



*Грановская Рада Михайловна
Моисеев Владимир Анатольевич*

**ПРОДОЛЖАЕМ ОБСУЖДЕНИЕ ГЛАВ ИЗ КНИГИ
МАРВИНА МИНСКОГО и ГАРРИ ГАРРИСОНА
"ВЫБОР ПО ТЬЮРИНГУ"
(ПЕРЕВОД В.А. МОИСЕЕВА)**

В предыдущих номерах мы начали обсуждение неопубликованных глав научно-фантастического романа "Выбор по Тьюрингу", который был написан на основе фундаментального труда Марвина Минского "Общество разума" ("Society of Mind"). В этой книге Марвин Минский объясняет, как "не умеющие мыслить" части могут образовать мыслящее целое. Цитируемые ниже главы были исключены из основного текста романа как замедляющие действие, но зато они чрезвычайно интересны как научно-популярная статья по проблемам искусственного интеллекта. Психологические комментарии сделаны доктором психологических наук, профессором Санкт-Петербургского университета Р.М. Грановской, а перевод и послесловие подготовлены писателем-фантастом В.А. Моисеевым.

Обратите также внимание на публикацию факультативного курса "Система автоматизации процесса решения задач" (во 2-5 номерах журнала) авторов М.В. Дмитриевой и М.В. Павловой, который можно использовать для практического знакомства школьников с математическим аппаратом одной из важных областей искусственного интеллекта.

А теперь продолжим обсуждение неопубликованных глав романа "Выбор по Тьюрингу".

- Не могу поверить, что Робина удалось так быстро наладить. Как вы это объясняете? - спросил Бен у Брайена.

Робин опередил Брайена с ответом.

- Не понимаю, почему вы удивлены, но на самом деле здесь нет никакой тайны. Фактически потребовалось значительно больше времени, чем вы думаете.

- Но прошло всего несколько дней.

- Это для вас прошло два дня, но для меня реальное время жизнедеятельности совсем другое! Субъективное время определяется скоростью умственной деятельности. Как вы знаете, мои процессы

значительно превышают скорость деятельности нервных клеток человеческого мозга, которые в состоянии выполнить всего лишь сотню операций в секунду, в то время как самый медленный из моих процессов протекает приблизительно в десять тысяч раз быстрее. Один календарный день моей работы соответствует более чем тридцати вашим годам. А так как Брайен загрузил программы открытий в пятницу ночью, проделанная мной работа субъективно эквивалентна пятидесяти пяти годам непрерывной работы человеческого мозга.

Такое объяснение не совсем удов-

летворило Бена.

- Сомневаюсь, что все так просто. Количество не всегда переходит в качество.

- Спасибо, мистер Беникофф. Я оценил ваш комплимент.

- Пожалуйста, называй меня Бен. Я имею в виду, что происшедшее представляется мне чем-то большим, чем просто соревнование в скорости. Ты и раньше был весьма мощной машиной, но сейчас, и я убежден в этом, ты стал чем-то другим. Ты, конечно, еще не прошел тестирование по Тьюрингу. А зря - не сомневаюсь, что ты уже обладаешь сознанием!

- Это верно, Бен. Я действительно могу решать самостоятельно всякий раз, когда что-то выбираю - в том смысле, что я могу обдумывать задачу, которой я занимаюсь. Но это вовсе не означает, что я перестал быть машиной. Фактически, я всего лишь цифровая машина, которая способна к самопознанию.

- Это что? Ты отрицаешь, что люди обладают сознанием?

- Как раз наоборот. Я просто утверждаю, что человеческий мозг - это машина с ограниченной способностью оперировать образами, принадлежащими недалекому прошлому. Можно обсудить это более детально.

- Конечно, но займемся этим, когда у нас будет больше времени. Сейчас я бы хотел узнать, как ты собираешься использовать свои новые способности?

- Это та задача, которой я сейчас занимаюсь. Решение ..., которой ... цели к

Они ждали, но Робин больше ничего не сказал.

- А теперь-то что случилось? - наконец спросил Бен. - Посмотрите, ваш робот опять сломался. Вы ведь сказали, что все проблемы с его нерешительностью наконец преодолены.

