

# ИСТОРИЯ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Николаев Д.А.

Кафедра прикладной математики

Липецк – 2018

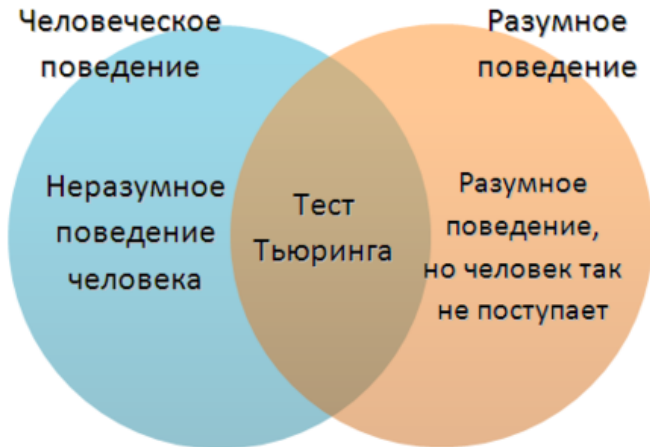
Под *искусственным интеллект*ом понимают науку, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются творческими.

## Этапы развития

1. Романтические этап (1940-1960)
2. Дедуктивный этап (1960-1990)
3. Индуктивный этап (1990-2005)
4. Онтологический этап (2005-2017)
5. Синтетический этап (2017-2030)

# Тест Тьюринга

*Тест Тьюринга* признает машину разумной, если она способна поддерживать взаимодействие с обычным человеком и тот не может понять, что имеет дело с машиной (1950).



Наиболее известным воплощением подхода к построению искусственного интеллекта является коннекционизм

1. однородность (элементы одинаковы и чрезвычайно просты, все определяется структурой связей)
2. надежность (элементы ненадежны, а система составленная из них весьма надежна)
3. эмерджентность (несводимость свойств системы к сумме свойств её компонентов)
4. голографичность (любой из компонентов системы может воспроизводить работу всей системы)
5. универсальность (способность решать широкий спектр различных задач)

Модели хорошо подходят для реализации в параллельных и распределенных вычислительных системах.

# Романтический этап (1940-1960)

Возникновение искусственного интеллекта (ИИ), тест Тьюринга, машинный перевод, распространение методов математики на биологию, социологию, психологию в рамках кибернетики. Построение простейших нейросетевых моделей, разработка методов структурной и параметрической идентификации с наличием или отсутствием обучающей выборки.

- 1943 Введение понятия искусственных нейронных сетей и их обучения (Маккалок-Питтс).
- 1949 Разработка алгоритма обучения без учителя.  
Разработка алгоритма обучения без учител (Хебб).
- 1949 Нейрокомпьютера. Разработка алгоритма обучения без учителей (Хебб).
- 1958 Разработка алгоритма обучения перцептрона с учителем для классификации (Розенблатт).
- 1960 Разработка алгоритма обучения перцептрона с учителем для прогнозирования и управления (Уидроу-Хоффа).

# Дедуктивный этап (1960-1990)

Снижение интереса к перцептронам из-за критики ведущими учеными. Перецение внимания на рекуррентные архитектуры для машинного зрения и ассоциативной памяти. Развитие численных методов и доказательной базы, способствовавших возвращению интереса к области в целом.

- 1969 Критика перцептрона из-за невозможности решения с его помощью некоторых важных задач (Мински).
- 1972 Разработка нейронных сетей, используемых в качестве ассоциативной памяти (Кохонен-Андерсон).
- 1982 Разработка рекуррентных нейронных сетей для распознавания образов (Литтл-Хопфилд).
- 1986 Разработка алгоритма обратного распространения (Вербос-Галушкин Румельхарт-Хинтон-Вильямс).
- 1989 Доказательство универсальной теоремы аппроксимации (Цыбенко-ХехтНильсен -Вильямс-Фунахаши).

# Индуктивный этап (1990-2005)

Разработка глубоких и гибридных архитектур, основанных на нечеткой логике и генетических алгоритмах. В задачах выявления скрытых образов и закономерностей в данных важное значение приобретает вопрос о представлении знаний, что приводит к использованию альтернативных алгебр.

