

Ордена Ленина ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
им. М. В. Келдыша
Российской Академии наук

А. В. Ворожцов

Критерии интеллектуальности искусственных систем

Москва
2011

АННОТАЦИЯ

Работа посвящена критериям интеллектуальности искусственных систем. Тест Тьюринга рассматривается как базовый тест, который следует разделить на более конкретные, чётко формализованные критерии интеллектуальности, определяющие наличие той или иной стороны “интеллектуальности”. В работе рассмотрено два новых критерия интеллектуальности — “Универсальный Игрок” и “Универсальный Ученый”. В качестве основного критерия интеллектуальности предлагается способность осуществлять метасистемные переходы на уровне мышления. Исследуется несколько возможных механизмов метасистемных переходов.

ABSTRACT

What are the criteria of intelligence of artificial systems? Turing Test (TT) was the starting point of testing intelligence. Now it is understood, that TT should be divided into separate tests measuring different aspects of intelligence. We propose two more tests “Universal Player” and “Universal scientist” aimed to show up capacity of artificial systems for making metasystems transitions. We consider some candidates for engine of metasystem transitions. One of them is “self-organization”: automatic adjusting of parameters to critical values.

Работа поддержана грантом РФФИ 04-01-00510.

1 Введение

Когда-то главной задачей науки об интеллектуальных системах (ИС) была задача создания искусственного интеллекта (ИИ) — искусственной системы, равносильной человеческому мозгу¹. Сегодня многие ученые относят эту задачу к разряду нерешаемых и предпочитают заниматься более конкретными задачами, связанными с созданием в той или иной степени интеллектуальных систем.

Нельзя сказать, что нет систем, равносильных по мощи человеческому интеллекту — есть программы, которые играют в шахматы, легко обыгрывают среднего человека и играют на равных с чемпионом мира. Есть также программы, которые с легкостью находят корни многочленов 100-й степени, аналитически решают дифференциальные уравнения, распознают образы, предсказывают землетрясения, анализируют состояние больного на основе показателей медицинских приборов и управляют самолетом лучше квалифицированного пилота. К тому же на компьютерах можно хранить в миллионы раз более информации нежели помещается в человеческом мозге. В этом нет ничего удивительного, как нет ничего удивительного в том, что человек способен создавать объекты больше его по размерам. То, что электронные помощники оказались быстрее, точнее и надежней человеческого мозга, не означает, что они умнее. Искусственные системы давно уже опередили человека по силе, мощи, ловкости, скорости движения, а сегодня они обогнали человека по скорости вычислений и размерам памяти.

Не поддается воспроизведению одно удивительное свойство человеческого интеллекта: умение увеличивать свои интеллектуальные способности с помощью метасистемных переходов². Это свойство имеет множество интерпретаций: способность решать новые задачи и разрабатывать новые методы, способность создавать метамоделли, способность к рефлексии, способность к индукции и дедукции, способность выявлять классы задач и находить общий метод решения.

Вопрос о возможности создания искусственного интеллекта в значительной степени философский. Большинство споров происходят из-за нечёткой терминологии и отсутствия общепринятого критерия интеллектуальности. Как только кто-нибудь заявляет, что создал интеллектуальную систему (обучающуюся, когнитивную, эволюционирующую или еще какую-нибудь), сразу же

¹Термин *интеллект* (*intelligence*) происходит от латинского *intellectus*, что означает ум, рассудок, разум; мыслительные способности человека. Соответственно *искусственный интеллект* обычно толкуется как свойство искусственных систем брать на себя отдельные функции интеллекта человека, например, выбирать и принимать оптимальные решения на основе ранее полученного опыта и рационального анализа внешних воздействий

²Термин “метасистемный переход” был введен В.Ф. Турчиным [?, ?]. Он означает переход от системы объектов к системе, элементарными объектами которой служат системы с предыдущего уровня.

найдется оппонент, который скептически скажет: “Каждый компонент вашей системы работает вполне определенным, заранее известным образом. Ваша система в целом есть не более, чем алгоритм; достаточно сложный, но все таки явно формализованный алгоритм, но никак не интеллект”. Другими словами, попытки создания ИИ как чётко формализованной искусственной системы обречены на провал уже потому, что многие в понятие интеллекта вкладывают нечто неформализуемое, иррациональное.

Отсутствие веры в возможность формализации мышления не есть прихоть гуманитариев. Для многих естественников это вынужденный вывод из опыта создания искусственных систем. Дело в том, что нельзя создать систему способную (потенциально) решать любую новую задачу, а именно это свойство связывают с настоящим Интеллектом, интеллектом с большой буквы (заметьте, что не каждый человек способен решить действительно новую задачу). Невозможно создать вычислительную машину, способную самостоятельно и неограниченно, так сказать, на лету, менять свою собственную архитектуру. А именно такого сорта задачи встают перед теми, кто пытается создать ИИ.

Утверждение 1 *Основное свойство интеллекта — это способность на уровне мышления осуществлять метасистемные переходы. Именно это свойство сложно реализовать в искусственных интеллектуальных системах.*

Искусственные системы как были, так и останутся бездушными машинами, без сознания и мотивации, без эмоций, без интуиции, без мгновенных озарений и без “общего видения ситуации”, без того, что англичане называют insight. И кажется, здесь следует поставить точку и в дальнейшем разговаривать не о ИИ, а о в той или иной степени сложных системах, решающих определенный класс задач. Например, можно развивать науки, связанные с нейросетями [?], генетическими алгоритмами [?], обработкой естественного языка, базами знаний и экспертными системами. Но в этой работе мы снова поднимем вопрос об ИИ в самом общем виде и обсудим следующие моменты:

- Интеллект и его компоненты, шкала интеллектуальности и критерии интеллектуальности.
- В чём заключается то непреодолимое препятствие, которое помешало создать ИИ?

