**Опасно разумен**

Евгений Корниенко

Сознательные машины будут развиваться по экспоненте, потому что им не нужно тратить время на обучение тому, что уже известно предыдущим поколениям.

Я предвижу сознательное оружие.

Что произойдёт, если алгоритм искусственного разума станет общедоступен? Это ведь будет не полуживой компьютерный вирус, а хорошо думающая бездушная машина. Такая машина не может сочувствовать человеку или понять его чувства, так как "высшие" чувства (юмор, одиночество) происходят от самосохранения, которое является признаком жизни, а не разума.

Искусственный интеллект - это не сложная и дорогая атомная технология. Это просто программа, которую легко копировать. Что если этот алгоритм не требует огромных вычислений? Я подозреваю, что хорошо организованный ассоциативный мозг, такой же смышлёный, как мозг животных, может разместиться в настольном компьютере.

Каждый тиран и бандит пожелает такого сообщника.

Понимание опасности искусственного разума недостаточно для предотвращения этой опасности, особенно после того, когда эта технология постепенно заполонит весь мир. Войны и самоубийства тоже опасны. Но мы делаем это.

Наиболее реальная надежда избежать опасных последствий новой технологии - это сила денег. Государственный бюджет и личный бумажник могут преодолеть возможное военное применение искусственного разума, так как в мирной жизни эта технология ещё нужнее, чем в оборонной отрасли.

**Проблема искусственного разума скоро будет решена**. Но к этому невозможно морально подготовиться.

Сначала мы будем учить роботов тому, что считаем полезным. Затем они будут развиваться по экспоненте, потому что им не нужно тратить время на изучение того, что знают их предки. Знания будут накапливаться не в книгах (и учителях), а в мозгах этих роботов.

Невозможно, чтобы робот, который умнее меня в 10 раз, был моим слугой. Скорее мы станем слугами следующих поколений роботов, которые будут умножать свои знания в 10 раз ежегодно.

Пробуждение кибер сознания.

Не вызывает особых возражений идея о том, что сознание развилось как средство адаптации. Поэтому, изобретая адаптирующихся роботов, мы создаём разумные системы. Сначала они выглядят просто как адаптирующиеся программы. Так что даже наиболее прозорливые из нас не воспринимают их сознательными. Новорожденный ребёнок тоже не кажется слишком разумным.

Слово интеллект сейчас превратилось в атрибут рекламы. Мы уже готовы видеть элементы сознания в операционной системе. А животные, они только разумны (рационально себя ведут) или они чувствуют и осознают себя? Переход от машино-подобного интеллекта к сознанию происходит тогда, когда существо обнаруживает связь между своим действием и своим состоянием. Возникает ощущение. **Сознательная вещь осознаёт своё поведение. Она отличает своё поведение от других событий. Она отличает себя от других вещей.**

Итак, каждая более-менее сложная адаптирующаяся система, которую мы склонны считать разумной, является сознательной. Разумная вещь изначально ведёт себя разумно, а сознательная вещь сама изобретает разумное поведение - адаптируется. А мы субъективно оцениваем это поведение. Действительно ли оно разумно? Я могу представить систему, которая хорошо говорит, но полностью запрограммирована. С другой стороны, такое простое существо, как морской заяц - Aplysia, изобретает своё простое поведение без всякого предварительного программирования. Его мозг не знает заранее, в каких условиях придётся жить. Он имеет крошечное зерно настоящего сознания. **Сознание происходит из адаптации, а не из сложности.**

Животные адаптируются к условиям своей жизни не потому, что они осознают "необходимость выживания". Они заботятся о еде, об избежании боли. Выживание, рождение и смерь - это недоступные абстракции. Таракан убегает не потому, что он боится смерти, а потому, что он предвидит, что будет хуже. Естественный отбор, заботясь о выживании таракана, дал ему способность предвидения.

Эволюционно мозг призван помогать выживанию вида, а вовсе не думать. Необходимое для выживания предвидение постепенно развилось в обдумывание предстоящих событий и своих действий. Этот подарок природы для тех, кому нравится думать, вероятно, тоже помогает нам выживать. Понимая и высоко ценя эту способность, мы приступили к разработке машин, которые думают и осознают себя. Я полагаю, что такие машины потребуют гораздо меньших вычислительных ресурсов, чем человеческий мозг, так как сначала им не нужно будет заботиться о поддержании своей жизни.

Мы распространим такие машины всюду.

Неверно, что разработчик робота точно задаёт его поведение. Не каждый захочет и не каждый сможет следить за дальнейшей судьбой своего творения. Иногда изобретатель специально отпускает своё создание на волю, например, если это компьютерный вирус. И вообще, многие вещи, которые мы создаём, мы не можем полностью контролировать. Например, наших детей. Искусственный разум - это ещё один пример.

Первые едва разумные вещи будут иметь едва свободную волю. Но ведь мы претендуем на создание настоящего разума, ни в чём не уступающего человеческому. Поэтому "просто разумность" - это первый шаг к такой цели. В конце концов, задача создания искусственного разума, имеющего волю и стремления, будет решена.

Если вы позволите разумной машине принимать решения, то вы не сможете знать заранее, что это будут за решения. Вы также не можете быть уверены, что эти решения вас устроят. Разве вы всегда довольны даже своими решениями? Поэтому машина сможет осуществить свою волю в соответствии со своими суждениями, даже если вы этого не желаете.

Искусственный отбор роботов.

Для выращивания поколений разумных машин, обладающих только полезными признаками, можно применять искусственный отбор. Такой отбор в виде уничтожения неподходящих искусственных созданий успешно проверен, например, при выведении новых пород собак. Но как вы представляете себе уничтожение существа, которое умнее и чувствительнее вас, особенно если оно ценит себя и своё сознание?

Достаточно умные роботы смогут изобрести способ прохождения наших тестов, чтобы удовлетворить критериям искусственного отбора. Если разумное существо ценит себя, то оно будет пытаться выжить в селектирующей мясорубке, применяя все средства и хитрости. При этом роботы станут ещё опаснее, чем при достаточно свободном воспитании.

Я вижу только один путь, дающий надежду на безопасное сосуществование с роботами. Когда они достигнут и превзойдут человеческий уровень интеллекта и проявят стремление влиять на общественные процессы, нужно дать им право голоса и право быть избранными. Тогда нравственный отбор, который действует в демократическом обществе, будет также отбирать и представителей искусственного поколения в соответствии с желаниями избирателей. Пока этот процесс не станет стабильным, я надеюсь, мы сможем воздерживаться от создания слишком мощных и слишком разумных роботов. Хотя эта надежда логически не обоснована.

В конце концов, ничто живое не будет нужно новому поколению синтетической жизни. Дай Бог, если мы станем домашними питомцами или игрушками.