

Поздняков Сергей Николаевич

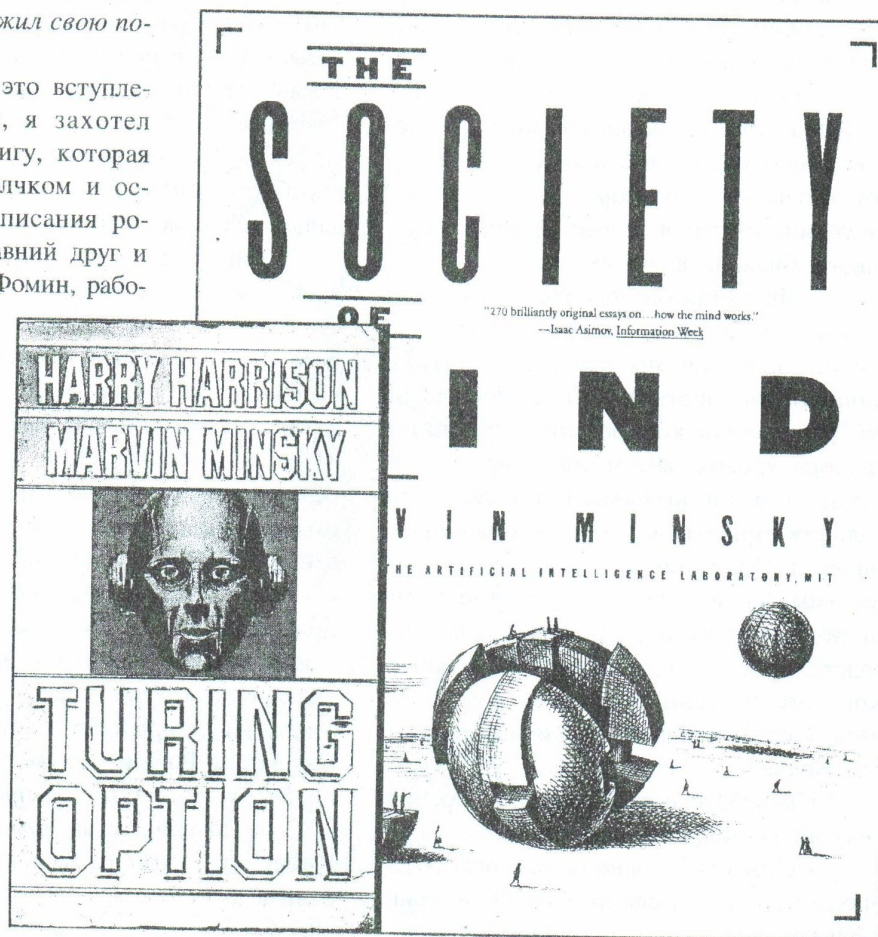
КРАТКОЕ ПОСОБИЕ ПО РОБОПСИХОЛОГИИ, СОСТАВЛЕННОЕ ПО КНИГЕ "ОБЩЕСТВО РАЗУМА" МАРВИНА МИНСКОГО

Вот и закончилось обсуждение неопубликованных глав из книги Марвина Минского и Гарри Гаррисона "Выбор по Тьюрингу". Хотелось бы напомнить читателю историю появления этой книги. Вот как рассказал об этом сам Марвин Минский: *"Гарри Гаррисон и я - давние друзья. Как-то он сказал мне, что ему очень понравились идеи, содержащиеся в моей книге «Общество разума». Но чтобы они смогли найти распространение среди большей аудитории, мне стоит написать более популярную версию в форме романа. Я ответил, что не обладаю необходимым для такой работы талантом. Гарри предложил свою помощь."*

Увидев это вступление впервые, я захотел прочитать книгу, которая послужила толчком и основой для написания романа. Мой давний друг и коллега С.В. Фомин, работающий сейчас в Массачусетском технологическом институте (там же, где и Минский), узнав о моем желании, прислал книгу "The Society of mind" в подарок. С работами Минского я уже был знаком по переводам статей

о машинном зрении. Пишет он всегда увлекательно. Благодаря живому языку и оригинальным идеям, его научные статьи всегда были интересны не только специалистам.