2 Тест Тьюринга “Универсальный собеседник”

Алан Тьюринг первый поднял вопрос об ИИ (“Может ли машина мыслить?”³, 1950 г.). В своей работе он попытался найти границы интеллектуальности

³Статья была впервые опубликована в научном журнале Mind, v. 59 (1950), pp. 433–460, под названием Computing Machinery and Intelligence и перепечатана в 4-м томе “Мира ма-

вычислительных машин и сформулировал первый критерий интеллектуальности:

Тест 1 (Тест Тьюринга) *ИИ должен уметь разговаривать как человек, и при разговоре с ИИ человек не должен догадаться, что разговаривает с искусственной системой, другими словами, ИИ должен уметь в разговоре имитировать человека.*

Вызов написать программу, способную вести беседу, был принят, и сегодня вы можете пообщаться с компьютерной программой по ICQ или в “чате” и даже не заметить этого. Но если ваша беседа затянется, то вы легко распознаете искусственного собеседника, поскольку он беден на сложные ассоциации, быстро теряет контекст разговора и не способен рассуждать абстрактно.

Тест Тьюринга (ТТ) включает в себе очень правильную и естественную идею. Мозг — это черный ящик, как он работает, наука на настоящий момент не знает (знает очень много, но воспроизвести целиком его работу пока не может). Нет необходимости накладывать какие-то ограничения на внутреннюю структуру ИИ, важно лишь его поведение — то, что он умеет делать. Что должен уметь делать интеллект? — конечно, рассуждать. Всякая мыслительная деятельность человека в момент своего выражения проходит через вербальную стадию, естественно требовать того же от ИИ.

Простота и естественность ТТ обманчива. Какого именно человека должен уметь имитировать ИИ? Имитировать лепет пятилетнего ребенка — это задача одной сложности, а профессора математики — совсем другой. Разговаривать “не о чём” не получится — именно беспредметным, несвязным разговором и выдают себя программы, имитирующие собеседника. Оказалось, что лучшая стратегия для искусственного собеседника не “расколоться” — это поменьше говорить и побольше задавать вопросов, то есть быть, в некотором смысле, искусственным психотерапевтом-слушателем, который позволяет собеседнику выговориться, изображает заинтересованность, хотя, возможно, даже не улавливает логику предложений, не говоря уже о сути, которую ему хочет поведать пациент. Такого сорта хитрости нас сейчас не интересуют. Очевидно, что ТТ слишком общий и охватывает широкий спектр тестов. В беседе можно проверить интеллектуальные способности самого различного типа — логику, геометрическое мышление, память, знание о мире — всё, кроме способностей, связанных с обработкой видео, аудио и другой невербальной информации. В этом смысле тест Тьюринга можно назвать тестом “Универсальный собеседник”.

Первое, что предполагает ТТ, — это понимание естественного языка. Это понимание может означать, что ИИ

тематики” Аж. Р. Ньюмена “The World of Mathematics. A small library... with commentaries and notes by James R. Newman”, Simon and Schuster, NY, v. 4, 1956, pp. 2099-2123), где опубликована под заголовком “Can the Machine think?” С этого издания был сделан русский перевод [?]

- 1) знает *структуру предложения и языка*, умеет осуществлять разбор предложения;
- 2) имеет *тезаурус языка*, то есть имеет словарь слов и знает правила склонения и образование различных форм;
- 3) имеет *смысловую онтологию языка*, то есть своеобразный толковый словарь слов и словосочетаний, который в действительности является набором связей – синонимических, антонимических, и различного рода функциональных связей типа “является частным случаем”, “содержит в себе как элемент”, “является обобщением”. Важно, чтобы эта смысловая онтология была *контекстно-ориентированная* — смысл слов и смысловые связи между словами могут сильно зависеть от контекста (темы) разговора. Для слов с неоднозначным смыслом эта онтология должна помогать по контексту разговора определять значение, в котором было употреблено это слово. Полезным было бы на уровне онтологии реализовать разбиение слов на рубрики — например, некоторые из слов относятся к рубрике “технологии” (слово “компьютер”), другие – к рубрике “спорт” (слово “дзюдо”), третьи с разной вероятностью встречаются в разных рубриках, а четвертые являются словами общего назначения (слова “был”, “может”, “большой”, “после”)

Самым объёмным и сложным для реализации является третий пункт. Все три пункта можно объединить в один — **знание устройства языка и смысловых связей между элементами языка**. После того, как в ИС заложено это знание (не видно причин, по которым этого нельзя было бы сделать), можно говорить о некотором интеллектуальном ядре, его использующем, и накладывать требования на содержание разговора ИС с живым человеком. И здесь мы встречаемся со следующей задачей — выделить ключевые пункты на пути приобретения ИС **общего знания об устройстве мира и способности проводить логические рассуждения**, или, грубо говоря, отметить несколько реперных точек на шкале “общечеловеческой разумности” между маленьким ребенком и условным профессором математики, физики, психологии, философии и других наук. Этих точек можно поставить много. Естественно их формулировать как способность отвечать на вопросы типа:

Понимание смысла: Какого цвета белая лошадь?

Арифметика: Сколько будет $13 \cdot 12 - 11$?

Простая математика: Простое ли число 113?

Логика: У Дон Жуана спросили “Верно ли, что если вы любите Бетти, то вы любите и Джери?”. Он ответил “Если это верно, то я люблю Бетти.” Любит ли Дон Жуан Бетти, если он сказал неправду?

