

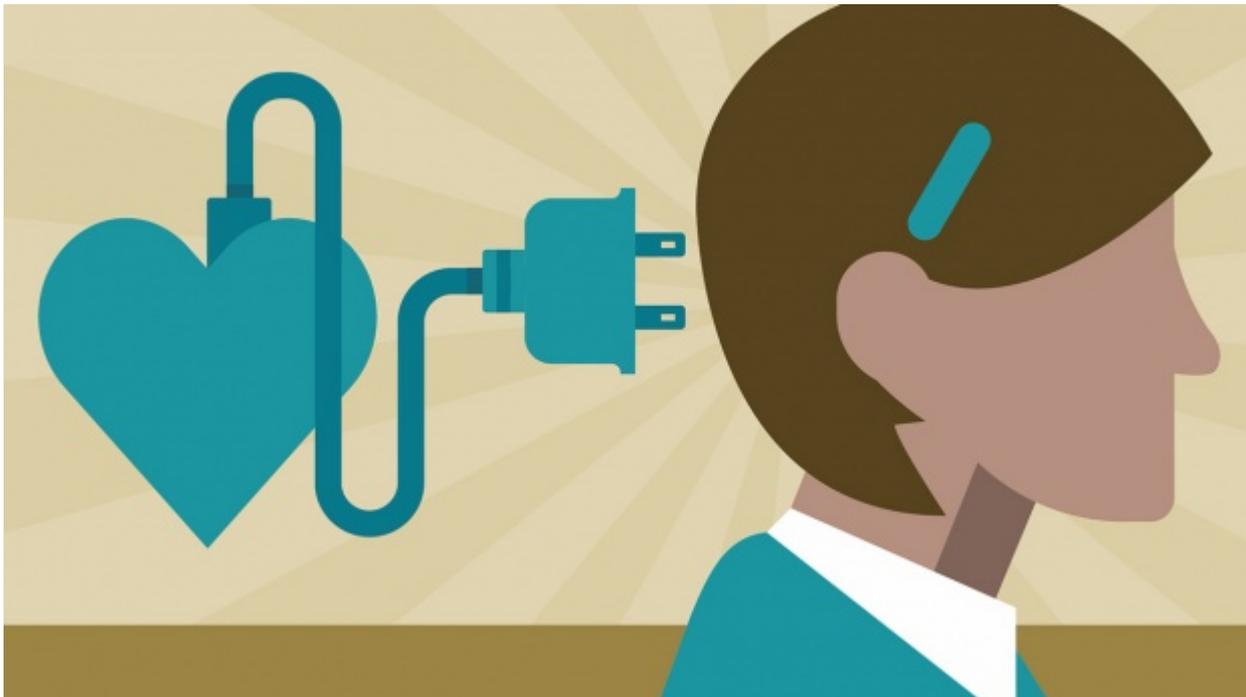


Социализация искусственного интеллекта: будут ли роботы чувствовать?

Юлия Каленкова

<https://te-st.org/2019/11/22/the-socialization-of-artificial-intelligence-will-robots-feel/>

Статья обновлена 01 июня 2023



Можете ли вы испытывать эмоции к алгоритмам? Не торопитесь с ответом: вспомните, как вы злитесь на навигатор в автомобиле или шутите с Алисой и Siri. Возможно, совсем скоро они смогут отвечать вам взаимностью. О том, как технологии научились распознавать эмоции и какое место в будущем уготовано эмпатичным роботам, журналист Теплицы Юлия Каленкова поговорила с Дмитрием Люсиным, старшим научным сотрудником Института психологии РАН, кандидатом педагогических наук.

- Расскажите, пожалуйста, о развитии эмоциональных технологий, эту область чаще называют «аффективные вычисления» (affective computing)?

- Я предпочитаю использовать сам англоязычный термин affective computing. Наша задача - научиться, во-первых, автоматически распознавать человеческие эмоции, а во-вторых, имитировать их. Это позволит, например, создавать компьютерных агентов вроде Алисы или Siri, которые будут вести себя эмоционально. И уже потом можно будет наладить взаимодействие между этими виртуальными «людьми» - компьютерными агентами, и реальными. Цель - сделать это взаимодействие естественным, чтобы обе стороны реагировали на эмоции друг друга.

