

ТРИЗ (теория решения изобретательских задач) и Искусственный Интеллект – перспективы взаимодействия.

Тезисы к докладу на Конференции «Искусственный интеллект – вчера и послезавтра»

Нижний Новгород, 30 ноября – 1 декабря 2019г

Бахтурин Дмитрий Александрович, руководитель ТРИЗ-Центра Госкорпорации «Ростех», Вице-президент Международной Ассоциации ТРИЗ (МАТРИЗ), ТРИЗ-специалист 3 ур (МАТРИЗ).

Контакт rozmysl@mail.ru

Введение. Общая рамка.

Возможность соотнесения разных сущностей предполагает, что между ними есть какая-то общность («общий член» в противоречии, как сказал бы логик, или «рамка», как принято говорить в методологии). Для столь разнородных сущностей, таких как ТРИЗ и ИИ, найти таковую было нетривиальной задачей. Ведь ТРИЗ как продолжение диалектической традиции, казалось бы, максимально далеко отстоит от «вычислимости», за которой – опять же с точки зрения ТРИЗ – стоит перебор вариантов в противостоянии которому он, собственно, и возник.

Тем не менее, подобная рамка была найдена¹, и благодаря этому ранее зафиксированный «разрыв» из непреодолимой пропасти превратился в «растяжку», открывающую интересные перспективы для взаимоотображения, взаимовыражения и развития с опорой на противоположное.

В качестве такой общей рамки выступила концепция «правдоподобных рассуждений», восходящая к известной книге Д.Поля «Математика и правдоподобные рассуждения». И ТРИЗ, и AI есть разновидность подобных рассуждений, т.е., с одной стороны, некоторых определений, посылок, заключений и выводов, которые, с другой стороны, претендуют на то, что бы иметь отношение к «правде», к некоторой реальности, и выполнять по отношению к ней определенную функцию. Вспоминается известный (11й) тезис Маркса «о Фейербахе» - «Философы лишь различным способом пытались объяснить мир, а дело в том, чтобы его изменить». Вот эта интенция на включение «рассуждений» в практику, претензия на роль в практике и есть то, что объединяет указанные в заголовке – столь разнородные – интеллектуальные направления².

Взаимоотображение (1). AI «глазами» ТРИЗ

В современной ТРИЗ разработано большое количество инструментов высокой степени общности, которые позволяют сделать продуктивные интерпретации AI, когда и если мы сможем подвести его под понятие «Техническая система». Такая интерпретация сразу же позволяет применить к AI целый ряд инструментов и поставить вопросы:

– о структурно-динамическом характере AI, о рассмотрении его через призму «ТРИЗ-Системного оператора» (надсистема-система-подсистема в прошлом, настоящем и будущем),

-- о текущем этапе развития AI в логике S-образных кривых³

¹ В дискуссии с А.Н.Горбанем, за что ему отдельная благодарность

² Отметим, что в современных условиях сама эта растяжка может быть объектом отдельных исследований

³ Это оценка не по Гартнеру, а по отношению к уровню функционального развития, или уровню развитости функционала.

-- о следующем, ближайшем этапе, на который AI будет переходить, и о его характеристиках.

С другой стороны, применение разработанных в ТРИЗ моделей обобщенных, так называемых «полных» технических систем, позволяет уточнить место и возможные функции AI в техносфере, а также границы его «сферы компетенций». Наиболее развитая концепция технической системы в ТРИЗ предполагает выделение собственно «машины» (того, что оказывает непосредственное действие на внешний материальный объект, осуществляет собственно «действие»), и системы управления. Далее утверждается, что техническая система всегда «не полна», всегда есть внешняя, не входящая в ее устройство функция (позиция), которую традиционно выполняет человек – это постановка целей. В качестве гипотезы можно предположить что AI, по претензии, «вычерпывает» определенный слой именно этих «человеческих» функций, связанных с целеобразованием, в отличие от того, что выполняет классическая АСУ.

Взаимоотображение (2). ТРИЗ «глазами» AI

ТРИЗ с точки зрения подходов AI есть также, безусловно, «искусственный интеллект». Об этом говорит хотя бы сам метод получения первых ключевых инструментов ТРИЗ – Матрицы Альтшуллера и Таблицы приемов разрешения противоречий. Создавая их, основатель ТРИЗ Г.Альтшуллер более 15 лет изучал базу патентов на изобретения, благодаря чему вывел 40 приемов разрешения технических противоречий, содержащихся в сотнях тысяч патентов, которые были им проанализированы. Чем не обучение нейросетки? Только не чужой, из чипов, а своей собственной «подкорки»? Другим аспектом представления ТРИЗ из «позиции» AI является скачкообразный, качественный характер ТРИЗ-решений, в противопоставлении к «полному перебору» и количественно-ориентированному выбору экстремума, характерному для «вычислимости». Метафоричность, ассоцианизм, метод «по аналогии» - вот что характеризует решения на базе ТРИЗ⁴ с точки зрения AI. Да, безусловно, развитие AI как бы «выбивает почву» из-под ног подобных «нестрогих» подходов, имеет предельной претензией полный отказ от «творческого» (читай – трансцендентного) акта поиска решения, но пока, судя по всему, до реализации подобных претензий еще очень далеко. Формально все препятствия на этом пути можно свести к «проблеме размерности», но фактически вопрос не только в наличии вычислительных мощностей, но и в наличии онтологий для представления ситуаций, задач и решений⁵. Вся эта проблематика в полной мере развернулась при создании и использовании компьютерных версий ТРИЗ, программных продуктов, призванных обеспечить поддержку в решении изобретательских задач⁶.