Представление о времени и пространстве: Можно ли вывернуть мячик наизнанку, не разрезая его?

Пространственное мышление: Что будет с листом Мебиуса, если его разрезать вдоль посередине?

Простая физика: С какой скоростью упадет камень упавший с высоты 10 метров?

Извлечение неявно заданной информации: Каждый день, начиная с 10 августа, господин Иванов делал утреннюю зарядку, пропуская только воскресенья, когда позволял себе немного полениться. Сколько раз господин Иванов делал зарядку, начиная с раннего утра понедельника 30-го августа до вечера 30-го сентября?

Психология: Девушка расстроена из-за двойки на экзамене. Какими способами её можно утешить?

С одной стороны, есть способы научить (в смысле “запрограммировать”) ИС отвечать на подобные вопросы, поскольку разработаны алгоритмы поиска ответа на логические задачи, а базовые знания наук структурированы и формализованы, а значит их можно заложить компьютер. Несколько сложнее дело обстоит с извлечением из данных неявной информации, но и эта задача разрешимая. В общем и целом не видно никаких принципиальных препятствий на пути создания ИС, способной отвечать на эти вопросы, хотя, безусловно, путь этот труден. Очень уж большой объем знаний придётся превращать в “понятный машине вид”. Часто за “обучением ИС” скрывается не более, чем прямое сопоставление ответа вопросу. А именно, в ИС закладывается огромное количество шаблонов вопросов и соответствующих им алгоритмов вычисления ответа. Когда ИС задаётся вопрос, ИС определяет контекст вопроса, ищет подходящий шаблон для этого вопроса и запускает алгоритм генерации ответа, разработанный специально для данного типа вопросов. В этом подходе число шаблонов может стать таким большим, что практически сравнится с числом возможных вопросов.

Тяжелый труд кодирования общего знания в виде множества алгоритмов поиска ответа на различные вопросы не даст того, что человек привык связывать с интеллектом. Полученная в результате система не будет включать в себя способности обучаться, творчески мыслить, изобретать, выводить новое знание, строить теории, объяснять физические явления и решать нестандартные задачи. Получится не “интеллект”, а то, что принято называть экспертными системами (ЭС). Экспертные системы, безусловно, могут сильно отличаться степенью интеллектуальности, архитектурой, и самой предметной областью. Но они не могут называться интеллектом, поскольку их знания — это кропотливый труд экспертов, которые заложили знания в ИС, разложили все по полочкам и запрограммировали алгоритмы поиска ответов на вопросы пользователей.

Утверждение 2 *Если есть чётко формализованное описание того, как*

должна вести себя система, то нет никаких препятствий кроме технических для построения искусственной системы с таким поведением.

ИИ, в отличие от ЭС, обычно представляют как младенца, начальные знания которого невелики и ограничиваются базовыми знаниями о мире, пространстве, времени, человеке; словарный запас ИИ-младенца тоже не должен быть большим. Главное, чтобы он умел самостоятельно обучаться, например, читать школьные учебники, “понимать”, что там написано, постепенно увеличивать свой словарный запас, расширять “смысловую онтологию” и переходить к более сложным учебникам. Другими словами, ИИ должен быть *самообучающимся*.

Об этом писал Алан Тьюринг[?]: “Пытаясь имитировать интеллект взрослого человека, мы вынуждены много размышлять о том процессе, в результате которого человеческий мозг достиг своего настоящего состояния. . . Почему бы нам вместо того, чтобы пытаться создать программу, имитирующую интеллект взрослого человека, не попытаться создать программу, которая имитировала бы интеллект ребенка? Ведь если интеллект ребенка получает соответствующее воспитание, он становится интеллектом взрослого человека. . . Наш расчет состоит в том, что устройство, ему подобное, может быть легко запрограммировано. . . Таким образом, мы расчленим нашу проблему на две части: на задачу построения «программы-ребенка» и задачу «воспитания» этой программы”.

Утверждение 3 *Интеллектуальность заключается не в способности следовать определенному сложному алгоритму поведения, а в способности самообучаться и придумывать новые алгоритмы своей работы для решения новых задач.*

Тест на интеллектуальность должен проверять не знания, а способность обучаться и извлекать новые знания из опыта.

3 Интеллектуальность как универсальный алгоритм

Идею самообучения полезно рассмотреть в контексте понятия “алгоритм”. Под *алгоритмом* понимают точное предписание о выполнении в определенном порядке операций для решения любой задачи из некоторого данного класса задач. Нахождение алгоритмов является естественной целью человека при решении им разнообразных классов задач. Отыскание алгоритма обычно связано с тонкими и сложными рассуждениями, требующими большой изобретательности, способности к анализу, индукции и абстрактному мышлению. Именно такого рода деятельность считается интеллектуальной.

Способность решать задачи, алгоритм которых уже известен, не является чем-то примечательным. Конечно, такие задачи тоже сильно отличаются по сложности, а точнее, по вычислительным ресурсам, необходимым для их решения, но понятно, что способность исполнять алгоритмы существенно отличается от способности придумывать алгоритмы.

Таким образом, мы можем дать такое определение: интеллект — это *универсальный алгоритм*, то есть алгоритм поиска алгоритмов решения новых задач.

Эта идея встречается в многих работах по ИИ, и принимает самые разнообразные формы. Она заключается в способности ИИ к *рефлексии* (однократной или многократной) или, другими словами, к *метасистемным переходам на уровне мышления*. Определяющими свойствами интеллекта считается способность анализировать собственную работу, вырабатывать новые более эффективные алгоритмы решения новых задач, способность обучаться, улучшать алгоритм самообучения, улучшать алгоритм улучшения алгоритма самообучения и так далее.

