

«Запрещенная» химия в теории ИИ

Якомаскина Т. А.

Мелитопольский Государственный Университет им. А. С. Макаренко, Мелитополь, Россия

«Forbidden» chemistry in AI theory

Yakomaskina T. A.

Makarenko Melitopol State University, Melitopol, Russia

Введение

Эта тема достаточно актуальна, поскольку в настоящее время нейронные сети стали одним из основных инструментов теории, и трудно назвать прикладную область, где не было бы хотя бы попыток их успешного применения. Они сформировались в сильное разветвленное научное направление, которое включает теорию искусственных нейронных сетей, нейроматематику и нейронные компьютеры, реализующие парадигмы нейронных сетей. Особое место занимали сети обратного распространения. Теоретически они являются наиболее разработанными и используемыми в химии и химической технологии.

Цель моего исследования заключается в том, чтобы изучить теории выдвинутые ИИ и сопоставить с реальными экспериментами.

Основная часть

Химия, известная из школьного курса, в основном рассказывает о том, что происходит в «нормальных условиях». Фундаментальное изменение этих условий может привести к изменениям не только физических, но и химических свойств, что помогает создавать принципиально новые материалы. Эта наука вступает в цифровую эпоху: новые вещества и явления теперь появляются не в пробирке, а в виртуальном мире с использованием искусственного интеллекта. Это оказалось не только быстрее и дешевле, но и привело к революционным скачкам. Артем Оганов — одна из главных мировых звезд новой химии, рассказывает об удивительных результатах своих исследований.

«Я кристаллограф-теоретик, занимаюсь разработкой новых материалов и созданием методов, которые позволяют открывать новые материалы. Вместе с сотрудниками я решил задачу, которую, как раньше считалось, решить невозможно — предсказание кристаллической структуры исходя из химического состава вещества. Эта задача равноценна поиску иголки в стоге сена размером со Вселенную — возможных вариантов расположения кристаллических структур в пространстве невероятно много. Но мы решили её и пошли гораздо дальше — научились с помощью компьютера предсказывать устойчивые химические соединения по набору исходных элементов.» — рассказывает Оганов, о своей нынешней работе.

Что побудило людей начать использовать этот тип исследований? Известно, что мантия Земли состоит в основном из силикатов магния и железного ядра. Также было известно, какие силикаты магния образуются при разных давлениях и температурах — никто не ожидал от них никаких сюрпризов. Однако многие геофизики с 1950 года очень хорошо знают, что на границе ядра и мантии, на глубине почти трех тысяч километров, находится слой средней толщиной 200 километров, который обладает необычными, очень своеобразными свойствами. Этот слой ведет себя не как однородное вещество, а как слоистая среда, и не совсем понятно почему. Силикаты магния, которые были известны в то время, вообще не слоистые, а имели однородную структуру во всех

направлениях. Этот слой очень неравномерен: в некоторых местах его вообще нет, в некоторых его толщина достигает 300 километров.

Оказывается, над внешним жидким ядром Земли возвышается нечто вроде подземных гор высотой в двести километров.

Источником знаний о том, что происходит на глубине тысяч километров, являются сейсмические волны. Таинственный слой, отделенный от остальной оболочки довольно сильным сейсмическим разрывом. То есть скорость прохождения сейсмических волн внезапно, резко меняется, когда пересекается граница между нормальной оболочкой и таинственным слоем.

И все эти аномалии никак нельзя было объяснить, пока не был обнаружен новый минерал с составом $MgSiO_3$, названный пост-перовскит. Доказано, что структура этого минерала объясняет почти все загадки, которые люди пытались разгадать десятилетиями.

Подавляющее большинство веществ в твердом состоянии имеют кристаллическую структуру. Есть также аморфные материалы или стекла, а есть квазикристаллы — недавно открытое состояние вещества, уже исходя из этого факта ясно, что оно встречается редко.