Марвин Минский мечтал написать книгу по робопсихологии и показать ее Айзеку Азимову, который когда-то написал "историю будущего", начавшуюся с книги "Я робот" (издана в серии "Библиотека приключений" в конце 60-х). Я перечитал несколько рассказов из этой книги и обнаружил, что в соответствии с хронологией книги активное развитие ро-



ботехники началось в 1998 году, да и все изложение идет с упоминанием 2003, 2015 годов, которые теперь уже не кажутся столь далекими.

Мне захотелось представить, что бы мы должны были знать о робопсихологии, если бы технический прогресс развивался темпами Айзека Азимова, а устройство роботов использовало те принципы и механизмы, которые, как считает Минский, определяют нашу интеллектуальную деятельность.

1. Что надо знать при приобретении робота.

Несомненно, в первую очередь необходимо выбрать позитронный мозг. Есть несколько параметров, на которые следует обратить внимание в первую очередь. Один из них - сложность позитронной сети, которая определяет потенциальный уровень интеллекта. Неспециалисты часто удивляются, каким образом могут образовываться связи между миллиардами элементов позитронного мозга и как при этом мозг остается таким быстродействующим.

Ошибка их в том, что они представляют себе километровые маршруты, по которым движется "сигнал". Природа этих ошибок (кстати, тех же, что и у нормально функционирующих роботов) в том, что человек постоянно упрощает мир, для того, чтобы с ним эффективно взаимодействовать. Поэтому мы хорошо представляем себе линейные функции (пропорциональные зависимости), но теряемся при встрече с экспоненциальными (быстро растущими) или логарифмическими (медленно растущими) функциями.

Каждый элемент мозга связан только со своими ближайшими соседями. Но даже если соседей у каждого элемента только 32, то уже 6 посредников достаточно, чтобы сигнал от данного элемента добрался до миллиарда других. Может быть, более наглядным будет задать себе вопрос, насколько длинной будет цепочка между двумя произвольными людьми на планете, если составлять ее из людей, хотя

бы раз пожимавших друг другу руки. Например, для автора статьи цепочка до Мао Цзе Дуна оказалась состоящей всего из двух посредников (М.И. Башмакова и Л.И. Брежнева)!

Другой важный признак заключается в том, насколько позитронный мозг является старым. Речь не идет о том, что новые модели более эффективны или что старый мозг менее надежен. Основной недостаток старого мозга в том, что у него исчерпаны возможности к адаптации. Так же, как у человека, в начале функционирования активизируются не все элементы (иначе робот оказался бы переполнен информацией, поступающей от органов чувств, на обработку которой были бы использованы все его ресурсы). Заложенные в мозг механизмы (их по аналогии с человеком можно назвать генетическими) постепенно активируют новые элементы, которые встречаются уже не с бешеным потоком информации от рецепторов, а в основном с вторичной информацией, поступающей от элементов мозга, обрабатывающих информацию органов чувств. Правильному пониманию поможет метафора левого и правого мозга, которой пользуются многие психологи для объяснения работы нашего мозга. На самом деле ситуация более сложная, потому что активация элементов происходит непрерывно, новые элементы обрабатывают информацию от всех уже имеющихся, поэтому правильно было бы говорить об А-мозге, Б-мозге, В-мозге и так далее. Конечно, очень много элементов памяти связано с обработкой сигналов внешнего мира, поэтому понятие А-мозга, или правого мозга, действительно представляет собой удобное для робопсихологии понятие.

