で判断すべき情報量が多過ぎて、人が直感的に判断・分類できないのに対し、画像的には実は複雑な特徴があって、それを機械学習で見ると非常に確度が高くクラスタリングできる、といったイメージか。
- A : ほぼそういうイメージである。どういう特徴がその疾病の細胞として定義されるべきかも分からないという状況に対し、多角度からその細胞を見ることによって判断を可能にするといったことを考えている。
- Q : 機械学習の原理から考えると、分類すべき対象が持つ特徴に対する重みづけは人間が手を加えないと、逆に効率的な分類ができなくなる。そこには人間の解釈がどうしても入るため、人知を超えるというよりも、お互いに補完し合うイメージだと考える。具体的には、きちんと学習させておけば、機械は人間よりはるかに速く正確に、AかBか、もしくはAでもなくBでもないという判断をできるが、機械がAでもなくBでもないという判断をしたものを人間がよく見ると、実は新しいCというクラスかもしれない。そこで人間が補完すると、今度は機械はA、B、C、あるいはいずれでもないという分類をデータに基づき行なう。機械がいずれでもないと判断したデータを人間が協力することで判断を可能にするという、コラボレーティブな関係性と考えられるのでは。
- A : 私の発表ではまさにCを「人の知識を超える」ということで意図していたが、私のもつイメージもコメントいただいた関係性により近く、コラボレーティブという言葉がより適切かもしれない。
- Q : 画像や蛍光染色は、ある解析結果について説明可能性を与えるツールであり、今回の2次元の画像を経由しない解析やラベルフリーの解析は、説明可能性に要する労力を排除して高速化したとも捉えられる。一方で、人間はバイアスがあって、同じ画像を見ても違った判断をしてしまう可能性があり、科学の営みにおいて、特に解析については人間は障害物で不要であるとも捉えられると感じた。
- A : ゴーストサイトメトリーでは、説明可能性を完全に排除しているわけではなく、画像化して説明可能性を付与することはできるが、高速化のため全ての細胞解析に対しては行っていない、という認識。一方、ラベル付けを行なう上で人がバイアスをかけやすい存在なのは確か。定量換算できるものは、できるだけうまく計測、解析のワークフローをデザインし、機械学習を始めとしたデータ解析ツールを適用する方が最終的にはよいと思う。それにより、可能性が爆発する対象を網羅的に探索することも可能であろう。しかし、科学のドライビングフォースは結局は人の気持ちや科学者として育ててきたバックグラウンドといったものであり、その違いがオリジナリティを与えていく根源だと思っている。そのため、人が要らない科学というのは結構つまらないものになるのではという可能性も考える。
- Q : 人間が要らないかどうかという論点に関して。心理学で有名な絵で、同じ画像なのに昼間は若い女性に、夕方には老婆に見えるという絵がある。これは確かに人間のバイアスであるが、現実には両方も正しく、人間が昼間と夕方それぞれもっともらしいフィルタをかけているだけである。情報とはまさに、老婆、あるいは若き女性なのかという解釈であり、これはある意味、人間のバイアスなしには、そ

## 2.7

細胞計測解析を人から解放し、人知を拡張する

の同じ画像に対して解釈を与えることができない、情報を抽出することができないということではないか。認識のモザイクというように、現実を見る人によって非常に多様に認識されるが、それぞれの人たちは全部間違いかというと必ずしもそうではなくて、その認識の総体が実は現実を表しているのかもしれない。

- A：角度は違うが私も同じようなことを思っている。人間は細胞をどうやって見たらいいのかよく分かっていないのに対して、細胞のデータを多面的に読み解く上で、あえて人間の知覚や認知のバイアスを入れてみることは、価値があるかもしれない。
- Q：一方、専門家が見て認識するというのは、その専門分野におけるある種の確立されたウィズダムに基づいた判断であり、ある意味客観的なフィルターが入っているとも捉えられる。
- A：患者の血液を採ってきて白血球の画像をたくさん撮りためて、それを病院の専門の技師などに疾病のラベルづけをしてもらい、それを教師データとして機械学習モデルを作るというプロジェクトに参画したが、その中で興味深い事象があった。同じ疾病の判断について、ある病院の中でトレーニングされている人たちの間では合致するのに対し、病院と病院の間、組織同士で結果が合わないケースが見られた。教科書が同じであっても、我々の能力は組織や文化に応じてトレーニングされている可能性もある。
- C：人間の認知に関連して、機械学習においてスパース推定を使った例が増えているが、スパース性というのは人間の認知にも当てはまる。物事を整理する際に、現象がスパースだと思い込むからこそ人間の限られた頭でも理解できる。自然科学において、コンパクトでシンプルなモデルのほうが真理に近いという側面があるのに対し、人間の認知的なバイアスが実は歴史的に大きな発見を生み出していると思っている。スパース推定を始め、人間が持っているヒューリスティクスや事前知識をうまく使って解析に当てはめるといことは、重要な役割を担うのではと考える。

## 2.7

細胞計測解析を人から解放し、人知を拡張する

## 2.8 機械学習、数理科学に基づく健康・疾患層別化と予測

川上 英良 (理化学研究所/千葉大学)

私は、理化学研究所の医科学イノベーション推進プログラムでチームリーダーを務めている。また、クロスアポイントメントで、千葉大学大学院医学研究院で人工知能医学教室の教授及び治療学人工知能 (AI) 研究センターのセンター長も務めている。このセンターは、医学部の中に作られた日本で初めての人工知能センターである。

簡単に経歴を紹介させていただく。私はもともと東京大学医学部医学科の出身で、医師免許を取った後そのまま基礎研究の世界に入った。高校時代から数学が好きで、数学オリンピックなどに出ていたという背景もあり、大学に入った頃から機械学習を使った医学研究をやりたいと考えていた。ただ、その頃はまだ人工知能や機械学習がいわゆる暗黒期と言われていた時代でそういった研究はほぼなく、まずは基礎研究でウイルスをテーマにして分子生物学の研究を始めた。6年ほど研究を続ける中でかなり計測技術が進歩してきて、データ駆動型科学が生物学、医学でも出てきていた流れの中で、北野宏明先生の研究室でシステム生物学の研究を始めた。これも基礎生物学研究で、特に出芽酵母をモデル生物としてその応答をパスウェイのネットワークモデルで表わすといった、データ駆動ではあるが、かなり基礎的なモデルを作った。また、転写因子の制御ネットワークといったデータを世界中から集めてきて、それに基づいてネットワークを作り制御因子の機能や役割を推定するアルゴリズムを開発するといったことをやっていた。

ここまでは、まだ実際の臨床とは全然関係なかったが、2016年から理化学研究所で始まった医科学イノベーション推進プログラムの中で、実際の医療データを使って疾患を層別化、予測するというプロジェクトに携わっている。実際の臨床データは多様で、計測データもきれいなものではないため、今までのシステム生物学的な手法だとうまくいかなかった。その多様性をどう扱うかというところで、機械学習に加え様々な数理科学を使うようになった。2019年から、千葉大学医学部ではより社会実装に近いところで研究を行なっている。

### 精密・個別化医療に向けた課題

医学において、いわゆる標準治療、標準医療というものがある。これはいわゆる Evidence Based Medicine とされるもので、もっと以前はそもそもエビデンスすらあまりなく、経験的にそれぞれの医師が治療をやっていた状況であったが、分子生物学の発展に伴い1970年代ぐらいから、分子的なエビデンスに基づいて治療をやっていくという流れができた。具体的には、何か原因となり得るような分子を見つけて、それに対して薬を作り大規模な臨床試験で検証するというプロセスを経る。その結果、例えば3割の患者がよく効き、5割はそこそこ効いていて、残りの2割は全く効かないという結果であっても、それは標準治療としてガイドライン化され、ある病気の患者に対してまずこれをやるという治療になる。ただ、これは不均一な患者に対してある単一の治療をやるということである。こうすると効果が非常に限られて、副作用のリスクもありコストも高くなる。今までは、効かなかったら次の薬を試せばいいというような試行錯誤で治療を進めていたが、最近出てきているオプジーボや抗体医薬といった高価格な薬に対してもそういうことをやってしまうと、医療費の高騰も問題となる。そこで、何かデータに基づいて患者をグループ分け、層別化し、それぞれのグループに対して最良の治療を最初から決めるという精密・個別化医療の考え方が出てきた。これは、患者に

とっても大変利益になると考えられるが、このグループ分けのところが結構大変である。

大変な理由としては、まずは分類・診断が明確でない疾患が非常に多いことである。例えばアトピー性皮膚炎は何か単一の疾病だと思われがちであるが、実はそうではなくて、様々な遺伝子要因や外的要因などが組み合わさって、最終的に見た目として似た病態をつくっているということが知られている。こういった疾患に対して単一の治療は効かない可能性が高い。しかし、疾病をさらに細かくグループ分けしようにも、そもそもその分類自体が医師ですらよく分かっていないという状況であるため、データに基づいて再定義していく必要がある。もう一つの理由は、時系列の情報や分析がないことである。データを基に層別化するとしても、多くの疾患データは、そもそも疾患発症後に病院で取られたものであり、疾患発症前のデータはまずない。また、疾患発症後の履歴や進行についても、あまり分析されることがなかった。これまでの医療、医学では、時系列の情報あまり扱われてこなかったという問題点があったため、これをどう扱っていくかを考えていかなければいけない（図2-8-1）。

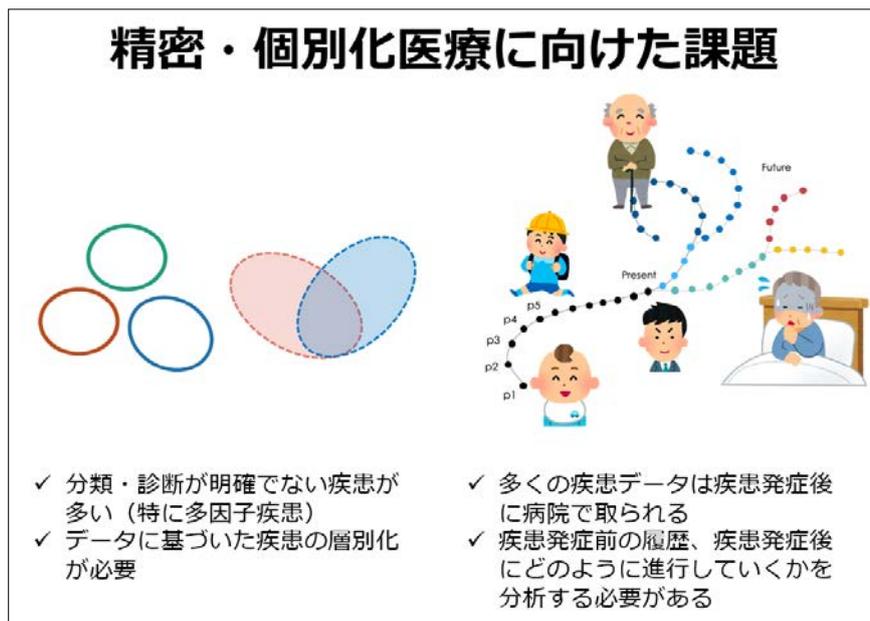


図2-8-1 精密・個別化医療に向けた課題

### AIによる医学研究事例

AIによる医学研究の具体的な事例を簡単に紹介する。1つ目は、医用画像解析で、今まで熟練した名医はできていたが研修医のような経験の浅い医師はできなかったような非常に難しい分類を、AIを使って行なうというものである。例えば、Natureに2017年に出た皮膚腫瘍の分類の研究では、13万枚もの皮膚腫瘍の画像を集めてきて、複数の臨床の医師がコンセンサスを取って、1枚1枚がんのタイプのラベルづけを行なっ

た<sup>20</sup>。このデータセットを深層学習で学習させることによって、皮膚科医と同等以上、90～95%の診断精度を達成することができたという論文である。画像というある程度決められた形式のデータに対し、人間が診断した場合は熟練したお医者さんでもやはり見落としや間違いが起きてしまうことから、AIによりいかに精度よく、再現性よく分析をするかというのが一つの課題になっている。もう一つの例は、日常の活動データに基づく疾患の検知を目指した研究である<sup>21</sup>。アップルウォッチを始めとしたウェアラブルデバイスにより、心拍変動や活動のデータが取得できる。そういったデータに基づいて、糖尿病や高コレステロール、高血圧といった、直接心拍変動と関係なさそうな他の疾患まで予測できる。具体的には、ある心拍変動のパターンを持っているからこの人は糖尿病である確率が非常に高い、といったことを予測できることが明らかになっている。こういった事例のように、いわゆるビッグデータと言われるような非常に高次元で複雑なデータ、場合によっては時系列データに基づいて、非常に高い精度で疾患の分類や予測を行えるようになってきたというのが、現状のAIを使った医療の方向性かと思う。

ただ、ここで問題になってくるのがデータをいかに集めるかという点である。アメリカはGAFAと言われる巨大なIT企業を中心とした企業主導で、すごい量のデータを集めている。グーグルは創業した2000年から、哲学を持ってずっとデータを集め続けている。また、中国は強力な国主導でデータを集めている。それに対して、日本がデータを今から集め始めて、データ量で勝負しようというのは筋が悪いと言わざるを得ない。

このような状況で日本が対抗していくことを考えると、クオリティーの高いデータを持続可能な形で収集して、システムを常に持続的にアップデートしていくという仕組みをつくる必要がある。あと、小規模なデータからも情報を抽出して、今まで誰もやっていなかったようなところから新しい知識発見をするという方向も非常に重要になってくると考えている。

### 事例①：疾病・健康状態の再定義

先ほどの皮膚腫瘍の研究は非常に精度は高く、医者と同等の精度で分析できている。データをたくさん集めれば予測精度は上がっていき、本研究では13万枚というほぼ世界で最高クラスの量を集めたのにもかかわらず、例えば、疾病Aと予測している中に疾病Bや疾病Cが少し混ざっていたりと、間違った分類が出てくる。これは、人工知能の精度やデータの量の問題と思われるがちであるが、ここで考えなければいけないのは、もともとの分類自体が本当に正しいのかということである(図2-8-2)。従来の医学に基づいてできてきた分類は、あくまでも経験的な分類である。分子的なメカニズムが分かっている場合がある一方、見た目での分類といったファジーな分類をされていることもあるため、それを教師データにして精度を上げていく方向性にはどうしても限界がある。そこで我々は、疾患や健康の状態というのをデータに基づいて再定義するというのを考えている。

その例として、我々が取り組んだ卵巣腫瘍の手術前診断の研究を紹介する<sup>22</sup>。卵巣腫瘍は、悪性腫瘍とその

- 20 Esteva, A. et al., "Dermatologist-level classification of skin cancer with deep neural networks", Nature, Vol. 542, pp. 115-118 (2017). DOI: 10.1038/nature21056
- 21 Brandon, B. et al., "DeepHeart: Semi-Supervised Sequence Learning for Cardiovascular Risk Prediction", Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18, New Orleans, 2-7 February 2018) , pp. 2079-2086.
- 22 Kawakami, E. et al., Clinical Cancer Research, Vol. 25(10), pp. 3006-3015 (2019). DOI: 10.1158/1078-0432.CCR-18-3378

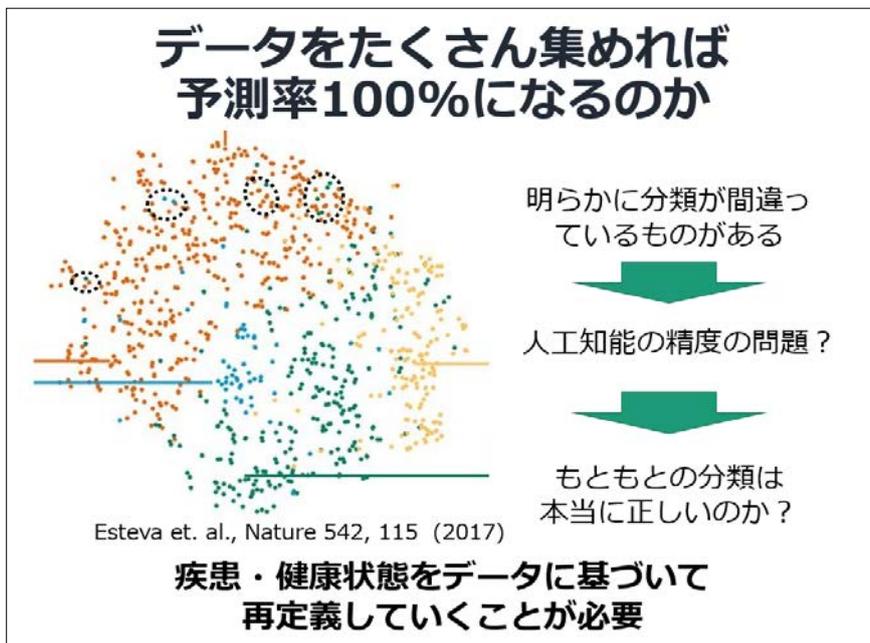


図 2-8-2 疾患・健康状態の再定義が必要な背景

まま放っておいてもあまり悪化しない良性腫瘍に分類され、悪性腫瘍はさらに、転移や播種がない早期がんと、転移があって放っておくと進行していつか亡くなってしまう進行がんに分けられる。こういった分類は、手術中に開腹して転移や播種を確認することで行なわれているが、これを手術前に血液検査だけで予測しようという研究である。血液検査はオミックスデータのような複雑なデータを使っているわけではなく、手術前にルーチンで行なわれる、30項目ぐらいの一般的な血液検査データだけを使っている。

まず、テストデータを用いてランダムフォレストという手法で良性／悪性を判定してみると、90%を超える非常に高い精度で判定ができていたということが分かった。従来のいわゆる統計的な手法である重回帰分析だと86%程度ということで、かなり精度が上がった。ランダムフォレストを使用したのは少数サンプルにも適用可能であることが一つの理由である。今回のような良性、悪性それぞれ100人、300人で全部で400人程度という比較的少数のサンプルに対しては、深層学習を使うよりランダムフォレストのような定義ベースの方法のほうが精度が良くなる傾向があることが知られている。実際、今回の例で深層学習を使ってもランダムフォレストほど精度は上がらなかった。あと、ランダムフォレストでは、例えばCRPが4.25より高いか低いかといった感じでデータを分割していくが、こういった決定木ベースのものを使うと外れ値に強い。例えば、CRPが200のような異常に高い値を示していた人がいた場合、回帰モデルのような統計的手法だとCRPの異常値だけで悪性と判定されてしまう可能性もあるが、決定木ではこれが4.25より高いか低いかで判定されるだけである。また、決定木では、「CRPが高くてリンパ球が高いと悪性」というように、複数の変数に基づいて条件分布していくため、複数の変数の相互作用を考慮して層別化が行える特徴がある。こうした特徴をもったランダムフォレストにより、実際の血液検査に基づいた良性／悪性の予測が比較的高い精度でできた。

一方、早期／進行がんの予測も同様にランダムフォレストを用いて行なったところ精度が69%程度で、ランダムフォレスト以外の手法を使ってもこれ以上改善しなかった。これだと臨床現場では全く使い物にならない。そこで精度が上がらない理由として、卵巣がんの早期がんと進行がんで似ている症例があるんじゃないか、そ

2.8

機械学習、数理科学に  
基づく健康・疾患層別化  
と予測

もそも早期がんと進行がんという診断自体が本当に確実なのかという点を疑った。早期がんと進行がんを分ける要因である転移や播種の診断においては、小さくて医者が見落としたり認識できていなかったりといったケースもありうるためである。教師あり学習ではこういったケースが見えてこないため、教師なし学習を試みた。

まず、似た症例をグルーピングする、いわゆるクラスタリングに近いようなことをやってみた。良性／早期／進行がんという医者の判定を教えることなく、30項目から成る各患者の血液検査のパターンの類似性というのをランダムフォレストに基づいて評価して、その類似性を用いて二次元の空間に多次元尺度法というもので埋め込んだ。その結果得られたマップに医者の判定のラベルを付けると、判定結果の情報なしでマップしたにもかかわらず、良性腫瘍と進行がんは非常にきれいに分かれた。一方、早期がんの分布は、良性腫瘍や進行がんとオーバーラップしている。特に進行がんとかぶっている患者は一定数いることから、これを見分けることはそもそも機械学習でどんなに頑張ってもほぼ不可能だとわかった。さらに、早期がんの中で良性腫瘍に近いタイプの患者は、5年以内にほとんど再発せず死亡はゼロと非常に予後がいいのに対し、進行がんとオーバーラップしている患者は、早期がんと判定されたにもかかわらず、5年以内に4分の1程度が再発が死亡しているということが分かった。こういった患者は転移や播種が見落とされていた、ないしは今までの治療だに対応できていなかった病態であることから、再発チェックのために定期的に検診を行なうとか早めに化学療法を始めるといった治療が考えられる。

このように、既存の医療をサポートするツールとして機械学習を使いながら、決して既存の分類を当てるだけではなくて、これまで見逃されていた新しい分類を探しに行くという方向性というのが1つ考えられるのではないかと思う。

## 事例②：ランドスケープモデル

他にも数理学や機械学習を使った研究を行っており、そのうちの幾つかを紹介する。その1つが、生命システムの状態を地形（ランドスケープ）として表現をする研究である。最近、機械学習の業界で機械学習モデル自体のエネルギーを考えるとといった動きが出てきているが、生命科学でも、こうした物理学のアナロジーで生命の状態を表現しようという話はかなり昔からあった。例えば、1940年ごろにWaddingtonという生物学者が提唱したエピジェネティックランドスケープという概念がある。細胞の分化において、例えばiPS細胞から血球の細胞、筋肉の細胞、骨の細胞といった様々な細胞に分かれていくが、分化過程が高低差のある地形上をボールが転がっていくようなイメージで表現できるのではないかという概念である。提唱当時はまだ遺伝子すら分かっておらず、あくまでもお絵描きとしてこういうふうに説明ができるというとても概念的な話であったが、最近では、シングルセル解析などの計測データから地形の再構成を試みる研究が出てきている。

最初に出てきたのが脳科学の領域である。ファンクショナルMRI (fMRI) により脳のいろんな領域の活動を経時的に測定できるが、データ列だけを見ても脳が今どういう状態にいるのかよく分からないということで、そのデータの解析にランドスケープを用いた、東大・合原研究室（当時）の渡部さんの研究例を紹介する<sup>23</sup>。まず、データ列をどこかの閾値で切って、活動しているか活動していないかといった形で二値化した。これを統計物理学の領域で使われているイジングモデルにより、各状態の観測頻度とエネルギーを対応づけた。地形上にボールをばらまくと谷に落ちるボールが多いように、よく出てくるところというのはエネルギーが低い、

23 Watanabe, T. and Ree, G., "Brain network dynamics in high-functioning individuals with autism", Nature Communications, Vol. 8, Article number: 16048 (2017). DOI: 10.1038/ncomms16048

逆にエネルギーが高いところはあまり出てこないというような関係性で、各状態の頻度を地形の高低差として表現したものである。例えば、ある脳活動の地形では、3つのlocal minimaと言われる安定状態がが非常によく出てくる一方、安定状態から離れた状態はあまり出てこなくて、3つの状態のどれかに大体存在しているということを表わすモデルになる。さらに、時系列データも地形の上での動きとして表現できる。

我々は、健康・疾患状態を地形として表現できないかと考えた。JMDC社より入手した250万レコードの健康診断の時系列データを基に、糖尿病にフォーカスして、どういった健康・疾患状態があるのかということを見た。糖尿病に関連しそうな10項目の検査項目を使って、それぞれの項目に基準値を設定して二値化した。二値化により、2の10乗、1024通りの状態が定義できるが、local minimaはおおむね9個に分かれるということが分かった。1番は全てのリスク因子が低い非常に健康な状態、2番はBMIや腹囲だけが低い肥満状態、3番は肥満プラス肝機能低下と、だんだん糖尿病に関係するようなものが多くなる。5番が血糖値やヘモグロビンA1cだけが低い純粋な糖尿病の状態で、さらに6番、7番、8番、9番となるにつれ、糖尿病に加えていろんなものが合併している状態を表わしている（図2-8-3）。

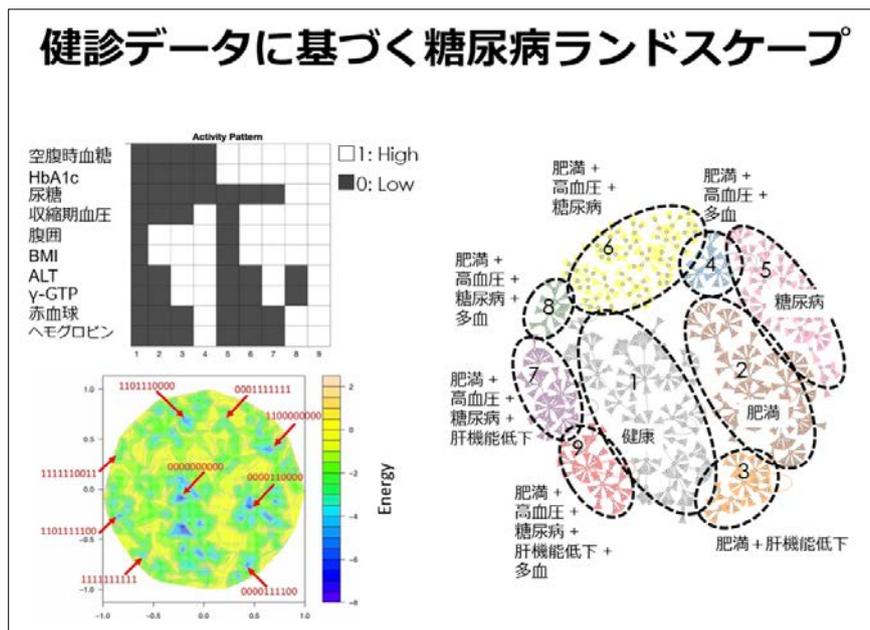


図 2-8-3 糖尿病ランドスケープ

この地形により、まずは、糖尿病の発症・進行過程が幾つかの状態に分類ができることが分かる。糖尿病専門の医者にもかなり納得してもらえことから、糖尿病の発症・進行過程をよく表現できていると考えている。さらに、それぞれの患者がどのような状態を経て発症したのかということが地形上で可視化できる。それにより、糖尿病と言われる1つの疾病であっても、発症・進行過程というのがそれぞれの人によって異なることが分かってきた。現在これを、マルコフモデルのような確率過程のモデルを使って、どういうふうに行進していくのかをモデル化し、さらにどういう治療をしたらそれぞれの進行過程が止められるのかということを分析しようとしている。

2.8 機械学習、数理科学に基づく健康・疾患層別化と予測

### 事例③：位相的データ解析

次は、結構いろんなところで使われるようになってきた位相的（トポロジカル）データ解析を使って、皮膚のバリア機能を推定した例を紹介する。皮膚のバリア機能の評価では、皮膚からどの程度水分が蒸散していくのかを評価する。この計測は結構面倒で、標準的な湿度・温度の部屋で30分ぐらい待機してもらった後で、専用の機器を使ってTEWLというものを測らなければいけない。毎回患者に来てもらう度に測ると結構な手間になってしまうことから、顕微鏡で取得した皮膚画像から直接TEWLを推定する試みを、慶應義塾大学とコーセーとの共同研究で行なった。

一般的に画像解析には深層学習が有効だが、画像の枚数が大量にあるわけではないため、画像から抽出した位相的特徴量の学習を行なった。位相的特徴量を抽出するには、まず画像の白い部分、皮膚の画像の場合だと皮膚の網目構造みたいなものをだんだん膨らませ、最終的に真っ白になるまで白い部分をだんだん増やしていく。この過程において、穴と言われる白で囲まれた黒い部分ができて、白い部分を膨らませていくとだんだん穴が細くなり、最終的に穴がつながって消える、といった感じで穴が出現・消滅していくが、そのタイミングなどを使って、画像の位相的特徴量を抽出できる。これは、トポロジーと言われるような画像の構造的な特徴であり、抽出に際して学習が不要なため1枚の画像からでも抽出ができる。ランダムフォレストのような方法を組み合わせれば、必要な画像枚数としては大体100枚ぐらいから、位相的特徴量に基づいた機械学習ができる。必要な計算リソースも少なく、さらに、特徴が数学的に定義されているので、深層学習により抽出された特徴と比較して非常に解釈もしやすい。この方法で、実際の画像から抽出した特徴を学習させたところ、非常に高いTEWLの推定精度を達成できた。これは、今、論文投稿中で、特許も出願している状況である。

### 医学とAI、数理科学の融合に向けて

このように、いろいろな数理科学、機械学習の手法を使いながら医学研究を行なっているが、こういった研究は当然ながらデータがないとできない。そこで、いろいろな臨床の先生や基礎生物学の先生、あと私の元々の専門であるウイルス学の先生たちと共同研究をやりながら、課題は何なのか、それに対応するデータというのはどういうものがあるのかということヒアリングしながら、そうして集めたデータに基づいて研究を進めている。

また、理論研究に関しても、データ同化の研究者や数理モデリングの研究者と共同し、いろんな数理手法を取り入れながら知識発見・問題解決を行なっている。そのうちの1つが機械学習と数理モデルの融合である。機械学習はデータから何か知識を見つけるのに対し、数理モデルは過去のデータから得られた既存の知識に基づいて作ったモデルでデータを説明する、という性質の異なるものであり、今まで機械学習と数理モデルはあまり一緒になることがなかった。数理モデルから新しい知識を発見することは難しいが、数理モデルと機械学習を組み合わせることで、数理モデルである程度その基本となるメカニズムみたいなものは仮定しながら、その背後にある多様性を機械学習で分析するという形で、説明可能な発見的な知識探索ができるんじゃないかということで、九州大学の岩見真吾先生らと一緒にデータアプローチで研究を進めている。