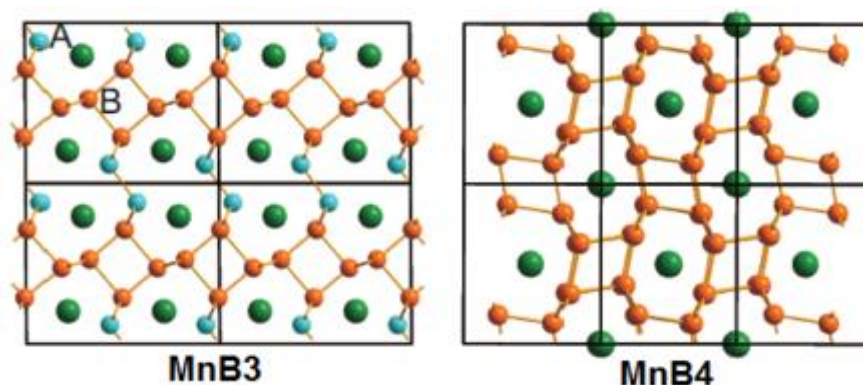
Кристаллические вещества — это материалы, которые имеют упорядоченную структуру. Вы можете взять параллелепипед микроскопических размеров и повторить его во всех направлениях, вдоль всех трех координатных осей — таким образом, воспроизводя всю трехмерную структуру кристалла: он состоит из таких параллелепипедов, которые идентичны друг другу. Металлы, керамика, глина, горные породы — почти все твердые вещества, которые известны и используются в технике, имеют кристаллическую структуру.

Существует целый ряд методов расшифровки кристаллических структур, хотя сделать это не всегда легко. Однако до недавнего времени это считалось невозможно предсказать, но это очень важно: зная структуру, вы можете не только предсказать огромный перечень его свойств, но и понять, интересен ли этот материал для практического использования, поскольку все они определяются его кристаллической структурой: теплоемкостью и электропроводностью, сверхпроводимостью, термодинамическими характеристиками, скоростью прохождения звука через вещество. Почему графит мягкий и черный, а алмаз прозрачный и сверхтвердый? Потому что атомы углерода расположены по-разному, и именно структура определяет свойства вещества. Что определяет биологические свойства белков? Пространственная структура. Как ДНК передает информацию? Благодаря своей структуре — двойной закрученной спирали. Таким образом, структура является основным носителем информации в веществе.

Существует ли уже нейронная сеть, которая может выполнять такие вычисления? Программа, конечно уже доступна и называется USPEX (Universal Structure Predictor: Evolutionary Xtallography).

USPEX решает ключевую проблему теоретической химии кристаллов, вопрос формулируется следующим образом: какую структуру будет иметь вещество с известным химическим составом при определенных условиях? В дополнение к прогнозированию кристаллических структур, USPEX также может быть использован для других целей: для прогнозирования структуры наночастиц, полимеров, поверхностей и 2D-кристаллов. Он может эффективно предсказывать стабильные химические соединения и структуру вещества, имея на входе только названия химических элементов.

Например, ученые с помощью алгоритма USPEX проанализировали, какие соединения марганца (Mn) и бора (B) стабильны, и в 2014 году опубликовали статью об этом. Теоретически они доказали, что экспериментаторы не обнаружили соединение — борид марганца с формулой MnB_3 , а также прояснили структуру MnB_4 . Затем оба положения были полностью подтверждены экспериментально.



Как работает USPEX? Суть поиска наиболее стабильной структуры сводится к вычислению состояния вещества, обладающего наименьшей энергией. Энергия в этом случае зависит от электромагнитного взаимодействия ядер и электронов атомов, составляющих исследуемый кристалл. Её можно оценить с помощью квантовомеханических расчетов, основанных на упрощенном уравнении Шредингера.

Таким образом, алгоритм USPEX использует функциональную теорию плотности, которая была разработана во второй половине прошлого века. Его основная цель - упростить расчеты электронной структуры молекул и кристаллов. Теория позволяет заменить функцию многоэлектронных волн электронной плотностью, оставаясь при этом формально точной. На практике это приводит к снижению сложности вычислений и, как следствие, времени, затрачиваемого на них. Таким образом, квантовомеханические вычисления объединены с эволюционным алгоритмом в USPEX.

Структуры с наименьшей энергией можно искать методом перебора: случайным образом расположите атомы друг с другом и проанализируйте каждое такое состояние. Но поскольку количество вариантов огромно, вычисление заняло бы слишком много времени. Поэтому ученым удалось добиться успеха только после разработки более хитрого метода. Сначала случайным образом генерируется небольшое количество структур и вычисляется их энергия. Система удаляет варианты с наибольшей энергией, то есть наименее стабильные, и генерирует аналогичные из наиболее стабильных, уже считая только их. В то же время компьютер случайным образом продолжает создавать новые структуры для поддержания разнообразия «популяции», что является необходимым условием успешного развития.

