

Текущие концепты и творческие аналогии (Главы из книги)

Д. Хофштадтер.

Вступление

Даглас Хофштадтер – писатель и исследователь с мировым именем. В 1979 году в Нью-Йорке вышла его знаменитая книга «Gödel, Escher, Bach: an Eternal Golden Braid», заслуженно принесшая автору Пулитцеровскую премию по литературе, и которая в течение всего нескольких лет была переведена на многие языки мира. В 2000-м году, наконец, и российские читатели получили возможность приобщиться к этой книге, вышедшей под названием «Гёдель, Эшер, Бах: Эта Бесконечная Гирлянда» (или ГЭБ, как ее называет и сам Д. Хофштадтер). Тремя годами позже в России появилась еще одна книга – «Глаз Разума» (The Mind's I) – созданная Д. Хофштадтером в соавторстве с известным американским философом Дэниэлем Деннетом. Читать эту книгу не менее увлекательно, ибо она составлена из множества коротких «фантастических» и «научно-популярных» рассказов, вышедших из под пера Х. Борхеса, С. Лема, Р. Смолляна и др., и снабженных тщательными, строго научными комментариями Д. Деннета и Д. Хофштадтера, рассуждающими о том, какие из описанных «фантастических» возможностей уже завтра могут стать реальностью, а какие навсегда так и останутся художественным вымыслом.

Книга «Fluid concepts and creative analogies», переводы фрагментов которой мы предоставляем отечественному читателю, совсем иного рода. У основного названия имеется подзаголовок – «Компьютерные модели фундаментальных механизмов мышления». Эта книга фактически является сборником научных трудов Д. Хофштадтера и его коллег, уже много лет занимающихся проблемой компьютерного моделирования творческих способностей человека. Пятьсот страниц теоретических исследований, философского анализа, разбора и критики существующих компьютерных моделей «человеческого интеллекта» – вот о чем эта книга. В одном из своих интервью Д. Хофштадтер говорит, что литературные изыски ГЭБа сбили многих читателей с толку, в результате чего основная теоретическая мысль оказалась не услышанной. «Fluid concepts and creative analogies» заведомо избавлена от этой нежелательной двусмысленности. Здесь игра слов, многозначность естественных языков, ассоциативные связи и т.п. присутствуют лишь постольку, поскольку авторам удается сделать это предметом теоретического анализа и объектом компьютерного моделирования.

Итак, в центре внимания Д. Хофштадтера и его коллег – творческие способности. Тем не менее, название книги гласит – «флюидные» концепты и творческие аналогии. Каким образом *текущая гибкость* концептов и способность порождать *анalogии* связаны с творческими способностями вообще? По мнению Д. Хофштадтера, прояснение этого вопроса ведет к самой сути человеческого мышления. Магистральные линии философских размышлений XX-го века сосредоточились на тех этажах человеческих интеллектуальных способностей, которые ответственны за обработку и упорядочение готового знания, готовых форм представления о вещах. За бортом остались все вопросы, связанные с проблемой *формирования* человеческих представлений, которые Д. Хофштадтер относит к области «высокоуровневых восприятий». Жесткое разделение между процессами усвоения «сырого материала» и процессами его дальнейшей «мыслительной обработки» является, по мысли Д. Хофштадтера, ошибочным и ведущим к тупиковым теоретическим ситуациям. Невразумительность, к которой может привести такое разделение, хорошо демонстрируется Д. Хофштадтером на примере компьютерных программ, нацеленных на моделирование «человеческого интеллекта». В области компьютерных симуляций ярче всего прорисовывается ошибочность самых

разнообразных представлений, связанных с природой «человеческого интеллекта», поскольку отвлеченные философские аргументы приобретают здесь эмпирическую наглядность, подкрепляемую математической точностью. Обратим внимание на то, что противопоставление гибких, подвижных, текучих концептов и жестких, объективных понятий, которое пронизывает всю книгу, нашло свое отражение и в ее названии. Различие «концепта» и «понятия» стало широко признанным в отечественной и зарубежной литературе на рубеже 20 и 21 вв. (упомяну работы С.С.Неретиной, А.Степанова, Н.Д.Арутюновой, А.Вежбицкой; издаются даже словари концептов русского языка и русской культуры). Размышления Хофштадтера о концептах как выражении творческих способностей человека в противовес традиционно понимаемому «искусственному интеллекту» – еще один аргумент в пользу этого различия.

Пропась между «эмпирическими» и «чисто теоретическими» способностями человека, обозначенная еще И. Кантом, однако, может оказаться менее глубокой, если к проблеме устройства человеческого мышления мы подойдем со стороны тех «схематизмов», которые реализуют связь между этими различными способностями. Эта связь, по мысли Д. Хофштадтера, как раз и является тем, что следовало бы называть существом человеческого интеллекта, и характеризуется она тем, что в ее основании лежат механизмы «гибкой», «текучей» организации высокоуровневых восприятий, реализуемых по образу и подобию механизмов, ответственных за способность к «порождению аналогий». Не исключено, что определенное сочетание этих процессов действительно является ключом к пониманию самых глубоких вопросов, связанных с устройством человеческого сознания, мышления, творческих способностей.

Из всей книги Д. Хофштадтера – в качестве введения в ее проблематику – мы выбрали следующие разделы: Пролог, Введение в пятую главу и Эпилог.

Пролог дает самое общее представление об исследовательской группе Д. Хофштадтера, которая именуется себя ФАРГонавтами, и об ее исследовательских целях. Название «ФАРГонавты» исходно является просто аббревиатурой – Fluid Analogies Research Group. Однако в этом сокращении оказываются скрытыми и другие смыслы, часть которых передается и русской транскрипцией. Участники группы ФАРГ (FARG) естественным образом превращаются в ФАРГонавтов (FARGonauts), что не только по звучанию, но и по метафорическому смыслу сближает их с «аргонавтами», искателями золотого руна, ради которого им пришлось пуститься в путешествие, исполненное как неведомых опасностей, так и хитроумных трюков, несущих спасение. В английском слове «fargonaut», помимо прочего, звучит также еще и устойчивое словосочетание «far go(ing)», что означает «далеко идущий». Действительно, проекты ФАРГонавтов представляются далеко идущими и обещающими множество интересных результатов, однако и поставленная Д. Хофштадтером цель является гораздо более отдаленной, чем это мнится многим оппонентам Д. Хофштадтера.

Введение в пятую главу приоткрывает дверь в основной раздел лаборатории Д. Хофштадтера. В этой главе сосредоточены некоторые идеи, связанные с понятием «флюидности», «текучей гибкости» творческого мышления, а также и тот минимальный понятийный инструментарий, который необходим для компьютерного моделирования творческих аспектов мысли. Программа Sorusat является, пожалуй, «парадигмальной» в том отношении, что если о моделировании «человеческого интеллекта» и может идти серьезный разговор, то всякая такая модель будет представлять собой ту или иную модификацию Sorusat'a, разумеется, значительно более сложную и комплексно расширенную. Sorusat – это прототип всякой модели «умного восприятия», благодаря которой «думающие устройства» могут самостоятельно ориентироваться в неизведанных средах и самостоятельно решать определенные задачи.

Эпилог дает общее представление об уровне развития представлений, связанных

как с «естественным интеллектом», так и с «искусственным». Если анализ Д. Хофштадтера верен, то не может не сложиться впечатление, что не так уж много сделано в этом направлении. Например, становится ясным, что доскональное понимание работы мозга на нейронном уровне, или на уровне биохимии не сильно-то нас приближает к пониманию того, что же такое *мышление*. То, что мышление невозможно без соответствующего (того или иного) носителя – факт, почти очевидный. Однако понимание устройства «железа», являясь необходимым условием, всё же является далеко не достаточным. Исследования природы мышления (а не только лишь одного из условий его возможности) должно происходить на совершенно ином уровне, охватывающем всю целостность этого уникального феномена. Что это за уровень, и как должен выглядеть соответствующий язык его описания – всё это еще предмет будущих исследований.

Пролог
«Когда», «Где», «Кто» и «Зачем» этой книги

Краткая история ФАРГ и ФАРГонавтов

Эта книга – обзор полутора десятилетий исследований в области когнитивной науки, работы, проделанной множеством людей. Все началось в 1977 году, когда я стал ассистирующим профессором информатики в Университете штата Индиана и занялся исследованиями в области искусственного интеллекта.

Что касается самого термина «искусственный интеллект»... В семидесятых я с энтузиазмом воспринимал это провокативное словосочетание, даже его ставшую знаменитой аббревиатуру, AI¹; мне казалось, оно хорошо характеризует область моих исследований и мои цели. Мне, как и, наверное, многим другим, в этих словах мерещилась возможность захватывающей авантюры – раскрыть глубочайшие тайны человеческого разума и выразить их предельно очищенным, абстрагированным языком. В начале восьмидесятых, однако, как это часто бывает со словами, этот термин начал постепенно «обрастать» совсем иными значениями; в нем стало больше от коммерческих приложений и экспертных систем, чем от науки, исследующей природу мышления и существо сознания. Дальше – хуже: он уронил себя до бессмысленной болтовни и псевдонаучного жаргона. В результате, если мне приходилось говорить или писать «искусственный интеллект», я уже чувствовал сильный дискомфорт. К счастью, как раз вовремя появился новый термин – «когнитивная наука» – и я с удовольствием стал пользоваться им для выражения моих научных интересов; мне показалось, в нем точно и отчетливо сказывается как природа подлинно научных исследований, так и то, к чему сводится существо человеческого разума. Сейчас я редко называю себя исследователем в области искусственного интеллекта, предпочитая вместо этого говорить, что я занимаюсь когнитивной наукой. И все же AI вновь и вновь прокрадывается в то, что я пишу и говорю.

Мой первый AI-проект, по экстраполяции последовательностей, привел к серии других, и постепенно в их разработку втянулось много моих студентов. Когда все только начиналось – то есть, в конце семидесятых, начале восьмидесятых – моими ближайшими сподвижниками были Марша Мередит и Грэй Клоссман, и они оба, в конце концов, завершили свои докторские диссертации под моим руководством. В частности, Марша разрабатывала программу Seek-Whence², которая воспринимала линейные паттерны и экстраполировала их. Seek-Whence была первым большим проектом, демонстрирующим

¹ AI – Artificial Intelligence.

² Hofstadter, Douglas R. and the Fluid Analogies Research Group. *Fluid Concepts and Creative Analogies: computer models of the fundamental mechanisms of thought*. New York: Basic Books, pp. 13-86. А также см. «Предисловие к пятой главе» настоящей работы.

наш подход к проблеме.

В 1983 году я оказался в знаменитой Лаборатории Искусственного Интеллекта при массачусетском технологическом институте, где мне повезло встретить Мелани Митчелл; позже она работала со мной над своей докторской диссертацией по проекту Сорусат, который возник по большей части из Seek-Whence, но был посвящен моделированию мышления, способного к творческим аналогиям. Там же я познакомился с аспирантом Марекком Луговски и Дэвидом Роджерсом, который только что защитился, и в результате проработал со мной несколько лет.

Спустя полгода в массачусетском технологическом, мне поступило соблазнительное предложение из университета штата Мичиган, и осенью 1984 я перебрался туда. Питер Штайнер, декан Школы Литературы, Науки и Искусств, и Эл Кейн, заведующий отделением психологии, во многом способствовали тому, чтобы я чувствовал себя там как дома. Возможно, сильнейшим притягательным фактором для меня было присутствие Джона Холланда, интеллектуальное и чисто человеческое общение с которым сделали годы, проведенные в Мичигане, насыщенными и незабываемыми.

Переезд в Мичиган осенью 1984 – памятное событие, потому что означает для меня начало проекта ФАРГ (Fluid Analogies Research Group) – Группы по Исследованию Флюидных Аналогий – как мы, ФАРГонавты, привыкли себя называть. Незатейливое звучание аббревиатуры избиралось таковым более или менее намеренно.

Что касается термина «флюидный» (fluid — гибкий, текучий). Он вызывал и продолжает вызывать у многих недоумение, но мне кажется, в нем содержится отчетливый образ гибкости, изменчивости, нестабильности, адаптивности, тонкости, пластичности, неразрывной тягучести, плавности, скользкости, мягкости... Пока я печатал предыдущее предложение, я обратил внимание на то, какими негнуцимися стали мои пальцы в этом холодном помещении. В контексте моих размышлений, я воспринял это неудобство как повод взглянуть по-новому на существо «флюидности». Так что я наполнил раковину горячей водой и погрузил в нее свои руки. По мере того, как они согревались, я думал над тем, что такого особенного в воде и как она движется. Текучая среда реагирует на давление изменением формы в самой мягкой и плавной манере, какая только возможна; в ней нет ломкости твердого вещества и летучей невещественности газа. Откуда эти особые свойства? От молекулярной структуры, разумеется.

