

Саморазвивающийся искусственный интеллект

Шнитко Артём Дмитриевич, учащийся 10-го класса;

Дёмин Сергей Александрович, учитель информатики

ГАУ КО ОО ШИЛИ (Калининградская обл.)

***Цель:** исследовать современное научное направление «искусственного интеллекта» и вытекающие из него подразделы.*

***Материал Исследования:** ресурсы интернета, книги об машинном обучении, Q-обучении.*

Понятие искусственного интеллекта.

Искусственный интеллект, что подразумевают люди под этим словом? Может быть несуществующего человека, способного думать? Вероятно, что многие люди далекие от темы информационных наук так бы и подумали. Свою историю, новое научное направление под названием «искусственный интеллект», берёт в середине 20 века. К этому веку уже было предложено множество предпосылок: психологи работали над теориями относительно работы человеческого разума, философы размышляли о процессе познания мира, были созданы первые компьютеры и, самое важное, Алан Тьюринг пишет статью, в которой задаётся вопросом «Может ли машина мыслить?». Интерпретировать его тест можно так: «Человек общается по переписке с одним компьютером и одним человеком (Рисунок 1). На основании ответов компьютера и человека он должен определить с кем разговаривает. Задачей компьютера же является – введение в заблуждение человека, заставив его сделать неверный выбор».

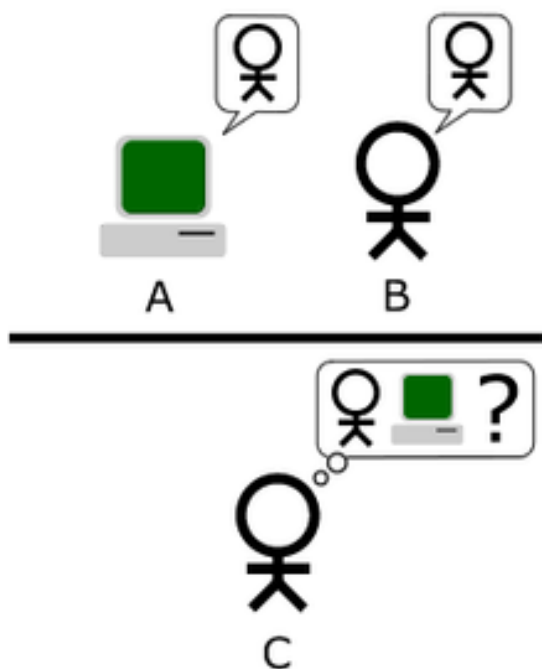


Рис. 1. Интерпретация теста Алана Тьюринга

Но даже этот прославленный тест имеет ряд своих весомых минусов: тест Алана Тьюринга направлен на антропоморфизм. То есть проверяется лишь способность искусственного интеллекта походить на человека, а не его разумность вообще. Тест не мог определить полный интеллект машины по таким причинам, что поведение человека иногда может не поддаваться разумному толкованию, а также, что некоторое разумное поведение не может быть присуще человеку.

Надо сказать, что упомянуто определение искусственного интеллекта впервые было Джоном Маккарти в 1956 году на конференции в Дартмутском университете. Но его определение не несло прямое понимание интеллекта у человека. Со слов Маккарти, ИИ-исследователи используют методы, которые не наблюдаются у людей, если это необходимо для решения конкретных проблем. Согласно этому, Маккарти говорил, что проблема заключается в том, что люди не могут в целом определить, какие процедуры называются интеллектуальными. Простыми словами, искусственный интеллект - всего лишь неживой алгоритм, неспособный самостоятельно думать. Можно

создать только что-то очень близкое к мыслительному процессу и обучению, как у людей.

Гипотеза «Сильного и слабого искусственного интеллекта».

Говоря об искусственном интеллекте, нельзя не сказать о ранее упомянутой в названии, гипотезе: теория сильного искусственного интеллекта предполагает, что компьютеры могут приобрести способность мыслить и осознавать себя как отдельную личность; теория слабого же искусственного интеллекта отвергает такое предположение. Термин «сильного искусственного интеллекта» был введён в 1980 году Джоном в мысленном эксперименте под названием «Китайская комната». В своём эксперименте, Джон Сёрл, ставил цель опровергнуть гипотезу «сильного» искусственного интеллекта и раскритиковать тест Тьюринга за его направленность на антропоморфизм. По той причине он выдвинул свои требования к созданию сильного ИИ:

- Планирование
- Обучение
- Сила воли
- Принятие решений в условиях неопределённости
- Сознание
- Мудрость
- Объединение всех этих способностей воедино для достижения общих целей и т.д.

На данный момент не существует искусственного интеллекта, обладающего всеми этими свойствами, поэтому данная вещь пока и остается гипотезой.

Q-Обучение или же один из способов машинного обучения.