- 1991 Сверточные нейросетевые модели с существенно глубокой архитектурой (ЛеКун)
- 1993 Гибридные нейро-нечеткие и нейро-генетические системы
- 1995 Нейросетевые модели в комплексных, гипер- и квази-комплексных алгебрах.
- 1999 Разработка нейросетевых моделей в неархимедовых алгебрах (Хренников).
- 2002 Эпоха дата майнинга, излечения скрытых образов и закономерностей из данных высокой размерности, и решения трудно формализуемых задач.

1. Большие данные выдвигают на первый план вопросы горизонтальной (а не вертикальной) масштабируемости. Достижение верхнего порога горизонтальной масштабируемости.
2. Встраивание в модель иерерхий понятий и признаков, анализ естественного языка.



1. Создание интеллектуальной системы сводится к выбору парадигм, может быть, их комбинации (гибридизации).
2. Использование гиперграфов и метаграфов для представления более сложных структур.
3. Интерес к гибридным и гетерогенным системам с ограничениями (способности обучаться, по типу входных данных, по типу решаемых задач, по масштабируемости, по кооперации между собой и с человеком)
4. Необходимость универсальных методов построения обучаемых (обучающихся) интеллектуальных систем без существенных ограничений, по аналогии с естественным разумом.

# КРИТИКА ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ

---

# Причины падения интереса

Достигнут предел полезности, определены границы возможностей, за пределами которых от Большинство задач, в которых ГО предложило принципиально новые решения (зрение, речь), в 2016-2017 годах вошло в зону уменьшающейся отдачи. Франсуа Чоле, Google, автор библиотеки Keras.

# 1. Зависимость от учителя

Невозможно обучить модель абстрактным связям или принципа работы посредством чёткого определения, не используя сотни тысяч или миллионы тренировочных примеров. Люди, даже в месячном возрасте, способны выучивать подобные абстракции, как через точное определение, так и через менее явные.

## 2. Приближенность решений

Хорошие аппроксимативные способности нейронных сетей являются следствием теоремы об универсальной аппроксимации действительно оказываются полезными во многих случаях. Но не в тех, когда требуется точное решение. Особенно это касается дискретных динамических систем, динамику которых нейронная сеть может моделировать лишь приближенно, основываясь на тех случаях, что фигурировали в обучающей выборке.

Также указанные замечания распространяются на продукционные цепочки логического вывода и лингвистические задачи, дискретные по своей природе, для которых невозможно даже составить выборку из-за неисчерпаемости вариантов. В тех случаях, когда испытывают нейросеть на ограниченной выборке, даже на ней ее предсказаниям часто нельзя доверять, а в тех случаях, когда результат адекватен на большей части задач выбранной области, в более редких случаях сохранить адекватность не удастся. В дополнение, внося шум в исходную

### 3. Неспособность к экстраполяции

Обобщение бывает двух видов – интерполяция между двумя известными примерами, и экстраполяция, требующая выхода за пространство известных тренировочных примеров. Чтобы нейросети хорошо обобщали, им обычно требуются большие объёмы данных, а тестовые данные должны походить на тренировочные — так, чтобы новые ответы были интерполяциями между старыми. В работе Крижевского, Сацкевера и Хинтона свёрточную нейросеть с девятью слоями, 60 млн параметров и 650 000 узлов тренировали на миллионе различных примеров, разбитых по тысяче категорий.

Такой подход с использованием грубой силы хорошо работал в сильно ограниченном мире ImageNet, где все стимулы можно развести по относительно небольшому набору категорий. Также он хорошо работает в стабильных областях вроде распознавания речи, в которых образцы неизменным образом сопоставляются с ограниченным набором категорий звуков речи — но по многим причинам ГО нельзя рассматривать (как это

## 4. Экспоненциальная неэффективность

Джофф Хинтон также беспокоился по поводу того, как ГО полагается на большое количество размеченных примеров, и выражал свои опасения в недавней работке по капсульным сетям со своими соавторами (Sabour et al., 2017), отмечая, что свёрточные нейросети (самая популярная архитектура ГО) может встретиться с «экспоненциальной неэффективностью», которая может стать причиной их гибели. Хороший кандидат — это трудности, с которыми сталкиваются свёрточные сети при обобщениях, связанных с новыми точками зрения (например, вид на объект с другой перспективы при визуальном распознавании образов). , но при других типичных трансформациях нам приходится выбирать между воспроизводством детекторов особенностей, расположенных на экспоненциально растущей решётке и увеличением размера размеченного тренировочного набора сходным экспоненциальным образом.