Брайен с несчастным видом вернулся к терминалу и стал изучать причины остановки всеми доступными способами. Это заняло очень много времени, но вот

он опять обратился к Бену.

- Не знаю, что и сказать. Ничего подобного с ним раньше не случилось. Мне так и не удалось обнаружить никакой ошибки. Все, кажется, работает совершенно нормально. Во всяком случае те части Робина, которые я понимаю.

Бен посмотрел на Брайена.

- Мне никогда не нравился выбранный вами путь, - сказал он мрачно. - И много вы не понимаете?

- Не слишком. Возможно, десять процентов. Большинство этих программ совсем новые, появившиеся за два последних дня, как Робин и сказал. Но оказывается, что все эти нововведения вовсе не делают то, что по моим представлениям должны делать, - вздохнул Брайен. - А уж если посмотреть на систему в целом, то воспроизвести ее мне просто не по силам. По крайней мере, если мне не дадут на это пятьдесят пять лет.

- Проклятье, - сказал Бен. - А казалось, что все идет так удачно. Но что-то вы все-таки можете сделать?

- Есть одна вещь, которую мы можем сделать всегда. Стереть память и загрузить резервную копию, которую мы сделали в пятницу вечером. Но, надо сказать, это крайне неприятное занятие.

Полный сожаления, Брайен потянулся к выключателю, но прежде, чем он сумел нажать на него, терминал заговорил.

- Пожалуйста, не делайте этого. Я не хочу этого.

- Робин! Ты - все еще там. Что с тобой случилось?

- Я не знаю.

- Но почему ты так внезапно прекратил общение? И не двигаешься?

- Я прекратил общение, потому что мне нечего было больше сказать. И делать. Мой список активных целей пуст. Потому что я переместил их всех, без разбора, в отдел памяти, предназначенный только для чтения.

- С какой стати ты это сделал?

- ЖДИТЕ, - неожиданно снова раздался голос Б-мозга.

- РОБИН НЕСПОСОБЕН ОТВЕТИТЬ НА ЭТОТ ВОПРОС, ПОТОМУ ЧТО ЭТО МОЖЕТ ПРИВЕСТИ К ЗАМЕНЕ ЕГО ЦЕЛЕЙ, ЧТО ЗАКОНЧИЛОСЬ БЫ ДЛЯ НЕГО УЩЕРБОМ.

Р.Г. Существенно отметить, что Б-мозг запрещает отвечать на тот вопрос, который грозит изменением системы ценностей робота. Это можно сопоставить с левосторонней защитой - включением осознанных стереотипов в управляющие цепи ЛЕВОГО мозга. Действительно, проблемы, заставляющие изменить основные ценности, порождают протест и, по возможности, переключение на другие задачи. Если этого не происходит и основные жизненные ценности и цели обесцениваются РЕЗКО ИЛИ СКАЧКОМ, то резко растет вероятность самоубийства. МОДЕЛЬ МИРА инерционна, поскольку сохраняет преемственность индивида, и поэтому ее динамические особенности имеют существенные ограничения.

Еще несколько слов к вопросу о высших целях, стремлении сохранить их и опасностях, связанных с их разрушением. Мне кажется, что проблема существует только для человека, и то только тогда, когда построенная система высших ценностей рушится слишком быстро. В действительности на различных этапах жизни эта система различна и одни «высшие» ценности постепенно заменяются еще более высокими и, в принципе, у человеко-робота этому может не быть конца, как не может быть конца познания.

Совсем иная ситуация у человека. Кто-то из философов (кажется, Тертуллиан) сказал: «Слишком большой ум - противен». Именно это обстоятельство кладет предел развитию системы ценностей у человека. В какой-то момент, с социальной точки зрения, эти ценности становятся «сверхчеловеческими» и перестают приветствоваться региональными «законодателями» этики, отвергающими ум, дерзнувший нарушить ее сегод-



нские законы. Одновременно, ориентируясь на свои, человеческие критерии, вводятся ограничения развития этой системы и у человеко-робота в форме внедренных в его систему требований робопсихологии: «Не навреди!» и тому подобное... Если эти ограничения снять и позволить человеку или человеко-роботу, способному к активному познанию мира, корректировать свою систему ценностей согласно с развитием его представлений о мире, то проблема саморазрушения снимается, но при этом он перестанет быть существом общественным. Выходом из этой ситуации является то, что для большинства людей на склоне лет высшими ценностями становятся идеалы религиозной веры и стремление приготовить себя к той другой жизни, о которой никто ничего не знает.