最後に、AIと医学、両方の専門性を持った人材の糾合・育成という点に関して、私が運営する理化学研究所のチームに、もともと数学科とか物理学科といった学部を出た後で医学の領域に移った、場合によっては医学部に再入学したような人が何人もいる。コロナ禍以降も、ドライの研究室のためオンラインで割といろいろな研究が進められるということもあり、こういった人たちがどんどん集まってきて20名を超えた。こんな感じで、医学とAIもしくは数理科学、両方の専門性を持った人たちをどんどん集めて、彼らがもともと持っていたバツ

クグラウンドを活用しながら、いろいろな数理科学、AIの手法を導入して研究を行なっている。また、千葉大のほうでも様々な専門性を持った人たちがどんどん参画してくれている。こういった若手の人たちに積極的にいろんなテーマに取り組んでもらうことで、新しい医学の領域を支えていくような若手人材を育成していかなければと考えている。

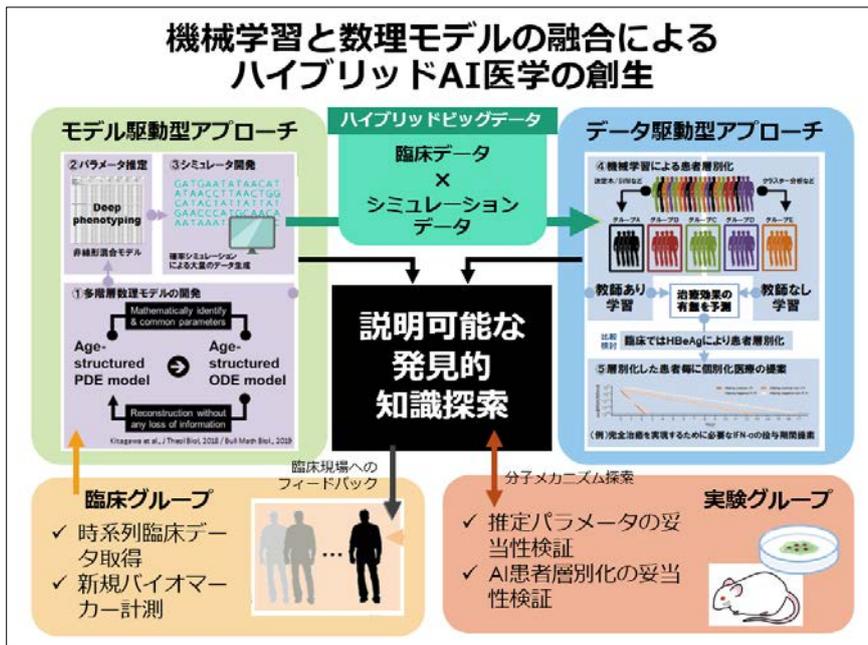


図2-8-4 機械学習と数理モデルの融合による医学

【質疑応答】

- Q : AIと医学の糾合によって新しい分野ができつつあると思うが、この分野におけるグラウンド・クエスチョンとは何か。
- A : 研究者により様々な考えがあると思うが、1つは、今までの医師の経験知をいかにAIに置き換えていくか、いかに再現性を持って誰でもできるようなものにしていくのかが大きなものかと思う。もう一つ大きな課題は、疾患の再定義だと考える。これまでの経験的な分類や分子的な分類では、多様性があまり考慮できていなかったことから、医学・医療データに基づいて、その中に潜むような多様性や非線形な関係性といったものを機械学習を使って見つけていくことが重要だと考える。
- Q : AIによる分類によって得られた新たな知見が再定義につながったような具体例はあるか。
- A : 私の卵巣がんの研究において、早期がんと言われていた1つの分類の中に、さらに予後が悪く死亡・再発しやすいグループと、放っておいても死亡・再発しにくいグループというのが機械学習によって見つかった、というのが実例として当てはまる。
- Q : 機械学習によって相関は見つかると思うが、どういう理由/機序でがんになったかという理解につながったような事例はあるか。
- A : 機序の理解というところまではなかなか難しい。卵巣がんの例だと、CRPやLDHのようないわゆる炎症に関連する指標が、早期がん、進行がん、早期がんの中で予後が悪いグループと予後がいいグループ

2.8 機械学習、数理科学に基づく健康・疾患層別化と予測

プの分類に寄与していたということが今回分かった。これを基に、メカニズムに迫るようなデータ取得と解析を現在行なっている。機械学習で分かることは、あくまでもこういった因子によって分けられるということであり、因子がどのように寄与してグループ分けが行われているのかとか、それがどういうメカニズムで起きているのかといったことは、やはりもう一回、基礎生物学や基礎医学に立ち戻って検証していかねばいけない。アーティファクトや計測的なバイアスもあるため、こうしたことを繰り返していくことによって、新しい分類が間違っていたことが判明することも当然あると思う。実際に治療につながったといった事例はまだ私たちも示せてはいないが、機械学習により仮説を見出し、さらに基礎生物学などで検証していくというような流れはできてきているのではないかと思う。

Q：機械学習と数理モデルの組合せにおいて主に行われているのは、データ同化のようにベースになる数理モデルがあって、それをデータで修正もしくは補強するといった形かと思うが、他の組合せ方はあるのか。

A：まずは、ランドスケープの例のように、数理モデルを用いて機械学習にけるパラメーターを生成するという形がある。もともとのデータをいきなり機械学習で層別化しようとする、背後の構造が全く分からないまま層別化してしまいすごくやりにくいため、二値化などにより一回パラメーターに落とした上で層別化をするのが有効である。

もう一つは、バーチャル・ビッグデータを作るといった方向性を現在行なっている。例えば何か1つ数理モデルを作ると、シミュレーター上でそのパラメーターをちょっとずつ動かしていくことで、いろいろな患者に対応するような架空のデータを作れる。ただ、例えばパラメーターの組合せがおかしいというような、現実ではあり得ないようなデータがたくさんできてしまう。それに対し、GANのような生成モデルを使うことで、時系列データのような割と抽象的なデータであっても、それが本物データと近いのかどうかということを機械学習で判定し、怪しげなデータをどんどん排除しながら本物そっくりのデータをつくっていくという、現実拡張型シミュレーションとも呼ぶことができるシミュレーションができるのではと考えている。

Q：シミュレーターで機械学習の訓練データをつくるということはいろんな分野でやられていて、自動運転のようにその物理モデルが非常にはっきり分かっている場合は、正解とそこから得られる観測との間の結びつきをきちんとつけられるのに対し、医療の場合は、中の仕組みや機序が完全には把握できておらず、もしかしたら違う結びつきをしてしまう可能性があると思うが、その辺りは大丈夫なのか。

A：実際のところ数理モデル自体が暫定的なものである。例えば、ウイルス感染シミュレーションを行なう場合、取りあえず暫定的に今まで分かっている免疫応答などを基にモデルを作るが、因子が足りないとか、パラメーターが不正確であるとかいった問題は絶対出てくる。そういったケースで、機械学習を使って取りあえず一回データを生成してみて、それに対して何か検証実験を行なうというようなサイクルを回すことでモデルを改良していく、という方向性を現在考えている。物理学であれば、変更するのは物理パラメーターで、物理モデル自体が変わることはほぼないが、生物学の場合はモデル自体が変わり得るため、そのモデル自体を機械学習と数理モデルを使って少しずつ作り変えていく、という方向性である。

Q：物理学の場合はモデルを原理から構成することができるため、もし実験事実と合わなかったときに、どこが間違いかという分析が比較的容易な場合が多い。生命科学・医学でも、何かしっかりとしたガイディング・プリンシプルに基づいてモデルを構成しないと、数理モデルそのものが非常に大きなバイアスを与えてしまうリスクがあると思う。

- A : まさにそのとおりである。例えば、我々と数理モデル研究者との共同研究では、免疫細胞の変化のシミュレーションで、もともと人口学や生態学で使われていたようなポピュレーション・ダイナミクスと呼ばれるモデルを使っているが、各要素間の相互作用などを非常にシンプルなものとして仮定してしまうため、本当にそれが正しいのかという検証が非常に難しい。一方、そこから出てきたパラメーターの分布や応答性などから、仮定した相互作用がおかしいということを推定する方法が幾つかあるため、そういうものを使って少しずつ作り変えていくことを考えている。物理学のように基本原理から説明するのは、現状個体レベルだとまだできないので、そこを観察的に少しずつ理解していこうと考えている。
- Q : 恐らくこれは本質的な学問の違いで、物理学の場合は要素還元的な見方が正当化されると思うが、生物学や医学の場合は、要素還元にするともそもそも生き物でなくなってしまうため、複雑なものを複雑なまま捉えるというのが本質であると思われる。機械学習そのものは非常に複雑な対象をそのままパラメーターと仮説空間の中で捉えることが可能だが、数理モデル化するときにはある意味トップダウン的な直感が必要である。その際、どういうガイディング・プリンシプルに基づいて数理モデルを作るかというところに、非常に工夫と慎重さが要求されると感じる。
- A : 最近の方向性だと、1個の数理モデルだけをこれは使おうという話ではなくて、幾つかの数理モデル、場合によっては数千もの数理モデルを並べて、その中から妥当そうなものを機械学習で選別していくといった方向性もある。数理モデルというのは、出力が間違っていた場合はそもそもの仮定が間違っているということもあり得るので、何か1個のものに固執するのではなく、いろいろなものを同時に評価していく。機械学習は、いろいろなものを同時に評価し、妥当なものを抽出するといったところに非常に向いているため、うまく組み合わせながら探索していければと考えている。
- コロナ禍においてSIRモデルが広く認知されることになったが、SIRモデルもいろいろな面で問題点は指摘されている。例えば、基本再生産数というパラメーターがあるが、最初は一定値を取るという仮定でモデリングするのに対し、それ自体が変わっていったらどうなのかという検討がなされた。このように、一応モデルとしてはベーシックなものはあるが、そこに他のさまざまな要素を取り入れて、ダイナミックに動くモデルとして拡張するというように、モデルはだんだん拡張されていくものだと思う。
- Q : モデル拡張に際して、モデルを詳細化すれば多分その予測値はより当たるようになると思うが、医学の世界では当たることの価値と、モデルがシンプルであることの価値のバランスは、どんな感じなのか。
- A : 医学の場合ではNが少ないことが多く、取りあえず当てるだけだったら当てられてしまう。例えば、今回のCOVID-19におけるSIRモデルだと、N=1のものが当たるかどうかといった話になってしまうが、そこをどう検証していくのかは結構難しい。私としては、そこから知識発見を行なうことが重要ではないかと思う。例えば、あるモデルを仮定したときに、非常に寄与するパラメーターが判明したら、そのパラメーターに対して介入を与えたときにどうなるのかということを検証していく、といったサイクルを回す必要がある。ただ、サイクルを回すのが難しいケースもあることから、そういう時はシミュレーション上で何回か検証した上で、実際のデータを使って、例えば何か別の新しい感染症が起きたときに、その基礎となるモデルを使いながらリアルタイムでその介入と検証を行なっていくといったことが考えられる。
- Q : 疾患の再定義のような新たな見方が出てきた時に、実際の医師の方の反応はどのような感じか。また、受け入れる上で、解釈性と現実のデータとの整合性ではどちらが重視されるのか。
- A : 場合によるものの、私自身がもともと医学出身のため、臨床の医者がどう感じるかというのは私自身もある程度感覚があることから、今までのケースでは比較的受け入れられていると感じる。もし、実際の

臨床の感覚とあまりにも乖離しているということがあれば、分類自体の再検討もあり得る。医師は自分たちの臨床にどう役に立つのかを求めているため、臨床現場と近いところで研究を行なうことはすごく重要である。臨床の感覚がないまま、新しい分類が見つかったので使ってくださいと言っても絶対使ってくれないし、なかなか普及しない。その分野の権威のような先生方と一緒にやらせていただいていることが、受け入れられやすい要因の1つかと思う。

受け入れに当たっては、臨床の医師にも理解できるような形で説明していくということも重要であると考えられる。卵巣がんの層別化の例では、例えばCRPが高い人たちがこういう風に予後が悪いといったことを一つ一つ説明していくと、納得してもらえることが多い。機械学習でただ単に予測精度を上げる方向性についてしまうと、何かよく分からないけれども当たったということになってしまうため、一つ一つの因子やその相互作用といったところに立ち戻ってちゃんと説明可能なモデルを提示することが重要であると思う。

Q：AIによって実際の医師と同レベル以上の診断精度が出てきつつあることで、現場の医師はそういうものの脅威を感じたり、AI診断への抵抗感があつたりはしないか。

A：昔のアメリカの事例で、自動の麻酔装置を造ったら麻酔科医の大反対によって潰された、といったことを聞くことはあるが、AI診断に関しては私の知る限り、どちらかというところとどんどん使っていきたいと思っている医師の方が多く、対抗意識を持っている医師はあまりいないと感じる。今までは人間が同じことを何回も何回も同じようにやらなければいけないし、さらに間違えると訴訟になる場合があるため、どうAIをうまく使っていくかということかと思う。医学部の中で研究を行なう上で、あなたたちの仕事を奪うものではなくて、あなたたちの仕事をより楽にしてサポートするものですよ、ということを経験している。

Q：糖尿病ランドスケープ上で健康・疾患状態の時間的経過を見ることができると思うが、医学的にどのように役に立つのか。

A：これまで糖尿病では、各状態にいる患者に対する治療方針については考えられていたが、そこに至るまでにどういう経過をたどって来たのかはあまり考えられてなかった。マルコフモデルというよりは履歴まで考慮するようなモデルになると思うが、現在ある同じ状態に分類される患者でも、健康な状態からいきなり発症した人と、ある程度別な状態に留まっていた人だと状態が違う可能性がある。また、時間経過を考えることで、ある状態にとどまりやすいのか、それともどんどん進行するのかということも見えてくる。このように、履歴や時間経過を考えたモデルによって、同じ状態にいる人に対してもさらに層別化でき、より個別化予測医療が進められるのではないかと考えている。

Q：糖尿病に直接は関与しない高血圧や肥満といった項目はどのように選択されたのか。また、糖尿病は発症のメカニズムが大体はっきりしていると思うが、高血圧のような原因が数多く存在する疾患を対象にした場合はどうなると予想されるか。

A：どの項目を使うかは自動的に決まるものではなく、糖尿病と関連のあるものを実際のデータから見つけてきて、我々で決めている。ただ、他の項目は一切関係ないと言い切れないため、どの項目を選んで使うかは試行錯誤が必要となる。また、ランドスケープモデルでは、例えば100項目をそのまま解析すると、状態数が2の100乗と膨大になり理解ができなくなる状態になってしまうため、そういった場合は次元圧縮により似たような項目をまとめて大体10～20次元ぐらいに落とした後、ランドスケープモデルで解析するというを行なう。

原因が数多くある疾病としては、現在心不全を対象に研究を行なっている。糖尿病の場合は、糖尿病

になる前の状態はそんなに出てこなかったが、心不全や高血圧ではそこに至るまでのルートが幾つか出てくるのではないかと予想している。

Q：イジングモデルを適用するために、ある変数や検査項目を0/1で表わすというのは非常に人工的な設定だと思うが、問題は生じないのか。

A：バイナリでいいのかという点は、議論があるところである。あと、今回は定常状態を仮定しているが、本来は新しい人たちがどんどんこの中に入ってきて、一定数は亡くなってこのモデルの中から消えていくという並行系のため、本来はそれに対応したモデルを使わなければいけない。あと、今回はエネルギーだけを考えたが、エントロピーを使うといったモデルの拡張の方向性というのは多くある。

こういったものを取り入れることで、今まで見えてなかった中間的な状態が見つかることがある。医学的な新しい分類になるには検証が必要であり、そのためには、例えば治療に対する応答性が違うであるといった医学的な意味づけをやっていくことが重要だと考える。モデルを作って終わりではなく、理解を進めるためのツールとして数理モデルを使っており、それがデータ駆動型研究の基本的な考え方だと思う。今までの数理モデルベースの研究は、数理モデルを過信してしまい、それありきで議論が進められてしまうことが多かった。場合によっては機械学習を入れながら数理モデルの検証を行ない、本当にそのモデルでいいのか常に疑っていくという観点で研究を進めている。

Q：数理モデルの中で、トポロジカルデータアナリシスやデータ同化は、普通の人間の観察では得られなかったような見方を与えてくれる。機械学習と数理モデルを併せた研究を行なう上で、こういった数理モデルは役に立つと思われるが、その辺りの期待について伺いたい。

A：数理モデルでも微分方程式みたいなものではなくて、物の特徴を見るトポロジカルデータアナリシスのように、幾何学的な発想というのはすごく重要だと思っている。私自身もCOVID-19のSIRモデルを作ってみたが、確定的に指数関数のようなものが出てくる一方で、実際のデータとはなかなか合わないし、社会の状況が少し変わるとモデルも変わってしまうといった面があるため、そういうものを幾何学的なもので抽出していくのは大変重要な方向性かと思っている。SIRモデルを始めとした微分方程式モデルは1つの原理に基づいて説明できるという強みがあるものの、あくまでも暫定的なものであり、背後にハイパーパラメータみたいなものがあるため、モデル自体が可変で他のいろんな社会的要因や環境要因、遺伝的要因などに影響される可能性があるため、モデルがどんどん複雑になってしまう。その性質を分析するにあたり、トポロジカルデータアナリシスのような解析がすごい重要になってくるんじゃないかと思っている。我々も時系列データに対してトポロジカルデータアナリシスを適用するというような方向性も進めている。

C：微分方程式に基づくとグローバルな状況というのはそれだけでは見にくいのに対し、トポロジカルデータアナリシスなどでは割と大域的構造というのが見れるということで、両方組み合わせるとするのは非常に有効な方法になっているんじゃないかと思う。さらに同時に、情報が大量過ぎて人間にはぱっと分らないようなことを機械学習的に見ていくという、そういう組合せは今後大事になっていくだろうと、今日ますます思った。

Q：日本人のデータを集めていくということの価値に関連して、体内の現象や症状は人種に依存するようなところは大きいのか。

A：人種の違いというのは当然あり、それだけであれば、単純に同じアルゴリズムをもう一回日本のデータで学習させれば済む話であるが、それだけではないと感じる。例えばICUでの病状推移の予測のようなアルゴリズムがあったとして、アメリカではかなり膨大なデータを使って作られるが、それをそのまま

日本に持ってきて使えるかという、実は使えない。その原因としては、どちらかという人種の違いよりは医療システムの問題で、例えば検査項目が違うとか保険適用の有無とか、そういった要因もある。人種の違いと医療システムの違いを考えながら、日本の臨床現場で使えるものを作る必要がある。

Q：臨床のデータを使って研究し、臨床の医師に説明していくという意味では、やはり臨床の経験がある方が研究に携わるのが非常に重要だと思う。一方、AI・数理モデルを使える医学研究人材を生み出していく上では医学部の教育もかなり重要だと思うが、教育自体が変わってきている面はあるか。

A：私が教育に携わる千葉大医学部では、カリキュラムを大幅に変えようとしており、AIやデータサイエンスをほぼ必修にしようとしている。そういう中で、ただ単にAIとかデータサイエンスを教えればいいのかというそうではなく、医学における問題解決ができるような医者を育てることを考えながら教えないとあまり意味がない。今カリキュラムをいろいろ作り、文科省のグラントに大学院プログラムの申請をしている。そういう中で、医学と数理学両方を分かった人がそういったカリキュラムを作り、医学部生とかに対してそういった素養を教えていくという方向性というのは、いろいろな大学で行われつつあると思う。

Q：先生のラボでは医学部卒ではないPhDが数多くいるとのことだが、医学部以外のドクターが医学研究に入ってくるのはハードルが高いと言われている。日本の中でこういった風潮を変え、活発化させていく上で、何かアイデアがあれば伺いたい。

A：医学の領域だと、実際に診療に携わっていてデータの提供元であるMD（Medical Doctor）が、ある程度尊重されるというところがあるが、MDが上だからほかの人は全部使い走りといった状況になるのは望ましくない。一時期、バイオインフォマティクスが実験研究者に使われるような状況になってしまっていたが、そうするとやはり裾野が広がらず、結果的にバイオインフォの人たちがあまりアカデミアに残らず、どんどん企業に出て行ってしまった。似たような状況が起こらないようにするには、まずはしっかりとそのポジションを作ることが必要である。また、論文の評価に関しても、ファーストオーサーだけが評価されるようなシステムだと、解析を行なった人間がちゃんと評価されず、ただのお手伝いになってしまう。今後は解析主導でやるような研究も増やしていくことが重要かと思う。あと、上にいるPIが両方を理解し、解析や数理の重要性というのを理解しながら研究を進めていくことにより、医学研究がキャリアパスとして確立するというのも重要である。私自身PIとして、若手の人たちをモチベートする、プロモートするようなことをやっていきたい。

## 2.9 因果推論 – 決定論的アプローチ –

陳 洛南 (中国科学院系統生物学重点实验室 教授)

従来のビッグデータは、データの相関から色々なものの関係を見いだして、科学研究などに応用してきた。問題は、多数の間接的な相関があるなか、直接的な相関を抽出し、そこからさらに直接的な因果性を捉えること。特にAIや機械学習は、基本的に相関を捉えているが、いかにして相関から因果に発展させるかが重要になってくる。

モデルとしては、統計モデルとダイナミカルモデルの2種類がある。統計モデルは今の機械学習の主流であるが、本稿では最近の新しい流れであるダイナミカルモデルを中心に、いかにダイナミクスから因果関係を見いだせるかについて紹介したい。

### 統計モデル

今のビッグデータは基本的には相関性を扱っている。たくさんのデータさえあれば関係・関連性は見いだすことができ、ディープラーニングなどに盛んに使われている。ただし問題は、直接的な相関と間接的な相関との区別ができないことである(図2-9-1)。

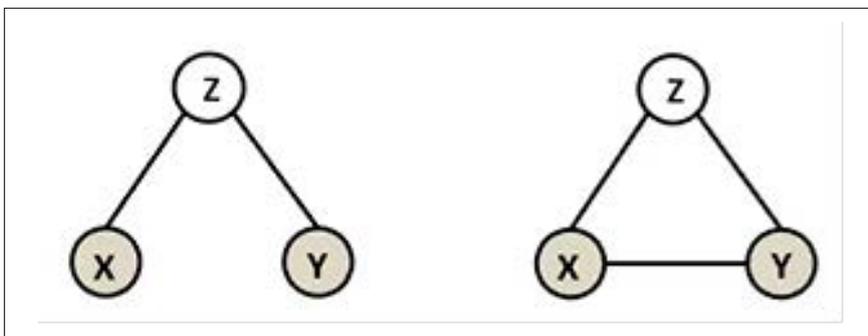


図2-9-1 XとYの間接的な相関(左)と直接的な相関(右)

この問題に対して、統計学ではpartial correlationやconditional mutual informationといった手法をもちいて第3の変数であるZの影響を評価し、本当のXとYとの関係を同定することができる。これによって、原理的には直接的な相関と間接的な相関との区別はできるが、ネットワークのように多数の変数がかかっている場合、第三者である変数の情報を除いていくと、本当のXとYとの関連性が見えてこなくなるという問題が発生する。

これに対して我々は、partial independenceという手法を4年前に提案した。従来用いられてきた条件確率の代わりに、2つの確率変数の相互依存の尺度を表す指標であるPMI(Part Mutual Information)を導入することで、多数の変数がかかる場合でも直接的な相関を見出すことができる。このPMIを様々な相関関係(線形、2次、3次、三角関数、指数関数、ランダム等)において、XとYとの直接的な相関を見出せるか評価したところ、CMI(Conditional Mutual Information)、PC(partial Condition)といった従来手法と

比べ良好な結果が得られた。また、PMIを実際の遺伝子ネットワークに適用し因果関係の推定を行った(図2-9-2)。

ただし、統計モデルでは因果関係を数学的に証明することはできない。確認実験を行い、推定結果を検証する必要がある。

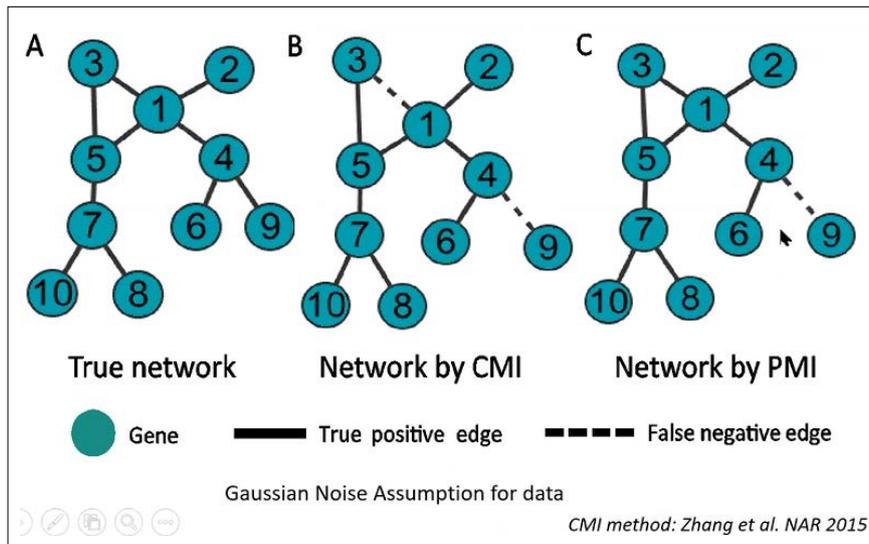


図2-9-2 PMIの適用例：酵母の遺伝子制御のネットワーク推定

### ダイナミカルモデル

ダイナミカルモデルにおいてXとYの関係は、差分方程式で表現される(図2-9-3)。

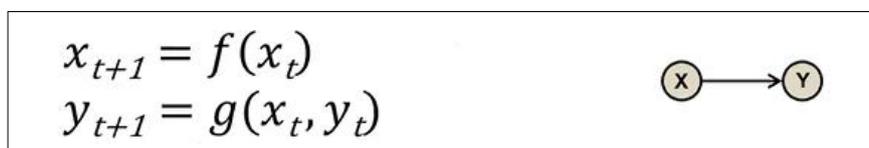


図2-9-3 差分方程式で表現したXとYの相関関係

この式は、YがXに依存する場合である。Xが変わると、それによってYも変わるということを表している。これが、ダイナミカルな因果関係を構成している。

ダイナミカルな因果関係は、大きく3つ分類される(図2-9-4)。1つめはgranger causalityに代表される「線形回帰」。2つめは、「情報理論」。時系列データを用いたtransfer entropyは、従来からよく使われている。そして3つめが、本稿で主に説明する「データの埋め込み」である。

「線形回帰」の代表例であるgranger causalityは、XYについて計測された時系列データから、以下に示す誤差によってXとYの間に因果関係があるか判断する。

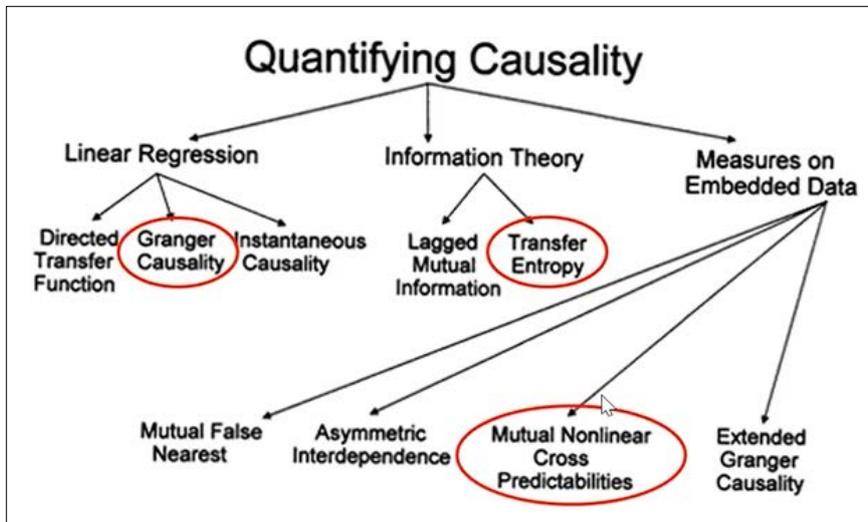


図 2-9-4 ダイナミカルな因果関係の分類

$$\text{If } \sigma(Y_{t+1} | \Lambda_t) < \sigma(Y_{t+1} | \Lambda_t | X_t)$$

$\Lambda_t = \{\text{past terms of } Y_t, X_t, W_t\}; \Lambda_t | X_t = \{\text{past terms of } Y_t, W_t \text{ excluding } X_t\}$

すなわち、現在のデータからYの将来の値を予測するとき、現在のデータにXの情報を含めることにより、含めない場合よりもYの予測値の誤差は減るのであれば、Xは原因でYは結果ということになる。この granger causality は非常によく使われるものだが、いくつか問題が存在する。1つめは非線形性の問題である。近年、非線形性に関するいろいろな拡張がされているものの、本来のモデルは線形であり、非線形の因果関係を捉えることはできない。2つめは、X、Yのデータ分離の問題である。XとYはきれいに情報面でデータから分離できるという仮説が要るが、非線形システムの場合、実現できない。3つめは、すごく強い因果関係はよく判定できるが、多数の変数がYの原因となり、それぞれには強い因果関係はない場合、判定が困難である。また、間接的な因果関係と直接的な因果関係を判定できないという問題もある。