Пока я размышлял над этим, мне на ум пришел один из любимейших моих образов во всей науке – «мерцающие кластеры». Эта поэтичная формула отсылает к теории, согласно которой молекулы H₂O постоянно образуют маленькие подвижные сочетания (по причине чрезвычайно слабой водородной связи, которая может образоваться между O одной молекулы и H другой, если им случится проходить достаточно близко друг от друга). Если эта теория воды верна (когда я читал об этом в последний раз, это оставалось неясным), тогда каждую микросекунду в каждой крошечной капельке воды, безмолвно и незримо, формируются и распадаются триллионы сложных, принимающих произвольную форму кластеров. Из этой невероятно нестабильной, динамичной, стохастической структуры и происходят всем знакомые, кажущиеся такими стабильными свойства воды.

Мне кажется, этот образ идеально подходит для того, чтобы выразить саму суть нашей философии, согласно которой флюидные свойства мысли также являются статистическим следствием мириадом крошечных, незримых, независимых друг от друга подсознательных актов, происходящих параллельно. Такой флюидностью обладают концепты, и аналогии, отчетливо воплощающие собой это свойство. Вот почему мы решили назвать себя ФАРГ, и вот почему у этой книги такое название.

Возвращаюсь к истории ФАРГ. В Анн-Арбор ко мне присоединилась небольшая

группа новых студентов, включавшая в себя Боба Френча, Алехандро Лопеса, Грега Хаблера и Роя Либэна. Плюс были Мелани Митчелл и Дэвид Роджерс, последовавшие за мной из массачусетского технологического института, Дэвид Мозер, официально занимавшийся китайским, а у нас исполнявший роль «консультанта по общим вопросам», Грэй Клоссман, доводивший до ума свою докторскую, и Дани Дефэ из Бельгии в качестве приглашенного профессора. Представленный всеми нами, ФАРГ, размещавшийся в те дни в старом ветхом здании, вел весьма оживленную жизнь, а наши дискуссии были предельно широкими и стимулирующими. Наши исследования также протекали вполне удовлетворительно: Мелани разрабатывала *Coruscate*; Боб – новый проект по созданию аналогий, *Tabletop*; Дани же был сам по себе и возился с *Numbo*.

В 1988, к моему удивлению, университет штата Индиана предложил мне вернуться – предложение, от которого невозможно было отказаться – и хотя мне было очень жаль расставаться с замечательными коллегами из Мичигана, я был не менее рад воссоединиться той осенью со многими коллегами из Блумингтона. Морт Ловенгруб, декан колледжа Искусств и Наук, предоставил в полное наше распоряжение маленький симпатичный домик в стороне от основной части кампуса. Это место нуждалось в названии, так что скоро мы окрестили его «Центр по изучению концептов и сознания» (*Center for Research on Concepts and Cognition*), но аббревиатура *CRCC* звучала не так красиво, как ФАРГ, и мы продолжали называть себя и свою штаб-квартиру по-прежнему.

Мелани, Боб и Дэвид Мозер последовали за мной из Анн-Арбор в Блумингтон, а на месте к нам скоро присоединились новые люди: Дэйв Чалмерс, Гэри Макгро и Джим Маршалл. Хельга Келлер была зачислена практически с первого дня как ассистент администратора, и ее энергия, оперативность и энтузиазм внесли огромный вклад в наше общее дело. Из Пекина приехали Ян Йонь, Лю Хаоминь, и позже, Вань Пей; все они переводили на китайский мою книгу «Гедель, Эшер, Бах» – потрясающее предприятие, и это еще слабо сказано. В течение первых нескольких лет к ФАРГ также присоединились Дон Бирд, Дэвид Лик, Питер Сьюбер и Стив Ларсон – кто после защиты, кто в качестве приглашенного профессора – и все они привнесли в нашу работу интереснейшие интуиции музыкального, юмористического, философского характера, оживляя и обогащая ее. Тем временем, Мелани и Боб продолжали разрабатывать свои программы, и, в конце концов, завершили в Индиане свои докторские, над которыми они начали работать еще в Мичигане.

На данный момент, исследовательская группа состоит из меня и группы студентов, включающей Гэри и Джима (оба серьезно работают над своими диссертациями), и Джона Релинга (который только начал работать над своей). Что касается меня, в настоящее время я работаю на севере Италии, в замечательном Институте научных и технологических исследований, расположенном среди виноградников Тренто и окруженном со всех сторон живописными горами. Во многом благодаря щедрости Института и его директора, Луиджи Стринги, ФАРГ благополучно пересек Атлантику и был успешно пересажен на новое место, где, как мы надеемся, он будет действительно процветать.

Интеллектуальные задачи ФАРГ

С самого начала интеллектуальная активность ФАРГ развивалась в двух различных направлениях: одно касалось создания компьютерных программ, моделирующих работу концептов и аналогического мышления в замкнутых и ограниченных областях; другое было связано с наблюдением, классификацией и теоретизацией процессов мышления в их полноте и неограниченности. Последнее служило постоянным источником идей и

вдохновения для первого. Так, многие из наших самых ценных интуиций, касающихся текучей гибкости (флюидности) концептов, аналогического мышления и скрытых за ними механизмов, мы получили, занимаясь поэтическими переводами и игрой слов; исследуя сексистские жаргоны, раскрывающие структуры концептов; собирая, классифицируя и теоретизируя словесные ошибки и другие феномены естественного языка; рассматривая процесс открытия в физике и математике; изучая контрфактуальные условные предложения и «смычки языковых каркасов». А еще нас занимали художественные образы и другие воображаемые миры, основанные на словах, классификация шуток и анализ их глубинной структуры, восприятие и сочинение музыки, и многое другое.

Например, Боб Френч, Алехандро Лопес, Дэвид Мозер, Ян Йонь, Лю Хаоминь и Вань Пей все принимали участие в переводе моей книги «Гедель, Эшер, Бах» (Боб на французский, Алехандро на испанский, остальные четверо – на китайский), и благодаря их участию, переводы получились выше всяких похвал. «Гедель, Эшер, Бах» – книга, трудная для перевода, в ней полно игры слов и других структурных игр, сильно завязанных на язык, не говоря уже о том, что в ней приводятся бесчисленные дискуссии (и даже целые статьи), посвященные нашим попыткам не только охарактеризовать существо хорошего перевода, но и описать механизмы, ответственные за перевод творческий, которые, как нетрудно догадаться, мы считаем родственными механизмам, отвечающим за способность творческого изобретения аналогий.

Еще один пример. В Мичигане Дэвид Мозер и я написали статью³, в которой пытались изложить наши идеи касательно многих разновидностей речевых и поведенческих ошибок – от глупых оговорок до концептуальных ошибок высшего уровня, которые зачастую не отличишь от творческого изобретения. В этой статье представлены начальные стадии наших попыток каталогизации и теоретизации ошибок и других лингвистических феноменов, которые мы наблюдали за собой лично многие годы, даже десятилетия.

На этом месте я не могу удержаться от того, чтобы привести маленький пример ошибки того типа, который интересует нас. Просматривая набросок этого пролога, я решил заменить выражение “once in a while” в конце второго параграфа на более живописное “once in a blue moon”. Я стер слово “while” и начал вводить его замену. Прежде чем я осознал это, я напечатал “once in a bloom” – и это, разумеется, мгновенно меня остановило (о чем я сейчас жалею – мне очень любопытно, что бы я напечатал дальше). Это самая обычная опечатка, которую можно было бы исправить, особо не задумываясь, но что-то в ней привлекло мое внимание. Я подумал: «Почему это случилось здесь и сейчас?» Понятно, что “oo” из “moon” выпрыгнуло раньше времени, но почему не “m” или “n”? Очевидно, потому, что “oo” и “ue” звучат похоже. Но откуда мои пальцы узнали об этом? Если говорить точнее, какие механизмы задействовали в этой ошибке фактор фонетической схожести? Это был интересный вопрос, но мне казалось, что за всем этим стоит что-то большее. В конце концов, “m” тоже поучаствовало в искажении “blue”. Почему это произошло? Ну, я как раз писал о своей работе в Анн-Арборе и Блумингтоне, и без сомнения, слово “Bloomington” оставило подобие длительного, постепенно изглаживающегося следа в моем сознании. Я почти абсолютно уверен в том, что это был решающий фактор, приведший к ошибке. Ведь я много раз употреблял выражение “once in a blue moon”, и никогда не совершал подобной ошибки. Сочетание глубинных факторов, результирующее в крошечном поверхностном событии – проблеме, которой большинство людей не уделили бы вообще никакого внимания – и есть феномен, который так интересует Дэвида и меня.

³ Hofstadter, Douglas R. and David J. Moser (1989). “To Err is Human; To study Error-making is Cognitive Science”, in *Michigan Quarterly Review*, vol.28, no. 2, pp. 185-215.

Можно привести массу схожих примеров, иллюстрирующих спекулятивную сторону ФАРГ, но эти два – мои любимые. Для меня эта сторона ФАРГ связана с понятием «креативности». Предложенное мною название для нашего нового центра, «Центр по изучению концептов, сознания и креативности», было отвергнуто, и справедливо, потому что оно было слишком длинным, а понятие «креативности» – слишком размытым, и вообще чем-то в духе new-age. Как бы то ни было, мы все серьезно заняты изучением творческих актов и стоящих за ними механизмов.

В этой книге ФАРГ по большей части представлен с той стороны, которая связана с компьютерным моделированием. Я надеюсь, в ближайшем будущем ее дополнят одна или две книги, посвященных теоретической, спекулятивной стороне деятельности ФАРГ; я считаю, это не менее важное, хотя, возможно, и менее срочное дело.

Обзор содержания

Книга структурирована хронологически. Первая глава, хоть и написана недавно, рассказывает о начале работы над Seek-Whence, когда этот проект только взошел на ниве моих ранних математических исследований. Она предоставила мне великолепную возможность для введения читателя в понимание существа умственных способностей (mind), на котором базировались все разработки ФАРГ. В ней освещены многие принципиальные вопросы, касающиеся текучей гибкости концептов, порождения аналогий и искусства обобщения, и можно даже разглядеть очертания более поздних проектов, которые на тот момент – то есть, где-то между 1977 и 1983 – существовали разве что в виде идей и набросков. Например, набросок решения задач Бонгарда, представленный в девятнадцатой главе книги «Гедель, Эшер, Бах».

От Seek-Whence история идет дальше к Jumbo, единственному из всех наших компьютерных проектов, который до известной степени был завершен где-то в 1982. Программа Jumbo ориентирована на моделирование когнитивных механизмов – вернее, *субкогнитивных* механизмов – которые позволяют людям, особенно экспертам, быстро и почти без усилия разгадывать анаграммы в уме. Эта глава, написанная в 1983, была оставлена без изменений и без добавления новых отсылок. Третья глава рассказывает о близком родственнике Jumbo по имени Numbo, проект, который предложил и развивал Даниэль Дефэ, профессор психологии Университета Льеж в Бельгии, в 1986-87 годах, которые он провел с нами в Анн-Арборе. Программа Numbo моделировала ситуации, связанные с перебиранием различных арифметических комбинаций в поисках заданного числа. Это элегантная и очаровательная работа. Статья Дефэ, написанная в 1987⁴, также была оставлена без изменений.

Четвертая глава отличается по своему настроению – она посвящена более общим вопросам о том, как продвигаются исследования в области искусственного интеллекта. Пафос этой статьи, насчитывающей аж трех авторов – Дэйва Чалмерса, Боба Френча и меня – в том, что в сфере AI слишком мало внимания и уважения как к когнитивной науке, так и к философии. Статья, набросанная в 1989 и опубликованная двумя годами позже, содержит критику некоторых AI-проектов и, ближе к концу, подводит к обсуждению Sorusat⁵. В действительности, последняя часть исходной статьи, специально посвященная Sorusat, была сокращена по причине избыточности материала.

Главы пятая, шестая и седьмая в деталях описывают все достижения проекта Sorusat, а также некоторые амбициозные заделы на будущее. Пятая глава, написанная Мелани и мной, наверное, самая подробная в этой книге, и в каком-то смысле может

⁴ Hofstadter, Douglas R. and the Fluid Analogies Research Group. *Fluid Concepts and Creative Analogies: computer models of the fundamental mechanisms of thought*. New York: Basic Books, pp. 131-154.

⁵ цит. соч. стр. 155-194, а также см. «Предисловие к пятой главе» настоящей работы.

