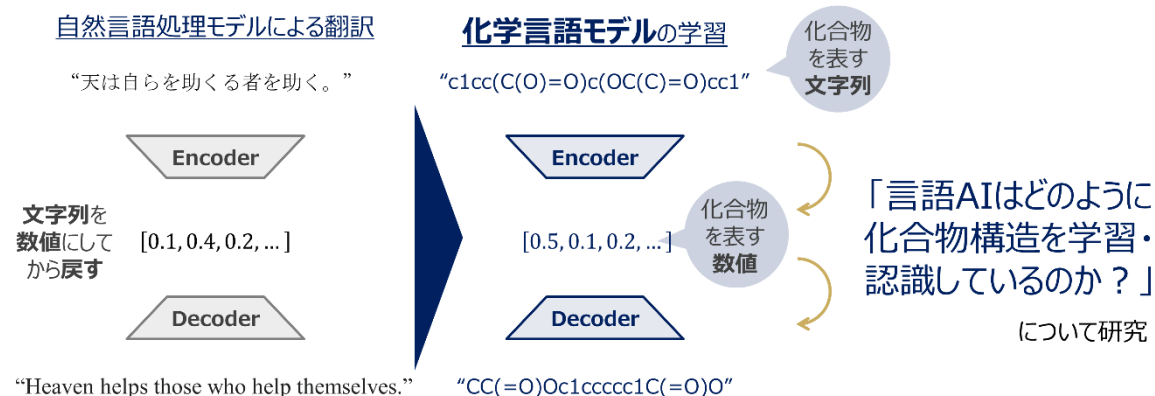


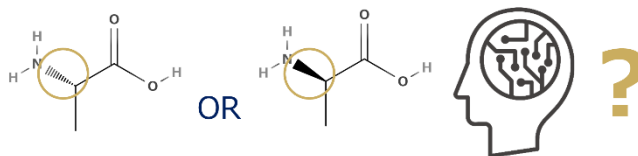
言語 AI が多様な化合物構造を学習する過程の特徴を発見 ——化学言語モデルとしての Transformer はキラリティの認識を苦手とする——

発表のポイント

- ◆ 代表的な深層学習モデルである Transformer が多様な化合物構造を学習していく過程を、モデルが認識する部分的な構造に着目して精査した。
- ◆ 多様な化合物構造を学習するには、キラリティを誤認識しやすいことなどを明らかにした。
- ◆ 「どのように言語 AI が化合物構造を学習・認識しているか？」というブラックボックス解明の一助になると期待される。



キラリティを誤認識
しやすいことを見



※キラリティ：化合物の立体的性質の一つ

概要図

概要

東京大学大学院薬学系研究科の吉開泰裕大学院生、水野忠快助教らによる研究グループは、深層学習モデルが多様な化合物構造を学習する際に苦手とする構造を明らかにしました。

深層学習モデルは現在創薬分野でも盛んに応用されています。特に近年は、自然言語処理（注1）を援用するアプローチが頻用されていますが、深層学習モデルがどのように多様な化合物構造を学習・認識しているかは明らかになっていませんでした。本研究グループは、代表的な深層学習モデルである Transformer（注2）が多様な化合物構造を学習していく過程を、モデルが認識している部分的な構造に着目して精査しました。結果、Transformer モデルが化合物の立体的な性質の一つであるキラリティ（注3）を誤って認識しやすいことなどを発見しました。本研究成果は2月16日、国際科学誌「Nature Communications」に掲載されました。

発表内容

〈研究の背景〉

東京大学大学院薬学系研究科 水野助教の研究グループは、データサイエンスにより医薬品の認識されていない作用（注4）を理解・活用する研究を進めています。機械学習をはじめとしたデータサイエンスは、薬学研究でも日常的に活用されています。一方、一般に機械学習モデルや統計モデルは化合物の構造を直接扱えないため、まず何らかの方法で化合物を数値へと変換する必要があります。多様な化合物構造を数値へと変換する方法として、2016年以降、自然言語処理の技術を援用した深層学習モデル（化学言語モデル）がよく用いられています。具体的には、言語の翻訳のように、化合物構造を表す文字列表現を入力し、同じ化合物構造を表す文字列を出力するタスクを学習したモデルなどが該当します。しかし、どのようにこれらの深層学習モデルが多様な化合物構造を学習・認識しているかについては明らかとなっていませんでした（注5）。

〈研究の内容〉

深層学習モデルは学習を重ねることで目的とするタスクの精度が向上していきます。そこで本研究グループは、学習進捗に応じて深層学習モデルの性質がどのように変遷するかに注目しました。Transformer は、2017年に開発されて以来、機械翻訳など自然言語処理の深層学習モデルとしてデファクトスタンダードとなり、他分野でも頻用される深層学習モデルです。本研究グループは、Transformer モデルの化合物構造の学習進捗をフォローしていった際、学習開始早々に部分的な文字列の学習は完了する一方、文字列全体の学習には時間がかかることを見出しました。部分的な文字列は化合物の部分構造を表すと推定されます。そこでモデルの入出力それぞれの部分構造を比較すると、学習開始早々に両者が一致することを見出しました。さらに学習経過ごとのモデルを用い、それぞれの化合物を数値化してその毒性や物性を推定したところ、学習経過に依らず、学習初期段階から概ね同程度の精度を示すことを見出しました。