Универсальность не спасает.

## 5. Непереносимость знаний на другие данные

отя ГО способно на некоторые весьма удивительные вещи, важно понимать, что слово «глубинное» относится к его технической, архитектурной особенности (большому количеству скрытых слоёв, используемому в современных нейросетях, поскольку их предшественники использовали только один слой), а не к концептуальной (репрезентации, выстраиваемые в этих сетях, нельзя естественным образом применить к каким-либо абстрактным понятиям вроде «правосудия», «демократии» или «интервенции»).

Не существует рекуррентных процедур адаптации даже в хорошо подходящих задачах.



## 6. Непредставимость иерархической информации

Лингвист Ноам Хомский не удивился бы проблемам, описанным Джиа и Лияном. По сути большинство текущих языковых моделей на базе ГО представляют себе предложения в виде последовательностей слов, при том, что Хомский давно говорит, что язык имеет иерархическую структуру, в которой более крупные построения рекурсивно составляются из более мелких. Проседает работа с естественными и формальными языками. Даже рекурсивные глубокие сети могут неплохо обобщать данные при небольших различиях в тренировочных и проверочных данных, но когда обобщение требует систематических композиционных навыков, RNN с треском проваливаются»

Те же самые проблемы могут появляться и в других областях, например, при планировании или при управлении моторикой, в которых необходима работа со сложной иерархической структурой, в особенности, если система обречена сталкиваться с новыми ситуациями.

## 7. Непрозрачность работы

Относительная прозрачность нейросетей, работающих по принципу «чёрного ящика», является одной из основных тем, обсуждаемых в последние несколько лет. В текущей реализации ГО-системы обладают миллионами и даже миллиардами параметров, определяемых разработчиками не в виде каких-то человекочитаемых списков переменных, используемых каноническими программистами, а в виде их географии в рамках сложной сети (значение активности узла  $i$  в слое  $j$  в сетевом модуле  $k$ ). Хотя для визуализации вкладов отдельных узлов сети в сложных сетях были сделаны определённые шаги, большинство наблюдателей признают, что нейросети в целом остаются «чёрными ящиками».

Какое это имеет значение в долгосрочной перспективе, остаётся непонятным. Если системы сами по себе достаточно надёжны и самостоятельны, это может и не иметь значения; если важно использовать их в контекстах более сложных систем, для возможности поиска ошибок это может быть критичным.

## 8. Неинтегрируемость с существующими знаниями

Основным подходом к ГО является интерпретационный, то есть, самодостаточный и изолированный от остальных собранных знаний, которые в принципе могли бы быть полезными. Работа с ГО состоит в том, чтобы найти тренировочную базу данных — набор входных данных, связанный с соответствующими выходными данными — и научить сеть всему, что требуется для решения задачи, выучивая взаимосвязи между входными и выходными данными при помощи хитрых архитектурных вариантов, а также технологий для очистки и дополнения набора данных.

Таким образом, например, система, используемая в работе Лерера с коллегами (Lerer et al, 2016), пытающаяся обучиться физике падающих башен, не имеет предыдущих знаний по физике (кроме того, что определено в свёртке). Законы Ньютона не вшиваются в её программу, вместо этого система строит их приближение (в рамках определённых ограничений), изучая последствия на основе сырых пиксельных данных. Как я от-

## 9. Не выделяет причинно-следственные связи

о, что причинно-следственная связь не есть синоним корреляции — это трюизм, однако это различие серьёзно волнует специалистов в области ГО. Грубо говоря, ГО выучивает сложные корреляции между входными и выходными особенностями, но не строит репрезентации причинно-следственных связей. ГО может легко выучить корреляцию роста и объёма словарного запаса в популяции, но ему будет не так-то просто представить способ, которым эта корреляция выводится из роста и развития людей (дети, вырастая, разучивают всё больше слов, но это не значит, что они растут из-за разучивания слов). Причинно-следственная связь была основой в некоторых других подходах к созданию ИИ (Pearl, 2000), но, вероятно, ГО не приспособлено к таким задачам, и решать их в этой области практически никто не пытался.