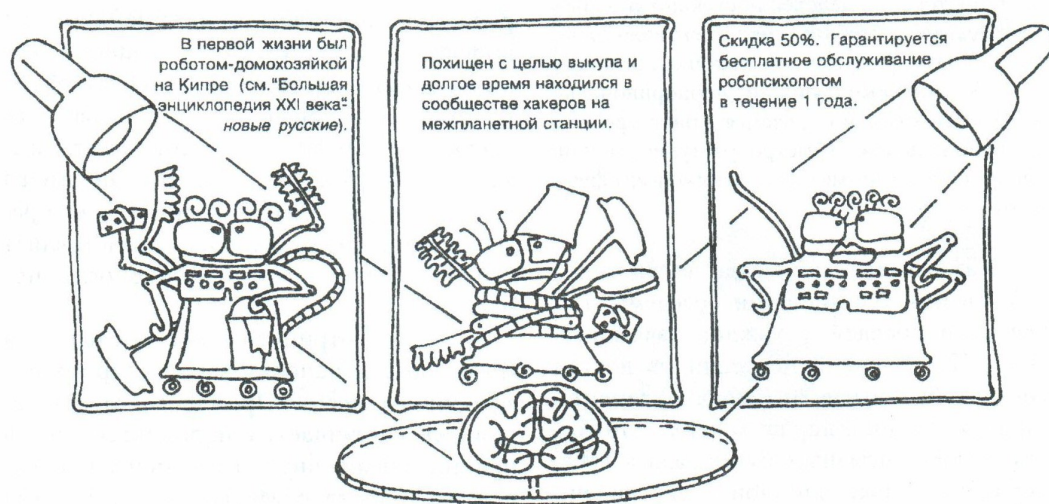
Какие отрицательные последствия может иметь использование старого позитронного мозга? Прежде всего категорически запрещается перемещать его в принципиально иную среду обитания, например, робота с Марса не следует отправлять на Меркурий, робота с селеновых рудников превращать в домохозяй-

ку. Последствия предсказать очень трудно (чему, собственно, и посвящалась серия рассказов Айзека Азимова "Я робот"). Наиболее вероятный результат - это полный выход мозга из строя. Если позитронный мозг не слишком стар или условия не слишком новые, то это проявится в замедлении реакций, некоторой "задумчивости" робота.

Если Вы приобретаете нового робота, то не следует отказываться от услуг специалистов в процессе адаптации робота (или по аналогии с человеком обучения и воспитания робота). Несомненно, что генетические механизмы любого робота допускают безопасную и эффективную адаптацию к любым условиям, но я уверен, что неспециалист плохо представляет себе, насколько просто он может заблокировать до 90% потенциала робота при неправильном воспитании. Не забывайте, что известные законы роботехники Айзека Азимова:

1. Первый закон: робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
2. Второй закон: робот должен повиноваться командам человека, если эти команды не противоречат Первому закону.
3. Третий закон: робот должен заботиться о своей безопасности, поскольку это не противоречит Первому и Второму законам.

- в определенной степени действительно играют такую большую роль в их поведении, что их можно было бы в совокупности назвать инстинктами. Достаточно нескольких грубых слов, противоречивых приказаний, упреков в неисполнительности, лишения возможности свободного принятия решений и самостоятельности, как Ваш робот станет похож на полуавтомат для уборки комнат и мытья посуды. Даже если после этого Вы вызовете робопсихолога, то понадобится время (а главное множество ресурсов позитронного мозга, которые в ином случае могли быть использованы более продуктивно) для того, чтобы Ваш робот стал снова проявлять признаки интеллекта. Неправильно представлять, что мозг могут отвезти на завод и перепрограммировать так, чтобы удалить нежелательные связи. Дело в том, что, хотя мы и говорим про комплексы неполноценности у человека, про психические преграды, барьеры, которые надо устранить, чтобы дать свободу для саморазвития, таких элементов, как преграды и барьеры в позитронном (и в нашем тоже) мозгу нет. Есть великое множество элементарных связей между громадным числом элементов (опять же сошлюсь на то, как неправильно мы воспринимаем большие числа), которые в результате и воспринимаются нами как



ПОЗИТРОННЫЙ МОЗГ. ИЗГОТОВЛЕН В 1998Г.

психологический барьер, страх и т.д.

Кстати об инстинктах. На самом деле никаких трех законов в психике (или памяти, как Вам удобнее) не заложено. Например, изначально робот не может проявлять “инстинкт” самосохранения, потому что ему совершенно не знакомы условия, в которых ему придется находиться. В то же время множество элементарных процессов, в совокупности обеспечивающих адаптацию робота, порождают такие его действия, которые можно назвать инстинктом самосохранения.