一方、「情報理論」の代表例である transfer entropy は、granger causality を非線形に拡張したと言えるものであり、非線形の因果関係を判定できる。ただし非線形の問題以外では、granger causality と同じ問題を抱えている。

「データの埋め込み」は、ターケンスの埋め込み定理を基礎として、これまで挙げたいろいろな問題の解決を目指すものである。ターケンスの埋め込み定理について簡単に説明する。複雑で多数の変数で記述される未知のシステムについて、1変数を長時間観測したデータが得られるとする。delayed embedding は、観測された1変数のデータを少しずつ遅延したものを埋め込むことで、高次元データを構成できるというものである。例えば図 2-9-5 では、1次元の観測データ  $y(t)$  から、 $y(t)$ 、 $y(t+T)$ 、 $y(t+2T)$  と時間をずらした3つデータを並べることで、3つの変数からなる3次元の時系列データを取り出すことができる。

2.9

因果推論  
— 決定論的アプローチ —

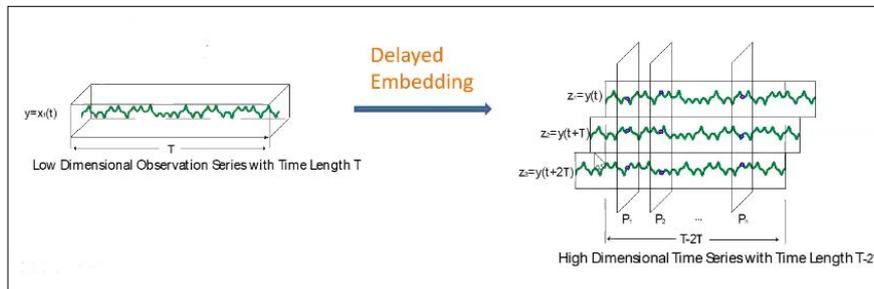


図 2-9-5 遅延埋め込み：1次元の時系列データから3次元の時系列データを取り出す

ここでは、情報量は全く増えていないが、1次元のデータから高次元にリフティングし、例えば3次元の位相空間  $Z_1$ 、 $Z_2$ 、 $Z_3$  にプロットすることで、複雑なダイナミクスを再構成することができる。位相空間にプロットされた軌跡は、元にあった未知のシステムでの軌跡と1対1対応をしており、元のシステムを再現できたことになる。すなわち、1変数から高次元のダイナミクスをトポロジカルに再現できるということである(図 2-9-6)。

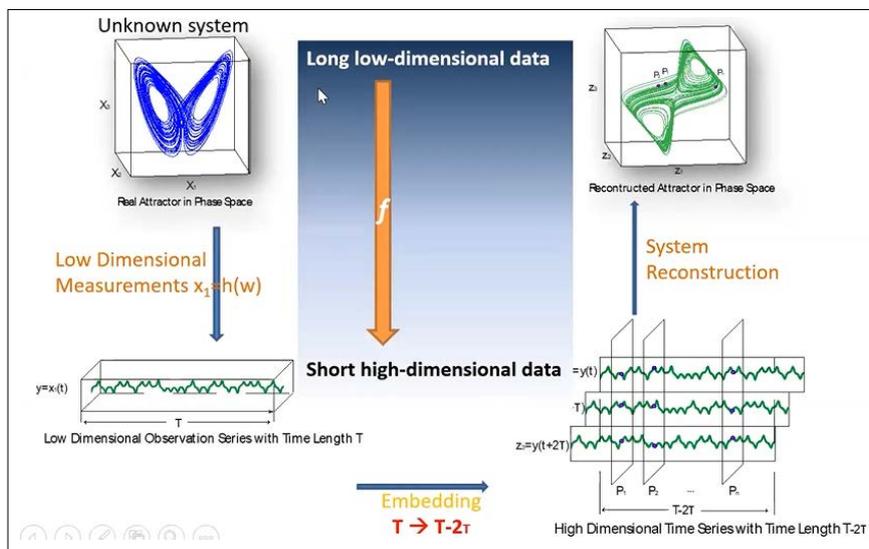


図 2-9-6 ターケンスの埋め込み定理

この理論をもとに、我々は1つの変数から他の変数を予測できることを RED: Randomly Distributed Embedding として示した。そしてそれが、2010年に合原一幸により提唱された Cross Mapping の基礎となっている。Cross Mapping は、2つの時系列変数間の因果関係を判定するもので、時系列変数  $Y$  の近傍が、それに対応する時系列変数  $X$  の近傍になるのであれば、 $X$  は  $Y$  の原因になるというものである。G.Sugihara らは、予測によってダイナミックに因果関係を同定しましょうという MCM: Mutual Cross Mapping (2012) を提唱し、さらに我々が提唱している PCM: Partial Cross Mapping (2020) へと発展している。

これら埋め込みの良いところは、非線形システムについて何も仮定はしなくても、非線形の因果関係を判定

しており、データの分離性についても仮定が不要な点である。先ほどのgranger causalityやtransfer entropyに代表される従来のダイナミクスによる因果関係判定方法の欠点が解消できている。この点が埋め込みの最も評価すべき点ではないかと思う。また、YからXへのembeddingが成立することは、XとYとの因果関係の証明になっており、数学の観点からも重要な手法である。

従来のMCMでは、直接的な因果関係と間接的な因果関係の区別はできないが、我々が提唱しているPCMでは、XとYに新たな変数Zを加え、YからXへの直接のcross mappingと、Zを経由したcross mappingとを比較することで、間接的な因果関係なのか判定し、ネットワーク構造を明らかにすることができる(図2-9-7)。

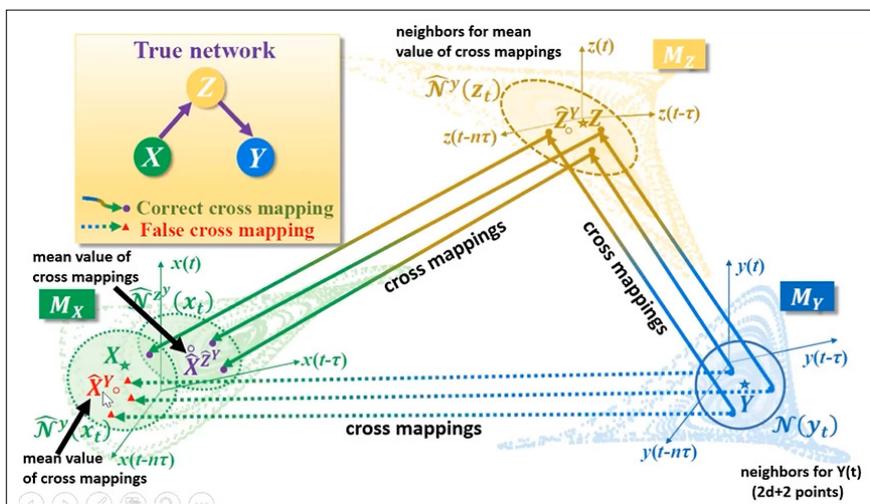


図2-9-7 PCM:Partial Cross Mapping

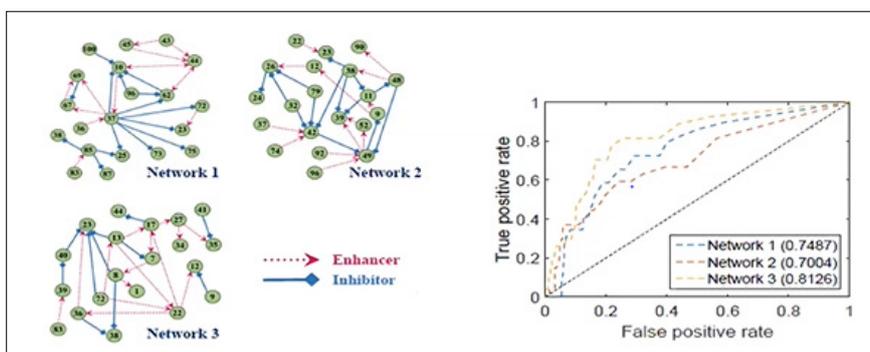


図2-9-8 PCMの適用例：概日リズムの遺伝子ネットワーク

このPCMを遺伝子ネットワーク(図2-9-8)、プランクトンの食物連鎖ネットワーク、香港における心臓病の発病と大気汚染物質(NOx、SOx等)との関係、といった実際のネットワークに適用し、ネットワークを上手く再現する結果を得ている。改めて強調しておきたいのは、今の機械学習で用いている統計モデルでは因果関係を明らかにできないが、ダイナミカルモデルを用いることで、本当の因果関係を数学的に証明できるということである。

### 埋め込みを用いた予測

時系列データの予測において、統計的手法が使われている。多数のデータを集めれば統計パターンを得ることができ予測ができるというものである。

ただし、データ量が少なく統計パターンがない場合、予測をする方法はほとんどない。我々は、データ量が少なく統計パターンがない場合について予測ができないか考えた。

変数が少ない長時間の観測データから予測を行う手法は色々ある。一方、変数が多く単時間の観測データから予測を行う手法は存在しない。しかしながら情報の観点においては、これら2つは非常に近いものである。

したがって、空間の情報から時間の情報に変換する手法があれば、短いデータから将来を予測できることになる。

また、アルファ基など、今盛んに使われているディープラーニングは大量のデータによる学習が必要である。あらゆる経験をするにより、正しい予測や決定はできるようになる。我々は、少ないデータから本当に経験してないことを予測できるかについて考えている。このために埋め込み理論が必要となる。

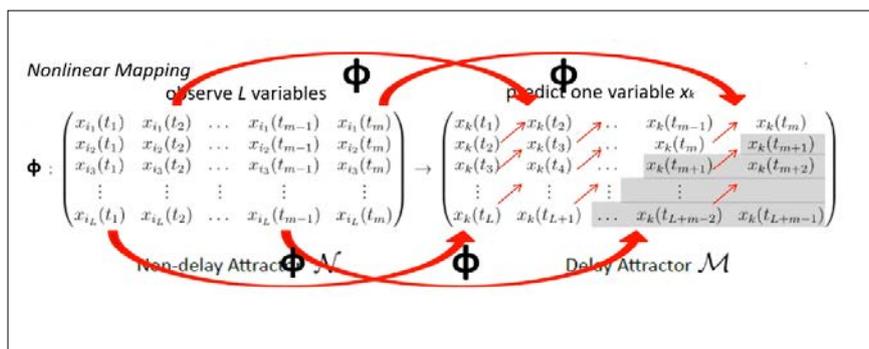
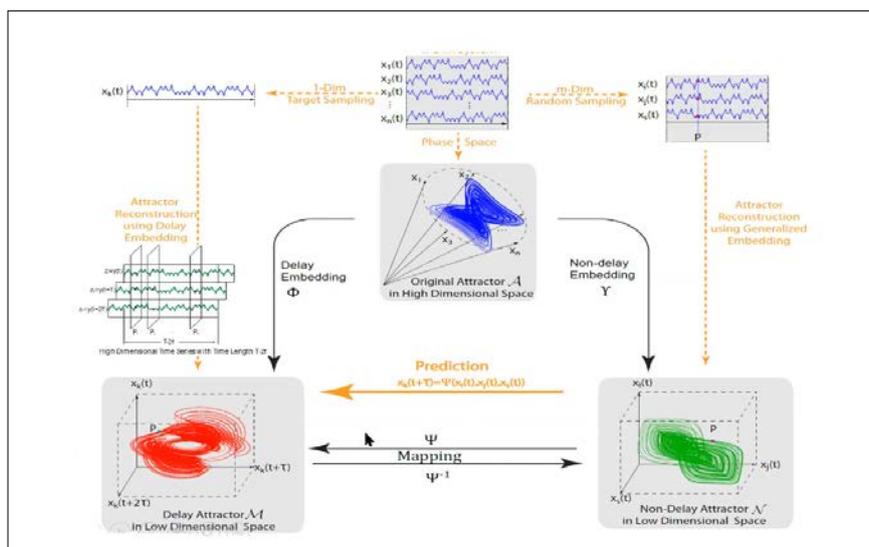


図 2-9-9 埋め込みによる変数の予測

このような問題に対し、我々は以下の様なアプローチを行った(図2-9-9)。すごく複雑な未知のシステムを考える。ただし、そこから多数の変数(例えば1万変数)を観測できるとする。

この1万個の変数について20タイムポイントを観測し、観測された変数のうち1つについて将来予測をしたとする。そこで、まず予測したい変数を1つ取り出して、その変数について時間をずらして埋め込むことで高次元のデータ(図2-9-9のDelay Attractor)を再構成する。

一方、1万個の変数から100個の変数を抽出し、時間をずらさずにそのまま高次元の空間に埋め込む(図2-9-9のNon-Delay Attractor)。このようにして、元のダイナミクスとトポロジカルに同相な2つの高次元データを取得することができるが、この2つは1対1の関係になっており、この2つを関係づける非線形の関数を見つけることで、取り出した変数について予測をすることができる。

さらに、我々は、これを埋め込みの理論を用いたAuto-Reservoir Computingという機械学習の方法を提案し、様々な系についての予測を行った。例えば、ネズミの概日リズムに関わる遺伝子の発現予測を行った。ここでは約2万の遺伝子のデータを22タイムポイントで計測し、次の11タイムポイントの予測を行っている。他にも台風の進路予測、風力発電施設における風力予測、海水温の予測、交通量の予測に適用した。少数のタイムポイントから非常に効率よく予測を行うことができている。

### 【質疑応答】

Q：ダイナミカルモデルはモデルフリーであり、予めモデルを仮定せずにデータから因果関係を推定できるということだが、なぜそうなっているのかというメカニズムを科学者は理解できるのか。

A：理解できないかもしれない。モデルフリーというのは、埋め込み等を使うと、コンピューターの中にはアルゴリズムにモデルがつくれるということ。ここはディープラーニングと似ている。だから、ディープラーニングがブラックボックスであるのと同じ意味で、ダイナミカルモデルもブラックボックスになっているといえる。

Q：埋め込みというのはディープラーニングと同じレベルの表現力を持っているものなのか。

A：ディープラーニングにはいろいろな埋め込みが使われている。例えば隠れ変数は一種の埋め込みであり、オートエンコーダーは高次元から低次元への埋め込みであると言える。ただし、これらにはダイナミクスが含まれていない。一方、埋め込みはダイナミクスを表現できる。また、今のディープラーニングは学習に大量のデータを必要とするが、少ないデータからの予測や、経験していないことの予測をするために埋め込み理論が必要となる。この埋め込みとニューラルネットを組み合わせることで、動的な学習が可能なニューラルネットを作ることができる。

Q：因果関係の推定により、因果構造を1つに定めることはできるのか。複数の構造が予測されることはないのか。

A：理論的には、全ての変数を観測できたなら因果構造を1つに定めることができる。ただし、全ての変数を観測できるとは限らないので、実際にはそこまでは言えない。

Q：埋め込みをするときの元のシステムが、1つのアトラクターに向かって収束するのではなく、2つや3つのアトラクターが共存している場合も扱うことができるのか。

A：それはできない。全ての観測データは同じアトラクターにあることを条件としている。さらに、高次元システムであってもアトラクターは低次元にあること、全ての観測データは同じシステムにあること、観測データの量または次元が大きいこと、が条件として求められる。

## 2.10 記号推論と科学

井上 克巳（国立情報学研究所）

「記号推論と科学」について、3つのパートに分けて話をしたい。まず、科学哲学と発見の科学について、AIによって自動化することについて。特にアブダクションという推論形式に注目した。次に、ダイナミックな環境における関係ダイナミクスについて。伝統的にAIで使われてきた知識表現・推論技術を用いた研究に続いて、ディープラーニングなどの機械学習と融合する最近の研究についても紹介する。最後に人工知能と科学の将来、それに向けて今何を研究すべきかに関して論じたい。

### 科学哲学と発見の科学

科学哲学は、科学における理論の合理性や限界について論じる学問であり、科学的発見や考察がいかになされるかを議論するうえで論理的な推論が欠かせない。古典的な推論の三分法として、演繹と帰納とアブダクションがある。演繹というのは昔からある数学的に正しい推論で、前提から必然的な含意を導くもの。帰納とは、いろんな個別事例から一般的な帰結を得るもの。アブダクションは、これらとは少し異なる観点であり、原因を導くということ。科学における帰納法はフランシス・ベーコンから始まり、これが近代科学の基礎と言われている。一方、デイビッド・ヒュームは帰納法を否定した。帰納法自体の確かさが証明できないので、帰納推論の例をいくら集めても、帰納法自体を正当化することができず、帰納法が正しいというのは循環論法に陥らざるを得ない。カール・ポパーは、やはり帰納主義を否定し、観測を重ねても理論の証明にはならないが、反例があればそれを棄却することができるので、推測と反例による試行錯誤で科学が進歩するという「反証主義」を掲げた。その後、実際に自分たちで実証していくことが大事だとする「ネオ・プラグティズム」、ポパーの意見をさらに発展させた「パラダイム論」などが登場している。

科学自体にAIを使うということは、コンピューターの出現時からの中心的課題である。最初のエキスパートシステムと言われているDENDRALは、科学者が行うような判断を自動化した。ダグラス・レナートはヒューリスティック探索を使った概念の発見に取り組んだ。今世紀になって、ロス・キングなど自動的な仮説生成に取り組むグループが出てきた。ここでは推論が重要な役割をしている。ロス・キングは、アダム、イヴといったロボットサイエンティストを作った。

現在のデータ駆動の科学、機械学習は、基本的に帰納主義であるが、ヒュームらによる帰納主義への批判に対して、どのように立ち向かえば良いのだろうか。人間だと予想外の結果から思いついたりひらめいたりするが、帰納主義ではこれらを反映することが難しいのではないか。それならば科学者がどうやって仮説を発見しているかということもAI的に考えて、それを実現してやれば何かできるのではないかということが、一つモチベーションとしてある。人間の発見に関係する思考能力に相当する仕組みというのが何で、それをコンピューターに組み込むということ。AIでは機械学習以外にも、例えば不完全な知識で推論するといった研究があり、このようなものを利用しない手はないと考えた。

### アブダクションによる仮説生成

古典的な推論の三分法の1つであるアブダクションは、アメリカの哲学者チャールズ・パースが再発見し、これこそが科学的発見の一番重要な推論であると論じた。パースのアブダクションの定式化は非常に単純であ

り、簡単な論理形式で実現することができる。図2-10-1はアブダクションを論理形式で書いたものである。背景理論Bがあり、観測をEとしたら、仮説Hというものを導入することで、 $B \wedge H$ からEを導くことができるならば、このHを求めることがアブダクションの目的となる。これを演繹的に逆計算することで背景理論Bと観測Eから仮説Hを導くことができる。この演繹計算はコンピューターででき、これを導く仕組みが、私の学位論文にて提案したSOL導出である。ロス・キングらのロボットサイエンティストにおいては、SOL導出を用いた仮説発見が使われている。

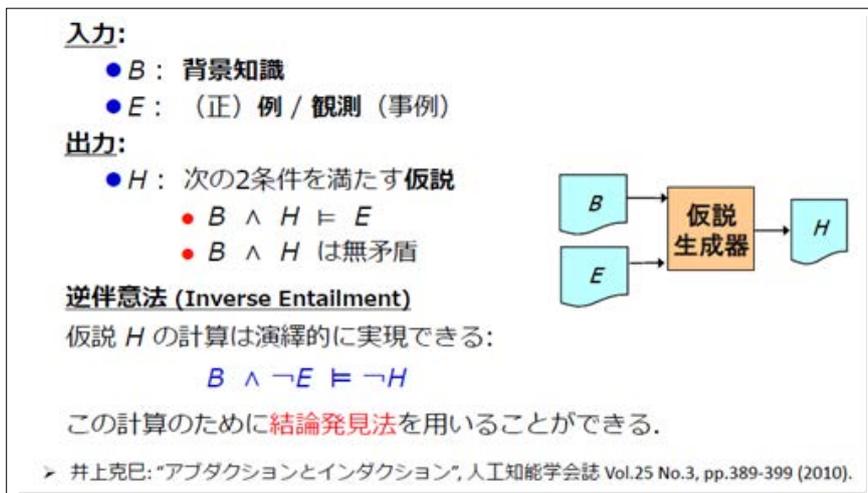


図2-10-1 アブダクション・機能における仮説生成の論理

2000年以降に出てきたシステム生物学は、複雑な現象を要素に分解する還元主義とは異なる立場を取り、生命・細胞・生体を系として考え、要素間の関係や相互作用を重視する。このため関係を述語で記述する述語論理と相性が良い。この述語論理による仮説発見をシステム生物学に応用し、2014年に「生物システムの論理モデリング」という本にまとめた。この論理モデリングには、次のような特徴がある。

- ・要素間の関係を制約や関係述語として宣言的に記述可能
- ・因果関係の第一近似として系の定性的性質が推論可能
- ・論理的説明により専門家の理解が容易
- ・背景理論(既存生物学知識, 既存データベース等)を伴った学習が容易
- ・不完全な理論における理論補完が可能
- ・論理的に可能な複数仮説の枚挙が可能
- ・能動的学習による実験計画作成が可能
- ・シミュレーションに代わる、推論によるモデル検査法が利用可能

我々が取り組んだ仮説発見の事例を図2-10-2に示す。アブダクティブな仮説発見器で得た仮説集合から、仮説のランク付けをすることで最尤仮説を得ている。

科学知識の生成・発展においては、図2-10-3で示すようなサイクルを形成することが非常に重要である。観測を説明するために仮説を発見し、その仮説から演繹的に予測が出て、予測を実験で検証し、帰納的に知識ができてくるが、そのときに、また別の何か説明できないことが出てきて仮説を修正するといったサイクルを

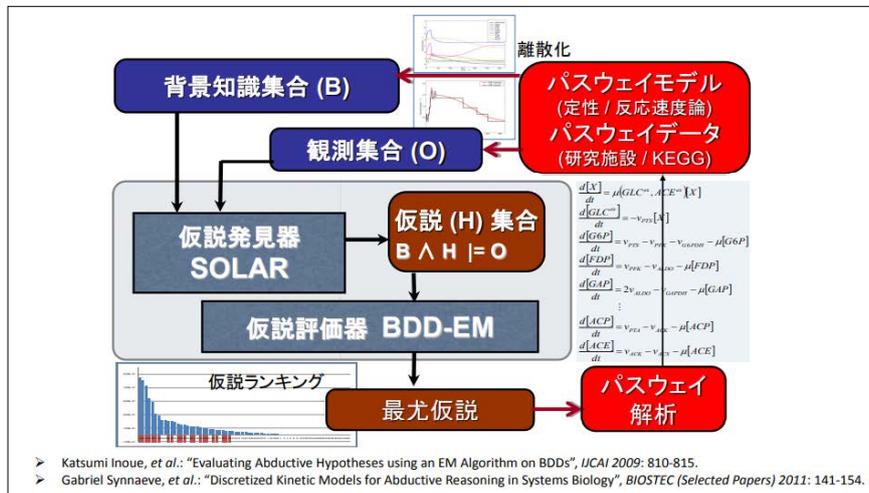


図 2-10-2 システム生物学における仮説発見

繰り返すことで、科学の理論は洗練されていく。1980年代のエキスパートシステムや、単純な帰納主義ではこうしたサイクルがないので、真の発見には結びつかない。キングらのロボットサイエンティストは実験計画を立てて仮説を修正するというサイクルを作っている。

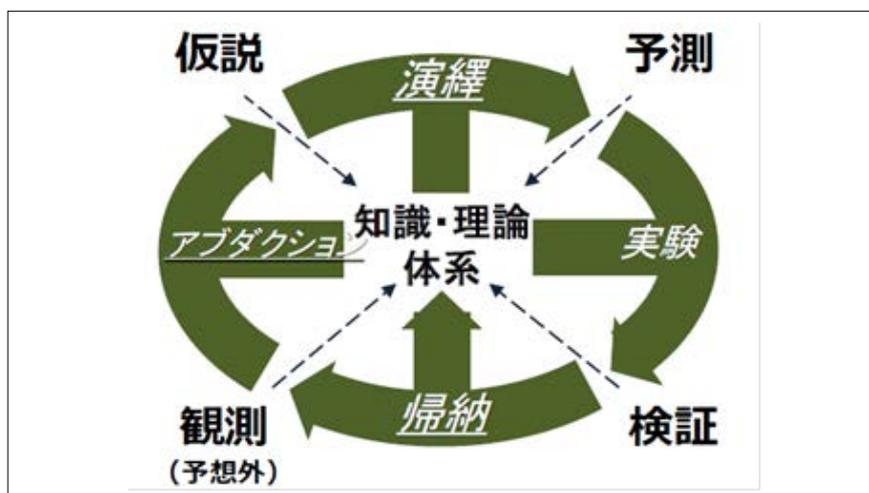


図 2-10-3 科学知識の生成・発展サイクル

メタレベル・アブダクション

システム生物学における仮説発見では専門的な背景知識を必要としたが、背景知識の獲得自体が容易ではなく、人間の思考過程をモデル化するという意味では少し弱い。このため、科学者が行う科学的な思考過程をさらに一般的に記述し、対象領域に依存しないような一般的な推論方式であるメタ推論を使うことで発見プロセスの自動化を目指した。メタ推論とは、オブジェクトレベルとそれに基づく推論のさらに上に位置づけられるメタレベルでの推論 (図 2-10-4) である。

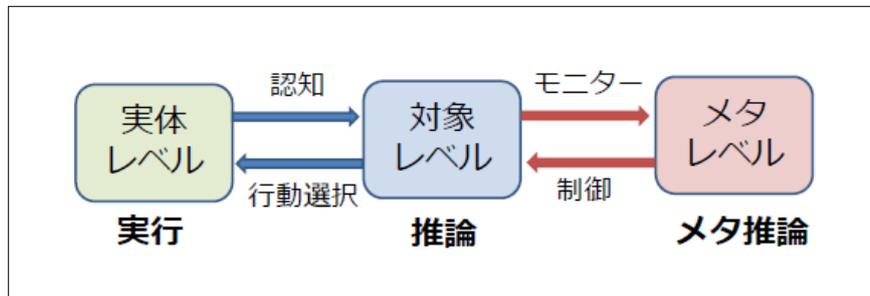


図2-10-4 メタ推論の図式化

このメタ推論をアブダクションでも行った。例えば、

$$\text{solve}(A) \leftarrow \text{rule}(A \leftarrow B) \wedge \text{solve}(B)$$

((A ← B) という規則があって B が証明できれば、A が証明できる)

というメタルールに対して、

$$\text{solve}(A) \ \& \ \text{solve}(B)$$

を事実としてアブダクションを行うことで（これをメタレベル・アブダクションという）、(A ← B) という規則を推論できる。このアブダクションをスキルサイエンスに応用して、チェロの演奏技法のコツを説明するという研究を古川康一（故人）と共同で行った。

メタレベル・アブダクションを生物学に応用した事例もある。生物のネットワークは、阻害する因子と促進する因子との両方がある非対称な因果ネットワークとなっているが、そこでのメタレベル・アブダクションにより、ある種のがんを抑える仕組みを解明するという研究を行った。また、Systems Biology Graphical Notation (SBGN) というものを背景理論、遺伝子発現データを観測結果とし、これらをリンクして仮説を導くといった研究も行った。

このようなメタレベル・アブダクションでは、アトムとしての事実を導くアブダクションをメタレベルで実行しているが、それをオブジェクトレベルに翻訳すると複雑な規則を導くことができている。

アブダクションの「観察事象」(Explanandum)と「説明」(Explanans)の関係を様々な分野の応用レベルで見ると、科学では「観測」から「理論・仮説」を導き、診断/医療では「病気の兆候や症状」から「病気の原因」、言語であれば「文書/発話」から「発話の理解」、シャーロック・ホームズであれば、「証拠」から「犯人」とか「動機」を導く、エンジニアリングな設計であれば「仕様」から「部品の組み合わせ」を導く、これらは全てアブダクションの例になっている。究極的には、芸術において「テーマ」を与えて「作品」を導くといったことも、広い意味でのアブダクションである。

### ダイナミックな環境における関係ダイナミクスについて

アブダクションによって、既にあるネットワークを補完しミッシングリンクを発見することはできるが、ゼロからスタートしてネットワークを作ろうとすると、アブダクションだけで構築するのは非常に無理がある。

そこで、複雑系で有名なブーリアンネットワークにおいて、状態の遷移列から帰納的にネットワークを導く研究を行った。遺伝子制御について、遺伝子をノードとし、前の状態と次の状態との関係をブール式で表すことで、状態の遷移をブーリアンネットワークとして表すことができる。

例えば、図2-10-5のように3つの遺伝子p、q、rがあり、qとrが発現していてpが発現していない状態011から、ブール式の論理に従い次の状態101に移るが、有限状態であれば、必ずアトラクターとよばれる

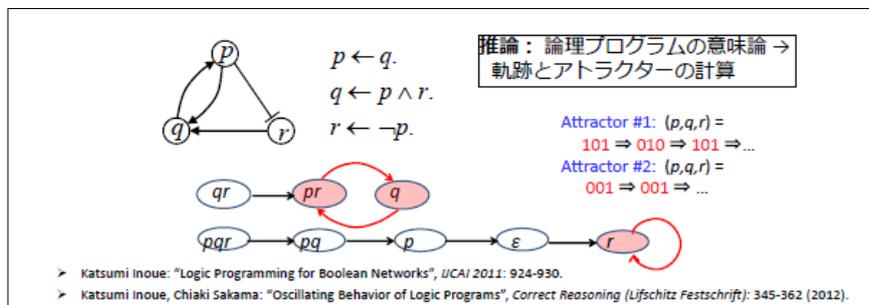


図2-10-5 3つの遺伝子のブーリアンネットワークの例

終点に辿り着く。

これを逆に利用し、状態の遷移列からダイナミクスを学習することで、先ほど述べたゼロからのネットワーク構築が可能となる。これは状態遷移からの関係ダイナミクス学習といって、物理系にしても生物系にしても、あるいはプログラム理論にしても、状態を伴って遷移していくような系であれば、全てこれで記述できる。時系列データから遺伝子制御のネットワークを学習したり、ロボットを適当に行動させてその遷移からロボットのアクション規則を学習したり、エージェントの振る舞いからその性質を見るとか、あるいは推論履歴から推論規則を学習するとか、プログラムのトレースからプログラムを合成するなど、様々な系に適用できる。