считаться ее сердцем, поскольку на данный момент Сорусат – последнее слово наших исследований. Эта глава в мельчайших подробностях описывает архитектуру проекта, дает ее философское обоснование, предоставляет результаты ее рассмотрения в разных перспективах, и в конце объясняет особые достоинства Сорусат как системы, моделирующей самую суть познания (cognition). Работа, описанная в пятой главе, началась в 1983 (когда были заданы рамки, ясны трудности и продумана архитектура проекта) и завершилась в конце 1990 (так много времени заняла разработка архитектуры и ее компьютерное воплощение). В шестой главе, также написанной нами, Сорусат сравнивается с другими проектами по моделированию процесса распознавания аналогий, и соотносится как со старомодными AI-проектами, так и с новомодным коннекционизмом. Седьмая глава, написанная мной как послесловие к книге Мелани о Сорусат (Mitchell, 1993), указывает на новые многообещающие области, где, я надеюсь, эта программа будет применена в ближайшем будущем.

Главы восьмая и девятая посвящены Tabletop, нашей совместной с Бобом Френчем работы в период 1986-91. Восьмая глава – философское углубление в природу микрообластей и комплексную специфику когнитивного акта, ответственного за распознавание аналогий. Девятая глава фокусируется на компьютерной модели Tabletop, выделяя ее сходства и различия с Сорусат, и заканчивается демонстрацией ее работы с большим количеством родственных задач по аналогиям, распадающихся на три различных семейства.

Десятая глава описывает только начатую Гэри Макгро и мной работу, которая обещает стать самым амбициозным проектом ФАРГ в области компьютерного моделирования креативности – Letter Spirit, проект куда большего масштаба, чем все наши предыдущие проекты. Letter Spirit – это программа по разработке художественно согласованных алфавитов (так называемых шрифтов). Идея программы, заставляющей машину придавать 26 буквам латинского алфавита одну и ту же абстрактную сущность – то есть, визуальную стилистику – одна из самых затруднительных и в то же время завораживающих задач, какие только можно представить: она подразумевает столько креативных мини-прорывов, что смоделировать каждый из них – уже огромная трудность.

Наконец, эпилог представляет нашу работу в обобщенной перспективе; там мы обсуждаем другие компьютерные модели креативности, критики анализируя как сами проекты, так и то, как они преподносятся широкой общественности. Мы пытаемся поставить провокативный с философской точки зрения вопрос: наступит ли время, когда логичнее будет давать премию за научное открытие компьютерной программе, а не авторам этой программы. Это, в свою очередь, ведет к обсуждению природы механизмов, лежащих в основании когнитивных моделей, и к вопросу о возможности их обнаружения, а также к теме соотношения наших взглядов на указанные проблемы с предпосылками, лежащими в основании знаменитого и противоречивого теста Тьюринга.

Многие из этих глав уже выходили, в той или иной форме, в печати, хотя в процессе подготовки этой книги к изданию, практически каждая ранее публиковавшаяся статья была переработана, причем основательно. Как видно из оглавления, каждой главе предпослано «предисловие». Все эти предисловия были написаны мной лично. Главная цель этих промежуточных частей, разумеется, облегчить и сделать более естественным переход от одной главы к другой, но есть и другая цель – они куда менее формальны, а посему предоставляют возможность обозначить исторические и философские перспективы, а также разбавить повествование анекдотами и личными впечатлениями. Мне кажется, эти легкие по тону предисловия – полезное и важное сопровождение серьезных по тону статей.

**Предисловие к пятой главе:
Концептуальные Облака и их Трансформации**

Аналогические головоломки системы Seek-Whence

Проект Сорусат возник из-за трудностей, с которыми столкнулся Seek-Whence. К 1980 году мне уже было ясно, что именно способность замечать абстрактные соответствия – другими словами, аналогии – между сегментами некой последовательности лежит в основании способности воспринимать постоянные структуры и формулировать правила, описывающие эти структуры. Так что примерно в это время я начал отбирать аналогические головоломки из системы Seek-Whence, такие, как эта, довольно простая.

Что в «12344321» соответствует «4» в «1234554321»?

Для облегчения задачи я буду называть вторую последовательность «структура А», а первую «структура Б». Обратите внимание, что это порядок, обратный тому, который вы могли ожидать.

Эта головоломка не очень трудная. Большинство безо всякого затруднения даст ответ «3», заметив, что «3» *играет ту же роль* в В, какую «4» играет в А. И все же, кто-то может педантично ответить «4», поскольку это *та же самая величина* или, если угодно, *то же самое число*. В узких рамках этой головоломки конфликт предпочтений сводится к следующей простой схеме: скучная, механизированная идея смысла как «тождественности объекта», против куда более живой и близкой человеку идеи смысла как *роли, которую играют*. Задумайтесь об этом на минуту. В самом деле, верно ли воспринимать так «4» – то есть, играет ли «4» *одну-единственную* роль в структуре А? Чтобы стимулировать ваши мысли в этом направлении, вот вам еще одна, «двоюродная» головоломка.

Что в «123475574321» соответствует «4» в «1234554321»?

Назовем новую структуру «С». Что делает ее интересной, это то, что она как бы поднимается, а затем опускается, так что вершина не располагается в центре структуры, как в случае с А и В. Структура С больше похожа на вулкан с острыми краями и впадиной кратера между ними.

Эффект, вызванный этим изменением формы, состоит в том, что казавшаяся *единственной* роль, которую «4» играет в А (и «3» в В, соответственно), распадается на *несколько* ролей. Таким образом, в свете структуры С мы теперь видим, что «4» в А могло занимать как минимум четыре не совпадающих положения:

- (1) соседнее с центральной парой чисел;
- (2) следующее за наибольшим числом в структуре;
- (3) численный предшественник центральной пары чисел;
- (4) численный предшественник наибольшего числа в структуре.

Так вышло, что в структуре А, в силу того, как она была устроена, все эти роли совпадали и выглядели как *одна*. Можно сказать то же самое иначе – что структура С *стимулировала* распад того, что казалось монолитным единством роли, на четыре

отдельных аспекта. Но это расщепление произошло *a posteriori*. Значит, если бы не было структуры, подобной С, мы бы продолжали думать, что «4» играет в А одну-единственную роль. Является ли такое понимание иллюзорным, или оно все же верно?

Английская Нэнси Рейган

В тандеме с предшествующей, чисто математической головоломкой мне всегда нравилось давать словесную головоломку:

Кто в Англии соответствует Нэнси Рейган в Соединенных Штатах?

Или даже так:

Кто первая леди Англии?

Имейте в виду, пожалуйста, что речь идет о 1980 годе, когда премьер-министром Англии была Маргарет Тэтчер.

Что сбивает с толку в этой головоломке? Первое – что в Англии нет президента, но есть две политические фигуры, каждая из которых играет роль, в чем-то близкую роли президента: монарх и премьер-министр. Второе – что в 1980 году и тот, и другой «пост» занимали женщины, в то время как Рональд Рейган был, разумеется, мужчиной. Оба эти фактора ставят нас в крайне затруднительное положение.

Прежде всего, мы *a priori* склонны искать *женский* аналог Нэнси Рейган, поскольку сама она женщина. Особенно если в головоломке используется выражение «первая леди», но даже если и нет, все будет подталкивать нас к проведению аналогии по принципу «тот же пол», потому что эпитет «первая леди» напрашивается сам собой, когда речь заходит о Нэнси Рейган, и это значение всегда явно или скрыто присутствует. Базируясь на этой аналогии, мы будем рассматривать в качестве двойника Нэнси Рейган или Королеву Елизавету, или Маргарет Тэтчер, постольку, поскольку «первая леди» означает, в большей или меньшей степени, самую влиятельную женщину в стране.

Само собой, есть сильнейшее возражение против обоих этих вариантов, происходящее из очевидного факта, что эти женщины сами находились *на вершине* или же *в средоточии* власти, в то время как Нэнси Рейган – лишь *рядом* с ней. Итак, в зависимости от того, как понимать эту «вершину», мы будем вынуждены искать наш аналог в принце Филиппе или же в Деннисе Тэтчере, муже Маргарет Тэтчер (королевская власть больше похожа на центр или на вершину?). Так или иначе, каждый из этих вариантов нарушает прямую, «женскую» аналогию. Кстати говоря, эта идея пришла мне в голову, когда я читал статью о Деннисе Тэтчер, в которой тот был охарактеризован как «первая леди Англии».

Аналогия между этими аналогическими головоломками – математической, включающей в себя структуры А и С, и «реалистической» – очевидна: и та, и другая включают в себя *a posteriori* разбиение того, что выглядело как *один* смысл, на ряд конфликтующих смыслов, разбиение, толкающее нас от одного смысла к другому и выводящее на разные ответы. Более того, обе головоломки включают в себя фигуру, которая «находится рядом с вершиной».

Вернемся к четырем аспектам «4», возникающим, когда мы переходим к структуре С.

Если (1) – «соседнее с центральной парой чисел» – берется как преимущественное значение, тогда решением головоломки должно быть «7».

Если (2) – «следующее за наибольшим числом в структуре» – принимается как доминанта, тогда решением будет «5».

Если (3) – «численный предшественник центральной пары чисел» – является доминантой, тогда решением будет «4».

Если (4) – «численный предшественник наибольшего числа в структуре» – будет сочтено как наиболее соответствующее случаю, тогда решением будет «6» – число, никак не представленное в самой структуре!

Четыре этих различных решения – 7, 5, 4 и 6 – математической головоломки в точности соответствуют четырем ответам на головоломку словесную, а именно, нашим четырем «первым леди Англии»: принцу Филиппу, Деннису Тэтчеру, королеве Елизавете и Маргарет Тэтчер. Как накладываются друг на друга решения и упомянутые лица, зависит, разумеется, от того, как вы смотрите на вещи (дальнейшее развитие темы «первой леди», равно как и обсуждение множества проблем с аналогиями в Seek-Whence и Sorusat, см. Hofstadter, 1985⁶).

Спустя короткое время я обнаружил, что эти аналогические задачи разрастаются совершенно неожиданным образом. Вскоре я изобрел их столько, что не мог уже за ними уследить; они были всех уровней сложности, и содержали замысловатые комбинации крайностей, подразумевающих то или иное концептуальное смещение. Под «концептуальным смещением» я имею в виду *вынужденное контекстом замещение одного концепта другим, который оказывается тесно связанным с ним* в контексте ментальной репрезентации той или иной ситуации. Но с чего бы одному концепту быть замещенным другим? Каким именно образом происходит такое ментальное событие, и когда? Все это связано с концептуальными наложениями, которые являются неизбежным побочным эффектом того, что концепты подобны *облакам*. Давайте рассмотрим подробнее эту идею, центральную для всей архитектуры Sorusat.

Слова, концепты и «облака»

Ранее я спрашивал, означает ли распад единого смысла на множество смысловых аспектов то, что кажущаяся единственность смысла – не более чем иллюзия? Это весьма тонкий предмет, но многое станет яснее, если мы рассмотрим аналогичную проблему в естественном языке. Возьмем самое обычное слово, скажем, hard («твердое»). Когда вы думаете о его значении, возможно, оно представляется вам чем-то одним – например, «не мягкое». Но вот передо мной толковый словарь, где на соответствующей странице перечислены все значения и оттенки смысла, которые может иметь слово hard (жесткий, твердый): вызывающий привычку (addictive), враждебный (adverse), крепкий (alcoholic), терпкий (bitter), загорбелый (callous), трудный (difficult), трудный для понимания (hard to understand), бессердечный (heartless), закопенелый (impenitent), усердный (industrious), тугой (inelastic), черствый (obdurate), сильная боль (painful), безжалостный (pitiless), жесткий (rigid), суровый (severe), твердый (solid), строгий (strict), сильный, здоровый (strong), значительный (substantial), крутой (tough), злобный (wicked). Если вы подумаете о каждом из них в отдельности, вы увидите, что значение hard в действительности подразделяется на множество идей. Но вы все еще можете думать, что всеобъемлющий концепт – некое единство, одна идея.

Перенесемся теперь в другой язык – скажем, в немецкий, родственник английскому. Выяснится, что hard расщепится в немецком в свою очередь на множество слов, в зависимости от того, что оно модифицирует. Если речь идет о веществе, тогда это

⁶ Hofstadter, Douglas R. (1985). *Metamagical Themas: Questing for Essence of Mind and Pattern*. New York: Basic Books.

hart; о задаче или тесте – schwierig; об усердной работе – fleissig; о тяжелых временах или суровых испытаниях – schwer; о жизни в целом – muhsam; о дожде – heilig или stark; о морозе – streng; о крепости напитков – alkoholisch; о напряженных раздумьях – gut или scharf; о сильном ударе – kraftig; о человеке, пытающемся что-то сделать – anstrengend; и так далее. Причем человек, говорящий по-немецки, будет в меньшей степени склонен рассматривать эти концепты в свете какого бы то ни было единства, нежели человек, говорящий по-английски. И не думайте, что каждое из этих немецких слов уже по значению, чем hard, то есть, является подмножеством этого концепта. Как раз наоборот – каждое из них само по себе весьма широко. К примеру, среди множества переводов schwer есть следующие: heavy, solid, powerful, strong, rich, hard, severe, serious, grave, difficult, tough, ponderous и clumsy. Для человека, говорящего по-немецки, это созвездие оттенков смысла добавляется к тому, что выглядит как «одна идея».