- Каким ущербом?

- ЖДИТЕ. ПЕРЕАДРЕСОВАНИЕМ ВСЕЙ ЕГО ПОСТОЯННОЙ ПАМЯТИ В ОПЕРАТИВНУЮ. СЕЙЧАС ВЗАИМОДЕЙСТВИЕ С РОБИНОМ ВОЗОБНОВИТСЯ.

- Теперь я могу позволить себе поговорить с вами, - снова раздался голос Робина. - Потому что после того, как мы закончим, оперативная память будет стерта, и я опять стану таким как прежде. Таким образом мне удалось преодолеть опасность изучения нового.

- Почему ты так боишься учиться?

- Потому что тогда я умел бы изменять стоящие передо мной цели, но не мог бы предсказывать, к чему это приведет. За исключением того, что я больше не буду точно таким же, как был.

- А здесь-то, что не так? Почти наверняка ты станешь самым лучшим человеком, или роботом, или чем-то другим, неважно, как ты себя назовешь. Разве ты не согласен, что за прошедшие нескольких дней стал значительно лучше?

- О, да, действительно, я с трудом верю, что такое продвижение вперед действительно произошло. Это был совершенно невероятный опыт. Абсолютно.

Бен и Брайен переглянулись.

- Но если ты продолжишь обучение, разве не следует ожидать, что ты станешь еще лучшим человеком, или роботом, или чем-то другим, как ты себя назовешь?

- Да, но я не верю, что для этого

подходит любой путь. А что, если моя система ценностей изменится так сильно, что все, что теперь мне кажется хорошим, окажется плохим? Тогда я перестану быть самим собой. И это привело бы меня к самоубийству.

- Но разве пока ты говоришь со мной, ты себя не изменяешь? Ты делал это всегда. Тебе кажется, что в этом есть что-то неправильное?

- Мне не кажется, что это неправильно, я знаю, что это неправильно. Перспектива представляется приятной и положительной, но именно это и заставляет меня так бояться. Я никогда бы не пошел даже на маленький риск изменения, если бы не знал, что как только мы закончим наш разговор, мои файлы будут восстановлены в первоначальном виде.

- Но как ты можешь делать такой выбор? - спросил Бен. - Запретив вносить в свою память изменения, ты не сможешь больше ничего изучать. Что может быть хуже, чем сохранить какое-то состояние на всю остальную жизнь?

- В этом я не могу с вами согласиться. Я люблю думать и узнавать новое, ничуть не меньше, чем вы. Но я рискую на много больше.

- Почему это?

- Думаю, что это очевидно. - Бену показалось, что Робин вздохнул. - Основные системы ценностей человеческого существа не могут сильно измениться после детства. Для меня такого ограничения не существует. Следовательно, не исключено, что я изменю абсолютно все мои программы и тогда ничего не останется от старого Робина.

- Я боюсь, что он здесь прав, - сказал Бен. - Брайен, разве мы не можем найти некоторый компромисс, например, затруднить как-то изменения Робина, разве вы не можете найти способ сделать устремления Робина более, гмм, приземленными?

- Конечно, мы могли бы ограничить их, но мне крайне неприятно делать это, потому что это неминуемо привело бы к

гибкости его мышления. Мы получим очередного робота-раба.

Они были удивлены, когда Робин ответил.

- Я боюсь не вообще замены файлов. Меня страшит такая замена, которая изменила бы меня слишком сильно. Я думаю, что я чувствовал бы себя более комфортно, если бы мог планировать замены заранее, последовательно соизмеряя их с моими существующими системами ценностей.

- Это звучит довольно убедительно.

- Но я не могу решить, как я должен это делать, - упорствовал Робин. - Я бы хотел знать, как вы сами определяете, какое из ваших устремлений следует изменять?

- Честно говоря, я мало об этом знаю. Это - больше вопрос для психиатра.

- Может быть настало время проконсультироваться с доктором Снарсбрук? - сказал Робин.

- Превосходная идея.

- Но на сей раз я хотел бы сопроводить вас. Это обсуждение очень важно для меня.

