

А.В.Разин

Этические возможности искусственного интеллекта:

В настоящее время активно обсуждаются вопросы ответственности искусственного интеллекта, в частности — вопросы применимости такой категории к действиям искусственного интеллекта в беспилотных автомобилях, в общении искусственного интеллекта с живыми людьми (роботы сиделки, роботы продавца и т.д.).

Действительно, возникают вопросы, кто отвечает за действия искусственной интеллекта - разработчики программ, этические комиссии, разрешающие его использование, сам интеллект, если учесть возможность его научения, и каковы формы такой ответственности, могут ли сами эти формы быть заложены в программу совершенствования искусственного интеллекта.

Несмотря, что современные системы еще достаточно далеки по своим техническим возможностям от актуальности некоторых подобных вопросов, это не означает, что данные вопросы не следует задавать.

В предложенном докладе, мы собираемся рассмотреть современные подходы к решению проблемы сознания, обсудить как функции сознания могут быть смоделированы в искусственных системах.

Основные проблемы, которые я здесь вижу:

Это возможность создания условий для формирования феноменального опыта в системах искусственного интеллекта.

Необходимость создания для искусственных интеллектуальных систем каких-то аналогов тела.

Формирования у таких систем заинтересованности в собственном существовании.

формирование эмоциональной жизни, на основе которой могут классифицироваться события, может определяться их значимость, соответственно - могут выбираться приоритеты при принятии решений.

Возможность вынесения этических решений, приобретение подобного умения невозможно без процесса саморефлексии. А последняя, в свою очередь, невозможна без взаимных оценок, а значит — без коммуникации машин.

Я не хочу сказать, что это все невозможно. Как раз наоборот, даже локализованный для решения каких-то узких задач искусственный интеллект может быть подключен к сетям. Через них он может вступать в коммуникацию с другими искусственными интеллектами и выносить оценки. Первоначально - на основе заложенных в него этических принципов. Но это не означает принципиальную невозможность самостоятельного решения, ведь и человек первоначально ориентируется на те нормы, которые он узнает от других людей, от общества (культуры).

Но нормы могут совершенствоваться, может уточняться область их применения. Принципиально на это может быть способен и искусственный интеллект. Но для этого надо дать ему право на ошибку. Надо создать комплекс эмоциональных переживаний в связи с нарушением моральных норм и принципов, негативных реакций со стороны других участников коммуникации.

На данный момент мозг человека во всей его специфике не может быть смоделирован в искусственных системах. Прежде всего, это определяется тем, что нейроны головного мозга это клетки, которые могут производить специфические белки, т.е. они могут быть специализированы. В то время как полупроводники, используемые в системах искусственного интеллекта, такой специализацией не обладают. Но это не означает, что мы не сможем создать сложные альтернативные носители, способные к специализации и способные по скорости своей работы даже превосходить биологические носители.