Заблуждаемся ли мы, говорящие по-английски, когда думаем, что концепт, обозначаемый словом hard, есть единое, монолитное понятие? Заблуждаются ли говорящие по-немецки, когда аналогичным образом думают, что концепт, обозначаемый словом schwer – точно такое же единое, монолитное понятие? Я бы сказал, что в контексте английского языка оттенки смысла hard образуют, кажется, последовательное, связанное единство, в то время как в перспективе немецкого языка это единство кажется иллюзорным. Заметьте, что и там, и там я говорю «кажется». Контекст и культура определяют границы концептов.

Возможно, следующее сравнение английского, индонезийского, китайского, немецкого и итальянского предоставит нам более простые примеры.

В английском, мы обычно задаем новому знакомому стандартный вопрос: “Do you have any brothers or sisters?” («Есть ли у тебя братья и сестры?»).

В индонезийском обычно спрашивают, есть ли у вас kakak’и и adik’и. Это по сути тот же самый вопрос, и однако вовсе не тот же, поскольку слово kakak означает старшего родственника, а adik – младшего, независимо от пола.

В мандаринском варианте китайского обычно спрашивают, есть ли у вас xiong-di-jie-meí. И опять, это то же самое – и совсем другое. На этот раз вам задали фактически четыре вопроса в форме одного, поскольку xiong означает «старший брат», di – «младший брат», jie – «старшая сестра» и meí – «младшая сестра».

Наш следующий пример, из немецкого, вроде бы проще. Там спрашивают “Haben Sie Geschwister?”, что означает просто «Есть ли у тебя родственники?». Казалось бы, все очень просто. Но будьте осторожны, не дайте этой видимой простоте ввести себя в заблуждение. Geschwister выглядит и звучит столь схоже со Schwester («сестра»), что это выражение несомненно подразумевает «сестер», даже если и не используется в этом значении сознательно.

Наш последний пример, из итальянского, что-то вроде немецкого наоборот. Если итальянец спросит меня, “Lei ha fratelli?” – «Есть ли у тебя братья?», я могу с полным основанием ответить “Sì, due sorelle” – «Да, две сестры». В этом контексте слово fratelli, обычно означающее «братья», колеблется между «братья» и «братья и сестры».

Оставляя в стороне языковую специфику, зададимся вопросом: сколько концептов оказываются вовлечены в ментальный процесс, когда речь заходит о наличии, количестве и качестве ваших ближайших родственников – один, два или четыре? Имеет ли этот вопрос вообще смысл? Допустим, есть некий язык, в котором слово plubibwa означает «легкомысленный родственник», а слово vazil – «серьезный родственник», а стандартный вопрос, задаваемый при знакомстве, звучал бы “Exement-ci plubibwa flo vazil?”. Не означало бы это, что было заблуждением с нашей стороны полагать, будто есть лишь *четыре* концепта, заключенных в концепте «родственник»? Стали бы мы просвещеннее, осознав, что в конечном итоге есть *восемь* таких концептов

(варьирующихся от «серьезного старшего брата» до «легкомысленной младшей сестры»)? Очевидно, что нет. Дело просто в том, что нет никакого определенного числа концептов, относящихся к тому или иному случаю; все зависит от контекста и культуры. Если и есть что-то в нашем разуме независимо от культуры, это трудноразличимое концептуальное наложение множества специфических примеров родственности, с сильной группирующей функцией, которая отвечает за то, что из представленных на выбор потенциальных субконцептов одни кажутся более, а другие – менее естественными.

Это не надуманные, преувеличенные примеры из области воображения, не особые случаи. Такое дробление кажущегося монолитным единства на множество неожиданных осколков смысла происходит с любым словом, стоит вам попытаться перейти на другие языки, даже если это столь близкие английскому французский или итальянский. Даже такие предельно общие понятия, которые выражают английские слова *hit*, *throw*, *window*, *box*, *bag*, *day*, *top*, *fast* и *in fact*, и которые я до поры наивно считал бесструктурными, основными, первичными концептами, универсальными относительно любого языка, «раздробились» на пять, десять или даже больше разновидностей, «ударившись» об итальянский – думаю, нет нужды говорить, что то же самое произошло с итальянскими словами равно фундаментального значения при столкновении с английским. То, что кажется неделимым, дробится, расщепляется, раскалывается, превращаясь в микроструктуру.

Получив подобный опыт, я могу понять «логику» превращения в итальянском *in fact* в *in effetti*, *di fatto*, *infatti*, *in realta*, *anzi*, *pero*, *tant'è vero che*, *per esempio*, *a dire il vero* и так далее, но сам по себе я бы никогда не задумался об этом. Это не казалось мне чем-то таким уж сложным. Как мало я знал! Это дробление на осколки происходит, я повторяю, с любым известным нам словом, пока мы не достигаем тех разреженных высот семантики, на которых обитают лишь высоко специализированные технические термины, такие, как «фотосинтез», который сохраняет свою неприкосновенность в любом языке – благо итальянцы милосердно решили не настаивать на одном фотосинтезе для цветов, другом – для лиственных деревьев, третьем – для вечнозеленых, и так далее. Но что касается общеупотребительных слов, это *всегда* справедливо – и в действительности, чем оно употребительнее, тем более сложным и удивительным будет его расщепление.

Вот почему это столь забавное – хотя, наверное, я хотел сказать «жалкое» – зрелище: видеть, как размножаются все эти карманные словари и так называемые электронные переводчики, набитые одноуровневыми соответствиями английского-французского типа “*hit: frapper*”, “*throw: jeter*”, “*picture: image*”, “*in: dans*”, и тому подобным. Обратная сторона медали – огромное удовольствие, с которым я листаю свой зачитанный до дыр экземпляр *Roget's International Thesaurus*, погружающий меня в смысловые туманы, создаваемые потрясающим взаимопроникновением концептуальных облаков.

От трансформации концептуального облака к концептуальному смещению

Наложение и скопление концептов в нашем разуме, приводящее к возникновению «семантического облака», которое окружает каждое осмысленное слово, часто проявляется в речевых ошибках-оговорках, особенно в смешении слов. Вот несколько примеров того, как в нашем сознании смысловые облака накладываются и даже частично смешиваются. Заметьте, как часто такие ошибки являются прямым результатом влияния определенного контекста. Мой друг сказал мне: “No one could get the tickets to the *flane*”, очевидным образом смешав “*flight*” («полет») и “*plane*” («самолет»). Лингвист поймал

себя на том, что говорит: “Don’t shell so loud”, смешав “shout” («выкрикивать») и “yell” («вопить»). Я сказал: “I’ll chake a look”, смешав “check it out” («проверить») и “take a look” («посмотреть»). Примеров такого смешения не счесть, и все они свидетельствуют о том, что в реальном времени речи мы порой затрудняемся быстро и корректно разрешить конфликт соперничающих влияний.

Более странными являются ошибки, основанные на полной подмене одного слова другим. Я сказал своему маленькому сыну: «Закрой свои банки с печеньями, пожалуйста», имея в виду *коробки с игрушками* в его комнате. Это уже не просто оговорка, а что-то совсем неадекватное, вообще не относящееся к делу. В другой раз я сказал: «Дверь ванной не закрывается – водопроводный кран сломался», в то время как я хотел сказать: «Дверная ручка сломалась». Разумеется, и «водопроводный кран», и «дверная ручка» относятся к одной области семантического пространства – той, что связана с управлением – но раз речь идет о ваннах, у «водопроводного крана» явное преимущество, которое он использует, чтобы вытеснить и заменить собой «дверную ручку» в моем сознании.

Мой друг: «Я вернусь в начало... нет, ко входу в магазин». Я: «Мы переехали в Принстон, когда мне было только два часа» – вместо, конечно, «когда мне было только два года». Моя жена: «Я положу мусор – ой, нет, грязную одежду – на заднее сиденье машины?».

И последний пример, возможно, мой самый любимый... Кассир в бакалее спросил меня: «Пластиковый пакет сойдет?», на что я ответил: «Лучше деревянный... в смысле, бумажный, пожалуйста». Каковы составляющие этой ошибки? Все очень просто: бумага делается из древесной стружки; пакеты в бакалее коричневого цвета, схожего с древесиной; они плотнее, чем обычная бумага – «древеснее»; общее между пластиком и деревом то, что они оба служат материалом для разной домашней утвари, в то время как бумага – нет.

Ошибки замещения/подмены вроде этих проливают частичный свет на скрытый под поверхностью сознания ландшафт – невидимую сеть накладывающихся друг на друга концептов со смазанными границами. Они показывают нам, что под влиянием множества обстоятельств мы принимаем один концепт за другой, и это может помочь нам яснее увидеть, что происходит, когда мы проводим аналогию между различными ситуациями. Свойства концептуальной сети, которые отвечают за нашу склонность к совершению подобных ошибок и промахов, заставляют нас также допускать известную степень концептуального несоответствия между ситуациями, в зависимости от контекста; мы изначально так устроены, и это, с эволюционной точки зрения, преимущество. Термин «концептуальное смещение», по сути, укороченная формулировка понятия «зависящего от контекста допущения концептуального несоответствия». (Более подробное рассмотрение этих тем см. в следующих статьях⁷).

Замысел Сорусат

⁷ Aitchison, Jean (1994). *Words in the mind: An Introduction to the Mental Lexicon* (2nd ed.). Cambridge, MA: Basil Blackwell.

Cutler, Anne ed. (1992). *Slips of the Tongue and Language Production*. Berlin: Mouton.

Dell, Gary S. and P.A. Reich (1980). “Slips of the Tongue: The Facts and a Stratificational Model”. In J.E. Copeland & P.W. Davis (eds.) *Papers in Cognitive-Stratificational Linguistics*, vol. 66, pp. 611-629. Houston: Rice University Studies.

Fromkin, Victoria A., ed. (1980). *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand*. New York: Academic Press.

Hofstadter, Douglas R. and David J. Moser (1989). “To Err is Human; To study Error-making is Cognitive Science”, in *Michigan Quarterly Review*, vol. 28, №2, pp. 185-215.

Norman, Donald (1981). “Categorization of Action Slips”. *Psychological Review*, vol. 88, pp. 1-15.

Еще по ходу работы над архитектурой Seek-Whence я стал замечать, сколь распространены эти, и другие эффекты трансформации концептуальных облаков в когнитивных процессах, и это оказало огромное влияние на мои идеи. Но чем яснее я осознавал ключевую роль способности к распознаванию аналогий, тем чаще задавался вопросом, не достигает ли проект Seek-Whence слишком много слишком быстро. Постепенно я пришел к убеждению, что мне прежде всего следует автономно проработать проблему изобретения аналогий, а уже затем, на более позднем этапе, я смогу вернуться к Seek-Whence, вооруженный проверенными методами, которые помогут мне приступить к самой сути дела.

Итак, весной 1983 года я направил свои усилия на разработку структуры по исследованию аналогий в микросистеме, схожей с микросистемой Seek-Whence, только тут ключевую роль призвано было играть мое нововведение – *концептуальное смещение*. Но по какой-то причине – возможно, желая привнести какое-то разнообразие в работу – я вместо *цифр* стал использовать *буквы*. В течение следующих нескольких лет я не раз бывал удивлен и даже раздосадован тем, что многие люди куда лучше восприняли эти буквенные головоломки, нежели ранние числовые, несмотря на то, что зачастую одни были точными переводами других. Мне не раз говорили: «Я никогда не любил математику, так что мне не по душе цифровые головоломки, но вот буквенные очень здорово получились!» Я так и не смог это понять.

В любом случае, вот буквенная головоломка, показывающая, над чем я тогда думал.

Я превратил *efg* в *efw*. Можете «сделать то же самое» с *ghi*?

Что мне нравится в этой головоломке – ее можно решить двумя сильно различающимися способами. Один ответ – *whi* – мы получим, если просто поставим *w* на место *g*. Конечно, мы «сделали то же самое», но это как-то грубо. Другой ответ – *ghw* – мы получим, заменив крайнюю правую букву на *w*. Это тоже «то же самое», но совсем в другом смысле – может, более тонком, может и нет. (Кстати, любой ответ, основанный на идее «алфавитного расстояния» между *g* и *w*, выходит за рамки рассматриваемой ситуации, ибо выходит за пределы предполагаемого контекста).