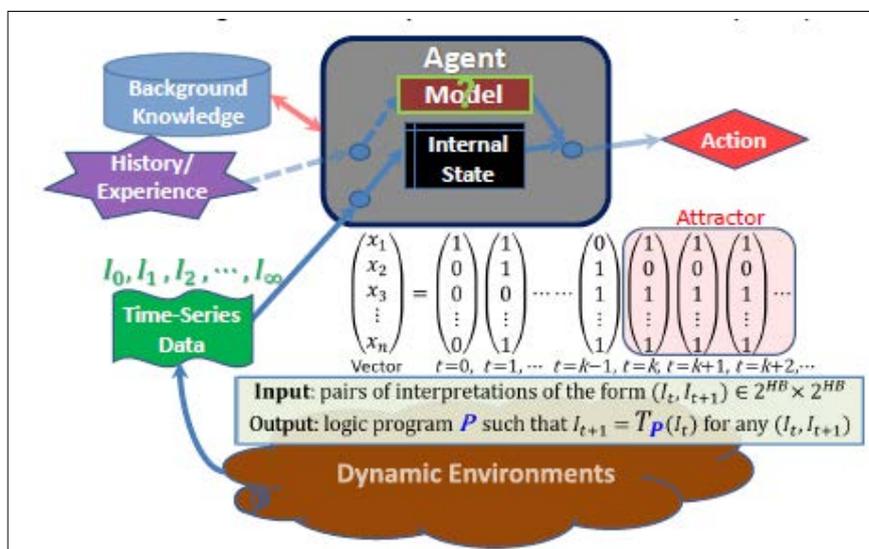


図2-10-6 解釈遷移からの学習：  
Learning From Interpretation Transition (LFIT)

図2-10-6に我々が提案した関係ダイナミクスを学習する方式を示す。これは解釈遷移からの学習：Learning From Interpretation Transition (LFIT) というものである。解釈遷移というのは、解釈というのはそれぞれの構成要素が論理でいうブーリアンであり、真か偽かという1/0からなるベクトルで表せ、これが時刻によって変化する。この変化を説明するようなプログラムを求めるのがLFITのタスクである。これは非常に一般的な枠組みになっている。

ただし、図2-10-6の遷移システムでは必ず前の状態に依存するということを仮定しており、非常に制約が強い。過去の状態にも依存するような遅延する系があっても良いし、1/0ではなくて多値や連続で状態が記述できて良いし、決定論的か確率的か、状態が同時に変化するかある値は変化しても別のは待っているという同期の問題もあり、この辺を考えた拡張によりバイオロジーや、物理系だとセルオートマトンの学習などに取り組んできた。

例えば生物系だと、イーストのセルネットワークは、12変数で記述される $2^{12}=4096$ の状態があるが、そのほとんどの状態から最後は同じアトラクターに収束するというロバストなネットワークを形成している。このネットワークはLFITのアルゴリズムを使うと瞬時に出てくる。ただし、組合せの問題なので状態数が増えると途端に計算爆発するが、ローカルサーチやサンプリングを使うことで、今ではヒトゲノムの数ぐらいいは対応できるようになっている。

**機械学習と知識表現の融合**

以上述べてきたことは記号処理をベースにしたネットワークの推論・学習である。記号処理は伝統的なAIで使われてきた。一方で最近のAIに関するホットな話題として、未知の状況においても機能するような「AIのロバスト性」ということが求められている。推論においてもノイズがある状況においても結果を出して欲しいが、これに関する技術としてジョン・マッカーシー以来論理ベースのAIでずっと研究されてきた常識推論 (commonsense reasoning) という常識を使った推論をマシンラーニングのコンテキストでやることも話題となっている。LFITの研究においても、観測である状態ベクトルは現実にはノイズを含んでいるため、ロバスト性を持たせる取り組みが必要と考えた。そこで、ニューラルネットワーク (NN) と解釈遷移からの学習 (LFIT) との融合を行った (図2-10-7)。ニューラルネットワークの入力を $p(t)$ 、出力を $p(t+1)$ として、NNに解釈遷移を学習させることにより、ノイズ頑健性が高まり、汎化性能が各段に向上した。

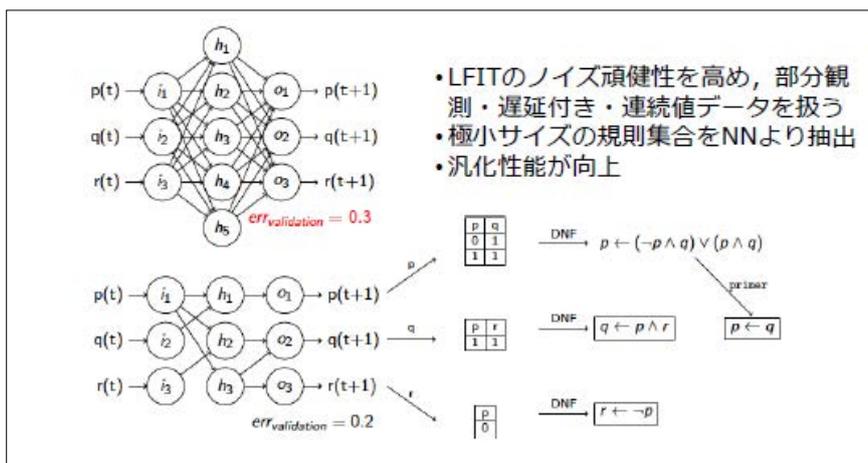


図2-10-7 NN-LFIT：ニューラル・シンボリックLFIT

機械学習と知識表現の融合に関する別の取り組みとして、少数サンプルからの学習を挙げる。体内時計に関する遺伝子の相互作用はサンプル例が少ないので、ニューラルネットワークによる学習は困難である。そこで、サンプルの学習ではなく、相互作用に関する論理プログラムを人工的に全部生成し、それをニューラルネット

トワークに学習させて、次に、少数のサンプル例をこれに当てはめると、学習した論理プログラムのどれかに適合するはずなので相互作用の予測が可能となる。先に挙げたNN-LFITでは一つのニューラルネットワークは一つのプログラムに対応していたが、少数サンプルと人工データからの表現学習では一つのニューラルネットワークが全ての論理プログラムを表しているということになる。得られた予測がどのプログラムから出てきたか分からないため説明性に欠けるという難点があるが、応用においては有用ではないかと考えている。

その他、ベクトル空間において論理計算を代数上の方程式に変換して、コスト最小化法により解くという研究を行っている。微分できない部分は閾値関数を使うことで、0/1の離散的な世界から連続値への橋渡しをしている。ロバスト性とスケラビリティがマルチコア環境で期待出来るなど、記号処理とニューラル処理の融合によるロバストAIの実現に向けていろいろな可能性が見えてきたところである。

今、注目しているのは、アブダクションを用いた新述語の発見である。1つの述語を1つの行列として、述語間の相互作用を学習し、アブダクションによって新たな述語(関係)を発見するというもの。これをフェイスブック上の知識グラフでやってみて、新述語を求めるといった研究をしている。

## 今後の展望

記号推論でどのように人間の知性が広がるかについて展望したい。最近のAIブームの立役者となったディープラーニングにおける学習対象はパターン認識が中心であったが、最近ではNeurIPSなどの機械学習研究を中心とする国際会議においても、機械学習と知識表現・推論との結合に関する論文がかなり増えている。また、人間が知性を獲得した一番の要因は記号を使ったことでもあり、記号をどうするかということを考えなければならない。そこでアブダクションが重要になってくるが、これまで手がけてきたようなきれいな記号だけの世界ではなくて、ノイズがあるような環境で使うことで何か発見ができるのではないかと考えている。

また、非人間的AI (Non-human AI) の活用が挙げられる。AIの中でもNNは人間を模倣することで進歩してきたが、人間の脳を必ずしも模倣していないAI、その中で推論系については人間よりも各段に進歩している。SATソルバーなど、人間の専門家が書いた専用プログラムよりも、高性能に問題を解くことが出来る汎用の推論システムが存在している。こうしたSATソルバーの内部では、同様の推論が出てきたら二度とやらないとか、ここが駄目だったら別の道に行ってみようとか、人間もやっていることに近いことを、人間の手に負えないレベルでやっている。

また、計算量のクラスでいうと、人間が解ける問題というのはクラスPという多項式オーダーに留まるが、SATソルバーをオラクルとして使うことでクラスNPの問題を解ける。さらに、その上のクラスPSPACEまで、QSATという限量化されたSATで解けるようになってきている。これにより数学の未解決問題が解けるようになってきた。さらに理論物理学の難問にも挑戦できるのではないかと期待している。

科学の在り方についても述べたい。物理学とか天文学とか生物学といった従来型のサイエンスではなくて、Technoscience、Transdisciplinary Scienceといった、いろんな分野が一体化したサイエンスがこれからの科学の在り方としてあげられる。また、AIをツールとして利活用するだけでなく、AI自体を研究対象とした科学として、従来の物理学や生物学からAIに貢献するような逆方向のアプローチもあって良いのではないかと期待している。さらに、オープンなサイエンスや世界同時参加型の研究が新しい科学のあり方となってきている。

また、喫緊の課題としてCOVID-19の問題がある。従来の時々刻々と出てくるデータと感染モデルなどを使ってパラメータ推定・感度解析・シミュレーションを行ったり、データ駆動科学としてAIによる機械学習

を取り入れたり、この辺りのことをするとCOVID-19に関して何らかの新たな発見が得られる。私の研究室では現在、重症になる人と無症状あるいは軽症の人はどこに違いがあるのかということを知りたいという動機から見ようとしている。

システム生物学に関する取り組みとして、COVID-19のためのオープンリソースの1つであるCOVID-19 Disease Mapは、文献等から得られた病態に関連した、シグナルや代謝や遺伝子制御に関する相互作用に関する知見を文献等から抽出し、データ統合を行って、SBGNを使ってネットワーク表現し、モデルの比較まで可能にしたようなプラットフォームを提供するといったことをしている。システム生物学的観点からは、このDisease Mapを使って、オミックスデータの解釈や仮説生成、予測といったことができる。このマップを使うと、各種の推論技法・ツールを特性や解析技法やインターフェースの観点から比較して見ることができるので、こういう研究を今後進めていけば良いと考えている。

また、レジリエンスについても述べておきたい。モシェ・ヴァルディヤクリストス・パパディミトリウはCOVID-19の下でレジリエンスの重要性について論じているが、こういう話は、我々も東日本大震災での経験から学んでいる。そのときは主に防災とか工学の観点からレジリエンスが語られていたが、情報システム研究機構のシステムズレジリエンス プロジェクト(2012-2016)では、科学的な観点からレジリエンスを捉え、どういう系がロバストであるかとかレジリエントであるか考えた。先ほど述べたTransdisciplinary ScienceとかMulti-disciplinary Scienceとして位置づけることができる。こういう観点は今の時代に改めて求められると考えている。

### 【質疑応答】

- Q：いわゆるタケズの埋込みと言われている力学系の理論では、決定論的な系であればずっと長く観察すればその状況が分かるかとされているが、アトラクターがちゃんと記述できるかという点で、状態遷移からの関係ダイナミクスの話と非常に相性が良いと感じた。そのあたりのことは、研究としてどういう位置関係になっているのか。
- A：無限の世界など、アトラクターも存在しないようなカオス理論みたいなものもあるが、有限の要素数、有限の状態決定論的であればアトラクターにいくことは保証できている。
- Q：確率と論理の結合による強力な推論、探索ツールも存在するという点だが、どのようなものか。
- A：確率論理プログラムでは、論理式や仮説に確率を割り振る機能を付加することで、パラメータ推定を用いて確率を求めることや、あるいは確率が与えられたときに帰結の確率を計算するといったことができ、2000年代の前半にかなり研究された。最近では、論理と確率はある種のニューラルネットワークを使って書け、それをプログラミング言語にしたものも出てきている。
- Q：アブダクションで仮説生成をするということが、研究の現場であまり使われていないように思えるが、もしそうだとすれば理由は何だろうか。
- A：知識表現・推論を重視してきたIJCAI等の伝統的なAI国際会議では、アブダクションは1970年頃から研究されており推論技法の一種として既に定着している。論理プログラミングの父でもあるロバート・コワルスキーはずっとアブダクションが認知モデルとしても重要であると唱え続けてきた。1990年頃からアブダクティブ論理プログラミングとして言語開発もされてきたが、2000年以降はツールとしてはあまり使われなくなっていた。理由の一つとして、アブダクションは組合せ問題の工学的な応用に使われていたが、その後、SATソルバーや解集合プログラミング(ASP)等の制約系ソルバーが非常に進化したため、それに取って代わられたという経緯がある。ただし、ロス・キングらのロボットサイエン

ティストや、古川 康一先生のスキルサイエンスでは、アブダクションが推論の中心であるため、こうした科学的な応用の開発現場では使われている。なお機械学習を中心とするNeurIPSのような国際会議でもアブダクションや常識推論に注目した研究が最近になって出てきている。そもそもアブダクションは科学哲学で主に議論されてきたため、科学の発展にとって大いに関係があり、今回のように科学とAIの関係をテーマにするのであれば、科学研究の現場で今後使われていくだろう。

- Q：メタレベル・アブダクションでは、同じ前提と結論から導き出されるルールは一意ではなく幾らでもあるのではないか。
- A：まさにその通り。マッチするルールは幾らでもある。だから、そういうところは制約で抑えるとかの対応が必要である。つまり、ここには可能性の道があるよ、ここにリンクを通してやれば証明がつながるよということなら確実に言えるが、それだけでは発見したことにならない。しかしその可能性を示してやることで、研究者がどこに注目すればよいか分かり、時として思いもつかないパスを発見することにつながる。
- Q：複数の仮説が、尤度によってソートできるとして、そのときに使う尤度のクライテリアというのはどのようなものか。シンプルなものの方がよりあり得るといったオッカムの剃刀的な科学の価値観が裏に入っているということか。
- A：複数の候補の中で共通している部分が多いと尤度が上がるという、非常に単純なやり方を使った。共通していれば、どの解を取ってきても妥当なものが含まれているだろうということ。ただし、一般的な仮説しか出てこなくて思いがけない発見にはつながらないという難点はある。
- Q：人間や科学者コミュニティが持っている、暗黙知やオッカムの剃刀といった発見に関するベースラインやクライテリアは、どうやってアブダクションのシステムに取り込むのか。
- A：仮定可能述語というものを設定することで、背景知識にこのようなクライテリアを組み込むことはできる。これは現実のアブダクションのシステムでも使っているものである。要するに、全て何でもいいとしてしまうと得られる仮説が多過ぎるので、「仮定できるものはこれとこれです」というサブセットを用意したり、仮定する専用の述語を用いたりすることで、背景知識的に考えることができる。
- その上で、どの仮説を選ぶかについては、我々はEM法を使った仮説評価を行ったが、この評価方法を入れ替えることで、選ばれる解が変わってくる。
- Q：北野宏明さんにノーベルチューリングチャレンジの話聞いたときにも、仮説がたくさん出過ぎて、それを絞り込むところがかなり難しいとのことだった。課題になっているのはどの辺なのか。
- A：仮説評価というのはエクストラロジカルなところがあるので結構難しい。ロジックで表したとしても、そのロジックで出てきた仮説がたくさんあると、いろいろな評価方法があって、そこが問題かなと思う。また、連続系あるいは微分方程式系の論理とどう橋渡しするかについては、いろいろ方法論が考えられていて、例えば微分方程式系をブーリアンネットワークに翻訳する技術があって、それでロジックレベルでの影響を見ることができるとは、それでどれくらい正確なことができるか分からない。そのような研究をもう少し進めようと考えている。
- Q：背景知識集合にニュートン力学を入れておいて、仮説として相対性理論が出てくるというふうにするためにはどうすればいいのか。
- A：ニュートン力学は相対性理論に包含されているので、ニュートン力学だけでは相対性理論は説明できない。全体を包含するような理論をつくると、それは広い意味のアブダクションになるが、それでニュートン力学が部分的に説明できるということを示せばいいのだと思う。究極的にはそういうことが

できるかもしれない。

Q：アブダクションで仮説を生成し、AIが人間が解ける範囲を超えた非常に高度な推論をし、ロボットサイエンティストが検証してということであれば、AIで閉じてしまっているようにも見えるが、将来的に科学の営みで人間はどういう役割を演じられるのか。当面、人間の手に残る部分は何か。

A：計算についてコンピューターは人間を超えているが、人間のツールとして使われている。この話もそれを拡張したようなものだと思う。今後も人間のツールとして残るとののではないか。アブダクションだけで人間の思いつきの様なものがどこまでできるのかまだ分からない。別のドメインでやったことを転移する「類推」という推論形態がある。もともとの仮説空間には入っていないかもしれなくて、仮説空間を加工してやるといったことをしなきゃいけない。当面、そこは人間が介在する要素かと思う。最終的にやっぱり人間が判断したい。信用できる機械があったとして、そこをどこまで許容するのか。今のところ、最終チェックは人間がやるということではないか。

## 2.11 機械学習によるデータ駆動型研究が拓く創薬と医療

山西 芳裕 (九州工業大学)

今日は機械学習と創薬について話をテーマを中心に紹介していきたい。私は10年ぐらい前に創薬に関する機械学習の本を出したことがあるが、時代はかなり変わったと感じている。

ここ最近、新しい薬を生み出すのが非常に難しくなっている。お金も時間もかかる。ほとんどが失敗に終わってしまう。動物ではうまくいったけれども人では有効性が示せないとか、薬の体内での吸収や代謝が悪い、いいものはデザインできたけれども大量生産するのが難しいとか、また、臨床試験まで行ったけれども、最後に毒性が見つかって頓挫するというケースがほとんどである。

そこで、ドラッグリポジショニングという概念が注目されている。既に承認されている薬とか過去に失敗に終わった化合物の新しい効能を見つけて、本来とは別の病気の薬として開発しようというアプローチになる。既存薬であるので、人に対する安全性とか作り方とかというのは確認されているので、高速で低コストで低リスクな創薬というのが可能になる。有名な例で、シルデナフィルという薬がある。商品名のバイアグラの名前のほうが耳にするとと思うが、これはもともと狭心症という心臓病の薬だったが、現在では男性機能障害の薬とか肺高血圧症の薬としても使われている。また、新型コロナウイルスの治療薬として、本来インフルエンザの薬であったアビガンが注目されたりしているが、それもこのドラッグリポジショニングに入る。これまで出てきた新薬の多くが実は既存薬の適応症の発見でもたらされたという歴史がある。例えば高血圧の薬が発毛薬になったりとか、緑内障の薬がまつ毛を伸ばす美容薬になったりとか、抗鬱剤が禁煙のための補助剤になったりとかいろいろ成功例はある。ただ、これまでの成功例というのは偶発の発見に大きく依存していたものがほとんどである。

2.11  
機械学習による  
データ駆動型研究が拓く  
創薬と医療

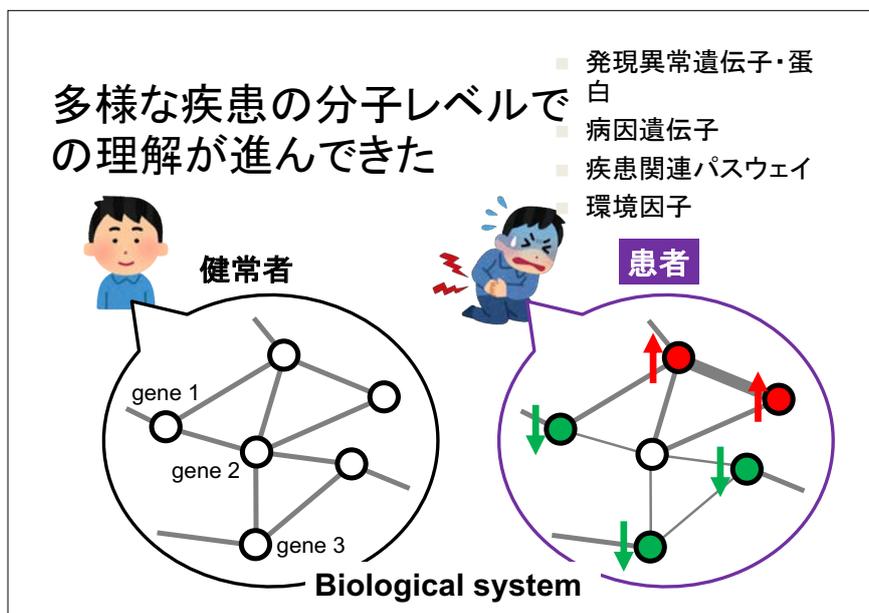


図2-11-1 多様な疾患の分子レベルでの理解が進んできた

そこで、我々のグループではドラッグリポジショニングの問題である薬と疾患の関係性を自動的に予測するような機械学習の手法の開発を行っている。この薬はこの疾患に対して効くことが分かっているが、こちらの疾患にも効くのではないかと、そういうのを自動的に一網打尽に予測してやろうと。これによって偶然の発見から脱却したい。

こういうことが可能になってきたのも、最近いろんな病気で分子レベルでの理解が進んできたというのもある。病因遺伝子とか環境因子とか、その病気に関係するような分子の情報が蓄積されてきている。この図2-11-1は正常人、つまり健康な人ではたくさんの遺伝子が協調して手を取り合って働いている様子を示しているネットワークであるが、患者の体の中では、ある遺伝子は暴走して、ある遺伝子は機能不全に陥ってという状態になって、その遺伝子ネットワークが破綻している状態と考えることができる。

この疾患の病態というのは、実は分子的な特徴が異なる疾患の間でも共通する場合がある。例えばバイアグラの例であると、男性機能障害と肺高血圧症、これは全く違う病気のように一見見えるが、PDE5というたんぱく質の異常発現が共通していて、それが同じ創薬ターゲットになって新薬開発に結びついたという経緯がある。

そこで、まず我々はいろんな病気の間での横断解析というのをやった。共通する分子的な特徴が多ければ多いほど近くに位置するように病気をマッピングした（図2-11-2）。

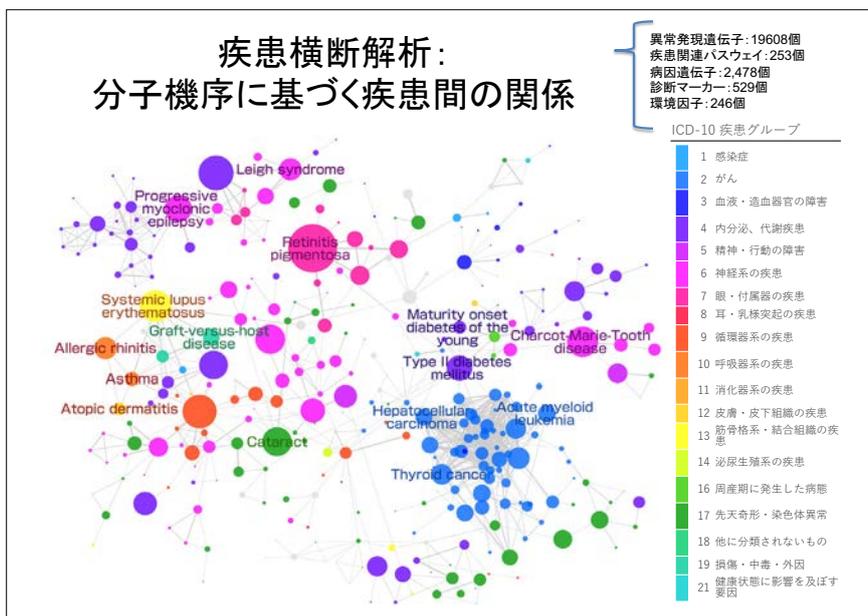


図2-11-2 疾患横断解析：分子機序に基づく疾患間の関係

1つの丸が病気に相当している。近くにあればメカニズムも近いのではないかとこのように考えられるが、この辺りは糖尿病が集まっていて、この辺りはがんが集まっているが、がんと糖尿病は近くにあることが分かる。これは一見全く違う病気のように見えるが、最近の医学的研究で慢性炎症という現象を通じて根底ではつながっているということが分かってきているし、両方に対して効くような薬剤、メトホルミンとかの発見なども報告されている。また、アレルギー性鼻炎、アトピー性皮膚炎、ぜんそくが近くに集まっているけれども、これ

も発症する部位とか臓器というのはばらばらだが、免疫システムの破綻という意味では共通性がある。なので、これまでの医学では人間が定義してきた臓器別の疾患分類、これが当たり前だったが、このデータから導かれるメカニズムによる疾患分類の可能性というのが示唆される。

我々はそういう疾患の類似性とかをうまく使って、また機械学習モデルによる治療薬の予測システムの開発を行っている (図2-11-3)。

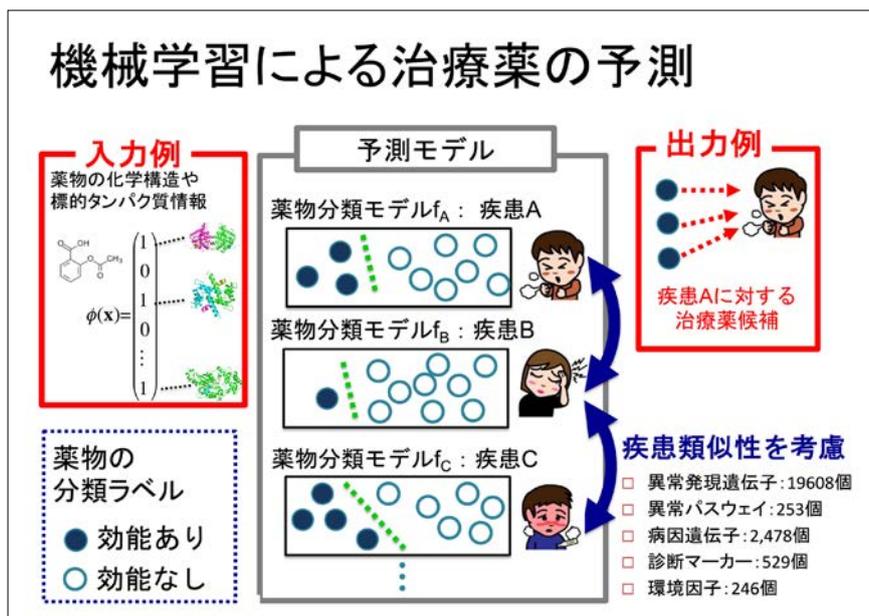


図2-11-3 機械学習による治療薬の予測

入力情報としては、薬の化学構造であるとか、それが体内でどういうふうな生体分子に結合するのかという情報を入れてやる。出力としては、この病気に対してはこの薬が効くかという薬のリストを出してやる。内部で何をやっているかというと、疾患ごとに予測モデルをつくっている。例えば疾患Aに対しては、これらの薬は効くか効かないか、疾患Bに対してこれらの薬は効くか効かないか、疾患Cに対してはこれらの薬は効くか効かないか、こういったものを予測するモデルを立てている。ただ、機械学習は学習データの質とか量にかなり依存するので、そういう学習データがあまりないとうまくいかない。例えば疾患Bにはあまり正例の学習データがないと。そういうときに疾患の類似性を見てやって、例えば疾患Aと疾患Bが似ていれば、この疾患Aの情報を疾患Bに転移してやることによって疾患Bの予測モデルの性能を高めてやろうというようなアルゴリズムの開発なども行っている。

ドラッグリポジショニングの手法は、世界中でいろんな研究者が研究開発していて競争が激しい分野だが、いろんな手法等を比べてやると高い予測精度を達成することが確認できている。実際には日本とかアメリカとかヨーロッパで承認されている8,000個の薬に対して、それらを解析してやって1,000種類ぐらいの疾患に対する効能を予測している。アウトプットとしては、薬物と疾患のネットワークという形で出てくる (図2-11-4)。青が薬、緑が疾患、グレーが既知の関係性で黄色が予測した関係性を表す。例えばこの辺りの薬というのは、本来はぜんそくの薬だが、アトピー性皮膚炎に効くのではないかと予測された例ということになる。