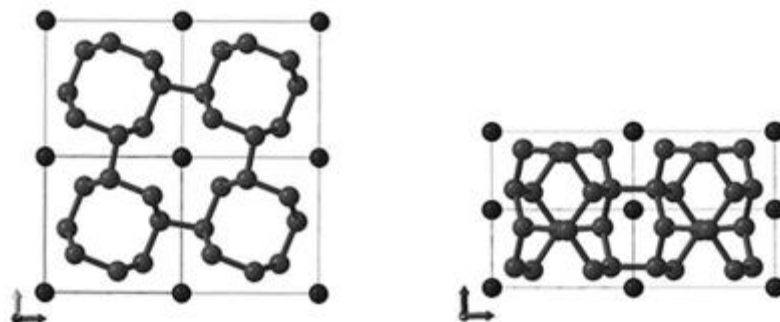
Таким образом, логика, взятая из биологии, помогла решить проблему, предсказав кристаллические структуры. Трудно сказать, что же в ней является «геном», потому что новые структуры могут отличаться от своих предшественников совсем другими параметрами. «Особи», наиболее приспособленные к условиям отбора, оставляют потомство, то есть алгоритм, который учится на их ошибках, максимизирует шансы на успех в следующей попытке. Система быстро находит вариант с наименьшей энергией и эффективно вычисляет ситуацию, когда структурная единица содержит десятки или даже первые сто атомов, тогда как предыдущие алгоритмы не справлялись и с десятью.

Как оценить ошибку, с которой работает USPEX? Вы можете взять задачу с заранее известным правильным ответом и решить ее 100 раз самостоятельно, используя алгоритм. Если правильный ответ будет получен в 99 случаях, то вероятность ошибки в вычислении составит 1%. Обычно правильные предсказания получаются с вероятностью 98-99%, когда количество атомов в элементарной ячейке составляет 40 штук.

Какие новые химические структуры были обнаружены? Оказывается, даже у «обычного» углерода есть свои секреты.

Углерод - очень интересный химический элемент, поскольку он образует обширный набор структур, от сверхтвердых диэлектриков до мягких полупроводников и даже сверхпроводников. К первым относятся алмаз и лонсдейл, ко вторым — графит, к третьим — некоторые фуллерены при низких температурах. Несмотря на большое разнообразие известных форм углерода, ученым удалось обнаружить принципиально

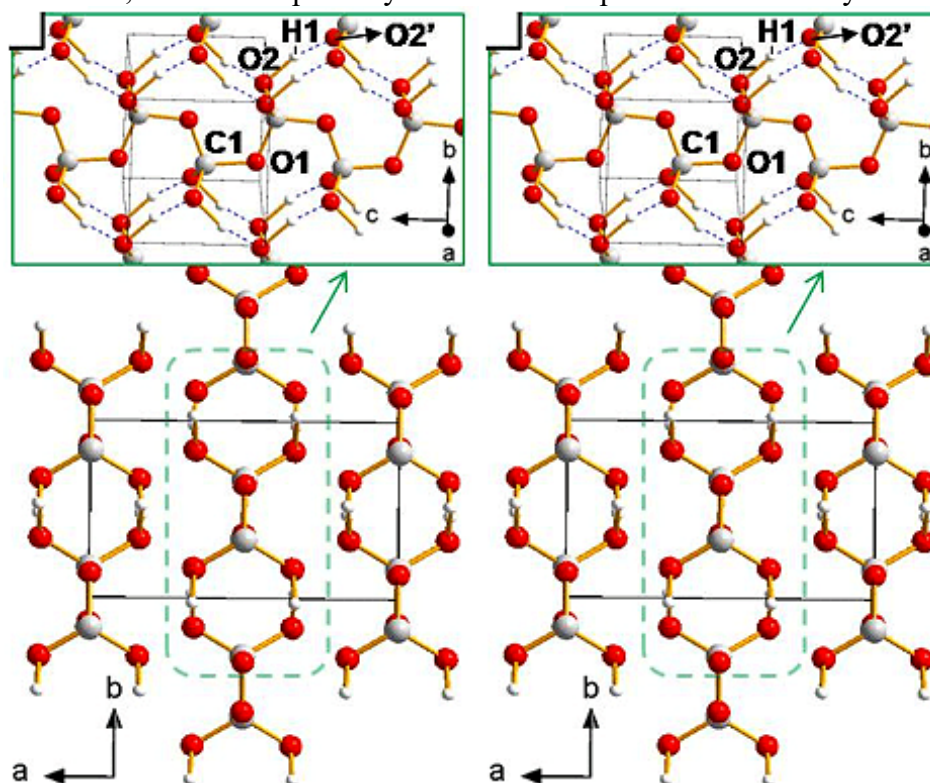
новую структуру: ранее не было известно, что углерод может образовывать комплексы «гость-хозяин».