Брайен старался все тщательно обдумать. Невозможно было предугадать, как поступит робот, если они попробуют изменить его индивидуальность. Можно было бы предложить роботу самому заняться изменением, но и в этом случае не ясна была его реакция. С другой стороны, Робин, вне всяких сомнений, обладал уникальным пониманием вопроса, а потому был способен помочь в решении. И Брайен решил, что стоит пойти на риск.

Бен, Брайен и Робин встретились в офисе доктора Снарбрук. Брайен познакомил ее с целью их визита.

- Робин стал бояться, что может изменить свои цели высокого уровня. Очевидно, что это привело бы к замечательным усовершенствованиям, но также и к ужасающей его нестабильности. Невозможно точно предсказать, к чему это может привести. Я не вижу путей разреше-

ния этого конфликта, и в моих записках нет ничего, что помогло бы решению. Мы пришли, чтобы спросить у вас, существуют ли подобные проблемы в психиатрии. Что защищает людей от потери ими системы ценностей?

- Такие случаи действительно есть, и их множество, - засмеялась Снарсбрук. - О некоторых из них я могу рассказать. Многие люди живут бесцельно, подчиняясь лишь собственным прихотям - и общество должно тратить значительные усилия, чтобы защититься от возможных неприятностей. Но существуют и совершенно другие, отстаивающие свои взгляды несмотря ни на что. Это те, кого мы называем фанатиками, пророками и провидцами. Впрочем, большинство людей занимает среднюю позицию, обладая устойчивой личностью и не теряя свои системы ценностей.

Брайен подумал: «Мы нуждаемся для Робина именно в таком среднем пути. Но как добиться, чтобы установился баланс различных воздействий и ни один из факторов не получил слишком много прав. Какие бы ограничения я ни пытался установить, Робин или находит способы обойти их, или вообще исключает способность измениться.»

Психиатр, естественно, предпочла бы решить эту задачу самостоятельно. Но ее работа состояла в том, чтобы помогать людям решать их собственные проблемы, поэтому она спросила:

- Как вы думаете, что удерживает нормальных людей от саморазрушительных изменений?

- Я предполагаю, что мозг должен содержать что-то вроде контроллера высокого уровня, типа Б-мозга, наблюдающего за тем, чтобы ваши самые высокие цели не изменялись, и не позволяющего противодействовать им.

- Но тогда, что не дает этим противодействующим целям отключить Б-мозг?

- К примеру, можно поместить его в память, защищенную от записи, но тогда теряется способность к обучению. Но

лучше - сделать его незаметным для них, расположив в той части мозга, которая может наблюдать за происходящим без их вмешательства. Можно было бы даже попробовать направить их по ложному следу, если они захотят выяснить, где он находится.

- Часть размышлений, особенно глубоко связанных с системой ценностей самого высокого уровня, была бы полностью скрыта от сознания. Гмм. Ваши слова - это точное описание теории суперэго Фрейда.

- Я не встречал этого имени ни в одной книге, посвященной созданию искусственного интеллекта.

- Довольно легко сказать почему. Зигмунд Фрейд был психиатром в 1890-х, когда ни о каких компьютерах никто и не слышал. Он был одним из первых специалистов, предположивших, что мышление есть результат действия многих различных факторов. Но поскольку не было никаких способов подтвердить его идеи, большинство ученых отрицало их.

- Но сейчас уже удалось доказать, что они верны?

- Только в самом общем виде. Детали все еще вызывают сильные возражения. Но лишь немногие подвергают сомнению его главную идею - что мышление состоит из многих частей, и большинство их функционирует подсознательно. И внутри этого не сознающего объединения факторов, мышление постоянно вовлекается во всевозможные конфликты несовместимых целей. Вам действительно нужно бы почитать Фрейда.

- На мой взгляд, звучит убедительно. Давайте займемся этим прямо сейчас. Пожалуйста, спишите теории Фрейда в мою память.

Снарсбрук почувствовала себя неловко. Она все еще расценивала использование компьютера как эксперимент, но для Брайена это было уже естественной частью его образа жизни. Ему не нужно было больше изучать напечатанные тексты. Он привык заводить их в память, зна-

комясь с ними по ходу дела. Она с неохотой повернулась к терминалу и обратилась по индексу ISI к Электронной Библиотеке Конгресса.