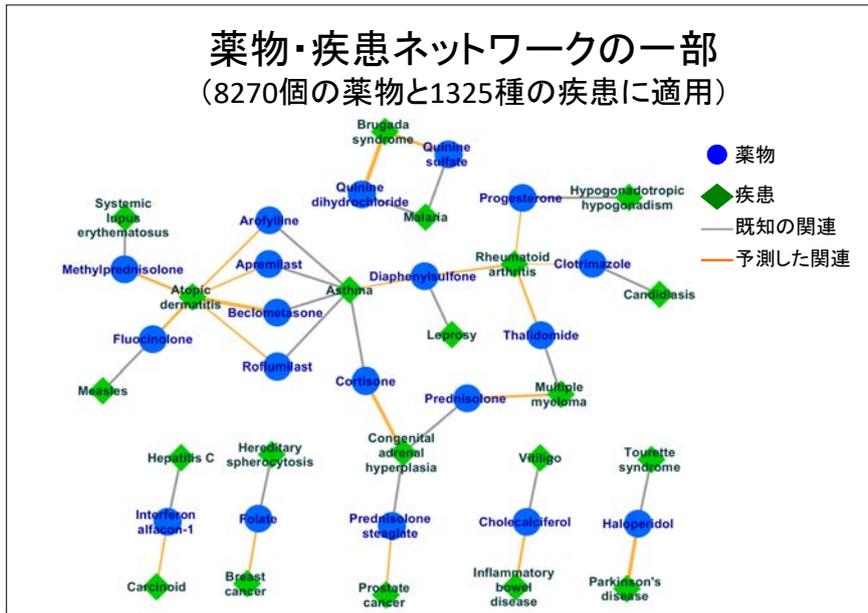


図2-11-4 薬物・疾患ネットワークの一部

これまでは薬と病気の間を予測するというをやっていたが、実際にはそれだけでは創薬の現場ではなかなか使いものにはならず、なぜ効くのか、なぜそういうふうな予測結果が出たのかというメカニズムの考察まで必要になってくる。その辺りを考慮した薬効予測の研究というも行っている。ほとんどの薬というのは体の中でたんぱく質などの生体分子に結合して、相互作用して疾患に対する効能を発揮する。ただ、薬の分子というのは狙っているたった一つの治療標的のたんぱく質だけに対して作用するわけではなく、それ以外のたんぱく質に対しても作用してしまい、オフターゲットの効果によって様々な副作用を引き起こしてしまう。ただ、ある副作用というのは別の疾患の患者さんにとっては効能になる場合もある。例えば血圧を下げるという薬は、高血圧の人にとっては効能だが、低血圧の人にとっては薬害、毒になってしまう。そういった薬のメカニズムというのをコンピューター上で表すことを考えた。

ここでは薬、たんぱく質、疾患のネットワークを予測するという問題と捉えてドラッグリポジショニングの問題というようなシステムの構築というを行っている（図2-11-5）。いろんな病気の空間がある。ここにいろんなたんぱく質の生体分子の空間もある。ここにそれを制御する薬の候補となるような化合物の空間もある。これらはどの薬がどのたんぱく質を標的として、どの疾患に効くのかという異なるノード間のエッジを予測するという情報学的な問題に落とし込んで解いてやろうというのがここでの狙いになる。承認薬、既に承認されている薬でも何でその薬が効くというのは分かっていない薬も結構ある。そういったゲノムワイドな薬とたんぱく質の相互作用を予測するための手法の開発というのを取り組んでいる。いろんなアプローチがあるが、まずは薬の化学構造をたんぱく質に結びつけてやるケモゲノミクスと呼ばれる枠組みについて紹介したい。

基本的な考え方としては、化学構造が似ているような薬は同じようなたんぱく質に対して相互作用するだろうというふうな仮定の下に予測を進めるようなアプローチになる（図2-11-6）。この場合、薬の化学構造の類似性、たんぱく質の類似性をどういうふうに捉えるかというのが重要になってくる。この辺りもいろいろやり方はあるが、薬の化学構造を二次元の数学的なグラフ構造として捉えて、グラフ構造の類似性を計算したりとか、たんぱく質のアミノ酸配列、これは文字列として捉えて、この機能部位を考慮しながら、たんぱく質の類似性

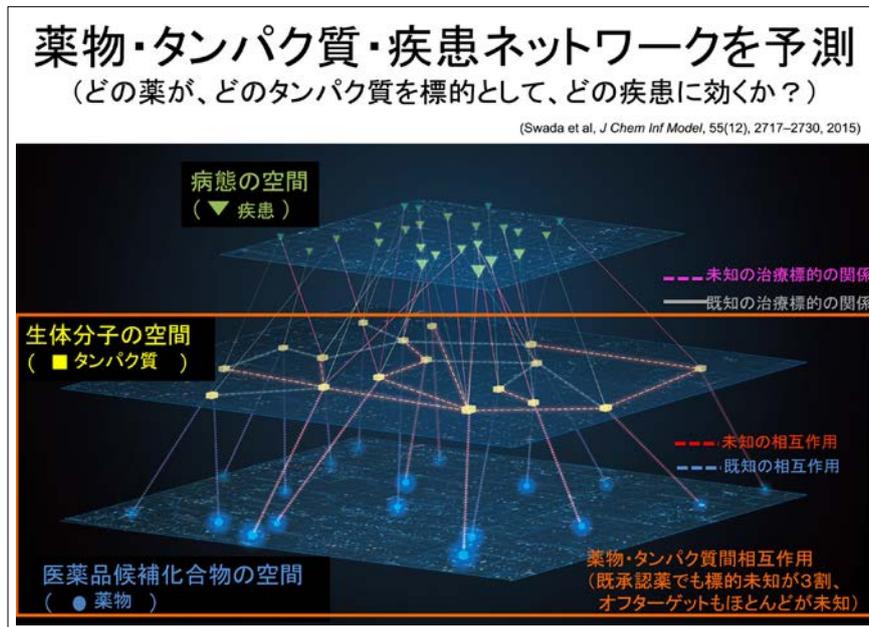


図2-11-5 薬物・タンパク質・疾患ネットワークを予測

2.11 機械学習によるデータ駆動型研究が拓く創薬と医療

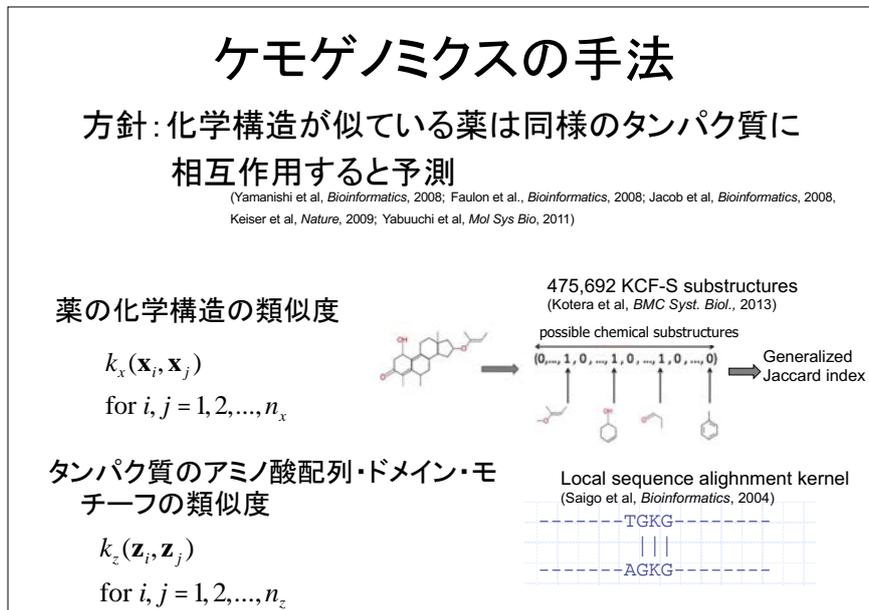


図2-11-6 ケモゲノミクス的手法

を評価するということを行っている。これを使って未知の薬とたんぱくの相互作用を予測したいというのが目的になる。

これを機械学習的に定義すると、これで薬をX、たんぱく質をZで表している。ペアをX、Z、これを入力するような予測モデルというのを考えている。カーネル法という機械学習の予測モデルというのは薬の類似性

2.11

機械学習による  
データ駆動型研究が拓く  
創薬と医療

とたんぱく質の類似性を入力するような形に書き換えることが可能になる。このモデルを学習するために薬の空間、たんぱく質の空間でデータがあったときに薬のたんぱく質のペア、これが相互作用するペアかそうではないかというのを分類するようにモデル学習してやって、実際に相互作用するのかどうか知りたい薬とたんぱく質のペアがあったときに、そのモデルを当てはめて、これらが相互作用するかしらないかを予測しましょうというのがここでのアイデアということになる。ここでも同様に8,000個の薬の解析というのをやってやった。ここでアウトプットは、薬とたんぱく質と疾患の3種間のネットワークという形で出てくる。さっきは薬と疾患だけだったが、仲介するノードとしてたんぱく質が入っていく。

具体的に見てやると、例えばここにピオグリタゾンという薬がある。これは、本来は糖尿病の薬である。これがMAOBというたんぱく質に対する相互作用が予測されたので、パーキンソン病に効くのではないかとというふうに考えられるケースになる。MAOBというのはドーパミンという物質を分解する酵素だが、これを阻害することによってドーパミンが増えて、パーキンソン病の治療効果につながるというメカニズムが考えられる。最近、ヨーロッパで大規模なコホート研究が発表されて、実際そのピオグリタゾンを飲んでいる人はパーキンソン病になりにくいというデータが出ているので、この予測結果を裏づけるエビデンスになるのかと思う。

これは医療の視点からとしても、投薬の最適化にもつながるのではないかとというふうに期待できる。よく高齢者とか複数の病気を持っている人とか、1日に10種類もすごいたくさん薬を飲む人とか、ポリファーマシーと呼ばれているけれども、普通はそれぞれの病気に対してそれぞれの薬を飲む。そうすると、飲む薬が増えるので副作用のリスクも大きくなってしまう。もし複数の疾患に対して効く薬が1個あれば、それを処方してやればいいと。それによって薬を飲む量が減って副作用リスクを下げることができるのではないかと期待できる。例えば私の友人で糖尿病とパーキンソン病、2つ持っている人がいるが、それぞれ別の薬を飲んでいるが、例えばさっきのようなピオグリタゾン1個だけ飲むことで処方できたらうれしいわけである。

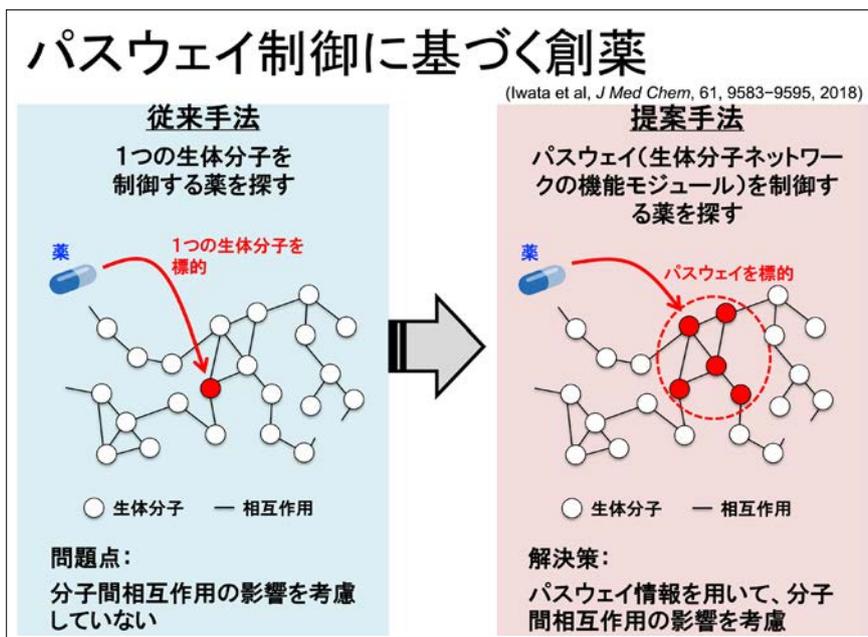


図2-11-7 パスウェイ制御に基づく創薬

次に、パスウェイ創薬というトピックに移りたいと思う。普通は創薬は1つの生体分子、1つのたんぱくに注目して、それを制御するような薬を探すというアプローチが一般的である。ただ、1つのたんぱく質の制御が必ずしも期待する薬効につながらない場合が多い。というのも、やっぱり体の中ではたんぱく質というのはつながっているので、注目する分子の周りの相互作用の影響というのは考慮できていない。そこで、我々はこのグローバルな生体分子ネットワークの機能モジュールであるパスウェイに注目し、これを制御するような薬を探索するというアプローチができないかと考えた（図2-11-7）。

ここではオミックスデータ解析と分子ネットワーク解析の融合を行っている。オミックスデータとしては、ここでは薬に反応する遺伝子発現のデータを使っている。ある薬を人の細胞にかけてやると、細胞の中で2万個ある遺伝子がどういうふうに活動が変化したかというふうなデータが得られる。これに代謝経路であったりとかシグナル伝達経路であったりとか、そういった生体分子のネットワーク、これを融合解析することによって期待する効果を持つような薬を探しましょうというのがここでの提案になる。遺伝子発現の情報を使うと何がうれしいかというと、薬を人の細胞にかけるとどういうふうな遺伝子が活動が上がっているのか下がっているのかというのが分かるので、これをさっきのパスウェイ分子ネットワークにマッピングしてやると、どういうふうなパスウェイが活性化しているのか、どういうふうなパスウェイが不活性化しているのかというのを推定することが可能になる。

ここでは、それを逆に用いて薬をスクリーニングができないかということを考えて。ここでは、実は抗がん剤ではないのだが、隠れた抗がん作用を持つような薬を見つけてやるという課題に取り組んだ。これは東大医科研の谷先生との共同研究になる。ここでがんに関係するパスウェイ、細胞周期のパスウェイとかp53のパスウェイとかを制御するような化合物を探してやるということを行った。予測スコアが高いもので抗がん剤じゃないものを実際実験して確認した。肺がんとか直腸がんとかすい臓がんとか前立腺がんとか、そういったものに対して抗がん作用があるかどうかというのを実験してもらった。そうすると、上位10個中半分ぐらいでヒッ

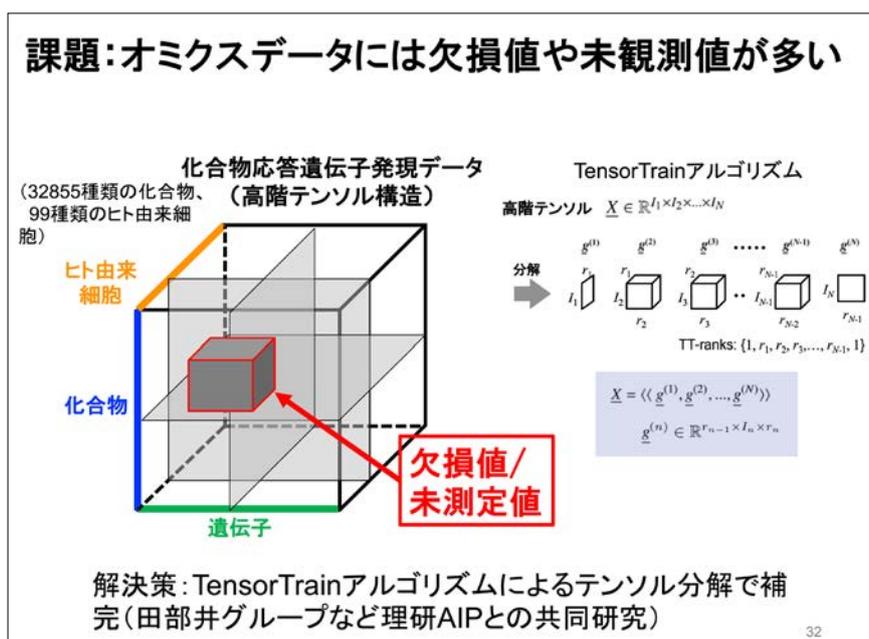


図2-11-8 オミックスデータには欠損値や未観測値が多い

トを見つけることができました。例えば精神病の薬である Penfluridol とか神経遮断薬である Promazine とかからも期待するような効果を見いだすような薬剤というのを見つけることができました。この最後の細胞というのは正常細胞である。がん細胞は減少する、がん細胞は死ぬんだが正常細胞には毒性がないような薬になっていたの、理想的なものを見つけることができています。

ただ、このオミックスデータ、遺伝子発現データのオミックスデータというのは有用だと言ったが、現実のデータの問題としては欠損値とか未観測値が非常に多いという問題がある。情報学的な視点から見ると、このようにテンソル構造のデータになるが、結構抜けが多い。

そこで、ここではテンソル分解のアルゴリズムの一つである TensorTrain アルゴリズムというのをうまく使って、欠損値を補完するような方法というのを開発している。それを用いることによって、従来の補完法よりもより精度高くデータを補完することができて、さらに、それを用いて疾患治療薬の探索を行うと精度を向上することができるということを確認できている。このような技術を用いて今、大学病院の先生とか製薬会社とかいろんな人と共同研究を行っている。

今度は再生医療に対する応用もできないかということ考えた。これらの低分子化合物、薬によるダイレクトリプログラミングの情報技術の開発というのをやっている。普通は再生医療では皮膚の細胞を iPS 細胞に初期化して、そこから分化させて目的の細胞に変化されるというアプローチが一般的かと思う。ただ、最近の研究で iPS 細胞を経由しなくても直接皮膚から目的の細胞に、例えば肝細胞に変換できるということが分かってきた。これを、ダイレクトリプログラミングと呼ぶ。ただ、方法としてはこれまで遺伝子の導入というのが一般的だった。それを今回我々の研究では低分子化合物、薬による誘導ができないかということに取り組んでいる。そういうふうな細胞変換というのは転写因子の組合せ、DNA に結合するたんぱく質には転写因子というものがあるが、この組合せで細胞変換というのは決まってくる。例えば皮膚から iPS 細胞に変換するときには、この Oct4、Sox2、Klf4、c-Myc という4つの転写因子が必要であるが、これを発見することによって山中先生はノーベル賞を取られている。ここの皮膚の細胞に例えば Hnf4a、Foxa という転写因子を入れてやると、これが直接、肝細胞になる。あと、Gata4 とか Mef2 とか Tbx5、これを入れてやると皮膚の細胞が心筋細胞に変化する。そういうふうに転写因子の研究が進んできている。ただ、この最適な転写因子の組合せが分かっていない細胞変換というのはまだまだたくさん残されているので、まずはこの細胞変換に重要な転写因子の組合せを見つけるような手法の開発というのを我々は行っている。最近、いろんなオミックスデータが手に入ることになってきているので、これから細胞変換に重要な転写因子セットを予測するような手法の開発というのをやっている。遺伝子発現のデータとか、転写因子のデータとか、エピゲノムのデータ、また、分子間相互作用のデータなどを使った方法を開発している。例えば皮膚の細胞から肝細胞への変換を予測するような転写因子を予測するということを行ったときに、このような転写因子が上位に上がってきた。スコアの高い順に並べているが、実際に妥当性を調べるために文献で検証してやると、この青いところというのは既報の転写因子に対応していた。つまり実験的に見つかった正しいと示されているものをうまく再現できていることが分かる。実際、肝細胞以外に軟骨の細胞への変換とか皮膚から神経細胞への変換、皮膚から心筋細胞への変換など様々な細胞変換に対して、結構高い精度で予測できるということが確認できている。最終的な目的は、この転写因子の導入というのを低分子化合物で置き換えたい。転写因子を使った導入だと、ウイルスを使って細胞に入れるので、ウイルスに起因する発がんリスクの問題を回避したい。また、生体内で目的の細胞が作れば、細胞移植を不要にすることができる。

分化誘導するための化合物は、未知の組織というのがたくさんある。今いろんな組織に対して、その細胞変換を実現するような化合物の探索というのをやっている。

2.11

機械学習による  
データ駆動型研究が拓く  
創薬と医療

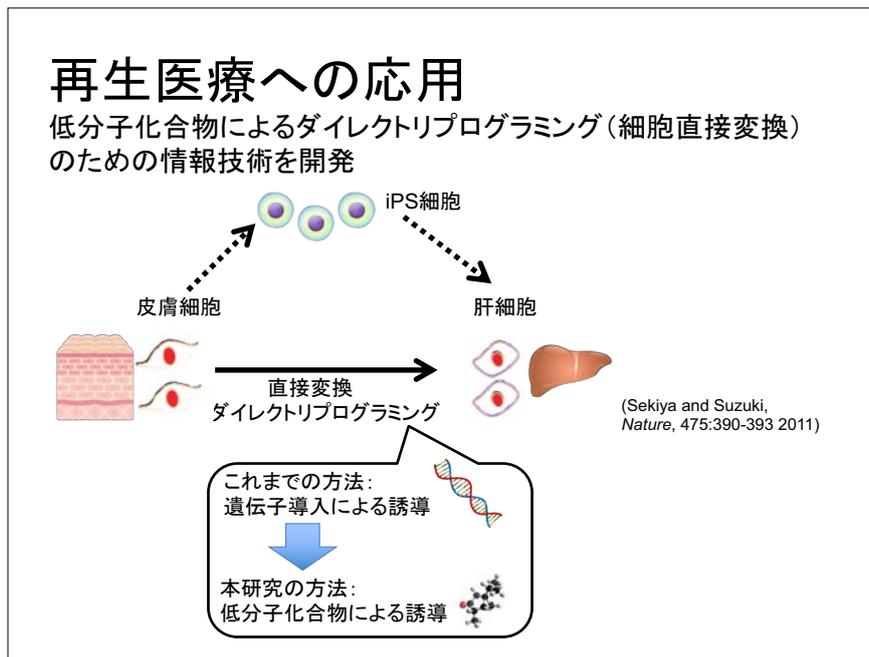


図2-11-9 再生医療への応用

2.11  
機械学習による  
データ駆動型研究が拓く  
創薬と医療

次に、漢方薬への応用について紹介したいと思う。これまで医薬品が中心的な話だったが、薬局とか日常的に手に入りやすい物質から新しい健康効果を見いだすことができないかということ考えた。漢方薬というのは日本では伝統的・独創的な医療体系で使っている人も多く、身近で有用だが、なぜ効くのかというメカニズムはほとんど分かっていない。ただ、300年ぐらい前に確立された方法なので、現代病、例えばがんとか認知症とか、そういったものに対する効能が眠っている可能性は大いに考えられる。ただ、漢方薬のメカニズムというのは非常に複雑である。通常の医薬品、我々が病院でもらうような薬、西洋医学に基づく薬というのは、有効成分の化合物というのは1つの場合が多い。それがどういふうなたんぱく質に結合するのかを見るだけで、ある程度薬効が推測できが、漢方というのは非常にたくさんの化合物が入っている。これらがどういふうなたんぱく質に対して作用するのかというのをそれぞれ考える必要があって、しかも、その標的分子というのは全然分かっていない。

そこで、まずは漢方薬ごとに標的たんぱく質群を予測するというを行った。これは富山大学の門脇先生との共同研究になる。漢方薬の成分化合物、人のたんぱく質、全てを解析してやった。あらかじめ数百万個の化合物、たんぱく質の既知の相互作用のデータから、これらの化合物が相互作用のペアかそうではないかを予測するような機械学習のモデルというのを構築する。ここでは、グラフ畳み込みニューラルネットワークというものをを用いて、この予測モデルを構築している。画像認識の分野で畳み込みニューラルネットワークの有用性が特に認識されているけれども、これを化学構造、薬とか化合物の化学構造に拡張したものがグラフ畳み込みニューラルネットワークと呼ばれるものになる。実際相互作用ペアが予測されてきた化合物、たんぱく質、これを漢方ごとにグルーピングしてやった。そうすることによって、漢方薬Aは実はこういうふうなたんぱく質に対して作用している、漢方Bは実はこういうふうなたんぱく質に対して作用しているんだと、漢方Cはこういうふうなたんぱく質に対して作用しているのではないかと、じゃ、こういうふうな疾患にも効くのではないかと、こういうふうな口ジックで効能予測を進めてやった。全ての漢方薬を解析してやった。

1つ例を挙げる。防己黄耆湯という漢方薬がある。これは痩せる効果がある漢方薬として知られている。ただ、漢方の医学書を見ると、非常に難しい説明が書かれている。読み上げると、虚証の人の肥満症に用いると。体力は中程度から虚弱で、色白で筋肉が柔らかくて、特に下肢がむくみやすくて水太りタイプで、疲れやすくて汗をかきやすい人に適すると書かれている。虚証とか実証とか寒証とか専門用語が出てきて非常に難解である。なので、これをサイエンス的な視点から漢方を解剖できないかということ考えた。

ここでは、その漢方の標的たんぱく質とか適応可能疾患の予測ということを行っている。まず、防己黄耆湯の例、同じ漢方の例だが、この左側に予測スコアが並んでいて、ここにたんぱく質の名前が並んでいて、ここにたんぱく質の機能が並んでいて、右側に適応可能疾患の候補が並んでいる。この肥満症が上位に上がってきている。これは肥満症の漢方なので、これは全然不思議ではないが、何で肥満症に効くのかというのが分かっていなかった。これは実際どういうふうな標的分子なのか見てやると、opioid receptorとかが挙がってきている。なので、こういうふうなたんぱく質を介して痩せる効果が生まれているのではないかという既知の効能のメカニズムを示唆することができたりとか、あと、glucosidaseに対する相応性はあるので糖尿病に効くのではないかというふうな推測できたりとか、様々な応用が考えられる。

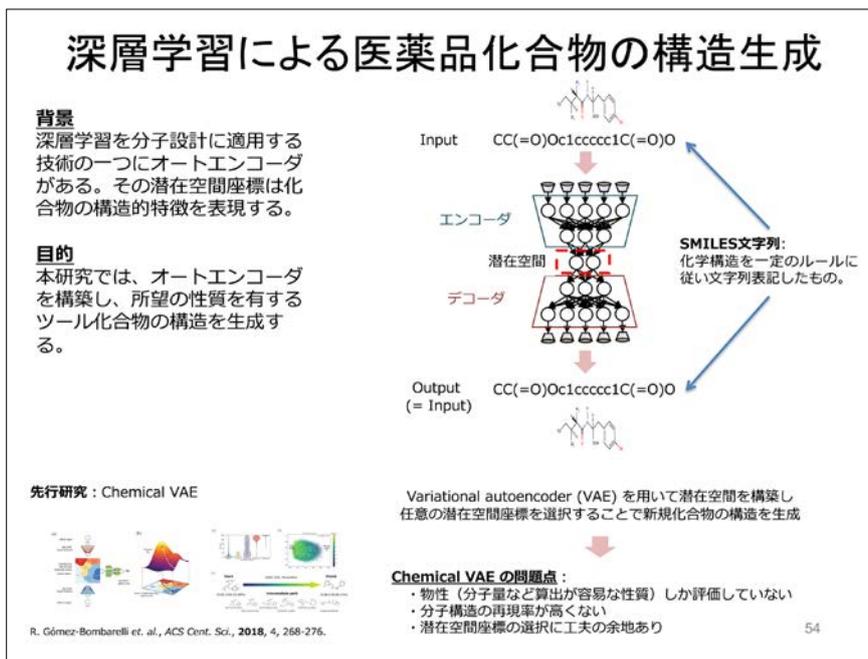


図2-11-10 深層学習による医薬品化合物の構造生成

ほかの例として腹痛の漢方薬で大建中湯というのがあるが、これは腸炎関連の大腸がんに対する効能と作用機序を予測できて、それをマウス実験で検証したという内容になる。左側の写真の左側は漢方を投与しないマウスで、右側が投与したマウスになる。ピンクになっているのは腫瘍、がんになり、これが縮小していることが分かるかと思う。腫瘍の数も減っていることが確認できる。なので、予測に基づく妥当性が確認できたかなと思う。現在もこのような技術を使って新型コロナウイルス感染症の治療薬の探索なども行っている。

最後に、医薬品化合物の分子設計の話をしたと思う。これまでは既存の薬とか既存の化合物の新しい効果を予測するという話だったが、存在しない化合物を設計するという話である。薬の候補となるような化合物

を設計するということは、ゼロから薬を作るという方針だと、考え得る化合物の種類というのは10の60兆個と言われていてほぼ無限の可能性がある。このように深層学習を用いてそういった医薬品化合物の構造生成をするような研究に取り組んでいる。

例えば、オートエンコーダーを用いた方法を紹介したい（図2-11-10）。オートエンコーダーというのは入力と出力が一致するように学習した深層学習のモデルになる。この潜在空間というのが化合物の化学構造の情報が凝縮された多次元空間に相当する。先行研究では、分子構造の再現率というのはあまり高くなかったり問題が多いが、今その辺の改良にも取り組んでいる。実際に我々が構築したオートエンコーダーだと、先行研究よりも4倍ぐらいの化合物を正確に復元することに成功している。また、医薬品を中心とする多彩な化合物を立体性を含めて再現することができるようになっている。

また、所望の性質を有するような化合物の構成を生成するということにも取り組んでいる。例えば理想の細胞状態を表すような遺伝子発現プロファイルがあったときに、これを模倣するような化合物を見つけたい。例えば、創薬標的のたんぱく質をノックダウン、過剰発現とかしたときの状態を模倣するような化合物をゼロから作ったりとか、そういうことを行っている。

最後にまとめになるけれども、機械学習によるビッグデータ解析によって化合物の新しい効能とか副作用の予測が可能であるという可能性を示した。その効能というのは、治療効果であったり健康効果だったりとか細胞の分化誘導能などに供する。また、承認薬以外の化合物とか漢方・生薬・食品・化粧品などのヘルスケアというのも可能になる。また、所望の生物活性を持つような医薬品化合物の新しい構造を生成することも可能であるということを示した。今日紹介した内容は、たくさんの方々との共同研究になる。この場を借りて感謝したい。