上記検討の最中、深層学習モデルの学習が確率的に滞る現象（注6）を見出し、原因の探索に取り組みました。化学言語モデルの入力となる化合物の文字列表現は、各文字が原子など何らかの化合物の要素と対応しています。そこでそれぞれの文字を除外した際の精度を調査することで、各文字の学習への影響力を評価しました。結果、化合物の立体的特性であるキラリティを表現する文字（@、@@）が見出されたことから、キラリティの認識がTransformerモデルの学習に大きく影響することが示唆されました。原子ごとの精度を評価した際にも同様にキラリティの学習が遅いことが明らかになっています。上記の結果は、同分野で頻用される文字列表現である SMILES 表現（注7）だけでなく、InChI 表現（注8）でも確認されました。

〈今後の展望〉

本研究により、代表的な深層学習モデルである Transformer が多様な化合物構造を学習する際の特徴の一端が明らかとなりました。大規模言語モデルが目覚ましい発展を見せ、言語 AI の知能に関する研究も盛んとなっている中、「どのように言語 AI が化合物構造を学習・認識しているか？」というブラックボックス解明の一助になることが期待されます。

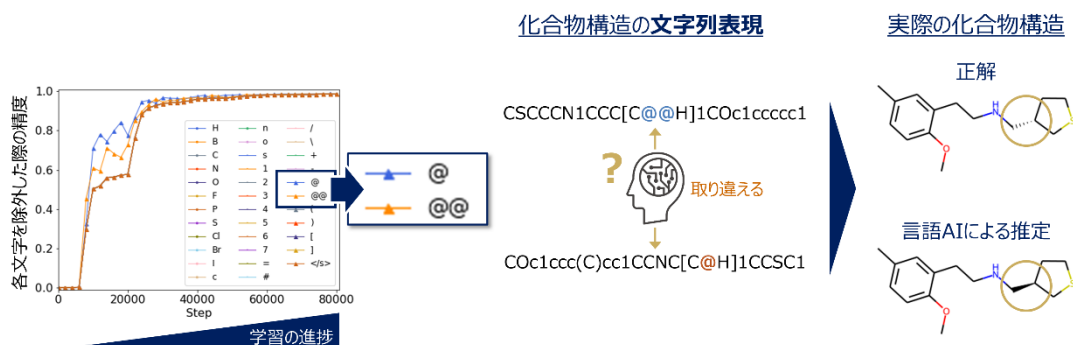


図 1：キラリティに関わる学習の進捗と誤認識による影響

発表者・研究者等情報

東京大学 大学院薬学系研究科

吉開 泰裕 修士課程

水野 忠快 助教

根本 駿平 博士課程

楠原 洋之 教授

論文情報

雑誌名：Nature Communications

題名：Difficulty in chirality recognition for Transformer architectures learning chemical structures from string representations

著者名：Yasuhiro Yoshikai, Tadahaya Mizuno*, Shumpei Nemoto, Hiroyuki Kusuhara

DOI: 10.1038/s41467-024-45102-8

URL: <https://www.nature.com/articles/s41467-024-45102-8>

研究助成

本研究は、国立研究開発法人日本医療研究開発機構 (AMED) 創薬基盤推進研究事業 (課題番号：JP23ak0101199h0001, 研究代表者：水野忠快)、及び AMED 医薬品等規制調和・評価研究事業 (課題番号：JP22mk0101250h0001, 研究代表者：水野忠快) の支援により実施されました。

用語解説

(注 1) 自然言語処理

ヒトが日常的に使っている自然言語をコンピューターに処理させる一連の技術。

(注 2) Transformer

2017 年に提案された深層学習モデルで Attention 機構に特徴がある。自然言語処理分野で提案されたものの、すぐに画像など他分野でも応用され、現在の深層学習モデルの基盤を支えるモデルの一つ。

(注 3) キラリティ

3 次元の物体等がその鏡像と重ね合わせることができない性質。キラリティを示す化合物は、当該化合物とその鏡像とで、タンパク質への結合親和性が異なるケースなどが存在する。

(注4) 医薬品の認識されていない作用

市場撤退につながる有害事象を引き起こす副作用や、ドラッグリポジショニングにつながる作用など、開発時には認識されていないもののその後明らかとなった作用。

(注5) End-to-end タスクの学習と表現学習との違い

発がん性の予測など、具体的なタスクを解くモデル(End-to-end なモデル)に関しては、どの部分を認識しているかよく研究されている。一方、多様な化合物構造自体を学習する際、どこを認識しているかについては不明であった。

(注6) 深層学習モデルの学習が確率的に滞る現象

一般に深層学習では、同じ構造を持ち、同じデータセットを学習していても、特別な操作なくしては学習回ごとに厳密には異なるモデルとなる。本記述は、学習回によって、学習が順調に進んだケースと進まないケースがあったことを意味している。

(注7) SMILES 表現

Simplified Molecular Input Line Entry System の略であり、化合物構造の代表的な文字列表現の一つ。化学言語モデルにて最も使用されている。

(注8) InChI 表現

International Chemical Identifier の略であり、化合物構造の代表的な文字列表現の一つ。

問合せ先

〈研究に関する問合せ〉

東京大学大学院薬学系研究科

助教 水野 忠快 (みずの ただはや)

Tel : 03-5841-4771 E-mail : tadahaya@mol.f.u-tokyo.ac.jp

〈報道に関する問合せ〉

東京大学大学院薬学系研究科 庶務チーム

Tel : 03-5841-4702 E-mail : shomu@mol.f.u-tokyo.ac.jp