Пожалуй, наиболее активно способность мышления к рефлексии задействуется у ученых при создании новых теорий (абстрактных, физических, объяснительных) и моделировании. Деятельность ученого включает в себя выделение существенных свойств, угадывание новых свойств, доказательство этих свойств, индуктивные рассуждения, логический анализ, распознавание ситуаций, выделение структур и закономерностей, и, что особенно важно, анализ методов мышления. До создания “искусственного учёного” науке о ИС еще далеко, но тем не менее этот вопрос уже можно поднимать и думать о путях решения. Мы рассмотрим его в части ???. Тест “искусственный ученый”, который там описывается, задает планку существенно выше теста Тьюринга. Если эта планка когда-нибудь будет преодолена, то у человеческой цивилизации начнётся совершенно новый исторический этап.

4 Модель “геном-алгоритм”

Прежде, чем говорить о метасистемных переходах в мышлении поговорим о метасистемных переходах вообще. В. Ф. Турчин в книге “Феномен науки” рассматривал метасистемные переходы, которые происходили в биологической эволюции и эволюции мыслительных способностей живых систем. Путь к интеллекту был долгим: химические элементы, самовоспроизводящиеся макромолекулы белков, хаотически движущиеся одноклеточные, организмы, способные совершать направленные действия, простые и сложные рефлексы, память и условные рефлексы, способность к моделированию и предугадыванию последствий действий (“Что будет, если ...?”), возникновение языка, способность к анализу (“Каковы причины того, что я наблюдаю?”), способность к самоанализу (“Как я анализирую и моделирую? Можно ли улучшить

методы, которыми я пользуюсь?”), и наконец последний уровень, где происходит регулярное структурирование и анализ накопленных знаний и методов мышления. Каждый из этих пунктов содержит в себе один или несколько метасистемных переходов.

Придуман ряд искусственных систем [?, ?, ?], которые моделируют той или иной метасистемный переход. Например, есть модель модель леса, в котором живут кролики и лисы. Изначально связь между возможными реакциями и раздражениями у животных случайна. Во время процесса “жизни” происходит размножение с рекомбинацией генов, мутация и борьба за выживание. Роль генотипа играет функция связи реакций и раздражений, назовем её *функцией мотивации*. В результате естественного отбора возникает некоторая “логика жизни”: лисы учатся бегать за зайцами и съедать их при первой возможности, а зайцы учатся убежать при виде лисы, прятаться в норах и только при крайней необходимости выбегать покусать свежей травки. Приобретение таким способом такой логики некоторые склонны считать переходом на новый уровень (зарождение логики), хотя скорее всего, это не более, чем подбор коэффициентов в функции мотивации. Эта модель не способна осуществлять следующие метасистемные переходы. В ней строго задан вид функции мотивации, не могут появиться новые рецепторы и новые типы реакции.

Ничто не мешает нам создать такую модель живой среды, где организмы обладают способностью тратить свои жизненные силы на развитие каких-то своих способностей (например, учиться быстро бегать или развивать нюх) и добавлять новый тип действий (например, макро-действий, составленных из композиций действий, которые они умеют делать на текущий момент). Это позволило бы зайцу вместо многоходового действия (сизу в норе – увидел траву – бегу к траве, стою у травы – ем траву, наелся травы – ищу нору, нашел нору – бегу в нору) обучиться выполнять только одно (вижу траву из норы – бегу до травы, кушаю и обратно). Возникновению у особей способности к сложным действиям уже можно назвать переходом на новый уровень. Но и эта модель “ограничена сама собой”. В ней не может идти речь о целеполагании и целенаправленных действиях.

Более общую идею можно сформулировать в терминах программирования. Назовем эту идею “геном-алгоритм”.

Рассмотрим модель, где у особей есть “программы жизни”, написанные на некотором простом специальном языке программирования. Причём чем сложнее “программа жизни”, тем больше ресурсов нужно организму для поддержания своей жизнедеятельности, а точнее, за выполнение одной инструкции программы особь должна платить некоторым количеством “жизненной энергии”. Особь может запустить команду “спать время X ” и во время сна восстанавливать свои силы. В модели чётко прописан механизм мутации и рекомбинации генов (программ), а также сам язык программирования и цена за

каждую конкретную команду. “Жизненная программа” представляет собой несколько функций — своеобразных обработчиков сообщений. Для каждой функции прописаны условия её запуска — условия на текущее состояние особи и внешние условия. Условия запуска функции плюс сама функция есть неделимый ген, и при рекомбинации генов внутренности функций не перемешиваются. Вот пример такого гена: “Если я бегу за зайцем, заяц еще далеко, а сил осталось немного, то лучше остановиться, притаиться за кустом и перейти в состояние ожидания добычи.” Мутация, в отличие от рекомбинации способна менять сами гены, как условия вызова функций так и логику самих функций, кроме того мутация способна дублировать случайным образом один из генов с небольшими изменениями.

Утверждение 4 *Идея “геном-алгоритм” заключается в отождествлении генома и “алгоритма жизни” и создании искусственной среды для естественного отбора алгоритмов. В результате эволюции в этой среде может получиться особь со сколь угодно сложным “алгоритм жизни”. Эта идея позволяет в одной модели реализовать несколько метасистемных переходов. На каждом системном уровне у особей имеются функции поведения (алгоритмы, отвечающие за распознавание ситуаций и реакции организма), которые оперируют (в смысле “вызывают как подпрограммы”) функциями предыдущих уровней.*

Возможно, модели, построенные на основе этой идеи, кому-то покажутся искусственными: если в них будут наблюдаться метасистемные переходы, то в значительной степени это будет заслуга создателя модели, который изобрёл подходящий “язык программирования” стратегии жизни, заложил правильные механизмы рекомбинации и мутации, придумал хорошие ограничения и, вообще, смоделировал подходящий мир с подходящими законами. Есть также пессимистичное мнение, что для создания модели, способной ко большому количеству метасистемных переходов необходимо приложить огромное количество усилий, и сама конечная модель будет сложной, громоздкой и неестественной.