### 【質疑応答】

Q：データは主に正事例が入ってくる問題が多いか。効果がないというものは分かっていないものが多い。

A：そうである。正例だけあるというデータが非常に多い。ただ、中には負例もある程度ある。例えば、あるたんぱく質に関しては正例も負例も結構そろっているというデータもある。よく研究されているものというのは新しい発見とかにはなかなか結びつかないので、科学者としてやりたいのは、あまりみんながよく研究されていないたんぱく質とか病気に対して新しい制御する化合物を見つけたり、新しい薬を見つけたりなので、そういう場合はやはり正例が一部しかない状況が非常に多い。最近ネガティブデータも重要じゃないかということで、ネガティブデータを収集するようなデータベースの開発が世界中でも認識されてきて、広まりつつあるかと思う。やはり負例のデータというのはそれぞれの研究者の頭の中で閉じている場合、眠っているものがほとんどである。特に創薬の分野ではネガティブデータを十分に取るというのは難しい課題ではないかと思う。

Q：正例の話でいうと、例えばこの薬、化合物はこのたんぱく質に相互作用するとか、ある病気に効くとかいう、それというのはある確率的というか、統計的なものなのかと思う。そこはどう扱われるか。何かその重みをつけてそのデータポイントを扱うとか、何かそういうのはあるか。

A：おっしゃるとおりそこは問題になって、同じ化合物とたんぱく質のペアのデータだが、ある人は相互作用するかないかという報告をしたものがあって、その親和性というのが重みに相当する。だから、連続値のデータと1/0のバイナリーのデータが混ざっている状況である。なので、それを一緒に扱うのがやはりなかなか難しい。連続値があったら、ある閾値を設けてやって、それが閾値以上の場合にはもう相互作用をみなして、それ以下は相互作用しないとみなして、1/0に直してから解析するというのを

行ったりしている。

Q：最後の分子設計の話は大変面白い話で、いわゆる潜在空間とおっしゃったが、薬の潜在空間というのはどのぐらいの次元数を持っていればよさそうな感じか。例えば薬からたんぱく質を通して疾患に至るようなメカニズムを考える際に、その薬と疾患の間の関係から実は中に潜在空間としてたんぱく質の種類というのを考えれば、どういったたんぱく質があるかということを知らなくてもできるかもしれない。

A：最初にこのVAEの枠組みで発表された論文では、256次元で構築していた。ただ、それが十分かどうかというのは全然分からなくて、我々はそれよりもっと大きい次元の潜在空間を構築して研究を進めている。512次元とか1,000次元で構築してやっているが、それが本当に十分なのか、最適なのかというのはまだちょっと分からない。

Q：そうなったときに、この次元の軸は何か。知られているどういうメカニズムに対応するとか、この領域はその解釈についてどういうやり方があるだろうか。

A：この創薬の分野ではまだそういう研究例がないが、ほかの領域の研究だと、 $\beta$ VAEとかいう方法があるみたいで、潜在空間の軸に意味を持たせるように学習したりとか、そういう研究も結構盛んに行われている。

Q：実際にこの潜在空間を自分でちょっと動かしてみて、どういう薬が出てくるかというのを見て、この軸はこういう軸なんだということがやってみると分かるという感じであるか。

A：そういう方向も今考えているところである。

Q：その際に解釈可能ではないけれども、結果に物すごく効いてくる軸とか領域というのが見つかったという、つまりある意味その新しいコンセプトを発見したみたいな例というのはあるか。

A：残念ながらまだ今のところない。ただ、創薬の世界というのは、1つのたんぱくに結合するからこの病気が治るという単純なものではないと思う。だから、潜在空間の一次元がそのたんぱくに対応していたら、それはそれで面白いが、還元論的な、1つの要素に還元して複雑な現象を常に説明できるのかというと、創薬の場合はなかなか当てはまらない。だから、得られた潜在空間の1つの軸というのが実は複数のたんぱく質の同時の制御によるものに対応しているとか、1つのパスウェイに対応しているとか、そういうふうな生体分子間の複雑な相互作用とかそういうものに対応していると面白いのかなと期待している。

Q：逆に言えば、創薬の世界はそのメカニズムがはっきり分からなくても効果があると分かればそれでいいのではないという立場の領域だと思っていかが。

A：昔はそれでも薬は承認されていたが、最近はずせその薬がこの病気に効くかというメカニズムの情報を書かないと承認されないようになってきているらしい。なので、やっぱり作用機序は明らかにすることというのは、製薬会社の中ではやはり重要課題である。

Q：ほかの西洋医学の薬もそうだと思うが、その患者さんとかその細胞の個性とか、そういったものの影響はどう取り扱うか。

A：最近の医学研究で精密医療というのが注目されているが、個別化医療にも対応すると思うが、同じ病気でも患者さんのグループがあって、患者さんのサブタイプごとに効く薬が違うから最適な薬を選択する必要があるという最適な医療、精密医療の重要性というのが叫ばれていると思うが、それを漢方では虚証とか寒証とか、そういう言葉で経験的に言われてきたのではないのかなと思う。だから、現代医学における疾患の層別化の患者グループ、疾患のサブタイプと東洋医学、漢方の研究における患者さんのタイプ、虚証とか実証とか、そういうのが何か関係性とかが明らかになると、東洋医学と西洋

医学を結びつける画期的な研究になって面白いのかなと思う。

- Q : そういったときに、最初の教師データとしてのところにお医者さんだったりラベルをしている。効く効かないの論文発表もそうだろうし、この病態をこの病気とみなしてというのも多分人間がある種ラベルしているように思うが、そこはもうちょっとデータだけから何とかみたいな、その辺りはいかがか。
- A : 西洋医学的なアプローチの場合には、その辺りは客観的な指標があるので、結構集まりやすいと思う。がんとかはいろんなサブタイプのデータが集まっているが、漢方の場合には、この漢方の専門医の特殊な診断とか経験的なものが大きくて、漢方の専門医というのは数があまりたくさんいない。だから、そういう方に大量のデータを、客観的なデータを出すようお願いしても、やっぱり限られてしまうのではないかなと思う。
- C : 私が気にしているのは西洋医学側の話としても、最初にその機械学習にける教師データにある種間違ってラベルされていたりとか、あるバイアス、一定の恣意的なものだったり、認知していないものだったりというのが入っているような気がするが、あまり避けられないと思う。
- Q : 西洋医学と、それから漢方で多分複雑さの程度が何桁も違うということだと思うが、そうすると、スパースネスが全然違う。そういうことを実際、機械学習をやっている感じられるか。
- A : やはり組合せのデータというのはほとんどない状況なので、この化合物とこの化合物の組合せがいいとか、この生薬とこの生薬の組合せがいいとか、そういうものが大量にあれば機械学習の教師あり学習でアプローチできるが、漢方の場合にはそれが無い。その辺りはちょっと限界を感じている。
- Q : ドラッグリポジショニングにしる漢方薬のリポジショニングにしる、何かデータに基づいて新しい見方が、従来と違った新しい見方が見えてきているというのは面白いと思う。特にコロナとかで結構注目、期待されていると思うが、AIなりこういった機械学習とかで可能性があるものが出てきたというような論文かなんかもちらっと見た記憶があるが、今実際どんな状況であるか。
- A : 非常にたくさんの方が取り組んでいる状況なので、皆さんそれぞれの視点でこれが見つかった、これが有効だ、この薬は効くかもしれないとかいろいろ発表されているが、客観的な指標で評価されているわけではないので、その中でもどれが一番有望なのかというのはなかなか評価が難しいのではないかと感じている。今我々は本当にそれが科学的エビデンスがどこまであるのかというのをウイルスの専門家の人と一緒に、ウイルスの増殖を止めることができるのかとかを実験で検証しようとしている。
- Q : 今回いろいろな取組を紹介していただいた中でも、その解釈性を確保するための考え方というのは何通りか整理できるか。例えば、何かある程度メカニズム的なものとか構造といったものを事前知識として持っていて、その中で可能性とか類似性の探索をやっているの、大体出てくるものは従来の解釈の上に乗っかっているものだったり、あるいは全然もともとブラックボックス的に何か出てきたものを後づけで解釈を人間がしていくとか、その解釈性と機械学習のアルゴリズムはどんな種類を使っているかというのにもよるが、関係というのは何通りかになるか。
- A : やはり目的によっても違うと思うが、ある程度確固としたエビデンスがあって、この生体分子のたんぱく質がこの病気のキーになっているから、これを制御するようなものを探したいというときには、教師つき学習的なアプローチになるので、そこに解釈性とかを加味しようとしても、やっぱり同じような結果が得られるだけになるので、意外な発見とかそういうのは難しいのかなと思う。また、指摘された後者のほうだが、ブラックボックス的なところがあって、後づけで何かメカニズムが分かったとか、そういうほうが新しい科学的発見につながるの、サイエンスとしては面白いのかなと思う。個人的な意見だが、私はデータ駆動型の研究の長所というのは人間のバイアスをできるだけ排除して、また、思い

込みもできるだけ取り除いた上で新しい発見を試みるようなところに重要性があるのではないかと考えているので、人間の視点からだと思いつかないようなものをデータ駆動で可能性を示唆することができたということが一番価値がある、期待できるようなものかと考えている。

Q：今まで薬は特にドラッグリポジショニングもそうだが、特に治療に関して、治療する効果だとか治療する薬を探すということがメインだったが、これからはだんだん予防というほうにシフトしていく中で、予防に効く薬であるとかそういったようなところが重要になってくると思うが、その中で今までのドラッグリポジショニングと同じような方法で適用していけるのか、それとも違う戦略を取らなきゃいけないのか、その辺はどのような考えをお持ちか。

A：やはり病気になってから治すというこれまでの医療ではなくて、病気になる前に予防するという予防医療とか先制医療的な考え方は重要になってくると思う。もしそういう予防医療、先制医療をやろうとしたら食品であるとか漢方とか、そういったものが最適ではないかと言われている。実際に富山大学のグループはそういう予防医療にかなり力を入れていて、特に富山の漢方というのを有効活用して、病気になる前の状態を漢方で制御するような研究に力を入れてやっている。実際、病気の状態になる前の状態のときに、その臨界点を示すようなバイオマーカーとなるような分子を見つけて、それを制御することによって病気になるのを防ぐ効果があるかどうかというのを今いろいろ研究しているような状況と聞いている。そこで数理的なアプローチというのがどうしても必要になってくるが、そこは東京大学の合原一幸先生らのグループとその辺りの数理モデルを組み合わせ、共同研究することによって今取り組んでいるようである。だから、漢方とか食品とかで疾患を予防するようなアプローチというのは、これからは重要になってくる領域だと思っている。

Q：機械学習オリジンでこういう創薬に取り組まれている研究者って、5年前とか10年前はほとんどいなかったと思う。今は少しは増えてきたか。なかなか大学でそういう人が増えなかったのは、1つはやっぱり医学部とか製薬会社と一緒にやらなければいけないという壁が結構あったとは思うが、その辺りも少し最近では低くなってきているか。

A：大学でも、増えてきていると思うし、製薬会社でもそういう情報系の人を採るようになってきているし、AI部門とかもできているところもあるので、人口的には増えているのではないかと考える。ただ、大学では特定の人のところ集中しているような印象はある。あと、昔は創薬関係のデータ、あまり公開されているデータとかなかったが、最近はずごく貴重なデータのように思われるものが公開データでも出るようになってきているので、その辺りもアカデミアの人が研究はしやすくなった要因ではないかなと思う。例えばアメリカのPubChemとか化合物のデータを扱ったデータベースがあるが、あれは2003年から2004年当初は1,000万個か2,000万個ぐらいしか化合物のデータはなかったが、今は1億個とかかなりのオーダーの化合物のデータが出てきているし、また、化合物の生物活性のデータとかも、イギリスのデータベースでChMBLとかあるが、そこにも生物活性のこの化合物がこのたんぱく質に対してどれぐらいのレベルで作用したとか、そういうふうなデータがかなり蓄積されてきて、それはここ10年でデータの量が何倍にもなっているので、その辺りはかなり機械学習的なアプローチをする人にとってはアドバンテージになっている。今日少し紹介した化合物に対する遺伝子発現のデータとかは、アメリカのブロード・インスティテュートで作っているデータベースで公開されたりとかしている。

## 2.11

機械学習による  
データ駆動型研究が拓く  
創薬と医療

## 2.12 プロセスインフォマティクスWSからの内容紹介

真子 隆志 (JST-CRDS)

CRDSでは「プロセスインフォマティクス」という調査チームが発足しており、その中で行われたワークショップの内容が、今回の「機械学習と科学」にも通じるところがあるので、簡単に紹介する。このワークショップは、単結晶成長など無機化学的な話をメインとするものと、触媒や有機化学を中心とするものの、2回行った。

まず、高橋啓介准教授（北海道大学）による「触媒インフォマティクス:材料組成とプロセス条件を考慮した触媒設計」である。普通の実験や計算というのは、例えばある物質を合成する、あるいは何かモデルを作り、そこから実験・測定や計算するなどして、物性・反応を見るという流れである。これを試行錯誤して物質をデザインする。一方のインフォマティクスは、これを逆向きに行うのである。つまり、ある物性や反応などを期待して、そうなるような物質の組成や構造などを見つけ出すというのがこのインフォマティクスの主な流れである。

マテリアルズ・インフォマティクスの源流は、マテリアルゲノムイニシアティブ（MGI）という、2010年オバマ大統領時代に遡ることができる。ここでは、材料・触媒インフォマティクスということで、大量に第一原理計算を行い様々なデータを貯めていくチームと、触媒に特に注力したチームでスタートした。基本的には、計算を大量に並列にやって、そこにデータ科学を適用するというようなことをやる。

その結果、確かにいろいろなことができるようになった。様々な分子配位に対してエネルギーをそれぞれ計算しておく、そのエネルギーの曲面が見える。そして、その高次元のエネルギー曲面から反応経路のようなものが計算できる。これまで、このようなことができるはずだとは思われていたが、このような規模で実際に行われたのはこれが初めての試みであった。ただし、確かに短期間で複雑な反応経路の作成には成功したが、これを利用した新しい触媒の設計には至っていない。

世界中で似たようなことが沢山行われ、計算結果のデータベースを作るなどといったことも様々なプロジェクトで取り組まれている。しかし、これは高橋准教授の言葉を借りると「架空材料候補の提案」がたくさんある、というだけである。得られたのは大量の架空の材料・触媒の候補であって、結局それらのほとんどが合成できなかった。そこでCRDSでは合成プロセスまで考えたインフォマティクスの必要性を考えて、このワークショップの開催に至ったわけである。繰り返しになるが、計算とその結果による候補がいくらでてきても、結局最後、物を作るところというのを議論しななければならない、というのが彼の主張である。

米国のマテリアルゲノムは、オバマからトランプへの政権交代の後にかかなり厳しく査定された。触媒の第一原理計算をたくさんやっていたグループは、最近では実験データの収集にも取り組んでいるという話である。その後の情勢変化により今後にもまた盛り返してきそうな気配はある。

さて、高橋准教授らは、メタンの酸化カップリング反応によってエチレンを作る反応について取り組んだ。反応はある触媒の上で起こるので、そのよい触媒を見つけるというのが研究のモチベーションである。1800件の文献の実験データから予測する試みで、触媒の組成や、実験条件など11の記述子を使いエチレンの収率を上げるような最適化を行った。これは、触媒組成と実験環境の同時最適化になっている。

最終的には、後述する谷池准教授の協力のもと、独自開発のハイスループット実験を行い、予測との差を比較している。過去30年分の文献から得られた1,800のデータを用いた予測は、交差検証10%となり回帰モデルがほとんど作れないという結果になった。一方で、ハイスループット実験装置では1日に約12,000点

のデータが取れ、高速生成された均質のデータを用いると回帰モデルが作れ、交差検証も90%と高い予測性能を示した。モデルはランダムフォレストであり、細かく見ていくと、組成の他にプロセス条件が詳細に考慮されていることも分かる。

詳しく見ると、転換率と選択率の関係が負の相関であることや、メタン酸素比と温度と収率の3次元曲面に、ハイスループット実験でない従来型実験が乗ることも分かった。実験手法を変えても、この予測モデルは有効であるということの意味している。

結論としては、実験と情報の研究者が断固たる覚悟をして、今まで自分たちがやっていたやり方を捨てて、新しい方法に移行できるかどうか重要だと指摘している。そして、第一原理計算を中心とする計算データだけを基にマテリアルズ・インフォマティクスをやっていると米国の二の舞になると警鐘を鳴らしている。合成できるようなインフォマティクスするべきである。

次に、谷池俊明准教授（北陸先端科学技術大学院大学）による、高分子の安定化実験の取り組みを紹介する。

基本的にポリエチレンやポリオレフィンなどのプラスチックは、そのままではあまり実用材料にならない。ある程度安定性を付与するために添加剤というのを入れている。添加剤を入れると、あるところまで安定性が上がっていくが、入れ過ぎると効能は飽和し性能低下を招く。従って、安定化剤の「量」を増やさずに耐久性を向上させることが重要である。そのためには、複数の安定化剤を配合し、相乗的な組み合わせを探索する必要があるが、多くの種類の安定化剤があり、最適配合の一般則はない。現在22種類程が市場にあるらしいが、この中から入手しやすい10種類程を選んで配合するとしても、400万種類も可能性がある。これを一つ一つ個別に実験で確かめていくのは大変なので、ハイスループット実験装置と遺伝的アルゴリズムで最適解を探索しようとした。

2.12  
プロセスインフォ  
マティクスWSからの  
内容紹介

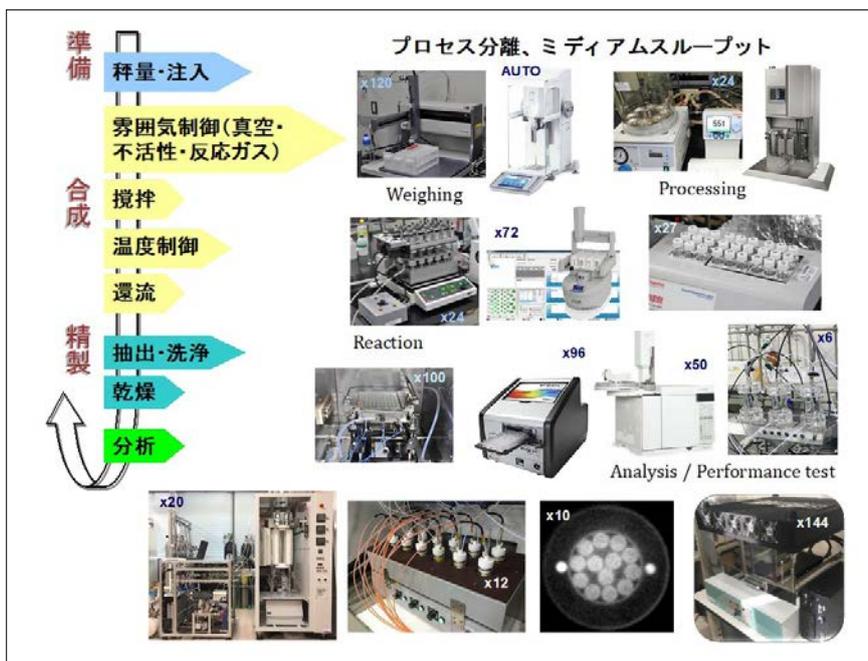


図2-12-1 準備・合成・精製プロセス

この研究は、数あるハイスループット実験の中で難しい部類である。というのも、ハイスループットの合成手法は様々なものがあるが、いずれも物性値の評価が難しい。唯一の例外は、ここで紹介するように色で見るという方法である。谷池准教授らは、プラスチックの劣化の評価を色で行えるように実験方法を整えた。普通は、力学的評価をするわけだが、それだと1つのサンプルについて何時間もかけて、それが割れるかどうかを評価しなければならず、とても時間がかかる。谷池准教授の提案方法は、酸素などを導入して温度を上げ、反応する際の化学発光で評価するものである。化学発光による酸化の検出は機械特性の低下のおよそ20倍の速さで評価でき、同時に100サンプルについての測定も可能となる。これにより、評価のスループットは2,000倍になるわけである。

安定化剤の配合の探索問題では、22種類の中から10種類を選んでも、組み合わせの数は膨大になる。まずランダムに対象配合を決めて、それでハイスループットの評価実験を行い、結果を見て遺伝的アルゴリズムで次の世代の組み合わせを作る。そして合成と評価実験を、耐久性が上がる方向に5世代程繰り返す。結果として得られた、150℃、200時間の耐久性というのは、プラスチックとしてはかなり良い性能である。配合の最適化に要した総実験時間としては、従来法では5.5年に相当するものを、この手法なら50日で測定が完了することになる。

高性能な配合の中身を後から確認すると、繰り返し使われる特定の安定化剤があることが分かる。これらは後から考えると、相乗的な組み合わせであると言えるのだが、探索の最初の段階ではこのような知識は必要ない。この実験では、乱数から始めても十分な性能の安定化剤の組み合わせを見つけることが可能で、かつ、組み合わせの相乗効果について配合の設計指針も導かれた。

高橋准教授との協働での固体触媒の開発については、先述のとおりである。最初に、文献データを使ってランダムフォレストを行うと予測性能が出ない。データ点数が少ないという問題もあるが、そもそもデータが特定の触媒に集中しており、分散が小さいという問題がある。そして、プロセス依存性があるのだが、論文に

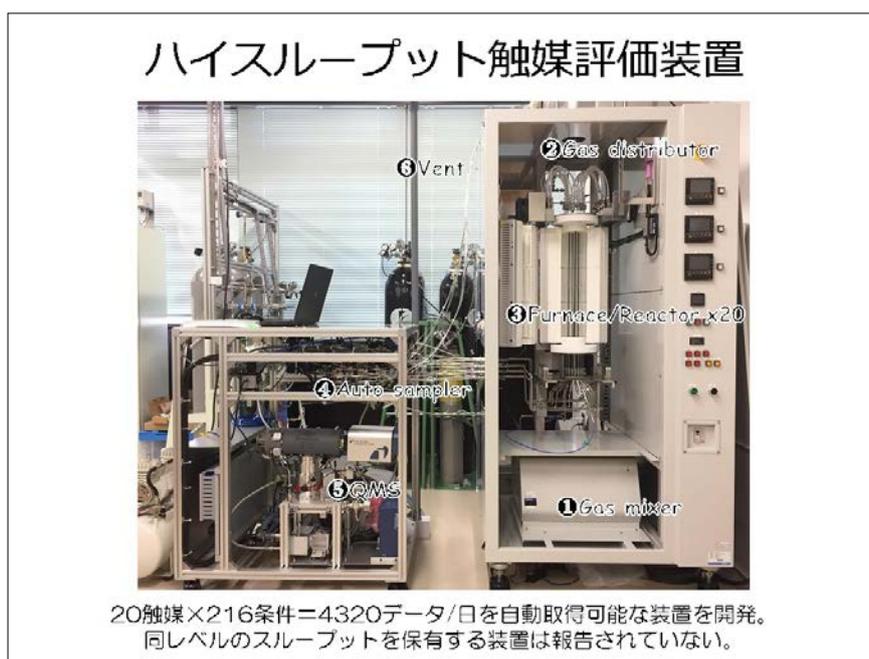


図2-12-2 ハイスループット触媒評価装置

は必ずしも書かれていない。また、おそらくこれが一番の問題だが、悪い結果となる材料についてのデータが無い。性能が出ない、悪い材料については論文に書く必要がなく、そのようなデータはそもそも文献から抽出することができない。一方で、研究者というのは失敗や悪い材料からも学習をしている。やはり実験を自分で行っていないことは、大きな問題だと言える。

谷池研究室にあるハイスループット触媒評価装置（図2-12-2）では20種類の触媒を、216条件で計4,320データを1日で自動取得可能である。これを使って、触媒を変えたり、条件を変えたりしながら毎日実験を行い、データをためていくわけである。ハイスループット実験によるものと、別のチャンネルで行ったものを比較し、最大誤差5%程度と、高い再現性も確認してある。

実際の作り方として、含浸や乾燥、サンプル管に詰めるなどのプロセスまで並列化の工夫を施し、複数の温度条件などで同時に測定するというようなことまですると、3日間の運用で59触媒、約12,000点のデータが取得可能となる。これは、触媒の良し悪しを人間が判別しない状態で、同じプロセスで広範な条件で齊一的に取られたデータである。これを散布図としてプロットしただけで、最高収率21%が得られる。現在知られている最大の収率は25%程とのことだが、それに匹敵するものを3日程度の実験で到達したことになる。散布図には、転換率と選択性のトレードオフも明確に示されている。ランダムフォレストによる回帰性能も $R^2=0.89$ と良い。一方で、意図的に収率が低いデータを除いて訓練したモデルの回帰性能は極めて低く、文献での報告が少ないネガティブデータの存在が、正確な回帰には必須であることが示唆される。

まとめると、ハイスループットの評価データは、とくに同一プロセスで齊一的に取得したビッグデータは、機械学習に非常に役に立つことが分かった。データセットの欠如は、マテリアルズ・インフォマティクス研究の致命的な問題である。谷池准教授の言では「ハイスループット実験は、マテリアルズ・インフォマティクスの果実を最大限に活かすための手段」である。短期間に均一条件で得られるデータセットを与えるハイスループット実験と、データ科学によって、従来探索できなかった広大な空間を対象とした研究が、短期間に実践できるようになる。

### 【質疑応答】

Q：過去の文献データをトレーニングデータとして使うのはダメだというのは、サイエンスにおけるシリアスな問題と思う。

A：どれほど論文をちゃんと見たか、というのが重要だと考える。徐先生の紹介例と、今回の話はその部分が違っていると思う。NIMSの例では、論文を読むためだけに専門家を何十人か雇って、何年間もかけてためた、ある意味厳選したデータである。単にテキストマイニングのようなレベルで持ってくるなどだと、上手く行かないのだろう。詳細は分からないが、ネガティブデータが無いという問題は、論文を使う以上は避けられない。

Q：12,000点の触媒の実験というのは、全空間を探索したわけではなく、その中をランダムにサンプルして実験をしたということか？

A：実験自体は全部行っている。59触媒の12,708点のデータを取得なので、これは全数を測定している。触媒の種類は59種類だが、それぞれに対して200種類ほど条件が振られている。プロセス条件として、温度やメタンと酸素の混合比などである。製造法の他に、実際に触媒として働かせるときの条件もある。

Q：ということは、この59種類の外にある触媒については、よく分からない、ということか？

A：基本的にはそうである。ただし、発見したと言っていたのは、その外側にある触媒かもしれない。ある組成とある組成の間のあたりにある触媒を新種触媒と呼んでいる可能性もある。

- Q：ランダムフォレストで合成条件を探す話で、専門家の人が、こういった物性探索をするときの経験則に比べて、新しい知見や思いもよらなかった方法論のようなものが、機械学習から得られたのか？
- A：少なくとも、テストデータに対して変な値を返さない、というくらいには正しく予測するモデルができています。これは、ある意味で、当たり前のことが正しく分類できているに過ぎない。従って、新しい知見があったかというよりも、何か実験する前には、まずこの予測モデルにかけてみるような使い方に、価値を見出すべきものだと思う。このモデルは、プロセスの条件まで同時に答えが出てくる。また、この場合ではそれぞれの記述子に対して条件で区分けしていくのだが、多次元空間であり、人間に一言で理解できるような形にはなっていない。もちろん、ディープラーニングが与えるよりもはるかに人間に理解しやすい形になっている。ランダムフォレストの分類は、得てして当たり前の分類をする。それはそうあるべきだろう。多くの場合、記述子を選んできている段階で、かなり物理が分かっている状況だろう。ただ、ごくまれに、理由が分からないが正しい答えを出すようなときがあり、これは人間にとってヒントになるだろう。
- Q：ハイスループット実験では、やはり評価がボトルネックなのか？
- A：そうである。谷池准教授のプラスチックの例は、本当に例外的である。光で評価するというのはすごく上手い。あの実験では、むしろプラスチックを合成するほうが、時間がかかっている。触媒の方は、評価装置を作るところまで力技のような印象である。一般的には、一個一個丁寧に測るのと比べて、ハイスループット実験では何かを諦めなければならないはずである。何を諦めるかの選定は、谷池准教授は「勘」だと言っている。触媒の例では、測定の絶対値の精度などを諦めているのではないか。多少の条件がずれた程度で探索から漏れてしまうようなピーキーなものは、そもそも必要ない、という割り切りもあっただろう。
- Q：機械強度のような本当に知りたい物性値と、ハイスループットで非接触で評価できる光学特性などの間には、一定程度相関があるというのは、原理的にも理解できるし、実験で確かめられてもいるのか？
- A：そうである。そのようなものを見つけるのが重要。第1回目のプロセスインフォマティクスのワークショップでは、シミュレーションを沢山やって、シミュレーターを良いものにしていけば、結晶成長などではうまくいくだろう、という結論になりかけていた。しかし、今日紹介した2名の話で、プロセスそのものをインフォマティクスの対象とするには、実験が大事だ、という方向になってきている。
- Q：基本的には、目的とするようなものが内挿になるように、実験データを取りに行くということか？ 最初20種類を決めてそのプロセスについての最適なものが見つかるが、では次になにをやったらいいかという実験計画を決めるのは、少し先の話か？
- A：この材料系ではそうだろう。徐先生の発表であった相分離や相境界があるような材料系では、そこを跨って、一緒にデータをとってはいけない。ここでの話は、組成を変えたときに違う相が出てくるような話ではない。ここは、1つの相の中でいろいろな組成を比べているということだと理解している。
- Q：材料の世界では、ネガティブな結果を論文として発表するというようなプラクティスは無い、ということか？ であれば、JSTのような機関が、少なくともJSTのグラントで行われた仕事に関しては、ネガティブな結果も発表しなさいというようなガイドをしてもよいと思うが、どうか？
- A：そうだろう。少なくとも、現状の材料分野に関しては、ネガティブな結果を発表するような文化はない。
- C：実験する前の、実験計画を作った段階で論文投稿し、ピアレビューをしてもらい、それを通れば、実験の結果に関わらずパブリッシュされるという取り組みがあるようだ。そういうことを考えているジャーナルが20本以上あるとNatureの記事に書かれている。こういうことを、JSTや文科省などが主導的に、