ТРИЗ и «работа над ошибками».

Как следует из установочных тезисов Конференции, проблема «ошибок AI» выходит на передний план, становится одной из ведущих тем для исследований и разработок. В этом направлении также видится целый ряд перспективных направлений взаимодействия подходов AI и ТРИЗ-инструментов. С одной стороны, в ТРИЗ есть оригинальная версия способа работы с аварийными ситуациями, браком, дефектами – известная как «инверсионный анализ». Вкратце суть ее сводится к тому, что вместо

⁴ Здесь важно отметить, что ТРИЗ предстает как «ассоцианизм» именно с точки зрения AI. По отношению к обычным способам решения зада типа «мозгового штурма» ТРИЗ есть логика, методология, технология. Достаточно вспомнить название одной из книг Г.Альтшуллера: «Поиск новых идей – от озарения к технологии».

⁵ В ТРИЗ, в отличие от «обычного» инженерного подхода, принята 3х ступенчатая схема работы: «исходная ситуация», «идея задачи», «идея решения». Можно сказать, что ТРИЗ это «интеллектуальное айкидо», в отличие от «интеллектуального бокса».

⁶ Среди профессиональных инструментов выделяются два – «Изобретающая машина» и «Новатор», но их рассмотрение выходит за рамки настоящих тезисов.

поиска причин «аварии» ее пытаются «изобрести», т.е. вместо вопроса «почему это произошло?» спрашивается «как это можно было бы сделать?»⁷. Этот подход в целом весьма эффективен, но по отношению к системам AI обладает одним принципиальным недостатком – а именно, «детерминистической» логикой, ориентацией на выстраивание причинно-следственных цепочек, поиском необходимых связей. Соответственно, для систем AI, работающих в модели «явных правил», инверсионный анализ вполне может быть пригоден. Но как быть с тем направлением в AI, которое называют «коннективизмом»? Основанном не на «явном знании», а на прецедентах? На мнении (докса, δόξα), а не на истине (алетейя, ἀλήθεια)?⁸ Ведь, продолжая эту метафору, попытка исправлять ошибки «доксического» AI с помощью детерминистских методов будут очень похожи на труды и мучения Сократа, пытавшегося наставить софистов на путь Истины.

Для этой второй модели AI и работы с его ошибками в ТРИЗ перспективными могут быть следующие подходы и инструменты:

1. Матрица Альтшуллера и список приемов разрешения противоречий. Как указывалось выше, этот инструмент также носит «доксический» характер, основан и обоснован на прецедентах, на анализе многих тысяч реальных патентов из разных областей деятельности. Ведь что такое «ошибка AI» с точки зрения ТРИЗ? Противоречие, причем без выявленного «общего члена» - каковые и составляют оси Матрицы Альтшуллера. В качестве гипотезы для исследования можно предположить, что подходы к разрешению противоречий в ТРИЗ могут быть использованы для разрешения апорий и парадоксов AI.

2. Список «геометрических эффектов»⁹. «Геометрическая» трактовка данных в AI может не учитывать некоторые эффекты и сверхэффекты, возникающие при их преобразовании. В качестве гипотезы можно предположить, что список эффектов в ТРИЗ и методы работы с ними могут вскрыть природу нежелательных эффектов в работе с данными AI, как то, что имманентным образом искажает базовую логику обработки и интерпретации данных.

3. Наконец, логику «инверсионного анализа» в исследовательском залоге можно объединить с подходами к анализу и снятию «когнитивных искажений», известных в психологии и логике – для последующего переноса в сферу AI. Поскольку сегодня в сфере AI систематическая работа над ошибками, в тч. их классификация и интерпретация, только начинается, то решение изобретательской задачи – «как сделать эту ошибку?», «как реализовать данное когнитивное искажение силами AI?» - вполне можно начинать на уже имеющихся, в определенном смысле «культуросообразных» образцах ошибок интеллекта.

⁷ Далее эта техника обустроена специальными алгоритмами и методиками, которые – в дополнение к классическим методам – весьма эффективно работают на практике.

⁸ Метафора, предложенная А.Н.Горбанем

⁹ Одна из баз данных ТРИЗ, содержащая также перечни физических, химических и биологических эффектов