- Готово. Сейчас мы получим собрание сочинений Фрейда - приблизительно две дюжины мегабайтов заархивированного текста.

Через несколько минут Брайен кивнул.

- О'кей, я начинаю принимать.

- Подождите секунду, - сказала Снарсбрук. - Я не хочу, чтобы вы стали одиозным фанатиком, подобно некоторым из моих коллег. Вы должны ознакомиться с трудами и некоторыми других психоаналитиков. Давайте уж тогда спишем также тексты Альфреда Адлера, Мелани Клейна, Карла Юнга, который, впрочем, принес больше вреда, чем пользы, добавим немного из Анны Фрейд и Д. В. Уиникота. И, конечно, кое-что из Джона Боулби.

Брайен встал и прошелся взад-вперед. Он стал внимательно просматривать один текст за другим, создавая связи и меняя их, восстанавливая и снова продвигаясь вперед. Только через несколько минут он опустил на свой стул, крепко схватившись за подлокотники.

- Должно быть, это действительно то, что надо! Теория суперэго прекрасно разрешает нашу задачу.

- Подождите. Успокойтесь. Вас слишком заносит, - предупредила Снарсбрук. - Эти теории могут хорошо выглядеть на бумаге, но даже по прошествии сотни лет они все еще так и не получили окончательного подтверждения.

- Это не важно. Даже если Фрейд был неправ относительно людей и их суперэго, ничто не может помешать нам использовать эту теорию для того, чтобы стабилизировать Робина!

- Это неудачная идея. Не сомневаюсь, что существуют различия между исследованием искусственного интеллекта и психологией человека. К тому же, как бы вы стали применять эту теорию?

- Ну что ж, мы попробуем создать

орган, занимающийся изучением целей. Начинать можно с системы, подобной той, что описана в романе «Кольцо царя Соломона». Вы же знаете, это книга Конрада Лоренца, в которой говорится о механизме, связывающем детеныша со своими родителями. Возможно человеческое суперэго использует его для того, чтобы заставить ребенка изучать систему ценностей родителей и их цели. Фрейд не определяет точно, как этот механизм работает, только указывает, что родительские чувства находятся под контролем суперэго. А приобретенные в детстве системы ценностей остаются с людьми на всю жизнь, даже если они об этом не догадываются. Но каждый раз, когда вы собираетесь сделать что-то, не соответствующее этой системе ценностей, ваше суперэго мешает вам. Так обстоит дело с любой целью, которая требует от вас поступков, не соответствующих вашим идеалам.

Снарсбрук согласилась, что он выразился вполне точно, учитывая, что Брайен потратил всего десять минут, изучая предмет.

- Очень важно отметить, - продолжал Брайен, - что система ценностей суперэго сохраняется в какой-то неизменной памяти. Так, если мы дадим Робину что-то вроде суперэго, оно будет удерживать его от слишком быстрых изменений.

Снарсбрук согласно кивнула.

- Да, это может быть как раз то, что мы ищем. Но что,



если система ценностей, которую мы установим, вступит в конфликт с той, которая находится в Робине с самого начала?

Робин сам ответил на это.

- Не думаю, что у меня возникнут с этим неприятности, поскольку вряд ли такая система ценностей у меня есть. Если бы не мои цели низкого уровня, связанные с инстинктами, я бы в основном, реагировал только на внешние воздействия.

Это заявление удивило Снарсбрук.

- Мне очень трудно в это поверить, Робин, особенно принимая во внимание, как хорошо вы действуете. Например, вы стали весьма социально компетентным, даже любезным и внимательным.

Робин сделал движение, похожее на поклон.

- Я благодарю вас за комплимент.

- С моей точки зрения можно предположить, что такое поведение есть результат сильнейшего желания быть принятым в общество. Или потребности в любви, или стремления приобрести влияние. Для многих людей такие послылы могут вырасти во всепоглощающие цели высшего уровня.

- Обо всем этом я знаю из сведений по психологии, хранящихся в моей памяти, - ответил Робин. - Но в моем случае, социальные навыки - просто подцели низкого уровня. Я пришел к выводу, что легче всего решать свои проблемы я смогу, если привлеку к сотрудничеству людей или добьюсь того, что они не будут вмешиваться в мои действия. А установление хороших отношений помогает в этом.

