

科学哲学から見た 科学の自動化

呉羽 真

大阪大学 大学院基礎工学研究科
特任助教 (常勤)
博士 (文学)

1

自己紹介

- 専門分野は、**哲学**。特に、**認知科学・AI/ロボット・宇宙開発などの、科学技術に関する哲学的問題**について重点的に研究
- 大阪大学知能ロボット学研究室 (石黒研究室) にて、**ロボットの哲学**の研究に従事 (2018.4~現在)
- 「科学の自動化」に関する研究で、**2019年度科学技術社会論・柿内賢信記念賞 (奨励賞)** を受賞 (2019.11)



クラーク著,
呉羽他訳(2015)



伊勢田・神崎・
呉羽編(2018)



稲葉他編(2020)

呉羽・久木田
「AIと科学研究」
を収録

2

科学の自動化がもたらしうるもの —Alien Science と Alienation of Science—

「異質な科学 (alien science)」…人間の科学とは異なる仕方世界を理解する科学の形態

- 科学哲学における「N=1」問題の解決?
- どこまで/どうやって実現できるか?

「科学の疎外 (alienation of science)」(呉羽・久木田 2020)…科学がその本来の目的 (ex. 理解) や価値 (ex. 創造性) を喪失する事態

- 例① 科学は人間に理解できないもの / 理解を目指さないものになってしまう?
- 例② 科学において人間が創造性を発揮する余地はなくなってしまう?

3

科学的理解の科学哲学

科学的理解に関する共通了解

- ① 理解は科学の中心的目的の1つ
 - 予測や制御の基礎としての実際的価値に加えて、好奇心の充足という知的価値をもつ (Hempel 1965; Salmon 1998)
- ② 理解は (典型的には) 説明によってもたらされる
 - 因果的説明による「メカニカルな理解」と、統合的説明による「コスモロジカルな理解」(Salmon 1998)
- ③ 理解はプラグマティックな (=主体の知識・能力・関心等に依存する) 側面をもつ (Hempel 1965)

科学的理解に関する議論動向の変化

- 従来…理解は説明がもたらす単なる心理的副産物 (Hempel 1965; Trout 2002)
- 最近…理解の再評価
 - 例: 理解とは、理論を特定の目的 (説明の構築等) に役立つスキルをもってのこと (De Regt 2017)

4

AI駆動科学と理解

機械学習技術の利用に伴う「ブラックボックス問題」

- 例) トゥルイユ「1つの興味深い可能性は、私たちは「理解可能な科学」の時代を終えつつある、ということだ」(Bohannon 2017)
- 理解をもたらさない科学は、「科学」の名に値するか? よく当たるお告げと何が違うのか?



Science 357
(6346), 2017

疑問: ブラックボックス問題は解決可能か? 本当に解決すべき問題か? → 関連する論点

- 理解可能性の基準は時代とともに変化してきた (例: 遠隔作用)
- 人間の思考過程もブラックボックス? (と言われるが、程度の差はある。コンピューターを用いた数学の証明に関する議論を参照)
- ブラックボックスを開けるXAIの研究も進められている (ただし予測性能と説明可能性の間にトレードオフがある?)

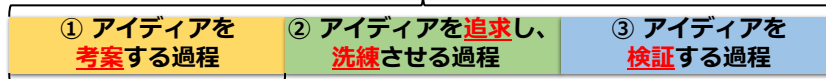
5

科学的発見の科学哲学

科学的発見…科学研究の最も創造的なフェーズ → これまで科学哲学はあまり論じてこなかった

- 「(狭義の) 発見の文脈」と「正当化の文脈」を区別し、前者を科学哲学の主題から除外
 - 例) ポパー『探究の論理』(Popper 1934): 科学的発見に論理はない