Стоит отметить, что эта идея витает в воздухе с 90-х годов, то есть это не та ситуация, когда создали искусственный интеллект, а затем вдруг решили «прикрепить» к нему социальные

навыки или способность к распознаванию эмоций. Мы не просто берем искусственный интеллект и обучаем его, а именно изначально создаем системы, направленные именно на взаимодействие с человеком.

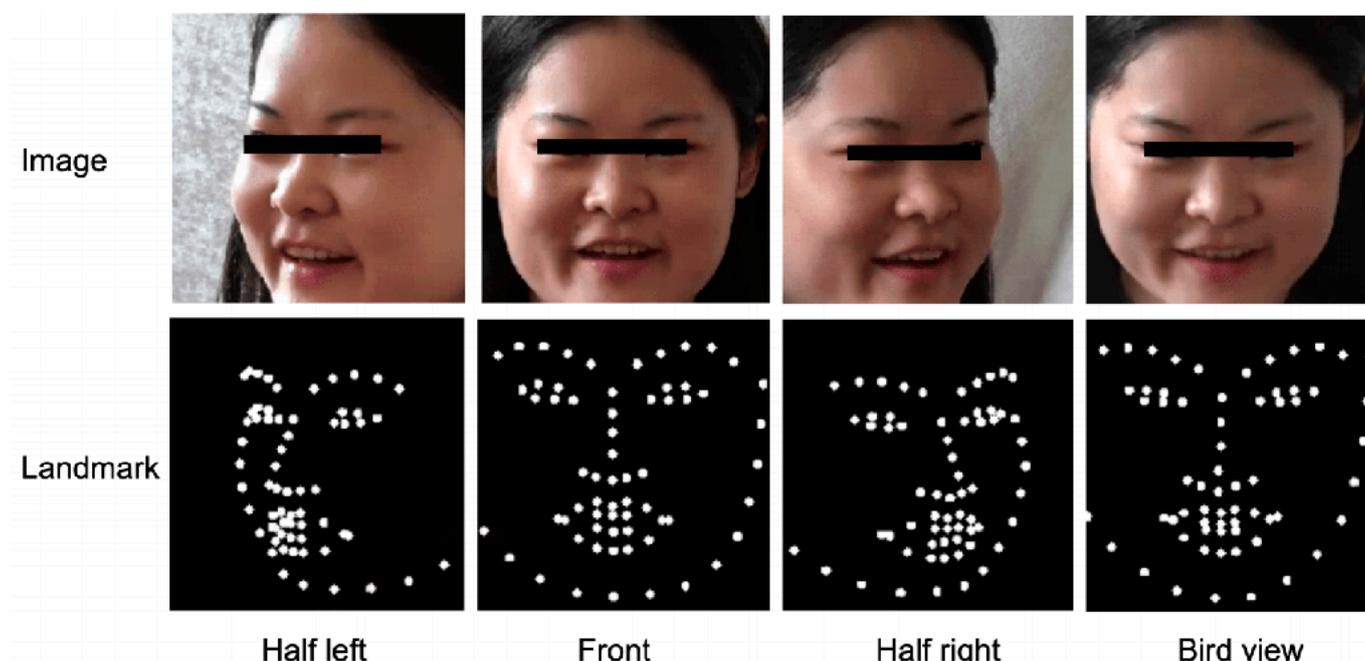
- А как выглядят эти технологии? Я представляю, как собирается база данных, допустим, по распознаванию лиц, а что насчет эмоций?

- Есть несколько подходов, и ни один из них не идеален. Проблема начинается уже с того, что даже люди по-разному понимают эмоции друг друга. Потому что мы получаем информацию об эмоциях из массы разных источников. Это мимика, жестикуляция, звучание голоса. Кроме эмоциональной экспрессии, имеет значение и смысл ситуации, в которой находится человек. Часто собранные из разных источников сведения противоречат друг другу. Таким образом, когда мы определяем эмоцию человека, весь вопрос в том, насколько точно можно угадать, что именно этот человек испытывает.

Кроме тех редких случаев, когда он однозначно обрадован или испуган, трудно определить конкретную реакцию. Порой ее не знает и сам человек. То есть в каком-то смысле мы конструируем эмоцию, и это уже, наверное, философский разговор.