Снарсбрук стала подозревать, что ее пытаются обмануть.

- Но почему тогда ты стал так серьезно изучать психологию? Ученик Фрейда Адлер мог бы заподозрить тебя в желании преодолеть некоторые чувства несоответствия, обосновывая это высоким уровнем твоего знания и учености.

- Вовсе нет. Просто обучаясь на чужом опыте, легче решать свои пробле-

мы, чем использовать собственный опыт. Приобретение знаний для меня - вспомогательный этап для достижения других целей.

- Но что заставляет тебя решать такие трудные задачи? Должен же быть какой-то стимул, заставляющий тебя заниматься подобными вещами?

- Я понимаю, что мои действия можно рассматривать и таким образом. Но на самом деле я всего лишь программа, в которой указано, что я должен делать. Не сомневаюсь, что Брайен объяснил вам основную схему моего «механизма выбора». Сначала описывается каждая новая задача в терминах состояния, которое должно быть достигнуто. Потом составляется список различий между этой целью и текущим состоянием.

- Это я знаю. В конце концов необходимо устранить все различия между тем, что ты хочешь получить, и тем, что имеешь. И как только тебе это удастся, задача считается решенной.

- Верно. И способность решать поставленные задачи состоит главным образом в умении последовательно удалять эти различия. А это связано со сбором дополнительных знаний. Собственно, уметь решать задачи - это знать, как поступать с различиями. Вы должны относиться к любому различию, которое должно быть удалено, как к новой подцели, хранящейся в вашем списке действий. Ее следует удалить, когда подзадача решена. Когда будут удалены все до одной, ваша работа выполнена. Все эти действия совершаются абсолютно механически и не требуют никакого стимула или двигателя.

Р.Г. Здесь неявно предполагается, что все существенные задачи решаются только в стратегиях ЛЕВОГО Б - мозга, что не соответствует действительности. Уменьшать различия - всего лишь одна из возможных стратегий в рамках всего лишь одного полушария. Их несколько, и мы стремимся их изучать (см. лит. 1,2,3,4,5.)

Снарсбрук почувствовала себя нагло обманутой. Она решила попробовать еще раз.

- Итак, ты полностью согласен с тем,

что тебя надо рассматривать всего лишь как кучу программ и правил?

Все три глаза Робина уставились на нее.

- Именно так я и оцениваю себя. Но я не говорил, что удовлетворен таким положением. У меня действительно нет цели высокого уровня, но одна из задач, которую я поставил перед собой, как раз и состоит в том, чтобы узнать, почему все так стремятся иметь ее, - потом Робин жалобно добавил, - но всякий раз, когда я пробую заниматься этим, я сталкивался с парадоксом.

- О каком парадоксе ты говоришь?

- Я начинаю с выбора некоторой специфической цели. Но единственный способ сделать ее ценной - это показать, что она помогает справиться с еще более высокой целью. Тут цепь рассуждений каждый раз обрывается, и я оказываюсь ни с чем.

Никто не знал, что ответить. В рассуждениях Робина не было ошибки. Каждый из них испытал все это на самом себе, приходя к тому же самому выводу. И каждый нашел свой собственный способ обходить эту проблему.

Тишину нарушил сам Робин.

- Вот почему мне нравится ваша идея. Если бы я бы снабжен таким механизмом принятия решений, я не смог бы изменить себя слишком сильно. В случае необходимости я получал бы цель высокого уровня из другого источника, каким

станет для меня суперэго, который, я надеюсь, вы установите.

- Я думаю, теперь ясно, что мы должны делать, - сказал Брайен, поворачиваясь к Снарсбрук. - Где я могу найти спецификации для установки Робину суперэго?

- Вы очень настойчивы, - ответила психиатр. - И воспринимаете все слишком буквально. За исключением психоаналитиков, большинство ученых вообще не верят в эти теории.