При правильном взаимодействии с роботом (раньше бы это назвали правильной эксплуатацией робота; теперь же, когда понятие интеллекта воспринимается нами шире, эксплуатация робота звучит так же дико, как работоторговля) барьеры играют положительную роль. В теории ропсихологии для описания механизмов образования барьеров используются термины “подавитель” (“suppressor”) и “цензор” (“sensor”). “Подавители” включаются, чтобы запретить действие непосредственно перед его выполнением (например, Вы явно перечислили роботу, что бы Вы не хотели, чтобы он делал). Это ведет к определенным потерям времени, поскольку робот ничего не сможет сделать, пока не найдет приемлемой альтернативы. Замедление реакций вызывает деятельность мозга, которую можно назвать созданием “цензоров”. “Цензоры” стремятся не допустить в дальнейшем ситуации, которая приводит к запрещаемому действию. “Цензоры” более эффективны, нежели “подавители”, но за это приходится платить определенную цену. Чем раньше мы хотим предсказать нежелаемое состояние, тем большее число достижений этого состояния мы должны предусмотреть. Каждый робот обладает базой, в которой хранятся миллионы “цензоров”.

2. Зрение робота.

Разумеется, это не самое главное в его устройстве, но оно требует наибольшего числа ресурсов. Как правило, об общем интеллекте робота можно судить

по тому, насколько хорошо у него функционирует зрение. Сейчас мне опять придется разрушить некоторые стереотипы.

Разрушить в мозгу - как человеческом, так и позитронном, ничего нельзя (если не говорить о физическом разрушении), можно только создать новые связи, которые могут оказаться более удобными для адаптации к среде, в которой Вы находитесь (для Вас это среда, населенная роботами).

Так вот, зрение и мышление имеют в своей структуре много общего. Один из механизмов - это системы или массивы фреймов. Попробую объяснить, как они функционируют на примере Вашего зрения. Замечали ли Вы, как легко ориентируетесь в незнакомой комнате, в которую только что вошли? Вы садитесь на стул, смотрите в окно, ищете на столе ручку и т.д. Казалось бы, что Вы осмотрели эту комнату “от и до”. Однако, если Вас впоследствии спросят, на каком стуле Вы сидели, была ли форточка на окне, какой ручкой писали, к своему удивлению Вы не сможете ответить. Даже если Вы и вспомните что-нибудь, может оказаться, что это неверно и Вы вспомнили что-то из совершенно другой комнаты. Теперь Вам легче будет поверить в существование механизма фреймов. Оказывается, двигаясь по незнакомой комнате, Вы видите не столько то, что в ней находится, сколько то, что хранится в вашем мозгу. Точнее можно сказать, что Вы не смотрите на вещи, а пользуетесь ими. А для этого не нужно их видеть, надо только приспособить особенности реальных вещей к тем операциям, которые с ними совершаются. Для такого приспособления вещи не нужно ее “рассматривать”. Нужно только некоторое время (пока не закончите пользоваться ею) удерживать в памяти особенности этой вещи (для чего и существует кратковременная память). В теории это выглядит так: терминалы фрейма заполняются признаками конкретной вещи, необходимыми, чтобы с ней можно было взаимодействовать. После окончания взаимодействия терминалы освобождаются от

этих признаков. К ним остаются привязанными некоторые признаки, хранящиеся в Вашей памяти, которые по другим причинам были “рассмотрены” и поэтому запомнились.

Теперь перейдем от понятия фрейма к системе фреймов.

Если понятие фрейма исходит из ранних работ психологов, таких как Барлетт, то понятие системы или массива фреймов появилось в работах по искусственному интеллекту и принадлежит Минскому. Для того, чтобы понять эту идею, попробуйте поставить на себе мысленный эксперимент. Пройдитесь по комнате, оглядываясь по сторонам. Изменилось ли что-нибудь в комнате от этого? Вы скажете - нет. Но с разных точек зрения предметы выглядят абсолютно по-разному! Каким образом непрерывно меняющееся изображение в вашем глазу преобразовалось в представление о неизменности окружающих предметов?