Взаимодействие гость-хозяин проявляется, например, в комплексах, состоящих из молекул, которые связаны друг с другом нековалентными связями. Это означает, что определенный атом/молекула занимает определенное место в кристаллической решетке, но не образует ковалентной связи с окружающими соединениями.

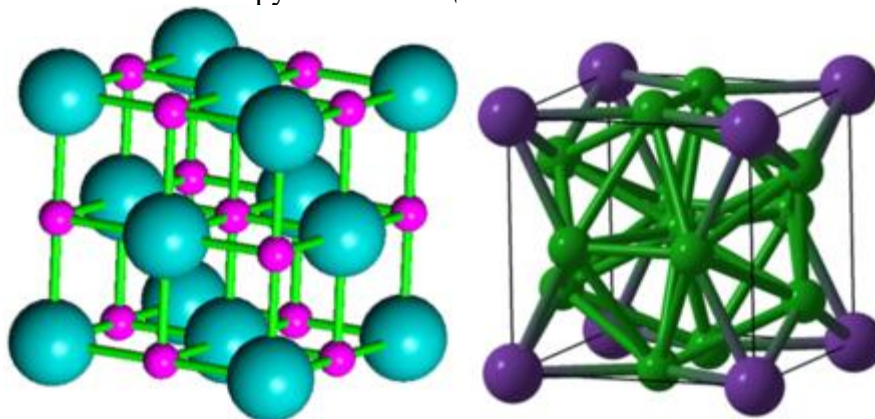
Новая химия под высоким давлением. Трудно предсказать, как вещества ведут себя под высоким давлением: большинство химических правил перестают работать, потому что эти условия сильно отличаются от тех, к которым мы привыкли. Тем не менее необходимо понимать это, если мы хотим знать, как устроена Вселенная.

Иногда ученые занимаясь фундаментальной наукой находят прямое применение полученным знаниям. Например, недавно была создана статья, описывающая результаты поиска всех стабильных соединений, полученных из углерода, водорода и кислорода при высоком давлении. Один из них был признан стабильным при очень низких давлениях, таких как 1 ГПа, и им оказалась углекислота H_2CO_3 . После изучения литературы по астрофизике было установлено, что спутники Ганимед и Каллисто (спутники Юпитера) состоят из воды и углекислого газа: из молекул, составляющих углекислоту. Таким образом, оказалось, что это открытие указывает на образование на них углекислоты.



Еще один пример необычной химии, который можно привести, относится к хорошо известной кухонной соли NaCl . Оказывается, если вы создадите в своей солонке давление 350 ГПа, то получите новые соединения. В 2013 году под руководством А. Р. Оганова оказалось, что если к NaCl приложить высокое давление, то некоторые

необычные соединения становятся стабильными — например, NaCl_7 и Na_3Cl . Интересно, что многие из обнаруженных веществ являются металлами.



Существует ещё множество открытий в этой области, которые, увы не смогут поместиться в эту статью, однако какой же можно сделать вывод из всего?

Выводы

Итак, мы смогли понять, что люди не смогут выполнить того объема работы, который возможен для искусственного интеллекта в наше время. Так же причина №2 весьма банальна: порой первоначальный вариант менее успешен, чем конечный, из-за чего ученые могут рассмотреть бессмысленным дальнейшее исследование и потерять, возможно, единственный удачный результат.

Как бы странно не звучало, но порой, нейронные сети мыслят креативнее ученых, придумывая с первого взгляда необычные формы и комбинации, о которых даже профессионал в той или иной сфере может не догадаться.

Вывод очевиден, создание новых нейронных сетей и ИИ, работающих в области науки положительно влияет на скорость и качество новых открытий, расширяя горизонты знаний всего человечества далеко за пределы нашей планеты.