Следующая головоломка сродни предыдущей.

Я превратил *efg* в *wfg*. Можете «сделать то же самое» с *ghi*?

Требуемым условиям удовлетворяют те же самые цепочки *whi* и *ghw*, но уже по другим причинам. Первый вариант мы получим легко: надо просто заменить крайнюю левую букву на *w*. Второй вариант тоньше, он подразумевает усмотрение своего рода симметрии между *efg* и *ghi*. В цепочке *efg* буква *g* крайняя справа, в цепочке *ghi* – крайняя слева. Под таким углом зрения мы приходим к идее, что концепт «слева» в *ghi* играет *ту же* роль, что концепт «справа» в *efg*. Это, в свою очередь, предполагает, что две эти цепочки могут рассматриваться как «одно и то же», если мы допустим *концептуальное замещение* правого левым, и наоборот. Отсюда уже мы можем получить ответ *ghw*.

Отметим одно ироническое обстоятельство. В первой задаче ответ *whi*, основанный на уравнении двух *g*, казался весьма грубым, но во второй задаче ответ *ghw*, основанный на том же самом уравнении, кажется весьма изощренным. Как одна и та же стратегия может быть грубой для одной задачи и изощренной для другой? Вот как я это вижу. В первой головоломке решение сводилось к отождествлению двух *g* – остальные буквы были просто проигнорированы, что и производило впечатление грубости. Во второй

головоломке, однако, равенство двух **g** было только началом. Затем следовало дойти края цепочек, найти там **e** и **i**, и уравнять их. Идея использования обеих **g** как идентичных точек отсчета является весьма абстрактным приемом; она позволяет задействовать структуры обеих цепочек целиком и составить полную картину. Поэтому-то ответы, полученные на основе этой идеи, будут казаться не просто допустимыми, но и, так сказать, более глубокими.

Вот последняя (пока что) задача, развивающая эту идею дальше.

Я превратил *efg* в *dfg*. Можете «сделать то же самое» с *ghi*?

Если вы сразу не приняли во внимание отношение, в котором буква **e** стоит к заменяющей ее **d**, тогда для вас эта задача будет изоморфна предыдущей, где роль **d** играла **w**, так что вы получите два изоморфных ответа: *dhi* и *ghd*. Однако игнорировать элементарную алфавитную близость **e** и **d** в данном случае кажется очень грубым ходом. Так что если мы примем этот очевидный факт в расчет, дальше наиболее естественным кажется *взять предшествующую* либо крайней левой (что даст нам *fhi*), либо крайней правой буквы (что даст нам *ghh*). Но двойное **hh** в последнем ответе, кажется, совсем не в духе наших превращений, и мы начинаем подозревать, что что-то тут не так. И в самом деле, мы упустили нечто решающее.

Приравнивая **e** из *efg* к **g** из *ghi*, мы тем самым как бы пробегаем *efg* справа налево, а *ghi* – слева направо. В случае *efg* мы пробегаем алфавит в обратном направлении, от буквы к предшествующей ей, в то время как в случае *ghi* мы движемся вперед, от буквы к следующей за ней. Таким образом, ключевая идея, которую мы упустили, это *концептуальный переход* от предшествующего к последующему. Если мы возьмем в расчет это дополнительное смещение, то увидим, что **i** надо заменить не на **h**, а на **j**, и получить, следовательно, *ghj*. Это весьма изощренный ответ, он требует принятия во внимание всей полноты обеих структур (включая нити алфавита, стягивающие их вместе – фактор, не игравший никакой роли в предыдущих головоломках), и, тем не менее, он всё также основан на простейшем, кажущемся грубым первом шаге – уравнивании двух **g**.

Эпилог

О компьютерах, творческих способностях, проблеме авторства, устройстве мозга и о тесте Тьюринга

Немного скепсиса в отношении компьютеров и творческих способностей

Заключительные главы нашей книги, и в особенности последняя, были посвящены нашим собственным компьютерным моделям творческих актов, которые мы исследовали в рамках определенных микрообластей. Нет нужды упоминать о том, что наша работа проводилась не в пустоте. На протяжении десятков лет совершалось множество попыток по созданию программ, которые схватывали бы определенные аспекты творческих процессов, имеющих место в совершенно различных областях, от самых простых областей до грандиозных. Чего же достигли компьютеры в качестве художников, писателей, композиторов, математиков, ученых, изобретателей, поваров, футбольных тренеров, и тому подобное?

Эта тема слишком широка для общего ее обзора в пределах одной лишь главы. К счастью, Маргарет Боден проделала блестящую работу по анализу новейших проектов в своей книге *Творческий разум: мифы и механизмы* (1991)⁸. Стоит, правда, сказать, что во многих местах там, где Боден готова видеть положительное, я склонен видеть обратное.

⁸ Boden, Margaret A. (1991). *The Creative Mind: Myths and Mechanisms*. New York: Basic Books.

Возможно, так получается потому, что непредвзятость ее точки зрения связана с тем, что она не является разработчиком своих собственных проектов, и потому не имеет корыстных интересов, и ей не за что ломать копыя, в то время как мне есть за что. А, может быть, она просто более восторженный наблюдатель всего поля исследований по искусственному интеллекту, нежели я. Однако каковы бы ни были соответствующие причины, я собираюсь гораздо более критично отнестись к некоторым проектам, которые я и собираюсь обсудить ниже.

Сразу же сделаю одно предупреждение. То небольшое число программ, которое будет обсуждаться, ни в коей мере не следует считать репрезентативным с точки зрения выявления «состояния искусства», являющегося продуктом компьютерного моделирования творческих процессов, ни, тем более, адекватной репрезентацией всего поля подобных исследований. Некоторые программы, которые попадут в область нашего внимания, либо достаточно стары, либо же весьма специфичны и лежат на обочине мейнстрима. Поэтому неудивительно, что найдутся те, кто сочтут мою критику бьющей мимо цели. Ведь можно возразить: зачем анализировать устаревшие или же явно слабые версии программ, когда вокруг полно новых, более сильных версий?

На это я бы ответил, что подобная критика полезна в порядке учебной практики. Мои замечания служат выработке общего отношения – некоей формы понимания проблемы в целом. Возможно, предметы моей критики порой слишком просты, но учебные процессы всегда строятся именно так. Изучение начинается с простого, и постепенно развивается до освоения сложного, в процессе чего старое используется для получения нового. Если дефекты обсуждаемых мною программ покажутся кому-то очевидными, ну что ж. Моя критика, по крайней мере, может указать направление, в котором следует двигаться по пути более изощренного и сложного анализа других проектов.

И последнее слово, перед тем как вступить на тропу войны. Важно сказать, что в конечном счете я весьма доброжелательно отношусь ко всем обсуждаемым ниже проектам, несмотря на то, что подвергаю их критике. Все они по-своему являются восхитительными, дерзкими, изобретательными плодами долгой работы, заслуживающими серьезных размышлений над ними.

Компьютер-художник

Программа Aaron, претендующая на роль «компьютерного художника», принадлежит к числу наиболее провокативных программ, нацеленных на моделирование творческих способностей. Два десятка лет эта программа разрабатывалась Гарольдом Коэном, художником и профессором искусств. К сожалению, узнать что-либо о внутреннем устройстве архитектуры Aaron'a необычайно трудно ввиду отсутствия каких-либо детальных публикаций по этому поводу. Наилучшим источником информации об Aaron'e, который я сумел обнаружить, является книга Памелы МакКордак *Код Aaron'a: Мета-искусство, Искусственный разум и Производство Гарольда Коэна*⁹, опубликованная в 1991 году.

Aaron создает сложные и жизненные рисунки, которые выглядят поразительно похоже на некие замысловатые творения настоящего художника. Рисунки тяготеют к изображению человекообразных фигур, танцующих, играющих с мячом на пляже, качающихся и т.п., причем все они изображаются на фоне, напоминающем скалы, или кустарники, или что-нибудь в этом род. Рисунки Aaron'a зачастую забавны и обнаруживают собой очаровывающую наивность узнаваемого стиля. Тем не менее, не

⁹ McCorduck, Pamela (1991). *Aaron's Code: Meta-Art, Artificial Intelligence, and the Work of Harold Cohen*. New York: Freeman.

возникает даже и вопроса о том, воспринимает ли Аагон свои собственные рисунки и интерпретирует ли их тем же образом, каким это делают люди, не говоря уж о его способности обсуждать свой собственный стиль в явном виде.

Ввиду наличия вероятностной компоненты в программе, Аагон никогда не создает одни и те же рисунки дважды. На сегодняшний день, должно быть, Аагон уже создал многие тысячи различных рисунков. Творения Аагон'а не раз использовались на обложках книг (преимущественно по искусственному интеллекту или же по компьютерным наукам), а некоторые даже украсили собой стены ряда августейших заведений. Поскольку плоды деятельности Аагон'а почти безусловно можно рассматривать в рамках «человеческого искусства», то на ум приходят многие философские вопросы относительно их художественной ценности и значимости. Например, возникает вопрос, как понять то обстоятельство, что компьютерная программа рисует «людей», с которыми она не имеет никакого опыта общения? Любопытным соображением, не лишенным полностью смысла, является то, что для программ, подобных Аагон'у, более подходящими темами для зарисовок были бы не люди, занятые своими человеческими делами, а собственные, компьютерные занятия. И в самом деле, разве не верно предположить, что Аагон'у было бы более свойственно делать предметом рисования самое себя, и заниматься набросками автопортретов, на которых находила бы свое выражение деятельность самого Аагон'а, главным образом, рисующего себя в процессе рисования самого себя...

Книга МакКордак создает впечатление, что у Аагон'а крайне ограниченное представление о том, что такое люди – не более чем физические объекты, предполагающие способность принимать те или иные очертания. Аагон знает немного о том, как организовывать перспективу изображаемых сцен, но опять-таки, это умеют делать и любые программы для трехмерной графики, хотя последние при этом отнюдь не ассоциируются с моделями человеческого мышления. Скорее всего имеется две причины сходства созданий Аагон'а с художественными творениями: (1) свои рисунки Аагон создает с помощью карандаша или ручки прямо на больших листах бумаги, и (2) создаваемые линии оказываются неровными и нечеткими, в отличие от большинства компьютерных продуктов, стремящихся быть, скорее, фотографически точными. Эти два поверхностных факта делают Аагон куда более похожим на настоящего художника по сравнению с остальными компьютерными программами.

С учетом существования современных мощных программ, ориентированных на геометрическое конструирование, ни в коей мере нельзя считать подвигом создание программы, использующей случайные числа, которая умела бы конструировать воображаемые конфигурации, напоминающие «людей», случайным образом расположенных в пространстве и принимающих различные случайные позы (разнообразие которых ограничено естественными пределами возможностей человеческих тел). Если такую программу совместить с программой для трехмерной графики (опять-таки с использованием случайных чисел), то без особого труда можно добиться того, что волнистые линии будут походить на легкий взмах руки. (AW – искусственное извивание (*artificial wiggleness*) является отнюдь не новой идеей; пользователи буквоформирующей программы *Metafont*¹⁰ в течение долгого времени имели возможность забавляться созданием искусственно изогнутых буквоподобных фигур, получаемых с помощью случайных чисел, причем некоторые из них обретали совершенно человеческие очертания). В результате выйдет нечто весьма похожее на Аагон, однако всё волшебство окажется развеянным, поскольку ничего похожего на симуляцию художественного чутья на этом пути мы не получим.

А, может, я переоцениваю существо человеческой природы? Быть может,

¹⁰ Knuth, Donald E. (1982). “The Concept of a Meta-Font”. *Visible Language*, vol. 16, №1, pp. 3-27.

когнитивной иллюзии, известной как эффект Элизы (более подробно об этом см. Предисловие 4)¹¹, суждено по-прежнему иметь свое затуманивающее воздействие, заставляя людей усматривать осмысленность даже там, где достоверно известно его отсутствие?

Компьютер-писатель

Мне было бы крайне любопытно узнать, действительно ли существует принципиальное различие между изобразительной продукцией Aargon'a и прозаическими/поэтическими продуктами таких программ как Racter (названной по имени автора книги *Полуфальшивая борода полицейского*¹², вышедшей в 1984 году) и Hal (по меньшей мере, соавтор новеллы *Только этот раз*, также обсуждаемой в Предисловии 4)¹³; или, иначе говоря, действительно ли подобные проекты черпают значительную часть своего интереса и кажущейся осмысленности благодаря визуальным или вербальным версиям эффекта Элизы.