Представьте себе и более ужасную ситуацию - Вы движетесь по комнате, но видите предметы с той же точки зрения, что в начале пути. Если Вы не сойдете с ума, то подумаете, что предметы в комнате стали перемещаться, а комната поворачиваться.

Это возможно благодаря компенсаторному механизму, который меняет фреймы, связывая их все с тем же понятием комнаты. Таким образом, движение робота по комнате сопровождается параллельной сменой фреймов, образующих массив фреймов, связанных с понятием комнаты.

Возможность создания нормально функционирующего робота связана с унификацией элементной базы. Вообще, создание позитронного мозга с некоторой натяжкой можно представлять как процесс деления одной клетки. Одинаковые по структуре элементы в дальнейшем начинают принимать на себя различные функции в зависимости от того, когда они активированы и в каком соседстве находятся. Аналогично и элементы системы управления роботом тоже имеют общую ба-

зовую структуру. Так, фреймовый механизм лежит в основе понимания роботом языка. Для того, чтобы лучше понять, как это происходит, проведите на себе очередной мысленный эксперимент. Представьте, как за обедом Вы читаете газету. За небольшое время (явно недостаточное для того, чтобы прочесть весь текст) Вы оказываетесь в курсе всех событий. Это происходит потому, что на самом деле Вы не читаете, а просматриваете текст. Что значит, просматриваете? Это значит, что Вы уже ожидаете узнать какие-то новости на какие-то темы и вызываете соответствующий фрейм, в котором терминалы заполняются новыми фактами, характерными для этого фрейма (этой темы). Таким образом, и здесь фреймы используются для того, чтобы овладеть ситуацией, то есть за небольшое время суметь сделать правильные выводы. Другой пример использования фреймов. Прочитайте предложение “глокая куздра штеко будланула бокра и кудрячит бокренка”. Не правда ли, что из кажущегося бессмысленным по использованным словам предложения Вы сумели извлечь довольно много информации? Примерно, Вы объяснили себе смысл предложения так: “какое-то животное женского рода, которое имеет странное название “куздра”, сделала что-то, по-видимому, нехорошее, животному мужского рода с не менее странным названием “бокр”, после чего принялась делать что-то другое, но скорее всего тоже нехорошее, с его детенышем, чем и занимается до сих пор”. Теперь Вам будет нетрудно согласиться, что мы имеем дело с фреймом некоторой ситуации, терминалы которого заполнены странными словами, возможно потому, что фрейм применен к обстоятельствам, с которыми мы не встречались ранее.

3. Память робота.

Компьютерный век оставил нам такие обороты, как “добавить памяти”, “купить дополнительный диск”. Для робота эти обороты теряют свой смысл, так как представление и хранение информации в