Кроме того, оппоненты могут заметить, что модели, основанные на идее “геном-алгоритм”, также как и остальные модели имеют ряд ограничений, от которых нельзя уйти. Эти ограничения заключаются в том, что организмы не могут менять язык, на котором пишутся их алгоритмы жизни, и не могут менять правила, по которым они тратят и приобретают жизненные силы. Это действительно так.

Но очень важно, что идея “геном-алгоритм” может породить модели, потенциально способные совершить любое число метасистемных переходов, и в этих моделях в принципе может появиться особь со сколь угодно сложным алгоритмом поведения. Например, могут образоваться кусочки кода, отвечающие

за сбивание особей одного вида в группы, коллективные направленные действия, заботу о детях, прогнозирование будущего и сложные поведенческие реакции. Вопрос в том, как долго придётся этого ждать. Чтобы увеличить скорость эволюции полезно сами алгоритмы рекомбинации и мутирования поместить в гены — то есть сделать алгоритм эволюции эволюционирующим, при этом необходим какой-то естественный механизм, который среди алгоритмов эволюции отбирал бы те, которые поощряют усложнение и многообразие форм жизни. Обнаружить такой механизм было бы величайшей находкой.

Опыт показывает, что три метасистемных перехода — это максимум, чего можно добиться от компьютерной модели. Это как раз связано с тем, что законы жизни на первом уровне подбирает сам человек, а он не может прогнозировать, к чему приведут эти законы после того, как в системе произойдут два или более метасистемных перехода. Попытки автоматически подбирать параметры так, чтобы статистические показатели (связанные со скоростью образования энтропии, плотностью вещества и их дисперсиями) соответствовали типичным значениям живых систем тоже не привели к успеху — исследователи построили множество “интересно живущих” динамических систем, но называть эти системы эволюционирующими и выделять метасистемные переходы сложно отчасти оттого, что не совсем ясны критерии метасистемных переходов. Первый метасистемный переход обычно связывают с образованием структур, второй — с взаимодействием этих структур и образованием структур второго уровня. Но что такое структуры первого и структуры второго уровней? После двух метасистемных переходов понятие “структура” становится совсем зыбким и третий метасистемный переход ускользает из рук исследователей.

Чтобы считать метасистемные переходы нужно очень чётко определить что это такое. Идея “геном-алгоритм” позволяет сделать шаг в этом направлении. Некоторые функции из генома-алгоритма являются низкоуровневыми, то есть используют только простейшие команды, изначально заложенные в язык программирования, другие — высокоуровневые. Описание высокоуровневых функции содержит вызовы других, более низкоуровневых функций. **Глубиной сложности функции** на некотором языке программирования естественно назвать глубину дерева разложения (сведения) функции к более простым. Но понятно, что любую функцию можно запрограммировать, используя только простейшие операции, и не выделять замкнутые логические блоки в отдельные функции. При определении уровня сложности функции (а не её реализации) естественно потребовать, чтобы число команд в определении каждой функции было ограниченным некоторым небольшим числом M . Тогда сложные функции мы будем вынуждены разбивать на высокоуровневые действия (этот подход связывают с процедурным программированием), создавать для них отдельные функции, а их, возможно, снова сводить к еще

более простым.

Приведём определения *глубина сложности* и *сложности реализации* вычислимой функции, которые аналогичны определениям глубины и комбинационной сложности булевых функций в заданном базисе.

Определение 1 *Глубина сложности функции f на некотором языке программирования — это величина, характеризующая глубину дерева разложения (сведения) функции f к более простым. Она равна минимально возможной глубине дерева сведения функций при условии, что определение каждой функции содержит не более, чем M команд.*

Сложностью реализации функций f при заданной глубине K назовём минимальный размер программы, реализующего данную функцию, при условии, что каждая функция этой программы (сама функция f , а также дополнительные функции, которые непосредственно или косвенно используются в её определении) содержит не более M команд, а глубина сведения не больше K .

В этих определениях мы предполагали, что граф зависимостей функций друг от друга не содержит циклов. Дать определение понятие уровня функций для произвольного графа зависимостей (например, допускающего рекурсивные вызовы) можно, но сложно и требует уточнения множества деталей.

Таким образом, если идет речь о эволюции некоторого набора функций, то можно дать вполне чёткое определение метасистемного перехода: метасистемным переходом называется увеличение глубины сложности функций на единицу. Значение глубины и сложности реализации функции будет зависеть от выбора языка программирования и параметра M . От этой относительности никак нельзя избавиться. Такая ситуация типична для теории сложности — нужно иметь эталонную вычислительную машину и “эталонный ассемблер”. Алгоритм следующей глубины сложности — это алгоритм, который оперирует алгоритмами и структурами данных построенными на предыдущих уровнях и при этом (чтобы не перескочить через уровень) достаточно прост. Например, если у нас есть несколько алгоритмов игры в шахматы, то алгоритм следующего уровня может быть таким: разобьем ситуации на дебюты, эндшпили и средние, исследуем в каких ситуациях какой из алгоритмов эффективнее, и каждый раз будем применять лучший из этих алгоритмов.