ネガティブな結果もポジティブな結果も、両方パブリッシュすれば評価しますよ、といったメッセージを出してもよいと思う。

- C : 米国NSFは、分野によってはあるが、データを所定のリポジトリに保管しないとグラントを受けさせない、などといったことを初めから言っている。JSTはそこまでドラスティックなことにはなっていないが、CREST・さきがけでは、データマネジメントプランの提出が義務付けられている。共有・公開するとメリットがあるという、インセンティブを考えないといけない。いまの論文誌、とくにNature、Science、Cellなどは商業誌で、読者はネガティブデータを求めてないので、インセンティブはない。
- C : NIMSのデータベースへの登録など、論文とは異なる形での取り組みも可能かもしれない。また、分野は違うが、ゲノム情報や天文台の写真データは、公開と共有が進んでいる。特に、天文分野では、アフリカなどでは、そういう公開データに依存して研究を進めているという話も聞く。そういった取り組みを、この材料分野にも利用していくというイメージかと思った。
- A : 失敗データというのは、同じようなことをやりたい人にしか役に立たず、それを共有するモチベーションは低いということだろう。材料屋は、得てして「こんないい材料作ったぜ」と自慢したいものである。失敗したことを言うモチベーションがない。しかし、これからは、何らかの仕組みを利用して、全データがたまっていき、しかもそれにインデックスがついているようなものを皆が収める、というのが必要だろう。
- Q : やはり、どちらかという、必要なデータは、自分で測定するほうが早い？
- A : 今回紹介した2人の研究は、まさにそうしたのである。論文データには無いから、全ての属性が分かっているデータセットを作ってしまったわけだ。しかし、誰もがこのようなことをできるわけではない。また、材料に関しては、論文の時点では測られていなかったことが、後から重要であった、というようなデータというのが結構ある。例えば、NIMSの超伝導体のデータベースでも、同じ金属組成の高温超伝導体でも転移温度がゼロから120ケルビンくらいまでばらついているようなものもある。これは酸化物の酸素組成が微妙に違っているためにそうなっているのだが、精度良く測ることは難しい実験なので、全ての論文で行われているわけではない。酸素組成がばらついているデータから回帰するのも問題だが、酸素組成が書いてある論文だけに限ると、今度はデータ数がものすごく少なくなってしまい、それも問題である。
- Q : 今日紹介された触媒のケースでは、そこはどうなっているのか？ 意図した材料が、意図した組成で作られているのか？ 意図したプロセス条件は、意図したとおりに効いているのかは、どうチェックするのか？ そこは飛ばして、エンドツーエンドで、結果からプロセスパラメーター予測しているような印象もあるが。
- A : 同じ作り方で作っているひとかたまりのデータなので、その意味ではコントロールされている。全体が同じようにシフトしていることはあるだろうが、振り返れば全部同じにずれているので、というのが重要なだろう。
- C : 自動化実験装置を作るノウハウのようなものが、結構サイエンスの足腰になっている印象がある。そこを強化する施策が打てるとよいと思う。
- A : こういう装置を作って売る、大学発ベンチャーはいくつかある。今日紹介した装置のコピーも、買いたい人はかなりいると思う。100台は売れないかもしれないので、大きな会社はやらない。あるいは、製薬会社などでは、すでに自前で持っている可能性もある。ケモインフォマティクスや、コンビナトリアル合成などは、そもそも創薬が起源。

## 2.12

### プロセスインフォ マティクスWSからの 内容紹介

- C : 自動化システムのベンチャーがあると良いと思う。また、パラメーターの組み合わせの探索や、実験計画、最適化のところでは、まだ最新のコンピューター・サイエンスの知見が入っていないような気がする。そこを繋げてあげると良いと感じる。
- Q : 合成はそれだけでも結構なデータ量があるように思う。どの材料が良いかという組成の話に加えて、合成は時系列である。どこで混ぜてとか、何を入れて、乾かして、など。そういったものを一緒にインフォマティクスとして扱えるとよさそうだが、次元も大きくなり、難しいところも出てくるだろう。CRDSのプロセスインフォマティクスの調査チームでは、どのあたりをフォーカスしているのか？
- A : 材料組成だけでも、パラメーターの組み合わせは膨大なのに、そこに作り方の軸を加えたらとてつもない数になるというのは、そのとおりだろう。なので、チームの調査としては、できるところとできないところの見極め、どのくらいの出来なさを加減なのか、などを見ている。少なくとも、幾つかの例については、組成と合成の両方を見ることはできる、ということが分かってきている。
- Q : アカデミアと産業界の温度差もあるだろう。自分のデータは出したくないが、他社のデータは欲しい、というような。アカデミアであれば、天文のように皆で共有、というのもありえるかもしれないが。
- A : それはワークショップを2回実施して、痛感している。予想はしていたが、「プロセス」という言葉ひとつとっても、アカデミアと産業界ではまるで違っている。また、今回は対象としていないが、鉄鋼業のような構造材料をつくるプロセスと、創薬で抗がん剤を作るプロセスは、全く違う話であろう。企業的には、新物質そのものよりも、プロセスを隠しておきたいということもある。化合物は売りに出されてしまえば、組成分析できる。しかし、どう作ったかは、相手の工場に踏み込まない限り分からないという部分がある。電気業界はどこの企業も競合他社の部品買ってきてリバース・エンジニアリングする専門の部隊を持っている。
- C : 金属の会社では、レアメタルのリファイナリーのプロセスでは、パラメーターのチューニングを何十年もやっているという人がいるという話を聞く。ここに、最適化問題や機械学習のテクニックを使うとどうなるか、という勘が聞く人がはいっていかないといけないだろう。三菱ケミカルや旭化成など、プロセスインダストリーで先端的な会社は、まさにそれを見始めているような気がする。
- A : 材料系には、その方面に弱い人が多い。これからますます必要になるだろう。

## 2.13 モデル駆動型アプローチからみた逆問題の諸相

山本 昌宏 (東京大学)

専門とする数学の観点から見て、何か問題を解くときには、やはりモデル駆動型のアプローチ寄りになることが多い。今日の発表は、まず数理モデルについて話した後、いわゆる数学でいう証明とはどんなものなのかということを考えてい。一種の数学者の心理というか、好みという部分にも触れる。また、一般的な話としての数理モデルと逆問題の関係、最後に個々の事例を紹介する。

有名な「ベンフォードの法則」というのがある。電話帳でも新聞でも何でも構わないが、数字の並びが出てくるようなものを取ってきて、そこに並んでいる多くの数の最初の数字を抜き出し、それらがどれ程の頻度で現れてくるかを調べる。すると、順番に1が出てくるのが大体30%、2が18%、3、4、5、6、7、8、9と頻度が続く。あくまでも、ある数字の並びの最初の桁である。ちゃんと1、2、3、4、5、6、7、8、9の順番に少なくなっていく。これは一種の経験則だが、様々な解析をすると、対数で表される式に結構よく当てはまることが分かる。これがベンフォードの法則である。

このような現象は、色々なところに見いだされる。興味深い応用に、米国での裁判の証拠としてベンフォードの法則が使われたというのがある。数値の不正操作に起因する裁判だが、データが自然なものであれば、最初の桁に9がたくさん出てくるとかいったことはあり得ない。従って、数の分布を見ればどこかで人為的な操作が加わっていないかが分かるわけである。この例は、一種の証拠として裁判に採用されたものである。

### ベンフォードの法則の1つの数学的な取扱い法

- データ駆動型サイエンス
- モデル駆動型アプローチ → 数学

「現象に伴って現れる数の分布は何らかの操作の結果である」と考える → モデル式の策定

**モデル式:** 偏微分方程式  
代数的操作(置換など)、図形的な変換  
試行(サイコロを振る, etc)

6

図2-13-1 ベンフォードの法則の数学的な取扱い法

さて、ベンフォードの法則はちゃんと説明、数学でいう証明ができるのかということが問題となる。もちろん統計的な有意性などを考えるのは1つの手であろう。しかし、ここでは、これを1つの例として数学的な証明、説明をどうするかを考える。

一見、数学におさまらないようなものをどう数学で取り扱うかという、常套手段としてどこかに構造や操作を導入するというのが考えられる。フィクションではないかと言われると、確かにそうなのだが、何らかの操作の結果であるというふうに想定するのである。このような手続きは、数学ではよくある。ある種の現象や操作の結果、今、目にしているデータが出ていと想定するわけである。これは、モデル式を仮定しているということになる。

モデル式といってもたくさん種類がある(図2-13-1)。数理物理で使われるのは偏微分方程式だが、確率的なものや、幾何的な変換もある。また、5次方程式の解の公式があるのかないのかというところから出てくるような、数を置き換える代数的操作もある。そのような操作やモデル式を想定し、その結果が今、目にしているデータだと考えるわけである。

ベンフォードの法則の場合には、非常に多くのデータセットからランダムに複数のサンプルを取ってきて、それを並べて、それぞれの数のアンサンプルとして見た結果、1が30%、2が18%、と続く特殊な分布になっているのだと想定するのである。多数のデータセットがあって、そこから一種の操作を施した結果を見ているというふうに考える。

そうすると、これは普通の確率論のロジックにおさまる。例えば、Theodore Hillが示したように、データセットとサンプル数が増加するほど、サンプル全体における一番上の数字の分布は、ベンフォードの法則に近づいていくということが定式化でき、この範囲ではきちっと証明ができる。このように、データを解釈・説明するときに、そのまま見るのではなく、どこかに操作や構造を入れて、その結果だとするのが数学では常套手段である。

**演繹的推論を可能にする数学の機能：  
武器としての数学独自の特徴**

**数学者の大多数が証明として認める手段**

- **背理法、不可能性の証明：**  
素数は無限個ある、無理数であることの証明、  
5次の代数方程式に解の公式はない、  
逆問題が安定になることは「不可能」である (II 節)
- **数学的帰納法**
- **一見自明に見える原理を基に理論を構築**  
ユークリッドの幾何学原論(定義—公理—証明—定理)
- ジョルダン(Jordan)の閉曲線定理
- 中間値の定理
- 鳩ノ巣原理

9

図2-13-2 演繹的推論を可能にする数学の機能

2.13  
モデル駆動型アプローチ  
から見た逆問題の諸相

次に、図2-13-2のように、数学における証明とは何かについて考えたい。大多数の数学者が証明の手段として認めるものがいくつかある。典型的なのは背理法である。これはある結論を証明するときに、その結論が成り立たないとして矛盾を導き出すものである。例えば、ユークリッドが素数は無限個あるということを背理法で証明した。背理法による証明は、技一本というような感じで鮮やかに決まる場合が多い。必ずしも背理法だけを使うわけではないが、不可能である、存在しないということの証明もできる。5次以上の代数方程式に解の公式はないというアーベル-ガロアの定理など。

数学的帰納法も、有効な証明方法である。また、ややレベルが違うものとして、構造や写像、変換の観点を入れ、そこから問題を捉え直して証明するというも行われる。さらに、全体の流れとしては公理のような自明と思われるような原理を基に、演繹によって推論を進めて定理を証明していくということもある。原理は随所で使われる。これは公理ではなく、普通にふと意識することなく使うような数学の定理である。それだけ見ると、当たり前という感じがするが、例えばジョルダンの閉曲線定理や中間値の定理、鳩ノ巣原理などがある。

背理法や数学的帰納法というのは数学者の大多数が証明として認める手段ではあるが、全ての人が無条件にこれを使うことをよしとするかという、必ずしもそうではない。ここのところを少し議論したい(図2-13-3)。

数学における証明がどのように理解されているか、ということは、粗く述べると2000年ぐらい前から考え方は変わっていないと言える。歴史的にはギリシャ数学的に公理、論理、定理と順番を追って証明していくアプローチと、エジプトやバビロニアなど、アルゴリズムとも言うべき数学のスタイルがある。

これらは、超越数学とアルゴリズム的数学と表現することもできる。超越数学というのは、一般の場合にこういうことが可能である、解が存在する、ということを示すものである。したがって、解をどうやって構成するかについては、あまりはっきりしない場合が多い。それに対して、アルゴリズム的数学は、具体的な問題が与えられたときに、こういう手順で解きなさいというような性格のものである。

2.13

モデル駆動型アプローチからみた逆問題の諸相

II. 数学で言う証明とは何か(断想風に)

「証明とは何か」で数学者の考え方には多少の違い: 数学内の分断

- ギリシャ数学 VS エジプトなどの数学 超越的数学 VS アルゴリズム的数学
- 形式主義 VS 直観主義
- ✓ 公理主義=証明なしでいくつかの公理を認め、無限回を含む推論で理論を構築していく
- ✓ 直観主義=背理法は認めない: 数学上の命題は正しいかそうでないかということは認めない。有限回の操作で証明できない命題はいろいろあるから
- ワイエルストラス VS クロネッカー
- ✓ クロネッカー: 「神は整数を創造したもうた、それ以外の数は人間が作ったものである」「無理数なんかない(数学上のたわごとだ)」カントールの集合論を全否定
- ✓ ワイエルストラス: 無理数が重要、さもないと微積分が成り立たない
- プラトン主義 VS 構成主義 数学理論は発見するのか vs 発明するのか? プラトンのアイデアに少しでも近づく vs 公理に基づいて理論を構築する
- 「天書の証明」: 難しい定理の素晴らしい証明を書いた天の本が存在する。数学者をそのページでもよいから、見つけるのが使命である(P. エルデシュ)

図2-13-3 数学で言う証明とは何か

歴史的な話として、形式主義と直観主義の対立がある。形式主義というのは、ユークリッドの幾何学のようなもので、証明なしでいくつかの公理を認め、無限回を含む推論で理論を構築していく立場である。これは、普通の数学のイメージだろうと思われる。一方で、直観主義というのは、有限回の操作しか認めない。そのことにより、色々なことが生じる。まず、背理法が認められない。それはなぜかという、全ての命題は正しいか、偽であるかの2つしかないという排中律を、全ての命題について有限回の操作で証明するのはほぼ不可能だからである。白か黒の2つの可能性しかないことを認めないので、それに依存している背理法というのはいえないことになる。

対立軸として、ベルリンの数学者カール・ワイエルストラスとレオポルト・クロネッカーの対立も有名である。クロネッカーは「神は整数を創造したもうた。それ以外は全部、人間が勝手につくったものである」という立場をあらわにしている。また「無理数は数学上のたわ言である」とも。したがって、カントールの集合論などはもう全否定である。あれは若者を惑わす有害なものであるとまで言ったらしい。

同僚のカール・ワイエルストラスは、正反対の立場である。彼は、そもそも微積分の基礎づくりをきっちりとした人で、数直線がつながっていないと話にならないのである。無理数を否定されてしまうと話が進まなくなるので、これはクロネッカーとはかなり違う立場である。

19世紀のクロネッカーの立場は、よく考えると古典ギリシャ数学の立場と似ている。ピタゴラスの定理で直角二等辺三角形を考えればルート2というのがすぐ出てくるわけである。昔のギリシャ数学は、無理数というのは基本的には認識し得ないものであるとする立場なので、これは困ったことである。

さらに、プラトン主義と構成主義という対立もある。プラトン主義というのはアイデアという絶対的真理がどこかにあって、それをみんなで探求していくのが哲学、科学の目標であると考えたものである。一方の構成主義はそうではなく、真理は人間がつくり出していくものであるとする。こういうフィルターで数学者を分類すると、多数がプラトン主義になるだろう。この軸は、別の言葉では「数学は発見するのか発明するのか」とも言える。

あるいは、有名な数学者であるポール・エルデシュは「天の書という神様が書いたすばらしい本があって、そこに数学の全ての定理のすばらしい証明が載っている。そのページのかけらでも見いだす努力をするのが数学者だ」などと言っている。これはまさしく、プラトンのアイデアを追い求めるという立場である。

超越数学とアルゴリズム的数学というのは、実際に数学を教えるときのカリキュラムにも影響する。単純な例では、面積とは何か、というのがある。正方形、三角形を出してくると、大抵の子供はすぐに計算できる。しかし、それは面積を定義したことになる。ただ単に、ある特殊な場合について面積なるものが求まるというだけで、複雑な図形では公式などないのだから。では積分で求めるとなると、今度はどういう場合に値がきちっと決まるのかという、なかなか悩ましい問題になる。突き詰めていくと、面積はいつも確定するとは限らない。選択公理といわれる約束を認めると、面積が確定しない集合が存在することが証明できてしまう。これは超越数学の困難である。

公式というのは単純過ぎるが、解法を提示して解決していくやり方は、アルゴリズム的数学といえる。従って、数学は2つの立場を使い分けているのである。証明は証明でも、その性格はどちらに寄っているかということで、見え方はかなり違って来る。

まとめると、大多数の数学者はギリシャ数学的形式主義、無限は認めて、絶対の真理を求めるという方向に立っていると見える。すると、いわゆる形式主義、公理に基づいて証明していく立場では、コンピューターを使うような証明は認め難いという心理が生まれるであろう。それと対応する直観主義では、アルゴリズムが確定して有限回の操作で証明するわけで、むしろそうならないといけないと考える。これはコンピューター・サイエンスの立場とも言えるだろう。クロネッカーは、数学の進化の立場からはとんでもないことを考えたと

もいえるが、じつは彼は、アルゴリズム的な数学の発想をしていたといえる。

一方で、直観主義だけだと背理法が使えず、これはとても不便である。集合論に関する公理も使えないとなると、いろいろと不都合なことが出てきてしまう。そこは、形式主義的にやりたいというのがある。従って、繰り返しとなるが、大多数の数学者は両者の立場を折衷している。心情的には、絶対的な真理があるという考え方に共感を持ちやすいのではあるが。

応用数学はアルゴリズム的な部分の多い数学なので、数学コミュニティの中で「あれは数学ではない」「ダーティー数学」などと、いわれない偏見もあったが、その背景にはこのような対立構造があるようである。

逆問題について考える。先程の写像の立場で書くと、ある求めたいパラメーターや原因  $x$  があり、それに变换を施した  $y$  というデータが出てくるものと一般的には考えることができる (図2-13-4)。

**III. 逆問題の原理と数理モデル**

- $y = Kx$
- $y$ : データ、 $x$ : 未知量(逆問題の解)
- $K$ : 未知量から観測データを対応させる写像

**逆問題 :**

- $y$  を用いて  $x$  を求める
- $y$  はもとの未知量  $x$  の情報を正確に反映していない : 情報が失われている !

14

図2-13-4 逆問題の原理と数理モデル

$y$  のみが観測できる量である。  $y$  が分かっているときに  $x$  を求めるのが逆問題である。普通の逆問題では、この  $y$  は情報落ちした量である。  $x$  はいろいろな情報を持っているのだけれど、変換を施すことによってある方向の成分が消え、それが  $y$  として観測されるのである。

これは数学的にはコンパクト性の問題である。要するに、  $x$  から  $y$  へ移るときに情報落ちをして、ある種の成分が消えたり極端に小さくなったりするのである。そういったものはきちんと定義でき、数学用語でいうと点列コンパクトという集合の中に入ってしまった。例えば、  $x$  は10次元ぐらいの情報があるけれど、  $y$  には1次元の情報しかないとかいったようなものである。データセットが無数個の違った方向を持つようなものであれば、これは点列コンパクトという概念を満たさないのだが、多くの逆問題では、データには有限個の方向しかなくなっている。

次元が落ちているという理解しやすいが、正確にはコンパクト性というべきで、これは次元とは異なる概

2.13  
モデル駆動型アプローチ  
からみた逆問題の諸相

念である。次元は代数構造であり、コンパクト性は位相構造である。これらは数学ではきちんと使い分ける。いずれにせよ、こういった構造を上手く考え、これを元にして考察するという流れである。

結局、逆問題は何をするかというと、情報落ちたデータから欠落情報を補って元のものを求めるということになる。これは補完ではなく、一種の補外をしていることになる。そうすると、恣意的な補外を防ぐためには、一定のモデルがあると安心できる。これが、言わばモデル駆動型のアプローチということになる。

モデル駆動型という言葉が的確かどうか分からないが、モデルを意識して解析していくということである。補外操作は、やり方によってはいろんなことができてしまうので、一定のガイドラインとしてモデルを設定するのである。

事例を幾つか紹介する。最初の例は、新日鉄（当時）との共同研究である。高炉（溶鋳炉とも呼ばれる）という、鉄鉱石から銑鉄を還元反応で取り出す非常に大きな設備が製鉄所にある。製鉄所の要のようなもので、熱効率を考えると基本的には10年位動かし続ける。生産量は、一番大きい君津にある高炉だと12,000トンほどである。

内部は高温で、何かトラブルがあると非常に危険である。その高炉の内部温度を常時モニタリングしているが、時々、測定温度が正常値を上回ることがある。そうすると、原因が分からなくても操業のペースを数日は抑える必要が出てくる。これは、かなりの経済的なダメージである。従って、これを事前に察知し、熱が出る前に薬を飲んで熱が出ないようにするという、そんなようなことをする必要がある。

高炉には熱電対という温度計が仕込まれており、常時、温度を計測している。これは高炉のサイズに比べそれほど多くない。つまり、データがそんなにたくさんはない。また、具体的には、大体300度ぐらいで推移しているのが、突然500度近くになるのだが、これは不安定なりに落ち着いた「安定な」状態で、すぐには

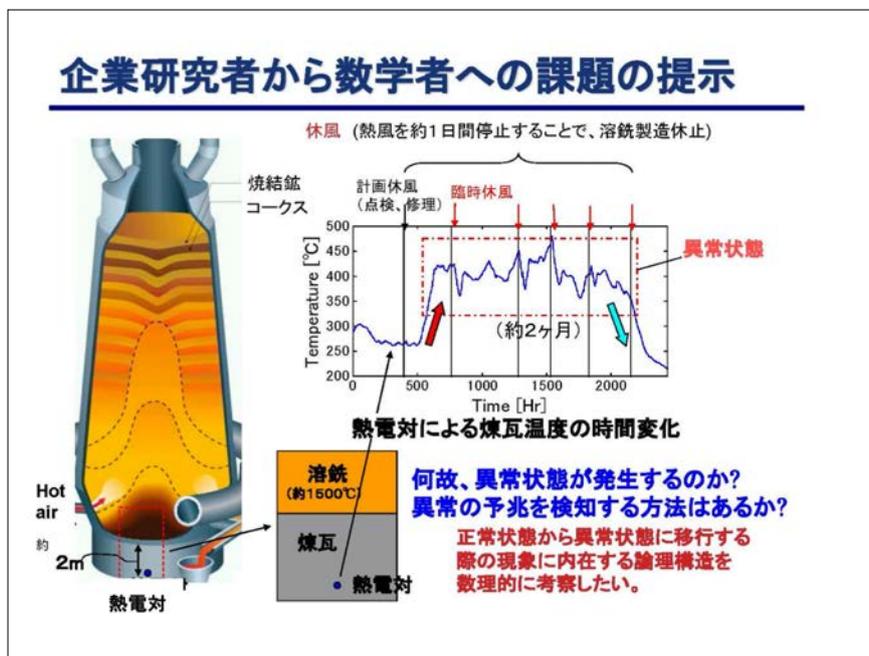


図 2-13-5 高炉内の熱流れを推定する逆問題

2.13  
モデル駆動型アプローチ  
からみた逆問題の諸相

解消せず、だらだらと続く。温度が上がっては操業のペースをダウンし、上がっては操業ダウンを繰り返す。このようなペースダウンを断続的にかなり長期的に行う必要があり、当然、生産量が減るので結果的にはかなりの経済的な損害となる。立ち上がりを検知して、立ち上がる少し手前のところでダウンすれば、こんなにこじれないのではないか、というのは自然な予想であろう。

そこで、高炉内の異常な熱流れに着目した。しかし、当然ながら操業中の高炉の中の熱流れというのは直接見ることができない。熱電対というかなりプアなデータから、熱流れを推定するということになる。これは、先述した逆問題である。熱電対なので、温度はポイントでの測定で、20個ほど設置したとしても、高炉の大きさに対しては大したデータの量にはならない。従って、どこかで補外する必要がある。推定手法の詳細はここでは述べないが、推定結果を見て異常状態を検知し、その手前で操業ペースを抑え、吹き込む風を減らすなどの操作をすると、この異常状態になることを防ぎつつ、安定的に操業を繰り返すことが可能になった。

次に、福島原発事故の後の農地除染での、逆問題の応用を紹介する。原発からセシウム137が含まれた粉塵が出て、地上に落ちて、雨水とともに地面に浸透し、農地を汚染する。深さ20センチ位までに浸透したセシウム137の濃度を決めるという問題である。この20センチというのは、ニンジンや大根など、農作物の根の深さを考えたものである。

データを取るのは大変である。羽田野祐子教授（筑波大学）との共同研究を行ったが、まず、ある一定の土地を仕切り横方向のセシウムの流入・流出をカットする。その後、人の手で掘っていくわけだが、1センチか2センチ掘っては、そこを計測するというのを繰り返す。従って、直接的なデータは、それほど多くは取れない。対象とする領域は、福島県全域などというような大きさと、農地の他に、山林や牧草地などそもそもデータが取れることが期待できないような土地も含まれる。従って、このセシウムの濃度を推定する問題も、典型的な逆問題と言える。

このために、モデル式を導入する。基本的には古典的な移流拡散方程式で、これは古くから知られていたものである。地面での境界条件が大事だが、ここは従来中に浸透していくセシウムの量が釣り合っているというような、ある種の保存則が境界条件として想定されていた。地表面でのセシウムの出入りは、地下全体のセシウムの総量が変わらないようなものになっているという条件である。これは、かなり古典的なモデルである。

ところが、実際には、雨が降ってセシウムが外に出てしまったり、あるいは、周りの土壌から雨水で流れこんだり、ということが起こる。つまり、保存則は成立していないのである。そこで、 $\beta$ という定数を入れて、セシウムの総量の増減も含むように既存のモデルを修正した。その修正モデルで予測を行うと、劇的に精度が上がった。

新しいモデルは特に、浅いところで実測値とよく一致していることが分かった。旧来モデルでは、地表面からの出入りはないものとしているので、浅いところは実際より大きな値が出てしまう。実際には、地表面から流出してしまっているのだから、その流出した分まで考慮する我々のモデルは、実測値とよく合うのである。

もちろんパラメーターを1個余分に導入しているのだから、それを決めなければいけない。この時、ある種、モデルの自由度と予測の精度の兼ね合いがあり、パラメーターを増やすと精度は理論的には良くなるかもしれないが、モデルの使い勝手は悪くなる。従って、なるべく少ない数のパラメーターで精度を上げるためにはどうすればよいか、という見極めが重要である。

また、このモデルを使うと、空間線量率の推定も可能となる。例えば高さ1.5メートル程のセシウム濃度は、地表面と特に浅い土壌の中に残留しているセシウムでおおよそは決まる。内部被曝による健康被害の危険性があるので、この1.5メートル程の高さの空間線量率を推定するのは重要である。土壌中のセシウム濃度が決まれば計算は簡単にできるので、浅い土壌中のセシウム濃度を我々のモデルで計算し、それを基に空間線量

2.13

モデル駆動型アプローチからみた逆問題の諸相

率を計算する。航空機による500～600メートルの上空のモニタリング結果と計算結果を比較すると、大ざっぱな傾向は捉えていることが確かめられる。データを取るのが難しいところも、推定できるという事例である。

最後に、おいしいカップヌードルの特徴づけという問題を紹介する。製造機から出てきた乾麺の画像A、B、Cがあったときに、しこしこした食感が際立つものとか、まろやかなものというのを、画像だけみて分類するという問題である。天井にとりつけられたカメラの大量の画像データをもとに製品を分類しないといけない。

統計処理もできるとは思うが、ここでは数学的なやり方で考える。画像の特徴量と食感の対応関係はそのままでは因果関係を記述するモデル式を創出することは難しい。このような官能試験の数理的裏づけは、様々な企業の方から提示されることがある。奥行きのある音を奏でる楽器の設計、内部がゴージャスに見える車の内装など、元々モデルがない問題である。このとき、数学では構造や変換を想定するというのは先述したとおりである。

今の場合、幾何学量を計算する。例えばこの麺の巻き方について、トポロジーでベッチ数と呼ばれる量や周期性を計算するなどして、それによる分類と食感との対応関係を作るのである。

### 【質疑応答】

Q：プラトン主義のところ、「ゲーデルの不完全性定理にも関わらず」という紹介があったが、もう少し詳しく説明いただけないか？

A：歴史的には、ゲーデルは、クロネッカーに続く世代の人である。さて、対立の1つの観点は、全ての命題は証明ができるはずである、ということであった。つまり、正しいか正しくないか、どちらかしかないで、背理法を使ってもいいということになる。普通に公理を設定したとして、その公理の体系の中で正しいか正しくないか、証明できない命題が必ず存在する、というのがゲーデルの不完全性定理の第1番目である。そうすると、形式主義というのは、実は根底から崩れてしまっている。めったにはないけれど、全ての命題がきちっと証明できるわけではない、というようなことを言ってしまうと、形式主義にとって大きな打撃となった。これは決定的なのだが、そこには目をつぶってしまっていたりする。ごく希にしか起こらないから、など、その辺りの心理はよく分からない。私も気にしていない。