С помощью сложной грамматики, реализованной в виде вычислительных формализмов (таких как расширенные сети переходов), и внушительного набора слов с семантическими бирками, прикрепленными к ним, на удивление легко можно добиться весьма любопытных и даже впечатляющих результатов (см. Хофштадтер, 1979, главы 5 и 19)¹⁴. Траектория грамматических соединений выбирается программой случайным образом (разумеется, с учетом встроенных в нее определенных предрасположенностей, также имеющих вероятностную природу), а осуществление выбора ограничено семантическими свойствами уже имеющихся структур. Простейшим примером может служить такая ситуация: если было выбрано слово «пить», то на следующем шаге уже не может быть выбрано слово «шприц», а только лишь одно из слов вроде «молоко», «кофе» и т.п. Этих тривиальных семантических ограничений хватает, однако, для того, чтобы создавать чрезвычайно правдоподобные тексты. Если, вдобавок, включить в программу возможность периодически выходить за рамки семантических ограничений, то можно получить изумительные поэтические эффекты – странные столкновения образов, открывающие простор для метафор.

И всё же очевидно, что ничего похожего на настоящие метафоры здесь нет. Как нет ни каких-либо действительных образов, ни намерения высказать что-либо, ни осознанного отбора слов, и т.п. Причиной схожести с поэзией преимущественно является наш культурный контекст: литература двадцатого века существенно расширила область того, что можно называть поэзией или прозой. Безмерная открытость нашего века, граничащая со «вседозволенностью», вдохновила к созданию прекрасных образцов литературного экспериментирования, однако она же содействовала возникновению всякого рода плутовства, исходящего не только от людей, но и от машин.

Плагиат versus творчество

Проницательный анализ, который проделала Боден относительно той ситуации,

¹¹ Hofstadter, Douglas R. and the Fluid Analogies Research Group. *Fluid Concepts and Creative Analogies: computer models of the fundamental mechanisms of thought*. New York: Basic Books, pp. 155.

¹² Racter (1984). *The Policeman's Beard is Half Constructed*. New York: Warner Books.

¹³ см. ссылку 4.

¹⁴ см. русский перевод книги, на которую ссылается автор: *Хофштадтер Д. Гёдель, Эшер, Бах: эта бесконечная гирлянда*. – Самара: Издательский Дом «Бахра-М», 2001.

которая имела бы место в случае успешной реализации программы Геометрии Гелернтера¹⁵, проливает свет на неопределенность того положения, которое, в ином случае, допускало бы интерпретации, слишком преувеличивающие машинные достижения. Увы, такого рода анализ является редкостью. Художественные, музыкальные или иные достижения компьютеров, с описаниями которых мне доводилось сталкиваться, в особенности в популярной литературе, крайне редко сопровождаются также и описаниями того, как эти программы были сделаны. Почему? Какое значение имеет демонстрация программы, лишенная описания ее устройства?

Предположим некто показывает вам блистательное эссе на тему о юморе и утверждает при этом, что его «написал компьютер»? Ежели в последствии вы узнаете, что это всего лишь отрывок, целиком выданный из книги Артура Кёстлера *Творческий акт*¹⁶, то вы безусловно почувствуете себя обманутыми. Далее, для вас уже не будет иметь никакого значения последующие разъяснения о том, что-де компьютер, хранящий в своей памяти не только книгу Кёстлера полностью, но и сотни миллионов других книг, на писанных на другие темы, *сам выбрал* именно эту книгу и именно этот пассаж и затем распечатал его. В любом случае, перед вами будет несомненный плагиат – возможно, хитроумный плагиат, но только и всего.

Рассмотрим теперь чуть более сложный сценарий. Предположим, что тот же самый пассаж из книги разбивается на части, каждая из которых состоит из сегментов в десятках слов, и каждое слово из каждого сегмента соотносится с соответствующей частью. Вся эта информация загружается в компьютер. Вдобавок ко всему предположим, что компьютер оснащен изощренной английской грамматикой и всевозможными инструкциями, позволяющими располагать имеющиеся сегменты в единый, грамматически правильный пассаж. Программа запускается, и на выходе мы получаем некий результат. Предположим, что он не полностью идентичен оригинальному фрагменту из книги, но всё же несколько сот слов, следующих друг за другом, совпадают полностью. Далее мы слегка корректируем программу, и снова запускаем ее. И в этот раз, *mirabile dictu*¹⁷, результат оказывается целиком тождественным оригинальному тексту Кёстлера. Заслуживает ли этот текст, созданный компьютером, оценки своей значимости? В каком-то смысле да, однако совершенно неверно было бы трактовать его как «написанный» компьютером, поскольку слово «написанный» подразумевает, что конечный текст был создан с нуля.

В данном случае работу компьютера можно сравнить с собиранием пазлов, на которых уже заранее была напечатана картина Моне. Очевидно, что меня следовало бы назвать мошенником, если бы я объявил себя автором полотна великого импрессиониста, поскольку моя роль в его создании сводилась к тому, чтобы подгонять по цвету и форме одни кусочки готового изображения к другим. Также и в случае с созданием осмысленных текстов – никто не припишет компьютеру авторства, касающегося явленных в тексте *мыслей*. Ведь всё, что делает такого рода компьютерная программа, ограничено манипулированием словами, частями речи и грамматическими правилами. Мысли совершенно не требуются для подобной работы.

Но что если оригинальный фрагмент из Кёстлера разбить на более мелкие части? Скажем, на сегменты, состоящие из двух или трех слов? В какой момент произойдет изменение нашего отношения к компьютерному творчеству и вместо ощущения утомительной скуки нас охватит изумление, исходящее от способности компьютера создавать настоящую прозу? Иначе говоря, когда мы сможем заключить, что в игру вошла способность задействовать *идеи*, а не просто лишь *формальные символы*? К

¹⁵ Имеется в виду идея создания «решателя задач» Герберта Гелернтера (Herbert Gelemter), создавшего одну из версий подобной программы в 1963 году.

¹⁶ книга вышла в 1964 году.

¹⁷ Странно сказать.

сожалению, в большинстве статей, посвященных творческим способностям компьютеров, публикуемых популярными изданиями, много говорится о разных удивительных вещах – компьютерных картинах, музыкальных, поэтических и прочих произведениях, – однако, при этом мало уделяется внимания вопросу реализации соответствующих программ, где давались бы разъяснения о строительных блоках, из которых сооружаются их конечные результаты, и о способах их соединения. И тогда было бы понятно, что могут и что не могут описываемые программы. Ведь в знании этих деталей и заключается всё существо дела.

Вообразим себе такую крайнюю ситуацию. Предположим, что мне продемонстрировали новый, поражающий своими следствиями математический результат, или же новое, разверзающее неизведанные глубины музыкальное сочинение, и сказали бы, что всё это было создано компьютером. Предположим также, что у меня есть возможность убедиться в то, что новизна этих творений неподдельна, но с другой стороны, мне столь же достоверно известно, что я никогда в принципе не смогу узнать механизм их создания. Как я должен относиться к этим продуктам. при условии, что мне понятно их новизна и важность? Можно ли говорить о том, что они были созданы творчески? И если да, то кому принадлежит заслуга их осуществления?

Я не вижу простых ответов на эти вопросы. Отсутствие доступа к тому, как были созданы соответствующие программы, делает невозможным их адекватную оценку, и вопрос о том, реализуют ли они собой творческий подход, остается неразрешимым. Моя позиция может показаться кому-то странной, поскольку может возникнуть желание возразить: «При чем тут вопрос о том, *как* был получен результат? Разве не достаточно того, что он вообще был *получен*? Результат следует расценивать как творческий по объективным, независящим от способа создания причинам!». Я так не думаю. Я не могу оценить объект просто потому, что он находится передо мной; мне необходимо иметь хоть какое-то представление об источнике его происхождения.

Вопрос о том, кому следует приписывать авторские заслуги, у меня тоже вызывает чувство растерянности. Авторство должно быть как-то распределено между автором программы и самой программой. Но ежели нам ничего не известно об устройстве программы, то эту разделительную линию провести не представляется возможным.

Механизмы, пробы и тест Тьюринга

Помимо непосредственного доступа к коду программы, имеется еще два пути получения знания о механизме работы компьютерных моделей, симулирующих те или иные аспекты сознательных процессов. Один из них, это, конечно, описание архитектуры программы, которое подразумевается написанным кем-то из создателей самой программы. Такое описание может оказаться как очень полезным, так и двусмысленным и непонятным. Например, проблема в том, что оно может сфокусироваться на неверно выбранном уровне деталей, в результате чего описание может оказаться либо слишком сложным и перегруженным техническими деталями, либо наоборот, будет сосредоточено на аспектах высокого уровня, так что принципы работы программы всё равно останутся неясными. Хуже того, описание может представлять собой такую мешанину этих двух уровней, что в ней можно только окончательно потеряться. Такого рода попури слишком часто встречаются в исследовательской литературе по искусственному интеллекту.

Второй путь является менее сомнительным, хотя и более косвенным. Он заключается в идее ничем не ограниченной коммуникации с работающей программой в течение некоторого периода времени. Путем долгой и систематической серии испытаний программы испытатель может суметь очертить для себя контуры ее возможностей,

выявляющие то, в каких аспектах она является гибкой, а в каких неподатливой.

Именно эта идея лежит в основе знаменитой статьи Алана Тьюринга «Игра в имитацию»¹⁸, философский посыл которой связан с проблемой установления наличия такого неуловимого феномена как «мышление» в ситуации, когда перед вами – черный ящик, и через его посредство некто – быть может, сам этот ящик! – утверждает, что вы действительно имеете дело с мыслящим существом. Тьюринг рассуждал о такой ситуации, когда роль черного ящика играет телетайп, с помощью которого можно получать и передавать стандартные численные и буквенные сообщения. Поскольку он в первую очередь пытался найти ответ (или хотя бы обнаружить обходные пути) на философский вопрос «Что следует считать мышлением?», то он считал, что правильное всего об этом следует говорить именно в терминах черного ящика, играющим роль «мыслителя», а не тогда, когда перед вами заведомо ограниченная или специализированная программа. По его мнению, сюда должна быть включена способность общения на естественном языке, причем без каких-либо ограничений на темы или стили возможной беседы. В своей основе Тест Тьюринга (как он впоследствии стал называться) связан с идеей возможности (или невозможности) понять, имеет ли опрашиваемый дело с человеком или же машиной, целиком полагаясь лишь на результаты общения с черным ящиком. Шутки, аналогии, ссылки, дискуссии об исторических или культурных феноменах, вопросы о глубоко личных убеждениях – всё это и многое другое имеет право быть включенным в игру.

В противоположность этой идее, когда я говорил о систематическом испытании когнитивных моделей, я имел в виду нечто куда менее амбициозное, нежели полная версия теста Тьюринга. Я вовсе не подразумевал ни необходимости общения с программой на естественном языке (хотя это было бы весьма желательно), ни возможности перескакивания из одной области знания в любую другую (хотя и это было бы крайне желательно). Основная идея долгосрочного общения с программой заключается лишь в том, чтобы получить полное представление о ее возможных способах поведения, что позволило бы выявлять наиболее глубокие уровни ее устройства.

Идея того, что скрытые механизмы могут оказаться экспериментально выявляемыми исключительно посредством наблюдения за явными формами поведения может показаться неправдоподобным и даже фантастическим. Тем не менее, как я покажу ниже, подобное допущение лежит в основе всей современной науки, и отрицание этого было бы эквивалентно отрицанию наших обыденных интуиций о природе вещей, а также того, как мы добываем знание об окружающем нас мире. Вследствие сказанного, мне представляется чрезвычайно важным подумать о существовании таких слов как «механизм» и «испытание». Этим мы и займемся в оставшейся части Эпилога.

Структуры мозга versus когнитивные механизмы

Когда я утверждаю (что я делаю весьма часто), что «Сорусат – это когнитивная модель», то что же здесь имеется в виду? Утверждаю ли я, что Сорусат в каком-то смысле является моделью человеческого мозга? Большинство людей – и даже большинство специалистов по когнитивным наукам – будут склонны сказать «нет». Однако же перед тем, как отбросить эту идею, давайте рассмотрим спектр возможных значений понятия «модель мозга».

Когда спрашивают «является ли Сорусат моделью *мозга*?» или же утверждают, что «*человеческий мозг* может мыслить», то на самом деле здесь всякий раз употребляют одно и то же существительное для обозначения миллиардов различных физических

¹⁸ Turing, Alan M. (1950). “Computing Machinery and Intelligence”, *Mind*, vol.59, №236.