5 Тест “Универсальный игрок”

Рассмотрим идею метасистемного перехода в алгоритме мышления на примере логических игр.

Не для кого не секрет, что есть программы способные играть в шахматы, крестики нолики, реверси и другие логические игры. В теории алгоритмов

есть направление, посвященное программированию игровых стратегий. Там рассматриваются эвристические оценки ситуаций, поиск в глубину с альфа-бета отсечением и другие более эффективные алгоритмы: *Negomax*, *SSS*, *MTD(f)* [?]. Компьютеры благодаря быстрдействию и безошибочности могут играть во многие игры лучше среднего человека. На настоящий момент лучшая программа, играющая в шахматы, играет на уровне чемпиона мира. Успехи компьютеров на игровом поприще заслуга не самих компьютеров, а инженеров и программистов, которые их спроектировали, формализовали стратегическое мышление и разработали “умные” алгоритмы.

Совсем другое дело, если компьютерная программа умеет играть в различные игры, способность играть в которые не была изначально заложена в неё человеком. Интеллектуальность в области игр может быть интерпретирована как способность придумывать стратегические алгоритмы для новых игр. Рассмотрим эту идею подробнее.

Программу, которая умеет обучаться новым играм и успешно играть в них, назовём Универсальным Игроком (УИ). УИ изначально не умеет ни во что играть, и необходима возможность учить его правилам новых игр, то есть необходим специальный язык описания правил игр, который УИ понимает. Такой язык несложно придумать, для этой цели подойдет любой функциональный язык программирования (*Lisp* или *Haskell*). На вход УИ получает описание игры на этом языке. Затем УИ даётся некоторое время “подумать” и “поиграть в эту игру с самим собой”, после чего он должен уметь играть в эту игру на достаточно хорошем уровне. Важно, чтобы игры были *стратегическими* и достаточно *сложными*. Сложность необходима для того, чтобы программа не могла перебрать за разумное время все возможные позиции и определить какие из них выигрышные, а какие проигрышные. Стратегический характер игры подразумевает наличие некоторых “стратегических моментов” типа “ценность фигур”, “двойная защита”, “укрепление флангов”, “вилка” и т. п. “Стратегические моменты” включают в себя идеи, которые позволяют, не продумывая возможные ходы на большую глубину, играть более менее разумно и малой ценой получать неплохой результат. Сложность программирования УИ заключается в том, что эти “упрощающие жизнь” идеи различаются от игры к игре: иногда, полезно стремится защищать фигуру столько раз, сколько на неё нападений, в другой игре важно поддерживать определенные структуры на “поле боя”, в третьем случае нужно как можно скорее занять определенные “стратегические позиции”. Самых фундаментальных идей можно насчитать порядка сотни и каждая из них с трудом поддается программированию в общем виде.

В индустрии компьютерных стратегических игр выработался специальный подход программирования стратегий. А именно, создается специальный *макро язык* для программирования поведения (логики жизни) “компьютерных ботов” (*бот* = искусственный игрок). Под понятием “макро язык” мы имеем

ввиду следующее:

- Это язык программирования высокого уровня, его команды соответствуют не элементарным действиям, а сложным составным действиям;
- Этот язык оперирует со сложными структурами и группами объектов, а не с элементарными объектами.

Например, на макро языке для военных стратегических было бы удобно записать в программе жизни боевой единицы строчки “заходи к сопернику с фланга” или “если соперников больше, то прячься за ограждениями” или что-нибудь в этом роде.

Каким-то образом человеческий мозг, накапливая опыт игры, сам создает для себя специальный макро язык, и начинает думать на этом языке, постепенно улучшая и дополняя его новыми “макро” объектами. Можно ли формализовать эту деятельность?

Перед обсуждением алгоритма конструирования макро-языка для стратегического мышления отметим два важных момента.

УИ не может быть абсолютно универсальным. Язык описания игр, каким бы хитроумным он не был, не сможет описать все возможные игры. Возможные игры двух игроков есть множество всех направленных графов, чьи ребра раскрашены в два цвета. Различных графов (с конечным или бесконечным числом вершин) несчетное количество, а множество слов в любом языке с конечным алфавитом счётно. Это значит, что любой язык игр описывает лишь счётное подмножество игр. Но дело даже не в счётности множества описаний. Очень многие игры имеют конечный, но очень большой граф. Например, граф игры “шахматы” содержит порядка 10^{43} узлов. И хотя формальный язык, описывающий конечные графы создать не сложно, правила шахмат на нём невозможно будет записать, так как не хватит памяти существующих компьютеров. Описание правил шахмат на человеческом языке имеет менее, чем 5000 слов. В шахматах нет необходимости описывать для каждой возможной позиции ходы, которые из неё можно сделать, достаточно описать как могут ходить фигуры и ряд других общих правил. Можно создать такой формальный язык описания игр, что с его помощью естественным образом будут описываться *все* известные логические игры.

Второй важный момент — это *иллюзия множества игр*, которой мы здесь умышленно поддаемся. Дело в том, что “правила игры” + “начальная ситуация в этой игре” можно воспринимать как начальную ситуацию в некоторой макро-игре. Макро-игра включает в себя множество игр. Информация о том, какая именно игра будет играть, заложена в начальной ситуации макро-игры. Получается, что от УИ мы по прежнему требуем умения играть в какую-то конкретную игру, и прилагательное “универсальный” оказывается сомнительным. Что ж, это действительно так. Но макро-игра включает в себе бесконечное количество игр, подобных шахматам, го, реверси, крестикам-ноликам. Для того, чтобы написать алгоритм для этой игры, нужно мыслить

и программировать на несколько другом уровне, нежели при программировании стратегии для отдельно взятой обычной игры.