Q：この、めったに起こらないことが、シリアスな結果に導いたという例は、これまで発見されていないのか？

A：微妙である。ある種の公理を余分に付け加えないと、回避できない、あるいは論証がたいへん不便になる、という例はある。しかし、今のご指摘は、もっとギリギリの例ということかと思う。何かを付け加えたら上手く行った例はたくさんあるが、付け加えないと大変なことになるという例は、数学基礎論の専門家ではない私には分からない。集合の要素の個数（濃度）についての連続体仮説をめぐる問題もあるが、これが真か偽であるかは、粗くいうと証明できないのであるが、それによって通常の数学の研究に重大な影響が出るとは考えにくい（今まではなかった）。パラドックスのようなものはたくさんあるが、「普通の」数学者の日々の研究に致命的な問題となったというのは、今すぐに思い浮かばない。

Q：数学の定理とその証明が書いてあるという天書。それに対して、私達が知っている領域というのは、そのごく一部だと思われる。その一部分が、今までの我々の数学によって得られた部分であるということに、何か理由があるのか？ つまり、私たちの持っている直感や認知バイアスなど、いろいろなものに引っ張られた結果が、今見ている数学なのか？

A：そうだと思う。プラトン主義の立場に立つと、我々人間の知性には限界があり、有限の人間の認識は、無限の真理のほんの一端にしか触れることができないので、おっしゃられた理由は特になく、ということ

ともなってしまふ。言い換えれば、天書全体をそもそも理解できないのが人間なので、我々の認識がごく一部であるとか、かなりの部分であるということすら認識できないわけである。これは、個人的には完全に受け身で、受入れ難い考え方とも思える。そうなると、構成主義のほうに傾く。しかし、これは可謬である。人間がすることなので、間違いがあり得ることも認める主義である。

あるときに証明したことが、間違っている可能性は認める。数学の内容を人間が創造していくというのが構成主義である。プラトン主義は、どこかに絶対的な真理があって、一部のみにしか見えず、全体は知り得ないが、それに少しでも近づくのが人間のサイエンス・数学であると規定する。従って、純粋にプラトン主義に立ってしまうと、ご指摘のようなことは、ほぼ何人も人間には知り得ないということになってしまう。神の御業は有限たる我々には知り得ないと。しかし、それはかなり不合理に聞こえるため、やはり構成主義に傾きたくなる。特に、応用数学や現象の解析などでは、少しずつ積み上げて、あるいは、自分でモデルを作って、ということになる。これは構成主義である。このように、心情的には永遠の真理を目指すプラトン主義に行きたいのだが、それだけでは窮屈なので、やはり構成主義的なところが混入する。しかし、ここの捉え方は、全然論理的ではない。

Q : アーテシュテイン「数学がいまの数学になるまで」という本を読んだ。ここでは、今の数学は、実は人々のバイアスから出てきた、というような話を書いてあり、衝撃を受けた記憶がある。(Z. アーテシュテイン著『数学がいまの数学になるまで』落合卓四郎監修、植野義明ほか訳、丸善出版、2018年)

A : このような考え方は、数学に絶対的な真理を認めてそれを見つけていく、というプラトン主義からは、出てこないであろう。これは、やはり構成主義的な発想やそれと関連した知識は絶対的に確実ではないとする可謬主義を背景にした考え方であろう。クロネッカーも直観主義も、基本的には構成主義寄りである。数学を絶対的な真理の体系と捉えて、まったく矛盾のない公理体系の中で全てを証明することは、なかなか難しいし、不都合なパラドックスも出てきてしまう。一方で数学の知識をそのときどきの時代で変わりうるものと捉えることに抵抗を覚えるのは数学者だけではないと思う。数学者の個々のレベルであるが、例えば、ある種の不都合なことが出てきたときに、数学者の態度として、「まだまだ我々の修行が足りないので、絶対的な知識体系としての数学理論をものにしていく努力を間断なく続けなくてはいけないのだ」と思うのか、「そもそも絶対的な真理はないのではないか」と思うのか、というぐあいにニュアンスが別れるのだろう。アーテシュテインの本に関する今のご質問は、かなり構成主義のほうに寄った感じがする。「人々のバイアス」に数学の進化のある部分を帰することは、割り切りすぎの可能性はあるが、数学の発展を理解するには有用な尺度かもしれない。

Q : 何が真理か、何が偽か、などの議論をしているのも我々人間である。人間、あるいは一般に生命体というのがどういうところに存在できるかという、地球環境のように、極めて安定したゆらぎの少ない環境の中に生きているわけである。これは宇宙でいうと、極めて特異的な場所である。我々は、そういう極めて特異的な環境の中で、いろいろな基礎的概念を形成しているわけである。物理の場合では、時空の概念などが典型例だが、環境に影響をうけて概念が形成されることはある。数学ではどうか？

A : まず最初のところで、地球外にそういう生物がいるかどうか、さらにおっしゃられるような特異的な環境でわれわれの数学が発生してきたという事実を認めるとして、それではそのような環境でなかったら数学はどうなっていたかに関しては、おおかたの数学者の立場としては、存在するかどうか証明されていないことを前提にしていたり、再現性がないことであるので、議論を避ける傾向がより強いと思われる。仮説を認めない(あるいは認めたがらない)数学者の性格からいって、概念形成における環境の影響を真剣に論じることをさけると思われる。しかし、このような議論においても、形式主義を重視し

2.13

モデル駆動型アプローチからみた逆問題の諸相

ている数学者と、直観主義的な立場の数学者では反応や思考がかなり違くと想像できる。数学者の間でも、この種の議論は平行線になるであろう。

数学者は、普段すごく論理的にガチガチなのだが、実際にそれを実施するときのマインドセットは、情緒的な面もある。語弊があるかもしれないが、自分が取り組んでいる問題に便利のように、主義を使い分ける、ということをする。これは、物理的な実在を対象としていないので、そこを使い分けてもあまり不都合は生じない、という考えに安住しているのかもしれない。こういう事情もあり、数学基礎論の専門家は知らないが、多くの数学者の間で、今のご質問のようなことが、真剣に論じられるというのは、めったにないと思う。

Q：要するに、その蓋然性が極めて高いと思うほうを信じる、というような立場か？

A：そうだろう。

Q：検証が極めて困難な例は議論しない、ということだろう。ところが、物理学者は、自然から、1つ深刻な教訓を学んだ。实在主義か、観測を通じて实在が現れるか、という。量子力学を作った人たちは、ほとんどすべて实在主義の立場であった。アインシュタインを含めて、観測などに関わらず、实在しないとおかしいと思ったわけである。検証できるテクノロジーがなく、長い間、水掛け論であった。しかし、ここ数十年での精密な実験により、観測によって現象が現れるということに1つも例外がないということが分かった。従って、例外がないから实在を信じていいとする考え方は、全部潰れた。従って、現時点において蓋然性は極めて高く、それ以外のことが全く考えられないということをもって、人間が真実だと思ったことに対して、おそらく神様は「それを決めるのはおまえではない」とアインシュタインにすら言ったのだと思う。地球環境が極めて異常な特異点であることを考えると、蓋然性だけに頼って真理かどうかを認めることを、本当に信じていいのかという気持ちになる。

A：物理的には、ご指摘のとおりだろう。数学では、「現象を観測で確かめる」というのが成り立たないところがある。成り立つとすれば、数学で保証されている予測結果を観測で確かめるということはあるが、それは、数学の使い方の問題であって、数学自体の問題ではないと解釈される。加えて、常識的に考えるとめったに起こらないと思われることが、実はほとんど全ての場合であった、というようなことは数学では多数の例がある。例えば、連続な曲線で、全ての点で接線を持たないような曲線は、普通は想像できないし、めったに無いだろうと直感的には思われる。しかし、位相空間の意味では、じつはそちらのほうが数ははるかに多いわけである。至るところは繋がっているのに、全ての点で接線がないという曲線は、当然グラフには絶対に書けない。しかし、集合全体から見るとそちらのほうが普通で、放物線などのほうが非常に特殊である。それが論理的な帰結であって、日常的な感覚では絶対にないわけである。また、0以上1以下の実数全体の集合のなかで有理数の集合と無理数の集合はどちらが大きいのかという話に関しても、古代ギリシア以来、作図などで「存在」が確かである有理数が、つかまえておけない無理数に比べれば集合論の意味（集合の個数＝濃度）でも、測度の意味でも（集合の長さを一般化した）も無視できる例外的な存在であるということもある。このような事実はたくさんあるが、このような事象から数学者は、他の自然科学者と異なるやりかたで、直観や感覚に向き合っている。物理での話と決定的に違うのは、やはり観測で検証するというのが、数学の普通の操作ではない、ということではないかと思う。

Q：自然科学でも、普遍主義と社会構成主義の対立といった話が出てくる。そういったときに、数学と自然科学の違いを感じるの、数学は、数学上の操作や証明などの手続きを数学の対象として扱える、ということがあるかと思う。自然科学は普通、科学のプロセス自体を科学しない。その観点で、この

構造的な対立を数学の問題として定式化して扱うという取り組みはあるのか？

A：それは、数学の中でも、数学基礎論と呼ばれる分野である。ただし、今日紹介した形式主義と直観主義の対立というのは1930年代の、それこそアインシュタインがそばに居た頃の論争である。彼は一步距離を置いていたらしいが。しかし、現代の傾向としては、対立をあまり表に出さず、融和というか、場合によって使い分けるといような風潮である。それには理由があり、形式主義だけでやっていると、ゲーデルの不完全性定理やパラドックスがたくさんあるので、数学としての基礎付けが緩くなってしまふ。一方で、直観主義にこだわると不便である。従って、今の風潮としては、形式主義と直観主義を使い分けるといのが、9割以上の数学者の立場ではないかと思う。大多数の数学者は、数学基礎論の結果や、論争とは関係なく自分の研究をやっているように思う。

Q：モデルが背後にあると仮定すると、説明もしやすいし、補外もできる。そのときに、そのモデルそのものが合っているかどうか、蓋然性と言うか、そこの正しさのようなものはどう判断するのか？

A：ご質問が「モデル」によって現象を定量的に解釈するための数理モデルを意味されているとすれば、現象の観測結果をよく説明し、予測でき、制御できるということであると思う。その意味では、物理的直感であって、これは数学自体の論点ではなく、数学「手法」の基礎となるモデルが現象の解析に役立ったかどうかの問題である。従ってモデルの検証は、物理などで考えられている考え方と大差はない。例えば、高炉の場合には、現場のテクニシャンの経験から、合っているかどうかである。ご質問でいわれる「モデル」を多分、個々の現象とは別個の数学的内容として解釈し抽象論展開の際の対象概念と解釈すると、それは現象から示唆は受けているものの、基本的にはそれ自体で存在し、最初に設定した定義、公理である意味自動的に展開していくわけである。そのような「モデル」の妥当性は、いかに多くのことを（現実世界にフィードバックして現象を解釈する場合も含めて）整合的に説明できるか、数学の多くの概念や理論を簡潔に統合できるか、いってしまえばその意味で（数学的に）美しい体系であるかに大きく依存するであろう。

繰り返してであるが、数学者はいつも非常に厳密に考えているように思われるかもしれないが、ある部分で数学者は「簡潔さ」とか「美」とかいう曖昧な基準や言葉使いをする傾向にあり、その意味ではいい加減になってしまうように見えるかもしれない。ダブルスタンダード、マルチスタンダードというべきか。形式主義と直観主義の分断についても、根底ではそうなっていると多くの数学者が認識しているのだが、あまり大っぴらには話題にしない話である。

Q：自然科学の歴史では、それまで持っていた理論が新現象に合わなくなってきたときに、パラダイムシフトが起きたりする。数学の世界では、そのようなことは生じないのか？

A：起きない。ある定理の50年前の証明の中に誤りが発見されたら、その定理は無効になるが、枠組み自身が無効になるということはない。現象からのサポートは、あればもちろんエキサイティングであるが、基本的には不要である。従って、観測技術の向上が決定的になるようなことは、狭い意味での数学ではほとんど起こらない。もちろん、時代によって、数学研究の方向性にも一定の傾向ははっきりある。たとえば自分の専門分野である偏微分方程式であるが、1960年代までは、線形偏微分方程式の一般理論の構築が主流であったが、1980年代をその知見を使った非線形偏微分方程式論の研究が盛んになり今に至っている。いまの修士の院生が、線形偏微分方程式の作用素論的研究をしたいと言ったら、おおかたの指導教員は別の課題を薦めるであろう。流行という言葉は適切ではないが、そのようなものはある。

Q：ゲーデルの不完全性定理やカントールの集合論の逆理などがあるにもかかわらず、便利だから、とい

## 2.13

モデル駆動型アプローチからみた逆問題の諸相

うのでやってよいのだろうかという疑問がわくが？

A：ご指摘のとおりである。深刻に考えると、本当にあやふやなものである。基礎はふらふらだが、みんな目をつぶっている。すぐ次の瞬間に大地震が起こる可能性はゼロではないけれど、皆普通に暮らしているような状態に近い。「起きない」と断言したのは言い過ぎで、「そういうことはありえないのだ」と目をつぶっているのだろう。もちろん、そのような変化はあった。19世紀にユークリッド幾何学の第5公準で、平行線の公準というものがあり、平行線が無限にある場合や1つもない場合など、そういったものを証明しようと長らく努力されてきていたが、誰一人成功しなかった。しかし、ある時、ロバチェフスキーやリーマンが、それも1つの矛盾のない幾何学の公理体系であるというふうに提唱した。それが一種のパラダイムシフトであったかもしれない。相対性理論の数学的な基礎づけにもなっている。しかし、その際にユークリッド幾何学が完全に捨て去られているということはない。この範囲では平たいと思えばいい、などとするわけである。このように、数学の場合には、捨て去られるということは非常に少ない。ほかの例であるが、集合論の観点であるが、「選択公理」というものがあり、全ての数学者が絶対的に認めているとはいえない。しかし、これはツォルンの補題とよばれる補題と同値であり、この補題はいろいろな局面でたいへん便利なので、一線の数学解析の専門家はその使用をまず否定しないであろう。

Q：自然科学におけるパラダイム・シフトは、枠組みの選択、という側面があるが？ 枠組みAという選択から、枠組みBという選択にシフトする。

A：それも、数学ではあまり行われぬ。ユークリッド幾何学を全部やめて、非ユークリッド幾何学に移るとか、双曲線幾何学に乗り換えるとかいうことはない。やはり、「この範囲ではこれ」などと使い分けるのである。自然科学の場合には、おそらく自然が1つだという信念があるからだと思う。そのことで、観測によって検証ができる。今できなくても、将来的にできる、ということが仮定されており、将来、それができた場合には現在の理論が書き換えられることがあるということが想定されているのであろう。一方、数学では、そもそも公理体系がただ一つではなく、「証明」ということの位置づけが普遍的に定まっているので、間違いがない限り、過去に証明された定理が時代とともに無効になることはない。ところで、そのようなことから、数学科で予算を請求するときは、図書が大事だと必ず言うのである。どんなに古い本でも無効になるということは、ほぼあり得ないし、逆に200年以上前に出版された本が現在の研究を活性化することも結構ある。このようなことは、他の自然科学では滅多にないだろう。

Q：文化的背景や環境的な背景が全く異なる宇宙人がいたとして、形式主義の立場に立つと、やはりその宇宙人は我々と同じ数学を見つけるのだろうか？ やはり人間にはバイアスがあり、そこに引っ張られているのか？

A：そう信じている人は居る。人間の持つ広い意味での文化で、宇宙人とコミュニケーションがとれるのは数学だけだろうとまで言うてしまう人もいる。それはロマン主義にすぎるが、私の感覚では、そう信じている人はかなりいると思う。もちろん、証明できないので、表立っては議論しないが。

A：今日紹介した話は、数学者の中で、そんなに特異な意見ではないと思っている。要は、信じたので信じる傾向にあるが、日常的にそれを原理にすえて実行しているかと言うと、必ずしもそうではないと思う。

# 3 | 総合討論

モデレーター 嶋田 義皓 (JST-CRDS)

CRDS で設定した3点について討論した。明らかになったポイントや重要な示唆を以下に示す。

## 3.1 現状

- ・ 一般には、「この仮説が成り立つとすれば、この実験の結果はこうなります」ということが分かっていた場合に、実際にその実験をしてそうだったとしても、その仮説が正しいとは必ずしも言えない。別の理由でその結果が出たかもしれない。統計で仮説検証しようとするなら、通常は帰無仮説を作って、それを証明する実験計画を作り、実験によってそれが否定されると帰無仮説が否定される、という流れになる。
- ・ 素粒子物理学などビッグサイエンスの場合には、通常、最初に理論があり、その理論に基づいて実験計画が立てられる。これは多分予算的な問題もあり、ビッグサイエンスには失敗が許されないというプレッシャーがあるために、ある意味で「分かったこと」を検証することになる。あるいは、蓋然性の高い仮説の中で、比較的夢のあるようなことを検証するというのが一般的な高エネルギー物理学の実験だろう。
- ・ 物理でも分野による。超伝導や超流動は全く理論なしに、たまたま実験がそれを発見した。物性物理学の分野ではよくあることで、量子ホール効果や高温超伝導の例もそうである。
- ・ 科学的発見のプロセスと、発見されたものを理解するプロセスは分けて考える必要がある。発見にもいろいろなものがある。高エネルギー物理学では、どれも間違っていないが自然がどれを選択するかが分からないというさまざまな理論があり、その中でもっともあり得そうなことはどれかというのを仮説から実験計画を作って試す。一方で、いわゆる世紀の大発見というものは、仮説が無い場合がほとんどである。なぜ「大発見」と呼ぶかということ、仮説がなかったからである。
- ・ iPS細胞の発見では、もともとの24因子から4因子を絞ったという流れなので、そこは発見があった後の理解のプロセスだと思う。基本的には、因子の組み合わせをつぶしていくプロセスである。しかし、そもそもなぜiPSを作ろうと思ったか、など、発想の部分は、発見プロセスだろう。現象が知られていない場合で原理的にはできるのではと誰かが思いつくところまでが発見のプロセスで、いったんそういう方針が決まってしまうと、仮説生成と検証のサイクルを回す理解のプロセスに入ってしまう。
- ・ それまで考えてきた仮説空間ではうまく説明できない実験結果が得られ、仮説空間を広げないといけないと認識した時点が、おそらく「発見」になっているのだと思われる。
- ・ 仮説を検証するループ（理解のプロセス）はサイエンスの中で重要なプロセスであるが、実験が精密化してくると、時々外れるものが得られる。こうして、専門家が理解している論理的なフレームに当てはまらないエビデンスが積み重なってくると、どこかでこの仮説空間を広げないといけなくなるのだろう。そして、その広げた部分まで含めてループが閉じた時に、その広げた部分が「発見」として認識されるのではないか。
- ・ 発見に関しては、セレンディピティーについても触れる必要があるだろう。
- ・ 数学の場合には仮説は「予想」と呼ばれる。予想よりももう少し広い「こういう枠組みがあって然るべきだ」というような問題提起というものがある。日本の数学者は割と未解決予想にチャレンジするなど問題を解く側で活躍してきたが、問題を作る側というのも大事である。数学の世界でも科学と同様に、仮説、実験、検証ということがぐるぐる回っている。

## 3.2 目標・研究開発課題

- ・「人間の認知能力を超えた」という仮説生成ではバランスを欠く。機械が得意な仮説生成、探索の部分と、人間が得意な部分というものがあるように思う。
- ・仮説空間をどのように設定するかということが非常に重要。仮説空間の境界が決まった後で、その中を探索するというのは機械が得意だが、「そもそもこの境界の仮説空間で良いのだろうか」ということは人間の発想に相当する部分である。機械で無限の探索空間を設定するわけにはいかないの、人の直感のようなもので「ここに行けそうだ」というのを決めないといけないだろう。
- ・仮説の前に「問い」の空間を決めるという話が抜けているのではないか。計画に基づいて実行された実験結果がうまくいっているかどうかというような解釈や、たまたま観測されてこれも考えなければいけないとか、あるいはそこでいったん解釈したものと過去に観測されているものとの間をつなぐにはもう少し仮説空間を広げなければならないなど、新たな問いや仮説空間が生み出されるプロセスを含める必要がある。
- ・人間にしかできない部分というのは、仮説の生成ではなく、「仮説空間の生成」だろう。その設定は人にしかできない部分で、仮説空間としては大きすぎると機械でも探索できないので、ある程度リーズナブルなサイズにしなければいけない。そこには人の直観が非常に重要だろう。
- ・仮説空間を探索して、例えばさまざまな材料の候補が出てきたとき「この材料が良い」というような評価は、やはり人にしかできないように思う。事前に「こういう条件を満たす材料が欲しい」という場合には問題にならないかもしれないが、思ってもみなかった性質を持つ材料が出てきたときに「これは良い材料だ」と評価するのは人間にしかできないだろう。
- ・こういった現在の情報技術の中でできない部分には人を配置して、それ以外のところはできるだけ自動化するような科学のやり方を考えるということになるだろう。
- ・北野氏（2.6節）はAutomationとAutonomyの違いで表現していた。Automationが力づくでいろいろ探索し、Autonomyの方は、その結果に対して人間の価値観に合うように評価していくというようなことを指していた。そして、そこをあまり人間に近づけすぎると、人間の認知バイアスなどと重なってしまうので難しいとも。
- ・評価にもできそうなものと、そうでないものがあるだろう。あらかじめ社会的に「こういうものは役に立つ」と分かっているタイプのものと、そうではなく人間の直観によって「これはすごく役に立ちそうだ」というのを事前に与えることが難しいタイプのもの。これらが混ざっているように感じる。
- ・徐氏（2.4節）が「AIは本を読んで勉強して欲しい」と述べていたように、データからしかできないというところがマテリアルズ・インフォマティクスのもどかしいところである。その意味で、物理など我々人間が少なくとも知っているようなところからスタートして先に進むというのが欲しい。もちろん、我々の知っていることが間違っていて、正しくないところからもスタートし得るのだが。
- ・仮説空間の探索という意味では、ここは探索しなくても良いということまで含めて、初期値というか境界を上手に与えないといけないだろう。
- ・研究には「ゴールオリエンティッド（ストラテジーオリエンティッド）」と「キュリオシティードリブン」の2つあるということ、科学プロセスのループの絵にどう折り込むかは重要なポイントの一つだろう。
- ・「科学の科学」の議論も大事だろう。ここの研究をどう進めたらよいかというもの難しい。科学哲学のようなアプローチもあれば、もう少し構成論的なアプローチも可能だろう。

### 3.3 推進方策

- ・メタ科学に関して、久木田氏（2.1節）の質疑応答でも議論になったように、科学で使われている道具箱というものが何かという整理をした方がいいように思う。中でも重要なのは「推論の仕組み」である。数学や統計という道具、あるいは論理学や演繹などがある、というような絵だろう。もちろん、科学の道具の中には、科学者たちがコラボレーションするための道具というのもあり、論文、国際会議、プレプリントサーバー、あるいはGitHubみたいなものなどがあるだろう。そういったもののランドスケープを作って見たら良いと思う。
- ・道具といった時には、何を「証明」と呼ぶか、何を「検証」とするかのような枠組みやプロトコルのような話も含まれるだろう。
- ・科学的発見プロセスの解明というものに対しても、やはり構成論的にアプローチしたい。それはやはり道具などの体系がどうなっているかということと表裏一体ではないか。学術に携わる人に対して、これまで日本では、「科学の道具箱にこういうものがあります」というようなことをきちんと教育してこなかったように思う。
- ・JSTなどがリーダーシップを取り、科学の方法論について教育を行うというのは重要である。科学の道具箱とは何かとか、人間のバイアスが入り込むのをどういうふうに取り除くかというようなものを教育する。
- ・データの共有のためのプラクティス作りも必要である。実験の事前登録制や、ネガティブな結果の共有など。
- ・国の基礎研究の体力としては、実験装置を作るエコシステムというのを国内の産業として育成しておくことは重要であろう。実験の自動化に関連して、何か実験装置を作りたいと思った時に、どの企業がどういうノウハウを持っているというようなことがどこかにカタログ化されていて、ワンストップでここに頼めばこういうのを作ってくれる、というようになっていけば、日本の科学がすごく進むように思う。
- ・また、論点として、CACMの記事に<sup>24</sup>書かれているような、さまざまなサイエンスの研究プロジェクトの中で有期雇用されているデータサイエンティストの問題にも触れて欲しい。日本こういうことが起きているかどうか分からないが、記事によると米国ではいわば新しいピペドのような形で使い捨てにされているという現状があるようだ。調査しておく必要があるように思う。
- ・橋本氏（2.3節）で紹介があった「ディープラーニングと物理学」のコミュニティのようなものが、化学や生物などの分野でも起きているとすれば、そういったコミュニティをできるだけ支援するべきだろう。起きていないのであれば、物理学ではこういう面白いのをやっているからこの分野でもやったらというようなことを言った方がよいだろう。JSTは色々なツールを持っているはずなので。
- ・数学を機械学習で進めるとするのは、実際には主流にはならないかもしれないが、結構面白いことも出てくるのではないかという気がする。数学者と呼ばれている人の中でも、偏微分方程式系など物理に近い方の研究をされている方が、機械学習にかなり注目している。
- ・医学や生命科学、生物学では、やはりデータベースを作らなければいけないというのがさまざまな文脈で出てくるが、いつも「誰がやるのか」という視点が欠けている。やはり研究者はそのところをやりたがらないので、大事だと言われていても、なかなかデータベースが整備されないというのが、日本の現状だと思う。

24 C. L. Borgman et. al., "Thorny Problems in Data (-Intensive) Science", Communications of the ACM, August 2020, Vol. 63 No. 8, Pages 30-32

ている。米国では、例えばBroad Instituteでは、あるがんの疾病例を聞きつけたら、そういうサンプルを集めてひたすら遺伝子解析をしてというデータを集める専用の部隊がいて、それがあるが故に、ビッグサイエンスでデータサイエンスができています。この視点で見ると、そこをやっている研究者のキャリアはどうなっているのだ、という問題は残っている。

- ・研究者のキャリアという面と、やらなければならないという面と、バランスを考えた仕組みを作ってゆく必要があると考える。キュリオシティードリブンだけではこういう部分は絶対にできないと思う。なんとかして機械だけでデータベースを作るみたいなのところまでいけると良いのかもしれない。本当にできるか分からないが、人手を介さないデータベース生成みたいなことが必要。
- ・データのクレンジングはドメイン知識がなければできないなど、悩ましい。できる人は限られているが、実際に取り組む仕事はキャリアにつながらない。日本では、逆にできる人が少ないというのが実態だろう。だから、できる人が増えたことでこの問題が顕在化してくるかもしれない。
- ・相関と因果の違いを国民に浸透させることができるか、というようなことが中教審の指導要領の改訂の際にも話題になった。これは結局未解決の問題であるが、こう言ったことをCRDSなどを通じて国の施策にインプリメントする良い機運なのだろう。長期的に見ても国のためになる提案という認識である。
- ・一流の研究者は、頭の中に大局観、情報の地図を持ちながらキュリオシティードリブンに研究を進めている。頭の中の地図とは、その分野の何が解けていて、何が解けてなくて、その解けていないものの中にどういう関係があって、どれが解けると全て解けてしまうか、といったような地図である。このような地図を作るのに、データサイエンスはものすごく役に立つのではないかと思う。
- ・このような地図は因果まで行かなくて相関が分かるだけでも、ものすごく大きな意義があると思う。さまざまな分野の間にどういう相関があって、どういう大小関係があるかなどという膨大なデータである。これまでごく一部のスーパー勉強家はこれを勉強するのだが、これからはその地図をもう少し多くの人が見えるようなデータベースとして整備されるようになれば、これは凄まじい価値を持つのではないかと期待される。
- ・情報地図に基づいてキュリオシティーを発展させていくというのが、よいサイエンスを進めるための助けになると思う。



総括責任者	木村 康則	上席フェロー	(システム・情報科学技術ユニット)
チームリーダー	嶋田 義皓	フェロー	(システム・情報科学技術ユニット)
チームメンバー	東 良太	フェロー	(システム・情報科学技術ユニット)
	島津 博基	フェロー	(ライフサイエンス・臨床医学ユニット)
	相馬 りか	調査員	(戦略研究推進部ICTグループ)
	福島 俊一	フェロー	(システム・情報科学技術ユニット)
	眞子 隆志	フェロー	(ナノテクノロジー・材料ユニット)
	的場 正憲	フェロー	(システム・情報科学技術ユニット)
	宮藺 侑也	フェロー	(ライフサイエンス・臨床医学ユニット)
	若山 正人	上席フェロー	(システム・情報科学技術ユニット)

## 俯瞰セミナーシリーズ報告書

---

CRDS-FY2020-WR-13

# 機械学習と科学

---

令和3年3月 March 2021

ISBN 978-4-88890-725-5

---

国立研究開発法人科学技術振興機構 研究開発戦略センター

Center for Research and Development Strategy, Japan Science and Technology Agency

〒102-0076 東京都千代田区五番町7 K's 五番町

電話 03-5214-7481

E-mail crds@jst.go.jp

<https://www.jst.go.jp/crds/>

本書は著作権法等によって著作権が保護された著作物です。

著作権法で認められた場合を除き、本書の全部又は一部を許可無く複写・複製することを禁じます。

引用を行う際は、必ず出典を記述願います。

This publication is protected by copyright law and international treaties.

No part of this publication may be copied or reproduced in any form or by any means without permission of JST, except to the extent permitted by applicable law.

Any quotations must be appropriately acknowledged.

If you wish to copy, reproduce, display or otherwise use this publication, please contact [crds@jst.go.jp](mailto:crds@jst.go.jp).

FOR THE FUTURE OF  
SCIENCE AND  
SOCIETY



<https://www.jst.go.jp/crds/>