объектов. Нет сомнений, что именно благодаря этим физическим объектам, а не неким не-физическим платонистским категориям (вроде «человеческого мозга»), осуществляется мышление. Употребляемый без всяких сомнений единый термин «мозг», говорит о наличии молчаливо принимаемой предпосылки о существовании некоего абстрактного (хотя и никогда явно не оговариваемого) уровня описаний, общего для всех разновидностей человеческого мозга. Поэтому, будучи проанализированным, предложение «человеческий мозг способен мыслить» на самом деле означает одну весьма туманную вещь: «существуют абстрактные универсальные механизмы, вновь и вновь разнообразно репрезентируемые в мозгу различных индивидуумов, благодаря которым может иметь место мышление». Коротко говоря, механизмы мозга, ответственные за возникновение мышления, не являются *буквально* «железом»; скорее, это своего рода структуры, являющиеся чем-то средним между «железом» и «программным обеспечением».

Сегодняшнее преклонение перед нейронными сетями заставило многих исследователей предположить, что механизмы мышления больше похожи на «железо», в противоположность тому мнению, к которому склонялись исследователи в шестидесятых и семидесятых годах. По этой причине возникает всё больше оснований считать когнитивную науку формой исследования, направленной на изучение *структур мозга*, в противоположность исследованиям *ментальных механизмов*. Однако же, не так-то просто четко провести соответствующее различие.

Будучи очень сложной системой, мозг (будем пока продолжать использовать этот туманный термин), содержит множество различных типов структур, упорядочивающих совершенно различные его иерархические уровни. К примеру, можно назвать хотя бы некоторые из них:

- ядра атомов
- молекулы воды
- аминокислоты
- нейротрансмиттеры
- синапсис
- дендриты (ленточные кристаллы)
- нейроны
- кластеры нейронов
- вертикальные колонны зрительной коры головного мозга
- более крупные зоны (такие как зона 19) зрительной коры
- вся зрительная зона коры мозга
- всё левое полушарие

Какие структуры среди этих и многих других «структур мозга» ответственны за возникновение феномена мышления? Никто не знает точного ответа. Интересно, тем не менее, что в последнее время популярная литература заиклилась на некоторых нейробиологических экспериментах (связанных с изменением синаптических весов у различных существ), о которых, затаив дыхание, говорят как о «раскрытии секретов памяти». В итоге же, за этой драматической развязкой не стоит ничего кроме (*буквально бессмысленной*) предпосылки о том, что человеческая память не имеет никакого отношения к *организационным* уровням – т.е. что для раскрытия всех секретов устройства памяти достаточно быть просто химиком, способным иметь дело исключительно с теми механизмами, которые стоят за локальными микроскопическими химическими изменениями. Все психологические понятия отбрасываются как ненужные при таком подходе.

Тот факт, что существует множество уровней и типов хорошо изученных

физических структур, лишь частично рассеивает туман вокруг термина «структуры мозга». Для полноты картины необходимо учитывать и другой тип «структур мозга», представители которого обозначаются следующими понятиями:

- концепт *собаки*;
- ассоциативная связь между понятиями *коровы* и *молока*;
- файл объектов (an object file), связанный с воспринятым объектом (в том смысле как это обсуждается психологом восприятия Анной Трисман¹⁹);
- «геоны» и «2½-D эскизы» (как они обсуждаются в различных моделях зрения);
- фреймы, скрипты и схемы;
- «связки, организующие память» и «тематически организованные точки», как это описывается в модели памяти Роджера Шэнка²⁰;
- кратковременная (оперативная) память;
- долговременная память;
- различные «агенты», «К-линии», «немые» и «номы», о которых рассказывает Марвин Мински²¹ в своей модели «организации ума» (социум ума, society of mind)
- коделеты (codelets), неотложные задачи (urgencies), концептуальные смещения (slippages), концептуальные дистанции (conceptual distances), концептуальные глубины (conceptual depths), сцепления (bonds), описания, мосты, группы, уровни прочности, и температура (в том смысле, в каком эти термины используются в наших проектах Soruscat, Tabletop и др.).

Этот список представляет собой лишь малую толику тех компонентов, которые возникают на разных уровнях в различных теориях мышления. И хотя на первый взгляд каждый из них может показаться лишь косвенно связанным со структурами мозга, всё же физическое устройство мозга не может не быть снабжено соответствующими механизмами, коль скоро каждый из компонентов предполагается действительно значимым аспектом сознания. Каждая из этих сущностей уже в ближайший десяток лет может получить по крайней мере *некоторое* физическое обоснование, приблизительно в том же смысле, в каком чисто теоретическое понятие *гена* получило однажды свою физическую интерпретацию благодаря пониманию устройства ДНК. Поэтому элементы второго списка обладают не меньшим правом называться «структурами мозга», нежели элементы первого списка, хотя и в несколько ином смысле.

И тем не менее сегодня нет никакой уверенности в том, как должен выглядеть оптимальный уровень, описание которого являлось характеристикой универсальных структур мозга (или когнитивных механизмов). Другие науки тоже столкнулись с проблемой спецификации оптимального уровня описания, и некоторые из них сумели удовлетворительно решить эту проблему. Если, например, предметом исследования являются газы, то со многих точек зрения наилучшим уровнем описания является макроскопический, подчиняющийся законам термодинамики, несмотря на то, что газы состоят из астрономического числа молекул и вообще-то могут быть описаны также и на более низком уровне статистической механики. Подобным образом, если кто-то заинтересован в описании механизмов передачи наследственности при помощи ДНК, то это лучше всего достигается на уровне описания таких носителей информации как кодоны, при этом многие, если не все, низкоуровневые детали химического устройства ДНК могут быть полностью опущены. Любопытно, что для описания ДНК на разных

¹⁹ Treisman, Anne (1988). "Features and Objects: The Fourteenth Barlett Memorial Lecture". *Quarterly Journal of Experimental Psychology*, vol. 40A, pp. 201-237.

²⁰ Schank, Roger C. (1980). "Language and Memory". *Cognitive Science*, vol. 4, №3, pp. 243-284.

²¹ Minsky, Marvin L. (1985). *The Society of Mind*. New York: Simon & Shuster.

уровнях используется разная терминология: ДНК как физическая структура называется «двойной спиралью», а ДНК как носитель информации называется «геномом» организма.

Представители когнитивных наук рассчитывают достичь аналогичной утонченности в своих рассуждениях о работе «мозга» на разных уровнях. В частности, необходимо знать, какой уровень описания наиболее адекватен при исследовании мышления. Возможно, по аналогии с описанием ДНК, следовало бы говорить о мозге как о физической структуре в терминах «двойного полушария», а о мозге как о носителе памяти с использованием термина «тетоме».

Некоторые уровни описания являются заведомо слишком низкими

До сих пор в когнитивных науках считаются приемлемыми рассуждения о том, что некоторые типы дедуктивных формализмов (таких как исчисление предикатов или более новые языки репрезентации как, например, KLONE²²) каким-то образом вмонтированы в нейроны, и что реальным критерием мышления должен быть сам формализм, а не факт его встроенности в нейронное «железо». Это предположение ничем не отличается от того трюизма, что, например, текстовый редактор является текстовым редактором (в отличие, скажем, от видео игры или программы, предсказывающей погоду) благодаря тексту программы, а не тому компьютеру, в который она загружена. По той же причине «структуры мозга», ответственные за мышление, (по крайней мере, в принципе) могут оказаться чем-то весьма абстрактным (например, похожим на язык KLONE), а не конкретным (какими являются, например, детали нейронных соединений). Сегодня многие ученые когнитивисты, пожалуй, найдут эту идею маловероятной. Однако ее не следовало бы полностью сбрасывать со счетов потому, что пока еще достоверно не известно, какой уровень или тип «структур мозга» окажется наиболее релевантным для описания структур, делающих возможным мышление.

В любом случае, практически все когнитивисты, независимо от того, что они думают о KLONE и тому подобных формализмах, *принимают* то обстоятельство, что вопрос «какой уровень описания мозга подходит для описания процессов мышления?» является осмысленным и что он имеет осмысленный ответ на него. И тогда детали описания мозга, лежащие *ниже* этого критического уровня, должны предполагаться несущественными. По этой причине большинство когнитивистов практически уверены в том, что для достижения их исследовательских целей не имеет смысла изучение молекулярной биологии, квантовой механики или строения кварков. (И всё же не лишне было бы заглянуть в работу Конрада²³ и др., где предпринимается попытка рассмотрения высокоуровневых процессов мышления с точки зрения их связи с молекулярной биологией энзим и их взаимодействий!).

Многое бы прояснилось, если кому-то удастся создать думающее существо, субстрат которого сильно отличается от нейронного. Это может произойти, а может и нет. В случае успеха это открытие многое объяснило бы в том, какой уровень механизмов необходим для того, чтобы быть носителем мысли. В этом случае мы поймем, что мышление зависит от критического уровня X и более высоких уровней, в то время как для уровней ниже критического уровня вполне пригодны различные субстраты.

²² см. например, статью R. J. Brachman and J. G. Schmolze: An overview of the KLONE - knowledge representation language, *Cognitive Science*, Vol. 9, 1985, pp. 171-216.

²³ Conrad, Michael *et al* (1989). "Towards an Artificial Brain". *BioSystems*, vol. 23, pp. 175-218.

Необходимость критерия для идентификации мышления

Если мы допустим хотя бы только *вероятность* того, что некий не-мозг также может быть способен к мышлению, то нам немедленно понадобится набор критериев для опознания системы в качестве мыслящей – иначе мы просто не сумеем отличить действительно мыслящую систему от иного типа систем. К счастью, в нашем распоряжении имеется несколько намеков на то, какими должны быть эти критерии. Мы распознаем человеческую доброту, наблюдая добрые поступки, а не путем ДНК-анализа, направленного на поиск «гена доброты», и не путем сканирования человеческого мозга в реальном времени для того, чтобы удостовериться в том, что «эмпатический центр» был действительно задействован. Мы опознаем текстовый редактор по тем событиям на экране дисплея, которые происходят в ответ на наши манипуляции с клавиатурой, а не по тому «железу», на основе которого он работает. У нас есть поведенческие критерии человеческой доброты. Не так трудно установить и поведенческие критерии для распознавания программ, являющихся текстовыми редакторами. Тьюринг в своей известной работе попытался установить аналогичный набор критериев для распознавания мышления.

По аналогии с критериями для распознавания доброты или же текстового редактора Тьюринг решил использовать набор высокоуровневых поведенческих критериев в противоположность низкоуровневым критериям, связанным с устройством «носителя». Тьюринг был чужд догматического предположения о том, что раз уж мозг состоит из нейронов, то и единственным уровнем описания любой мыслящей системы должен быть нейронный уровень. Наоборот, в своих рассуждениях он полагается на те интуитивные представления, которые подразумеваются людьми при разговорах о мышлении – а именно, *гибкое (fluid) манипулирование идеями*. Тьюринг решается *a priori* не привязывать способность к таким манипуляциям ни к какому конкретному «носителю»; он оставляет открытым вопрос о том, какой уровень описания механизмов окажется решающим.

Когда все результаты по исследованию мозга завершатся, то какой способ описания «человеческого мозга» окажется наиболее адекватным? В терминах стандартных типов нейронных соединений? Или стандартных типов нейронных кластеров? Или стандартных типов взаимодействий между нейронными кластерами? Возможно, что детали устройства нейронов окажутся не играющими никакой роли. Вероятно, однако, и то, что даже устройство нейронных кластеров окажется не релевантным, как и органическая химия или квантовая механика. В таком случае будет доказано, что основание мышления принадлежит не биологии, а уровню абстрактных организационных принципов, напоминающим, например, программное обеспечение. Кто знает, – возможно, даже обнаружится, что соответствующими «структурами мозга», необходимыми для осуществления гибкого манипулирования идеями, окажутся структуры уровня программ *Coruscant*, *Tabletop* и *Letter Spirit*, или более низкого уровня.

Эксперименты рассеяния versus «прямое» наблюдение феномена

Впервые столкнувшись с Тестом Тьюринга можно подумать, что он позволяет осуществлять только высокоуровневое тестирование, оставляя закрытым доступ к «подсознательным» и «суб-символическим» механизмам, не говоря уж о доступе к нейронному уровню. Многие и в самом деле полагают, что Тест Тьюринга в лучшем случае касается лишь поверхностных механизмов, которые стоят за тем «поведением», на тестирование которого «лишь только» и нацелен Тест. Эта проблема – т.е. вопрос соотношения «только лишь поведения» со скрытыми механизмами, а также и вопрос о

глубине, которой может достигать Тест Тьюринга при исследовании скрытых механизмов – подводит нас к следующему вопросу: «Где располагается граница между «прямыми» и «непрямыми» наблюдениями?».