И дальше все это переносится на affective computing. Изначально пытались делать по-простому, думали, что эмоция проявляется в экспрессии. И если мы правильно научим компьютер воспринимать и анализировать выражение лица, то он определит, какую эмоцию испытывает человек.

Такие попытки предпринимаются до сих пор, но без большого успеха – это всегда очень грубый результат. Но, в принципе, один из подходов в affective computing – анализировать выражение лица. Здесь мы имеем до некоторой степени априорные представления, какое выражение лица какой эмоции соответствует.



Датасет для распознавания эмоций на лице. Изображение: neurohive.io.

Современные технологии легко научились распознавать конфигурации лица. И если бы действительно каждая из конфигураций соответствовала определенной эмоции, задача была бы решена. Проблема в том, что нет прямого соотношения между конфигурациями лица и тем, какие эмоции человек испытывает.

Более продвинутый подход состоит не в том, чтобы просто анализировать лицо или звучание голоса, а в том, чтобы использовать машинное обучение. В этом случае мы собираем как можно больше информации о человеке. То есть мы берем видео, которое различает движение. Потом с помощью специальных программ по этому видео различаем движение лица, движение тела, звучание голоса и другие параметры. Видео дается аннотаторам – живым людям, которые определяют, какая эмоция в тот или иной момент переживается персонажем.

Нам нужен очень большой объем таких данных, и это отдельная сложная задача. И вот после того, как огромное количество видео оценило большое количество людей и определило, какие эмоции испытывают персонажи, мы «скармливаем» эту информацию нейросети. Она обучается распознавать эмоции не хуже самих аннотаторов.

За 21 год исследований Beyond Verbal Communications собрала данные о том, что от 35 до 40% эмоциональной информации содержится в вокальных интонациях. Продукт компании анализирует десятисекундные записи и извлекает из них данные об эмоциях, настроении, речевых привычках. Компания гордится базой в 2,3 млн образцов голоса на сорока языках мира. Стратегия проекта – на пересечении цифровой медицины и встраиваемых виртуальных ассистентов: определение эмоций по голосу снимает проблему множества слепых пятен, присущих классическим VPA (вроде Amazon Echo), ускоряя внедрение умных помощников.

Различные компании сейчас бьются над тем, чтобы создать большие дата-сеты (dataset, наборы данных), размеченные живыми людьми, аннотаторами. Остается только научить системы учиться, в частности, распознавать эмоции не хуже, чем живые люди. В целом это удается.

- А какие сложности есть? Как, например, учитывается контекст проживания эмоций?

- Возможный путь состоит в том, чтобы использовать не только информацию об экспрессиях, то есть о выражениях лица, движениях, звучании голоса, но и еще другую информацию – в том числе о ситуации, в которой находится человек. Когда распознают эмоции не автоматические системы, а мы, живые люди, учитывается не только эмоциональная экспрессия, но и логика ситуации.

Так, если человек находится в ситуации, где его кто-то обижает или дразнит, то каким бы ни было выражение лица и как бы ни звучал его голос, мы будем предполагать, что он испытывает что-то негативное, обижен или раздражен.

И если даже ничто не изменится в лице человека и его голос не дрогнет, мы предположим, что он скрывает свои отрицательные эмоции. Это и есть логика ситуации.

В идеале наши компьютерные системы должны уметь ее анализировать. Не просто ориентироваться на экспрессию, но и понимать контекст. К эмоциональной экспрессии мы пытаемся в affective computing добавлять анализ ситуации, анализ социальных взаимодействий, реакции других людей, помимо нашего основного персонажа. Это нам помогает более точно определить эмоцию, но процесс сложный и трудоемкий.

- А социально-культурный аспект как-то учитывается? Я верно понимаю, что не везде та же улыбка означает радость...

- Да, вы правы, и это одна из проблем в этой области. Действительно, эмоции в разных культурах и даже в разных социальных группах выражаются по-разному. То есть мы должны не просто брать аннотированные наборы данных, но и разделять их, учитывая разницу менталитета.

Тогда мы могли бы уже систему автоматического распознавания эмоций учить не просто каким-то усредненным образом, а именно культурному опыту. Не уверен, предпринимались ли такие попытки. Частично ситуацию спасает то, что обычно требуется определить эмоции определенных групп людей в определенных ситуациях. Этого достаточно для решения прикладных задач.