Утверждение 5 *Макро-игра — это игра, включающая в себя счётное множество игр. Начальная ситуация в макро-игре включает в себя описание правил некоторой игры и начальную ситуацию для этой игры.*

Утверждение 6 *Метасистемный переход в программировании игровых стратегий соответствует созданию алгоритма, успешно играющего в макро-игру.*

Соответственно получаем такой тест на интеллектуальность:

Тест 2 (“Универсальный игрок”)

Напишите программу, успешно играющую в макро-игру.

5.1 Макро-язык стратегического мышления

Рассмотрим путь создания Универсального Игрока, в котором явно присутствует этап конструирования макро языка мышления. Уточним, что имеется в виду под термином “явно”. Например, в нейросетевых алгоритмах игры в шахматы программирование стратегии игры превращается в подбор большого количества коэффициентов — весов в нейросети. Из набора чисел сложно вычленишь какие-либо структуры с конкретным физическим смыслом, хотя вполне возможно, что в нейросети неявным образом присутствуют такие понятия как “эндшпиль”, “вилка” или “двойная защита”. Тоже самое касается генетических алгоритмов, в которых рассматривается по сути *один алгоритм* с большим количеством численных параметров, которые и определяют генотип. В алгоритмах, основанных на эвристических оценках и поиске в глубину, макро язык в явном виде также отсутствует.

Но есть ощущение, что именно макро язык является тем неиссякаемым источником метасистемных переходов и когнитивности человеческого мышления. Поэтому имеет смысл рассматривать алгоритмы, в которых он присутствует в явном виде.

Итак, нам нужно формализовать процесс поиска разумом структур и ключевых стратегических моментов, то есть процесс создания эффективную терминологическую базу для обдумывания ходов. Это очень сложно. Например, исследуем, каким образом человек в игре в шахматы дошёл до понятия “вилка”. “Вилка” — это ход, в результате которого атакуются две фигуры соперника. Если рассматривать только задействованные три фигуры, то выгодность “вилки” очевидна — на следующем ходе мы увеличиваем преимущество в одну фигуру. Но дело в том, что на доске присутствуют другие фигуры, и соперник может ответить на “вилку” нападением на ферзя. Предположим, что ферзя

можно безболезненно увести в сторону. Но следующим ходом соперник может поставить шах и нам снова придется защищаться. Таким образом взятие фигуры в результате “вилки” может быть отложено соперником на несколько ходов вперед, а простое продумывание ходов на небольшую глубину может просто не дойти до этого момента. Это явление называется *эффектом горизонта* — не смотря на то, что возникла очевидная “вилка”, дерево ходов конечной глубины может не содержать вершины, в которой эта вилка наверняка “стрельнула”, эта вершина находится за “горизонтом продумывания”. Таким образом, “вилка” может быть очень тщательно зашифрована в полном дереве ходов. Важным шагом по выявлению понятия “вилки” было исключение из рассмотрения неактивных фигур и получение ситуации, в которой на втором ходе происходит неизбежное взятие фигуры. Упрощение ситуации путем рассмотрения небольшой части “поля боя” оказывается полезным во многих играх. Видно, что это есть частный случай фундаментальной идеи, взятой человеком из общего жизненного опыта: “жизнь” объекта определяется его взаимодействием лишь с небольшим числом “соседних” объектов, а другие объекты слабо на него влияют, по крайней мере, в ближайшем будущем. Назовем эту идею *локализацией пространства и времени*. Важно отметить, что выделение шахматных фигур в отдельные объекты дано нам а priori и было осуществлено еще на этапе формулировки правил игры. Упрощение ситуации за счет локализации пространства и времени дает нам также технику подсчета числа нападений и атак, которая позволяет “свернуть” продумывание дерева ходов (где происходит взаимное уничтожение одинакового количества фигур обоих игроков) в одну эвристическую оценку. Кроме локализации здесь включается еще одна фундаментальная идея — *группировка объектов в один макро объект*, свойства которого каким-то образом наследуются от объектов, которые он в себе объединил. В данном случае мы по сути связываем с каждой клеткой шахматного поля макро-объект, состоящий из всех фигур, бьющих данное поле; его сила равна количеству этих фигур, а ценность — сумме ценностей фигур. При группировке объектов мы встречаемся с рядом тонких моментов: Как выбрать принцип группировки, чтобы это дало полезный для стратегического мышления результат? Какие свойства и каким образом наследовать макро объекту от входящих в группу объектов? Как быть с новыми свойствами, которые не присутствуют ни у одного из объектов, входящих в группу (как догадаться до того самого “качества”, в которое перешло “количество”)

Анализ того, как человек развивал свое стратегическое мышление на примере игр, приводит нас к следующему выводу: для развития макро языка мышления человеку необходимо знание фундаментальных законов жизни, некоторых общих принципов природы. Это в определенной степени противоречит эксперименту, так как трудно поверить, что новорожденный ребенок изначально обладает таким знанием.

Есть масса работ, посвященных изучению законов эволюции человеческого языка [?, ?, ?, ?, ?, ?, ?, ?, ?, ?]. И во многих из них рассматривается такое понятие как *универсальная грамматика* [?], впервые введенное Хомским [?] как некоторое начальное знание языка ребенка, которое позволяет ему по примерам выучивать сложный человеческий язык. Очень многие исследователи приходят к выводу, что эта универсальная грамматика очень объёмное и сложное знание. Более того в [?] явно высказывается мысль, что задача обучения языку, решаемая детьми, вычислительно очень сложна: “Children (...) perform mental computations that are virtually impossible for state-of-the-art computers.” Попытки математически формализовать задачу обучения детей языку делались, и здесь интересны работы, где показывается невычислимость или *NP*-сложность задачи обучения языку по примерам в определенных классах грамматик. [?, ?, ?, ?, ?, ?]. Другие исследователи верят, что эволюционный подход способен дать объяснение возникновению и обучению сложным языкам практически “с нуля” [?, ?, ?].