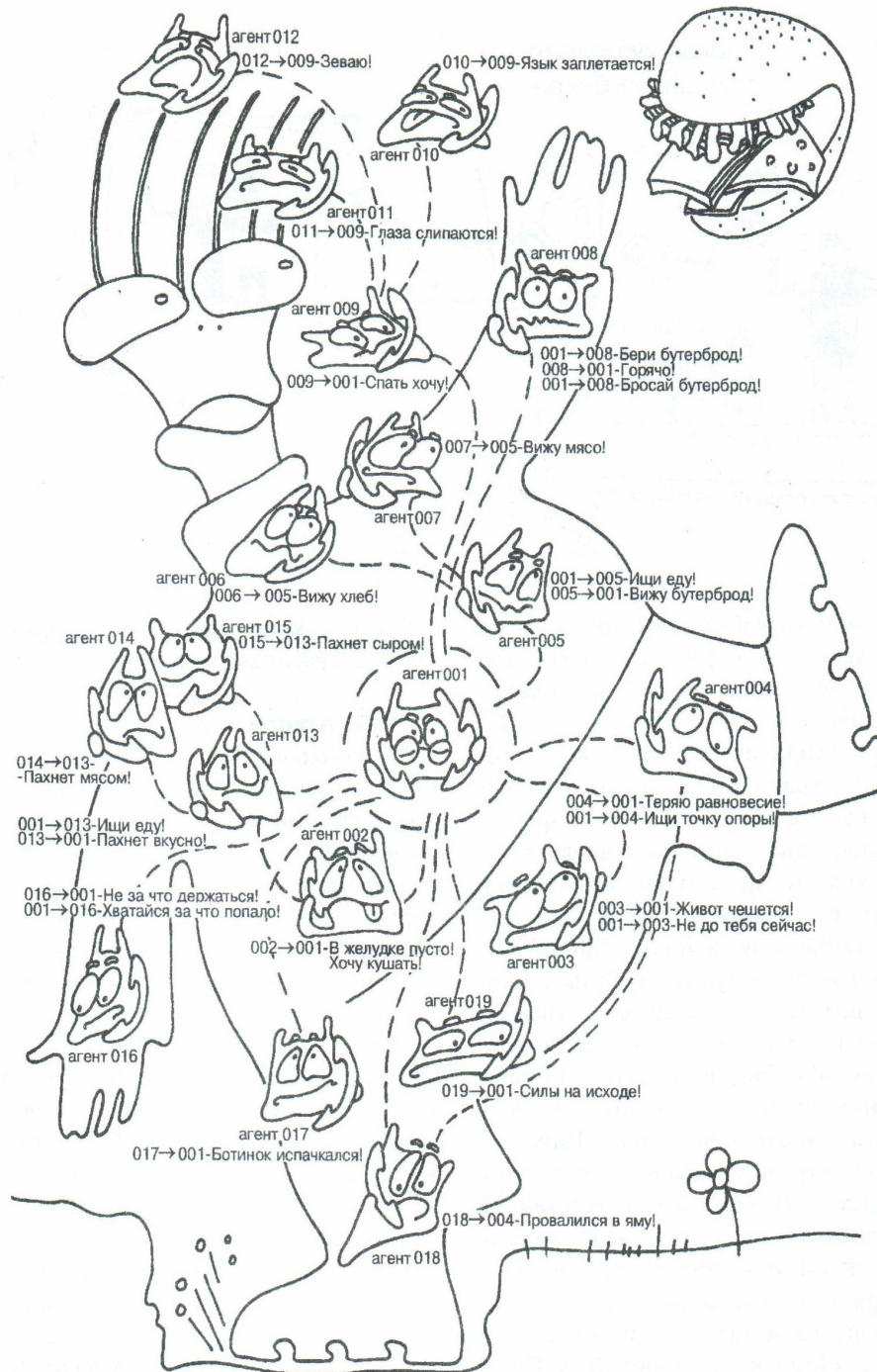


позитронном мозгу существенно отличаются от хранения информации в компьютере. Вообще, для робота правильнее употреблять термин "хранение знаний", поскольку только специализированные роботы могут хранить информацию в привычном для нас смысле (то есть в форме, по которой однозначно восстанавливается ее исходный вид). В роботсихологии используется термин "линия знания" (knowledge-line или K-line). "Линии знаний" создаются в процессе функционирования робота. (У людей мы называем их "удачными мыслями" или "плодотворными идеями"). Они позволяют запомнить "умственное состояние" робота в момент решения поставленной задачи. Впоследствии при встрече с аналогичной задачей произойдет активация соответствующей "линии знания", которая в свою очередь активирует все агентства, использованные ранее при решении аналогичной задачи. Замечательную метафору придумал один из учеников Минского Кеннет Хааз. Представьте, что Вы ремонтируете велосипед. Прежде чем начать, Вы мажете ладони красной краской. Тогда все инструменты, которые Вы использовали при ремонте, окажутся помеченными красной краской. Если впоследствии Вам снова понадобится отремонтировать велосипед, Вы

сможете сэкономить время, отобрав предметы с красными метками.

4. Агенты и агентства.