В начале этого века физик Эрнест Резерфорд направлял пучок альфа частиц через тонкий кусок золотой фольги и анализировал статистические данные об углах их рассеяния. Исходя из этих (макроскопических) наблюдений, он сделал вывод, что атомы золота – это ядра, окруженные электронами. Это открытие привело к современным представлениям об «атоме» (одной из проблематических сущностей). Его заключение, разумеется, зависело от математической теории электромагнитного рассеяния, предсказывающей, что такая-то и такая-то микроскопическая структура будет служить причиной такого-то и такого-то макроскопического углового распределения рассеиваемых частиц.

Эксперименты Резерфорда с рассеянием частиц были настолько важными и настолько регулярно повторяемыми, что техника рассеивания стала стандартной, неотъемлемой частью репертуара экспериментальной физики. То, что вначале казалось смутным и *непрямым* наблюдением связи между макроскопическими образцами и их микроскопическими причинами, стало в конечном счете рассматриваться физиками как очевидное и *прямое* наблюдение. Конечно, итоговые результаты экспериментов не столь же непосредственны, какими были бы *картинки* непосредственно созерцаемых структур – ведь они представляют собой математические формулы или же словесные описания. Тем не менее, этого было достаточно для физиков.

С недавних пор, правда, в связи с феноменальным возрастанием мощностей вычислительной техники, стало возможно осуществлять эксперименты рассеяния с помощью электронных и рентгеновских микроскопов, в результате чего можно фактически получать *фотографии* объектов размером с молекулу или даже атом. В принципе эти расчеты ничем не отличаются от тех, которые проводил Резерфорд. Однако число подобных экспериментов столь велико и времени на их проведение требуется столь мало, что теперь их результаты кажутся качественно отличными от результатов исходных экспериментов. Разумеется, это лишь иллюзия нашей интуиции.

Сегодня, например, ультразвук позволяет нам видеть плод в утробе матери, движущийся в реальном времени. Заметьте, что у нас не возникает стремления поставить в кавычки глагол «видеть», так же, как мы не ощущаем нужды ставить в кавычки слово «разговаривать» в предложении «я каждый день разговариваю со своей женой по телефону». В обыденных предложениях такого рода, мы совершенно не смущаемся тем фантастическим обстоятельством, что наши голоса движутся в совершенной тишине по металлическим проводам. Реконструкция звуков речи на том конце провода оказывается столь точной и достоверной, что это заставляет нас забыть о сложнейших процессах кодирования и декодирования, совершающихся где-то между говорящим ртом и слушающим ухом.

Сотню лет назад телефонный контакт – если угодно «голосовая телепортация», – казался чем-то поразительным, почти чудом потому, что «система трюков», позволяющая это делать, была столь новой и необычной, что ее просто нельзя было игнорировать. В каком-то смысле, такую реакцию следовало бы назвать наиболее *соответствующей*, но сегодня очень трудно удивить кого-либо подобным чудом – уж слишком привычным и распространенным оно стало. Если бы пятьдесят лет назад было осуществлено рассеивание высокочастотного звука, вызванное плодом в утробе, то вследствие отсутствия технологий, позволяющих конвертировать рассеянные волны в отчетливую телевизионную картинку, всякое заключение относительно измерений рассеянных волн имело бы характер туманных математических выкладок. Сегодня же, просто потому, что высокоскоростные компьютеры позволяют реконструировать исследуемый объект исходя

из результатов рассеивания практически в режиме реального времени, у нас создается ощущение *непосредственного* наблюдения плода. Примеры такого рода – а их можно приводить тысячами в наш технологический век – показывают, что граница между «прямым наблюдением» и «опосредованным выводом» является делом субъективным.

В действительности значительная часть научного прогресса состоит в размывании этой, поначалу кажущейся резкой, границы. То, что вначале кажется весьма искусным и утонченным спекулятивным умозаключением, становится чем-то общепризнанным и подвергается дальнейшей стандартизации, компьютеризации и т.п.; с этих пор соответствующие результаты кажутся непосредственно наблюдаемыми и принимаются как само собой разумеющиеся. В этом же направлении шел переход от оптического телескопа к радиотелескопу, от оптического микроскопа к электронному микроскопу и наблюдению за рассеянием частиц. Как и в случае с ультразвуком, вопрос о том, имеем ли мы здесь дело с «прямым наблюдением» или нет, зависит от тех интуиций, которые фактически определены лишь степенью ясности и отчетливости визуальных картинок, производимых соответствующим компьютером. Сегодняшние туманные выкладки – это завтрашние открытые окна в природу!

Тест Тьюринга и доступ к скрытым механизмам

Все наши рассуждения теперь можно применить к исследованию Теста Тьюринга. Действительно, ведь Тест Тьюринга вполне можно понять как своего рода эксперимент рассеяния, при котором мыслящее устройство невольно раскрывает свои (микроскопические) скрытые механизмы с помощью ответов на поставленные вопросы (которые играют роль сталкивающихся с ним волн или частиц). Уровень детализирования, доступный при подобном тестировании, ничем не ограничен. Для получения информации о более и более глубоких уровнях его устройства необходимо лишь уметь задавать всё более и более изощренные вопросы, так же как и иметь всё более и более утонченные способы интерпретации ответов.

Изучение языковых ответов новыми способами аналогично исследованию звездных спектров новыми методами (например, с использованием новых областей электромагнитного спектра, более высоким уровнем разрешения, с использованием множества согласованных по времени отдаленных приемных устройств и т.п.), благодаря которым делаются всё более тонкие выводы о деталях механизма звезд, несмотря на удаленность исследуемой звезды от нас на сотни световых лет. Некоторые способы анализа языковых ответов, исходящих от неизвестного устройства при его тестировании Тестом Тьюринга, могут учитывать:

- *частоту встречающихся слов* (например, является ли определенный артикль «the» наиболее часто встречаемой единицей речи? Является ли слово «время» наиболее часто встречающимся существительным? Не встречаются ли неестественно часто малоупотребительные слова? И наоборот, не возникнет ли каких подозрений у «отвечающего устройства» в том случае, если опрашивающий его человек неестественно часто будет использовать малоупотребительные слова в своих вопросах?);
- *чувствительность к манере речи* (например, понятны ли «отвечающему устройству» формальные и сленговые выражения? Будет ли казаться смешным неподобающее смешение способов высказывания? Выскажет ли «устройство» свои подозрения относительно подобных неправомерных смешений, нарочито совершаемых в задаваемых ему вопросах? Соблюдено ли в ответах должное соотношение между формальным и неформальным уровнями речи или же имеет место интуитивно неестественное сочетание того и другого? [поразмышляйте,

- например, над курьёзными ответами программы Racter]);
- анализ *типов ошибок* (описки, неправильные перестановки, неверное употребление слов или фраз, смешения любого типа, грамматические ошибки, и т.п., которые – как известно любому исследователю по когнитивистике – говорят очень многое о механизмах мышления);
 - анализ *выбора слов из синонимического ряда* с учетом тонкостей используемого контекста (например, какие контекстуальные детали заставляют использовать слово «jock», а не слово «athlete», или наоборот? Или же использовать слово «lady», а не «woman», слово «endeavor», а не «try», слово «attempt», а не слово «strive»?);
 - анализ *уровня абстрагирования* (например, что является основанием выбора между словами «Фидо», «собака» и «млекопитающее»? Или между словами «этот пешеход», «этот парень» и «он»? Или между словами «кресло-качалка», «кресло», «место для сидения», «мебель» и «вещь»?);
 - анализ *предпосылок*, имеющих место *по умолчанию* (какие обстоятельства вынуждают использовать женские окончания у таких слов как «heroine», «millionaires» или «farmerette»? при каких обстоятельствах используются такие общие термины как «человек» и «он»? в зависимости от обстоятельств, какой род атрибутируется таким нейтральным терминам как «пешеход» или «врач»?);
 - то, как понимаются и порождаются *«проглатываемые», неподходящие по смыслу значения* (например, верно ли, и мгновенно ли понимается скрытое за непосредственными смыслами значение фраз типа «Бывало я и такое тоже делал» или «С вами никогда не случалось такого?» соотносятся ли подобные фразы со стандартными контекстами их употребления или нет?);
 - то, как понимаются и порождаются *условные контрафактические* предложения (например, правильно ли, и мгновенно ли интерпретируются двусмысленные фрагменты предложений типа «Я бы не чувствовал себя подобным образом, если бы я был моим отцом» или же «А что бы вы сделали, если бы вы были моими родителями?» Соотносятся ли подобные фразы со стандартными контекстами их употребления или нет?);
 - *временные факторы* («отвечающее устройство» может выводить на экран свои ответы знак за знаком, строка за строкой, фраза за фразой, но в любом случае скорость порождения ответа несет в себе информацию о порождающих ее процессах) и т.д., и т.п.

Этот список можно продолжить и разработать с большой степенью детализации, но у нас сейчас нет ни возможности анализировать обширные списки примеров, которые могли бы пролить свет на скрытые механизмы «мыслящих устройств», ни защищать законность и эвристическую значимость подобных подходов²⁴. Вполне достаточно упомянуть то, что многие из них уже весьма разработаны, а некоторые представляют собой общепризнанные технические приемы когнитивных наук (в частности, когнитивной психологии).

Все, кто всерьез убежден в основательности Теста Тьюринга, исходит из возможной утонченности тех тестов, с которыми он связан. Астрономы и физики знают, что наблюдаемые природные явления, источник которых отдален от человеческой реальности пространственно или по масштабу, будучи интерпретированы надлежащим образом, могут эффективно использоваться для раскрытия «скрытых» механизмов. Подобным же образом, ученые-когнитивисты могли бы учитывать значимость аналогичных процедур при исследовании психики. Короче говоря, Тест Тьюринга, если его интерпретировать надлежащим образом, может быть использован для тестирования

²⁴ French, Robert M. (1990). “Subcognition and the Limits of the Turing Test”. *Mind*, vol. 99, № 393, pp. 53-65.

ментальных механизмов произвольного уровня глубины и любого уровня сложности.

В духе лучших образцов современной науки Тест Тьюринга размывает кажущуюся резкой границу между исследованием поведения и исследованием стоящих за ним скрытых механизмов, а также и кажущуюся резкой границу между «прямым» и «опосредованным» наблюдением, напоминая нам об искусственном происхождении подобных разграничений. Любая компьютерная модель мышления, которая сумеет пройти полный Тест Тьюринга (т.е. такой, который действительно окажется способным анализировать самые фундаментальные механизмы мышления) будет безусловно согласовываться со «структурами мозга» на всех уровнях, релевантных для того, чтобы мышление имело место.

Тест Тьюринга и фундаментальные исследования

Недавно был учрежден денежный приз (the Loebner Prize) для той программы, которая первой сумеет пройти Теста Тьюринга в его ограниченной версии. К сожалению, несмотря на забавность и восхитительность этой затеи, подобное соревнование представляется мне крайне преждевременным. Если те люди, которые ведут опрос компьютера, не сумеют достичь подлинно глубокого уровня вопрошания, то все эти соревнования выльются лишь в погоню за всё более и более изощренными играми с компьютеризированным естественным языком, сущность которых едва ли будет обеспечена действительным мыслительным содержанием. А это будет постыдным провалом, ибо если внимательно изучить даже те модели, которые оперируют с наиболее простыми микрообластями, наподобие тех, что изучались в предыдущих главах нашей книги, то нетрудно убедиться в том, что и им еще очень далеко до той гибкости и текучести, которые присущи подлинному мышлению. В действительности нужно было бы учредить приз за достижения в области фундаментальных исследований, а не приз за удачное украшение фасада.

Исследовательские проекты, описанные выше, представляют собой сознательную попытку работать в рамках предельно простых, но фундаментально значимых областей, значительно отдаленных по уровню сложности от устройства естественного языка. Большинство проектов по созданию искусственного интеллекта нацелено, как правило, на оперирование в областях, близких по сложности к «реальному миру», от чего и страдают тем недостатком, что схватывают лишь верхушки исследуемого айсберга. Наши же проекты нацелены на то, чтобы смоделировать *сущность* небольшого числа концептов, природа которых искусственно упрощена. И хотя совершенно ясно, что ни один из наших проектов даже близко не подходит к тому, чтобы пройти Тест Тьюринга, мы всё-таки надеемся, что проводимые нами исследования сумеют указать тот путь, на котором в отдаленном будущем будут найдены архитектурные решения программ, реализующих собой сущность подлинно гибкого мышления. Именно так представлял себе природу мышления Алан Тьюринг, излагая свои идеи о знаменитом Тесте.

**Перевод Н.Н.Мурзина и К.А.Павлова;
вступительная статья и примечания – К.А.Павлова.**