広義の「科学的発見」



狭義の「科学的発見」

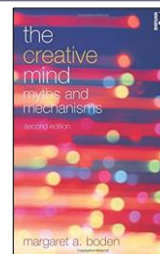
- 例外) 科学哲学者兼科学者による科学的発見の方法の分析
 - アブダクション (Peirce 1878)
 - 美的判断に基づくアイディアの結合 (Poincare 1908)
 - ヒューリスティクスを用いた問題空間の探索 (Simon 1977)
 - アナロジー (Holyoak & Thagard 1994)

6

AI駆動科学と創造性

創造性の哲学では、科学的発見を含む創造活動の本質についての分析が行われてきた

- 例) ボーデン (Boden 2004)
 - 「創造性」の定義: ①新しく、②驚異的で、③価値のある、アイデアや人工物を生み出す能力
 - 創造性の分類: 結合的 / 探索的 / 変形的



Boden(2004)

疑問: 力ずくの探索による科学的発見は創造的か?

- 人間の場合には創造的とは言われない
 - 例) ダイソン「科学は、1粒の砂の中に1つの世界を見たり、1本の野の花の中に1つの天地を見ることができるときに、最も創造的である。重いハードウェアや巨大な機械も科学の一部ではあるが、最も重要な部分ではない」(Dyson 1988)
- 人間のとは別種の創造性もありうる?
 - 創造的か否かは、プロダクトの特性だけで決まる or プロセスの特性にも依存する?



ダイソン(1990)

7

提言

- 科学の自動化は、「科学的理解とは何か」、「科学的発見/創造性とは何か」、といった問題を改めて考え直す機会を提供する
 - ➔ 哲学・心理学・認知科学などの知見が役立つかも
- 科学の自動化は、科学と社会の関係性を巡る課題を生み出す
 - ➔ 様々な分野の科学研究者、そして市民がどう考えているか調査を
 - ➔ 科学政策、科学教育、科学コミュニケーション、哲学…などの研究者も加えた議論を

8

文献①

- Boden, M.A. 2004. *The Creative Mind (2nd Edition)*, Routledge.
- Bohannon, J. 2017. 'The Cyberscientist', *Science* 357(6346): 18-21.
- Chiang, T. 2002. 'The evolution of human science', in T. Chiang, *Stories of Your Life and Others*, Tor Books. (T・チャン「人類科学の進化」『あなたの人生の物語』所収,古沢嘉通訳, 早川書房, 2003.)
- de Regt, H.W. 2017. *Understanding Scientific Understanding*, Oxford University Press.
- Dyson, F.J. 1988. 'Science and space', in his *Infinite in All Directions*, Harper & Row. (F・ダイソン「科学と宇宙」『多様化世界』所収, 鎮目恭夫訳, みすず書房, 1990.)
- Hempel, C.G. 1965. 'Aspects of Scientific Explanation', in his *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press. (C・ヘンペル『科学的説明の諸問題』長坂源一郎訳, 岩波書店, 1973.)
- Holyoak, K.J. & Thagard, P. 1996. *Mental Leaps*, MIT Press. (K・J・ホリオーク, P・サガード『アナロジーの力』鈴木宏昭・河原哲雄監訳, 新曜社, 1998.)

9

文献②

- 呉羽真・久木田水生 2020. 「AIと科学研究」稲葉振一郎ほか編『人工知能と人間・社会』所収, 勁草書房.
- Peirce, C.S. 1878. 'Deduction, induction, and hypothesis', *Popular Science Monthly* 13: 470-482.
- Poincaré, H. 1908/1952. *Science and Method*, Dover Publications; original published in French in 1908. (H・ポアンカレ『科学と方法(改訳)』吉田洋一訳, 岩波書店, 1953)
- Popper, K. 1934/1959. *The Logic of Scientific Discovery*, Routledge; original published in German in 1934. (C・ポパー『科学的発見の論理』上・下, 大内義一・森博訳, 恒星社厚生閣, 1971)
- Salmon, W.C. 1998. 'The importance of scientific understanding', in his *Causality and Explanation*, Oxford University Press.
- Simon, H.A. 1977. *Models of Discovery and Other Topics in the Methods of Science*, D. Reidel.

10