И как же тогда создать похожий на человеческий разум искусственный интеллект? Один из вариантов ответа – это саморазвивающийся интеллект, который способен получая абсолютно новую для него информацию, без

учителя, понять её и использовать в дальнейшем, но даже в наш век информационных технологий не смогли создать настолько идеальный интеллект, потому что на деле имеется множество подводных камней и почти неразрешимых вопросов в этой теме. Близким ко всему этому является способ машинного обучения ИИ, его идея заключается в том, что интеллектуальная система в процессе работы самостоятельно получает знания, обучаясь без учителя и распознает образы во входном потоке информации. Говоря простым языком, ИИ подобно человеку, чем дольше существует и получает различной информации, больше узнаёт о мире и может отдать в ответ. У машинного обучения существует множество способов его реализации, но главной темой этой статьи станет способ Q-обучения. Q-обучение относится к виду обучения с подкреплением. Для обучения машины этот метод использует систему поощрений и наказаний. Если искусственный интеллект делает что-то хорошее, например проходит через какую-то нужную нам вещь или идёт в правильном направлении, мы награждаем его, а если делает что-то плохое, например, врежется в стену, то мы наказываем его. В начале искусственный интеллект ведёт себя хаотично, набираясь опыта, чтобы понять, что стоит делать, а что нет, подобно взрослению ребёнка, который в раннем детстве не знает как правильно себя вести, а когда подрастёт, то в осознанном возрасте делает всё по закону. Для этого используют данные алгоритмы (Рисунок 3).

1. **Initialization** (Инициализация):
 1. for each s and a do $Q[s, a] = \text{RND}$ // инициализируем функцию полезности Q
2. **Observe** (Наблюдение):
 1. $s' = s$ // Запомнить предыдущие состояния
 2. $a' = a$ // Запомнить предыдущие действия
 3. $s = \text{FROM_SENSOR}$ // Получить текущие состояния с сенсора
 4. $r = \text{FROM_SENSOR}$ // Получить вознаграждение за предыдущее действие
3. **Update** (Обновление полезности):
 1. $Q[s', a'] = Q[s', a'] + \text{LF} * (r + \text{DF} * \text{MAX}(Q, s) - Q[s', a'])$
4. **Decision** (Выбор действия):
 1. $a = \text{ARGMAX}(Q, s)$
 2. $\text{TO_ACTIVATOR} = a$
5. **Repeat**: GO TO 2

Рис. 3. Алгоритм Q-Обучения

Применить его, конечно, возможно только в ситуациях вида марковского процесса принятия решений (Рисунок 4), где происходит последовательное принятие решений с моделью, где решения в ситуациях частично случайны и частично контролируемы лицом, принимающим решения. Чтобы определить марковский процесс принятия решений, необходимо задать 4-кортеж $(S, A, P^*(*, *), R^*(*, *))$, где:

- S – конечное множество состояний,
- A – конечное множество действий,
- $P_a(s, s') = \Pr(s_{t+1} = s' \mid s_t = s, a_t = a)$ - вероятность, что действие a в состоянии s во время t приведет в состояние s' ко времени $t + 1$,
- $R_a(s, s')$ - вознаграждение, получаемое после перехода в состояние s' из состояния s с вероятностью перехода $P_a(s, s')$.

В современном мире эта спецификация используется во многих областях.

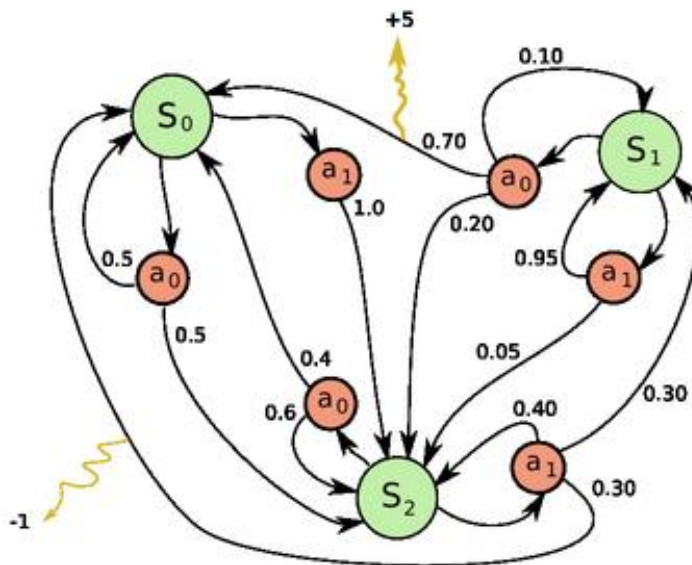


Рис. 4. Марковский процесс принятия решений с 3 состояниями и 2 действиями

Вывод.

Итак, в результате изучения темы искусственного интеллекта, машинного обучения и одного из способов машинного обучения, как Q-Обучение, можно

сделать вывод, что создать сильный искусственный интеллект трудная задача и пока никому не удалось её решить. Оправданий этому много, например тот факт, что нейросвязи мозга человека устроены другим способом, чем алгоритмы машины, использующие входящую информацию не по собственному желанию, а исключительно в определённом и узком круге возможностей, относительно человеческого мозга, записанных в этом самом алгоритме.

Литература:

1. Искусственный интеллект. — Текст : электронный // wikipedia.org : [сайт]. — URL: https://ru.wikipedia.org/wiki/Искусственный_интеллект (дата обращения: 21.02.2022).
2. Лонца, А. Алгоритмы обучения с подкреплением на Python / А. Лонца. — : ДМК Пресс, 2020. — 286 с. — Текст : непосредственный.
3. Чем отличается сильный искусственный интеллект от слабого. — Текст : электронный // neuronus.com : [сайт]. — URL: <https://neuronus.com/news-tech/1048-chem-otlichaetsya-silnyj-iskusstvennyj-intellekt-ot-slabogo.html> (дата обращения: 21.02.2022).
4. Марковский процесс принятия решения - Markov decision process. — Текст : электронный // star-wiki.ru : [сайт]. — URL: https://star-wiki.ru/wiki/Markov_decision_process (дата обращения: 21.02.2022).