Например, нас интересует удовлетворенность обслуживанием клиентов банка, и мы хотим определить ее по выражениям лица или другим экспрессивным признакам. Берем dataset, состоящий из клиентов этого банка, которые приходят именно в этот банк для получения услуг. Тогда наш набор людей, «носителей» эмоции, наш культурный и социальный контексты достаточно жестко фиксируются. И тогда уже проще научить систему распознавать эмоции.

Но эмоции всех людей на Земле и в любых контекстах действительно распознавать проблематично. Мне пока неизвестно, чтобы это делалось. Каждый из нас ограничен той культурой, теми представлениями, которые у нас есть. И поэтому когда мы видим проявление эмоций у людей из другой культуры, мы легко можем ошибиться. Самые очевидные вещи мы, конечно, воспримем правильно, но тонкости – нет. Этим ограничены и искусственные системы.

- Если у нас все же будет искусственный интеллект, способный к эмпатии, наделенный эмоциями, какой у него будет статус? Можно его будет приравнивать, допустим, к гражданину?

- Если коротко, я не знаю. Такими вопросами должны заниматься специалисты по этике. Казалось бы, так как это искусственный интеллект, который не испытывает эмоции, а просто имитирует их, то и проблемы нет.

Можно привести аналогию с куклой. Если в ответ на голос хозяина она начинает плакать или улыбаться, мы понимаем, что это всего лишь движения механизма. Имеет ли смысл говорить о защите прав этой куклы?

Но есть другой аспект – как человек относится к этому роботу или искусственному существу. Уже минимум 20 лет в психологии изучают отношение людей к искусственным агентам.

У нас легко возникает эмпатия по отношению к ним. При этом роботы совершенно не обязаны быть даже человекоподобными. Некоторые шутят с Алисой или ругаются с навигатором в автомобиле. И встает вопрос: что делать потом с этими роботами, когда они выходят из строя? Какие этические проблемы связаны с самими роботами, а какие – с людьми, которые взаимодействуют с роботами? Если человек ведет себя жестоко по отношению к роботу: ломает, бьет его, как к этому относиться?

Если ребенок проявляет жестокость по отношению к кукле или взрослый по отношению к роботу, то, возможно, у них есть психологические проблемы, которые потом будут переноситься на живых людей.

И еще: допустим, человек жестоко относится к животным и получает наказание. Это санкция за то, что он доставил страдания живому существу или за нарушение общих правил поведения?

- Можем обратиться к Уголовному кодексу: жестокое обращение с животными в целях причинения ему боли наказывается штрафом в размере до восьмидесяти тысяч рублей. То же деяние, совершенное с публичной демонстрацией, уже влечет штраф в

размере от ста тысяч до трехсот тысяч рублей. Получается, общество наказывает не только за страдания существа, но и за пропаганду насилия.

- Это интересный аспект. Таким образом, возможно, в будущем будут наказывать за жестокое обращение с роботами не потому что тем больно или обидно, а потому что это нарушает какие-то представления о человеческом поведении.

Но правильно ли это? Вообще, существует этическая максима (правило поведения, которое можно принять как хорошее руководство к действию. – Прим. ред.): человек имеет право делать все до тех пор, пока это не наносит вреда другим. И если человек, допустим, не издевается над кошкой, а лишь делает вид – он несет за это ответственность?

- Вероятно, будет расценено как пропаганда насилия, провокация...

- Есть такое ощущение, что это что-то, что на самом деле некоторым людям хотелось бы делать, но нравственные и общественные нормы не позволяют. И тогда, видимо, если будут санкции за жестокое обращение с роботами, то суть будет не в том, что роботам больно. А в том, что если можно по отношению к роботам себя так вести, то, возможно, люди потом будут вести себя так же плохо по отношению к живым существам.

Однако эта логика небезупречная. Может быть, вымещая свою агрессию на роботах, человек потом не будет срывать ее на других людях. В любом случае мы не можем отдать на откуп роботу решение этических вопросов, с ними нам придется разбираться самостоятельно.

Еще по теме

- Кто исповедует датаизм, и как роботы стали священниками
- Этика и математика: зачем кодировать мораль
- Три неудобных вопроса о киберэтике