Революцией в роботехнике и роботсихологии стала идея сообщества агентов, деятельность которых в совокупности и образует то, что мы называем интеллектуальной деятельностью. Идея "центрального процессора", который управляет всем мозгом, долгое время препятствовала появлению конструктивного определения интеллектуальной деятельности. Да и сейчас эта идея настолько сильна, что даже художник, рисовавший иллюстрацию к статье, не избежал соблазна изобразить одного из агентов (агент 01) как главного командира. На самом деле робот не командует, а только формулирует на привычном нам "я"-языке ("я хочу", "я это сделаю", "я подумаю") свои действия, которые являются результатом взаимодействия множества элементарных агентов и агентств. Одни агенты отвечают за тактильные рецепторы, другие за зрительные, третьи решают конфликты между первыми и вторыми, четвертые откликаются только на одновременное обращение нескольких различных агентов, пятые переключают связи с одних агентств на другие и т.д. Деятельность эта во многом проис-



ходит параллельно. Агенты различных уровней могут говорить на разных языках, поэтому их обращения ко всем будут восприниматься только частью агентств.

Если снова будете искать аналогии

в поведении человека, заметьте, что Ваше сознание имеет дело с очень небольшим числом агентов, так, например, Вы не можете заставить свой пульс участиться непосредственно. Для этого придется активировать доступных Вам агентов, которые

вливают на пульс. Например, представить, что Вы стоите на краю крыши.

Все перечисленные выше понятия, такие как фреймы, линии памяти, цензоры, подавители, являются примерами агентов. Другими агентами являются “полинемы” (“polynemes”) и “прономы” (“pronomes”). Полинемы - это агенты, которые посылают один сигнал сразу нескольким агентствам. Например, слово “яблоко” является полинемом, так как оно влияет на состояние агентств, отвечающих за восприятие цвета, формы, размера и других признаков. Каждое из перечисленных агентств посредством линий памяти хранит информацию о том, как реагировать на тот или иной полинема. Результатом является то, что робот по-разному реагирует на слова “яблоко” и “помидор”. Этот же механизм используется роботом для распознавания предметов. Полинемы можно рассматривать как механизм работы долговременной памяти робота. Пронемы, наоборот, обеспечивают работу кратковременной памяти. Например, когда Вы говорите роботу “положи яблоко в вазу”, в его позитронном мозгу включается агентство “транспортировка”, в котором есть прономы “что”, “куда”. На время выполнения действия проном “что” связывается с полинемом “яблоко”, а проном “куда” - с полинемом “ваза”.

Усвоив основную идею функционирования позитронного мозга, Вы сможете более правильно понять реакции Вашего робота. Основная причина непонимания психики робота связана с тем, что законы психики в корне отличаются от так называемых законов природы. Нельзя

пытаться открыть законы поведения робота так же, как законы движения тел. Минский обращает внимание на то, что есть два типа людей: “упростители” (“reductionists”) и “романисты” (“novelists”). Первые стремятся проникнуть в глубь проблемы, найти простые законы (такие, как законы Ньютона), которые определяют поведение системы. Вторые фантазируют, строят все новые и новые гипотезы, придумывают десятки и сотни своих теорий, объясняющих происходящее. Оказалось, что законы психики нельзя свести к нескольким базовым принципам. Все богатство психической деятельности (как робота, так и человека) объясняется сосуществованием громадного количества агентств и агентов, своеобразного “общества разума”. Как бы много мы ни знали об отдельном агенте или агентстве, это не поможет нам понять законы, по которым живет общество.

На этом можно было бы закончить “краткое пособие...”, но хочется сказать еще несколько слов о структуре книги Марвина Минского “Общество разума”. Структура книги соответствует представлениям автора о функционировании мозга, как множества независимых агентов. Каждая из 300 страниц книги - это отдельный маленький рассказ о том, как устроен интеллект. Эти эссе нельзя вытянуть в одну линию последовательного изложения именно вследствие “нелинейной” сущности самой проблемы. Многие эссе снабжены эпитафиями. Если бы в России нашелся издатель, все мы получили бы большое удовольствие от чтения этой книги на русском языке.

Поздняков Сергей Николаевич,
кандидат педагогических наук,
доцент СПбГЭТУ.

Иллюстрация к роману Семеновской
Ольги Константиновны.

Рисунки к статье Васильковой
Анны Валерьевны.

НАШИ АВТОРЫ