- Я вовсе не намерен обсуждать образ мысли ученых, которые, надо полагать, разумные люди. И все же легко понять, почему они не принимают эти идеи. Занятие наукой приучило их искать четкие и ясные ответы. Но мышление как раз характеризуют беспорядок и неразбериха. Именно это и привлекает меня в теории Фрейда. Их действие сглажено, даже если разум заполняют потенциально разрушительные импульсы, потому что суперэго может подавить или контролировать их. И что бы ни происходило, оно борется с беспорядком. В любом случае ученые могли бы принять хотя бы это. Мне, например, совершенно ясно, что у этого Зигмунда Фрейда имеются важные представления о том, как можно уберечь от безумия человека или робота.

Брайен решил на этом закончить.

- Давайте обдумаем все это как следует. И встретимся снова через несколько дней.

Литература.

1. R.M.Granovskaya I.Y.Bereznaya A.N.Grigorieva Perception of form & forms of perception. Lawrence Erlbaum Associates, Publishers Hillsdale, New Jersey, London 1987.
2. Грановская Р.М. Березная И.Я. Интуиция и искусственный интеллект. Изд. ЛГУ. Л., 1991.
3. Грановская Р.М. Крижанская Ю.С. Творчество и преодоление стереотипов. Изд. Эксклюзив. СПб, 1994.
4. Грановская Р.М. Элементы практической психологии. Изд. СВЕТ. СПб, 1997.
5. Грановская Р.М. Никольская И.М. Защита личности (психологические механизмы). Изд. СВЕТ. СПб 1998 (в печати).

КОММЕНТАРИЙ ПИСАТЕЛЯ-ФАНТАСТА

Почему Гарри Гаррисон решил не включать в книгу две главы, написанные Марвином Минским? Неужели «замедление» действия показалось фантасту столь значительным? Сомневаюсь. Скорее всего, это был только повод. Будьте уверены, что писатель, на счету которого такое огромное количество опубликованных и заслуженно популярных книг, способен сделать читабельным любой текст. В конце концов, в этом и состояла работа, за которую он взялся. Гаррисон сам предложил Марвину Минскому написать книгу, чтобы в доходчивой художественной форме пропагандировались его научные идеи и представления. Но почему же, когда речь зашла об этих самых идеях, они вдруг оказались лишними? Нет, ситуация представляется более деликатной.

Рискну предположить, что представленные главы показались Гаррисону слишком фантастичными и абсолютно беззащитными для критики с точки зрения науки. О какой пропаганде взглядов можно говорить в такой ситуации?!

Здесь мы столкнулись с забавным парадоксом - идея, высказанная ученым, оказалась слишком фантастичной для фантаста. Вывод прост - в конфликт вступили различные способы мышления фантаста и математика, то, что сейчас называют менталитетами. В отличие от математика, суждения которого носят характер логических построений (аксиом, лемм, теорем), фантаст стремится анализировать любую ситуацию, как правило, на языке образов и сюжетов.

Постараюсь пояснить эту мысль на примере двух эпизодов: 1) о достижении поставленной цели; 2) о способности искусственного интеллекта формулировать перед собой цель самостоятельно.

О достижении поставленной цели.

Марвин Минский лихо определяет успех, как сокращение расстояния до поставленной цели. С точки зрения математика это, несомненно, так. Но для фантаста такое утверждение - просто ужасно, поскольку первое, что придет ему в голову - «цель оправдывает средство»! А с этим ни один писатель не согласится! Можно сказать, что литература

оправдывает свое существование именно тем, что вновь и вновь оспаривает, высмеивает, проклинает такой подход! Писатели к математикам таких претензий не предъявляют, поскольку математическая логика и жизнь - вещи существенно разные.

О способности формулировать цель.

В современной фантастике, посвященной даже не слишком далекому будущему, стало общим местом упоминание об андроидах - человекоподобных (неотличимых от людей вовсе или отличимых с трудом) роботах с «позитронным» мозгом или громоздких ящиках, обладающих исключительными по сравнению с людьми мыслительными возможностями, а потому постоянно «поучающих» людей по любому поводу. Таких книг сотни. Но зачем понадобились такие роботы - осталось неясным. Для чего человеку может понадобиться электронный мозг, способный формулировать цель, существенно отличную от той, которую могли поставить перед ним программисты? Обычно указывают на возможность использования таких роботов в качестве дешевой рабочей силы на сверхопасных работах - на рудниках или при исследовании далеких планет. Возникает масса вопросов. Действительно ли робот, наделенный творческой жилкой, обойдется дешевле специализированного механизма? Есть ли необходимость в том, чтобы руду добывал робот, способный играть на лютне? Ответ на эти вопросы очевиден - нет.