## 10. Неадаптируемость к изменениям мира

Логика ГО такова, что она, скорее, будет хорошо работать в очень стабильных мирах — например, в настольной игре го, правила которой не меняются, и хуже работать в таких постоянно изменяющихся системах, как политика и экономика. Касаясь применения ГО для таких задач, как предсказание стоимости акций, есть большая вероятность, что этот подход повторит судьбу проекта Google Flu Trends, изначально отлично предсказывавшего эпидемиологические данные по поисковым запросам, а потом совершенно пропустившего такие вещи, как пик сезона гриппа 2013 года.

!!! Отсутствие инкрементальности обучения.

## 11. Аппроксимативность решений

Частично в результате других описанных в этой секции проблем, ГО-системы хорошо работают на большой части задач выбранной области, однако их легко обмануть.

Растущий набор работ демонстрирует эту уязвимость — от лингвистических примеров Джиа и Лияна, упомянутых выше, до широкого спектра демонстраций в области компьютерного зрения, когда ГО-системы путают изображения, раскрашенные в жёлтые аклеенные наклейками указатели парковки с теми, при том, что в остальных случаях результаты работы выбранных систем впечатляют.

Из недавних ошибок можно упомянуть слегка повреждённые знаки «стоп» из реального мира, которые система спутала со знаками ограничения скорости (Evtimovo проблемах британской полиции, чья система с трудом отличала изображения обнажённых тел от песчаных дюн.

Возможность обмануть ГО-систему, вероятно, впервые была упомянута в работе Церады (Szegedy et al. 2013). Целенаправ-

## 12. Ограниченная применимость

К примеру, к пониманию естественно построения соответствия между большим конечным набором предложений и большим, конечным набором других предложений. Это, скорее, разметка соответствия между потенциально бесконечным набором входящих предложений, и таким же по объёму набором смыслов, многие из которых могли ранее не встречаться. В такой задаче ГО становится квадратным колышком, который забивают в круглое отверстие — грубым приближением в случае, когда решение должно находиться где-то в другом месте.

В языке главное — это обобщение. Как только я услышу предложение вида «Джон пилкнул мяч Мэри», я могу сделать вывод, что грамматически корректно будет сказать, что «Джон пилкнул Мэри мяч, а Элиза пилкнула мяч Алеку». Точно так же, сделав вывод о том, что означает слово «пилкнуть», я смогу понять смысл будущих предложений, даже если я не слышал их ранее.

работе видно, что проблемы обобщения за пределы простран-

# Негативные эффекты обучения

Ошибки округления/отбрасывания. Переобучение, переподгонка — нежелательное явление, возникающее при решении задач обучения по прецедентам, когда вероятность ошибки обученного алгоритма на объектах тестовой выборки оказывается существенно выше, чем средняя ошибка на обучающей выборке. Переобучение возникает при использовании избыточно сложных моделей.

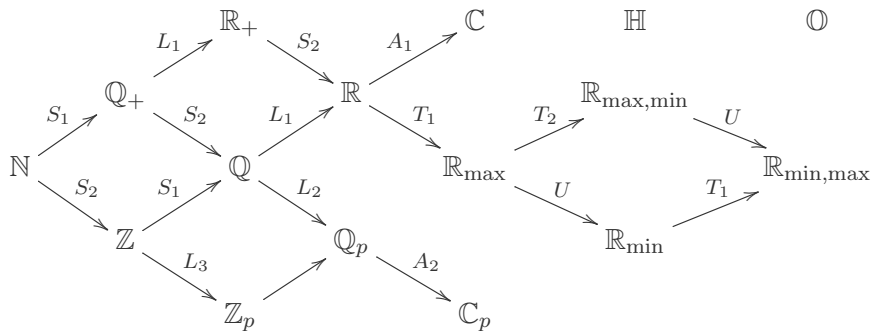
Недообучение — нежелательное явление, возникающее при решении задач обучения по прецедентам, когда алгоритм обучения не обеспечивает достаточно малой величины средней ошибки на обучающей выборке. Недообучение возникает при использовании недостаточно сложных моделей.



# ОСНОВЫ АБСТРАКТНОЙ АЛГЕБРЫ

---

# Расширение алгебраических структур



СПАСИБО ЗА ВНИМАНИЕ!