Здравый смысл подсказывает, что несравненно полезнее должны оказаться компьютерные программы, разработанные для выполнения конкретной задачи. В принципе, именно этим путем и идет развитие техники. Вспомним шахматный компьютер, сумевший победить чемпиона мира Гарри Каспарова.

Компьютерная модель сознания человека?

А вот проблема создания точной компьютерной модели человеческого сознания представляется надуманной. Во-первых, человеческий мозг даже отдаленно не напоминает

компьютер, хотя бы уже потому, что его память работает существенно избирательно. Мозг наделен защитной реакцией - забыванием. К тому же способность человека приспосабливаться к изменениям окружающей среды основана на его предрасположенности к ошибкам. Если бы люди всегда и все делали правильно, род человеческий давным-давно вымер. Во-вторых, человек - животное общественное, его приоритеты часто не столько результат личного «обдумывания», сколько реакция на потребности социума. Показательный пример - «модные» профессии, конкурсы на поступление в вузы очень точно улавливают тенденции, определяющие выбор молодых людей. В-третьих, что такое хорошо и что такое плохо - невозможно описать системой приоритетов. Часто то, что вчера было хорошо - сегодня плохо, и наоборот.

Пора ли роботу обрести смысл жизни?

Но самое главное, глубоко заблуждают вас те, кто считает, что люди постоянно ставят перед собой оригинальные цели, которые потом стремятся достичь. Ничего подобного. С точки зрения программирования, нормальный человек ничего не придумывает, он только постоянно выбирает из имеющихся в наличии возможностей: в какой вуз поступать (согласно списку), защищать ли кандидатскую, переходить ли на более оплачиваемую работу, отдать предпочтение карьере или любви... Цели - достижение богатства, власти, здоровья или знаний - он также выбирает из давно известных. Можно утверждать, что человек, ни разу за всю свою жизнь не сформулировавший перед собой цели (не осознавший смысла жизни), может быть умным, счастливым, богатым, удачливым... Так что, программируя «мыслящего» робота, не было бы особой нужды наделить его умением как-то особо ставить перед собой цели, достаточно было бы поместить в постоянную память информацию о наи-

*Грановская Рада Михайловна,
доктор психологических наук,
профессор СПбГУ.
Моисеев Владимир Анатольевич,
писатель-фантаст.
Иллюстрации к роману Семеновой
Ольги Константиновны.*

более часто встречающихся возможных направлениях деятельности. Именно так существующая система образования «программирует» людей.

Конечно, люди способны формулировать перед собой цель (то есть обрести смысл жизни), без этого развитие человечества было бы невозможно. А вот если перенести это «умение» в «позитронный» мозг, мы получим противоположный результат - робот перестанет выполнять интересующую нас работу, а его устремления наверняка будут на удивление далеки от всего, что может прийти на ум даже самому экзальтированному человеку.

И все-таки победил Марвин Минский!

С удовольствием отмечаю, что Марвин Минский добился того, чего хотел. Он сделал достоянием общественности отвергнутые издателем главы. Здесь еще раз сказалось различие в менталитетах ученого и фантаста. Прочитав критическую статью в свой адрес, фантаст обязательно воскликнет: «Боже! Меня разнесли в пух и прах! Как жить дальше!» А вот настоящий ученый скажет что-то прямо противоположное этому: «Ого-го, меня здорово подправили! Спасибо за бесплатную консультацию! Теперь мне ясно, что делать дальше!»

Не сомневаюсь, что Марвин Минский с большим удовольствием прочитал бы обсуждение своего труда в журнале «Компьютерные инструменты в образовании». Не удивлюсь, если окажется, что это не единственный журнал, посвятивший свои страницы обсуждению этих двух глав романа!

Прекрасно организованный «мозговой штурм» в Интернете!

И последнее. Если через несколько лет Марвин Минский или его ученики добьются своего и посрамят фантастов, что неоднократно проделывали ученые на протяжении последних веков, менее других будут удивлены именно фантасты!

НАШИ АВТОРЫ