Вообще, с языком связано несколько великих, неразгаданных тайн:

- Как возникает и эволюционирует язык?
- Как дети учатся языку?
- Каковы механизмы эволюции языка мышления?
- Каким образом человек по наблюдаемым явлениям находит законы природы (язык описания природы)?

Это важнейшие вопросы когнитивной науки — науки о приобретении новых знаний. Нас здесь интересует третий вопрос. Но может оказаться, что ключ к ответу на все эти вопросы один и тот же. Вопрос “Каким образом человек в игре в шахматы дошел до понятия *вилка*?” оказывается ничуть не проще вопроса “Каким образом был открыт закон всемирного тяготения?” Эволюция видов, эволюция языка, возникновение логики и интеллекта — всё это могут быть явления, приводящиеся в движение одним и тем же механизмом.

Во многих работах, посвященных исследованию природы мышления, в выводах пишется о чем-то неизвестном, которое может дать ответ сразу на все вопросы когнитивной науки. Роджер Пенроуз в своей книжке “Новый ум короля” [?] связывает это неизвестное с теоретической физикой и неоткрытыми физическими законами и даже даёт ему название — правильная теория квантовой гравитации.

В работах Малинецкого Г. Г.⁴ [?, ?] часто высказывается мысль, что динамический хаос и самоорганизованная критичность (см. [?] и другие работы Подлазова А. В.⁵) может объяснить многие явления, связанные с эволюцией, жизнью и развитием науки. Ниже будет рассмотрена самоорганизованная

⁴http://www.keldysh.ru/departments/dpt_17/gmalin.html

⁵http://www.keldysh.ru/departments/dpt_17/Podlazov.html

критичность как явление, способное пролить свет на природу двигателя метасистемных переходов.

6 Тест “Универсальный Учёный”

Создание ИС заменяющей учёного является на настоящий момент сверх целью. Промежуточным этапом на пути к этой цели могут стать *искусственные научные консультанты* (ИНК) — экспертные системы в области естественных наук. Но если последние, возможно, будут созданы, искусственный учёный кажется принципиально недостижимой целью.

Тем не менее сформулируем тест, определяющий возможность ИС вести научную деятельность. Искусственного “Универсального Ученого” (УУ) следуя традиции Теста Тьюринга будем рассматривать как черный ящик с вербальным каналом ввода-вывода информации. Причём естественно этот канал разделить на два:

- канал, по которому УУ задают вопросы и получают от него ответы — объяснения физических явлений, описания новых законов и технологий.
- канал, по которому УУ получает информацию об окружающем мире; можно считать, что на другом конце этого канала находится группа экспериментаторов, которые в соответствии с указаниями УУ проводят физические эксперименты и сообщают ему результаты.

Тест 3 (“Универсальный Учёный”)

Напишите программу “Универсальный Учёный”.

Например, УУ должен по наблюдению движения светил на небе обнаружить, что земля круглая, а также открыть закон всемирного тяготения. Конечно, неплохо было бы получить от УУ и общую теорию относительности.

Попытки более чётко формализовать этот теста выявляют целый ряд деталей, требующих уточнения. Обратим внимание лишь на одну из них. Открытие нового физического закона очень часто подразумевает увеличение или изменение терминологической базы, и об этих изменениях УУ должен будет нам рассказывать. То есть в УУ должен быть заложен мощный мета-язык — средство создания новых языков, которого должно хватать для введения новых физических и математических терминов. Но проблема в том, что, видимо, нет универсальной грамматики, которой бы хватило на все случаи жизни. И может так случиться, что понятия “алгебраическое уравнение”, “когерентность” или “лазер” не будут сводиться к изначально заложенной в УУ грамматике.

Таким образом, прежде чем переходить к штурму самой сложной программистской задачи создания искусственного учёного, необходимо решить загадку, как дети “с нуля” (с универсальной грамматики) обучаются новому языку.

Чему равен этот “нуль”? Можно ли его формализовать процесс обучения детей человеческому языку как эволюцию динамической системы с фазовым пространством равным множеству грамматик, начальное состояние которой есть грамматика (модель) с конечным описанием?

Все сформулированные вопросы так или иначе сводятся к двум

- Можно ли формализовать априорное знание?
- Вычислим ли процесс получения эмпирического знания?

Существование искусственного УУ равносильно положительному ответу на оба вопроса.

7 Самокритичность как двигатель метасистемных переходов

Есть ли рациональное объяснение тому, что мир богат сложными живыми системами? Почему метасистемные переходы происходят так часто? Почему эволюции хватило 2 млрд. лет для получения человека? Как возникает и эволюционирует язык?

Есть ли панацея, которая избавит наш мир от чудес и даст нам ответы на эти вопросы?

Кроме упомянутой книжки Роджера Пенроуза “Новый ум короля”, где эта панацея называется “Правильной Теорией Квантовой Гравитации” (ПТКГ), есть целое научное направление, призванное находить рациональное объяснение чудесам. Это теория самоорганизованной критичности [?, ?]. В ней рассматриваются динамические системы, снабженные естественными механизмами изменения параметров, которые стремятся установить эти параметры в *критические значения*. Это такие значения, при которых возникает динамический хаос, сложные структуры и “жизнь” системы наиболее активна и разнообразна.

Например, угол склона в куче песка естественным образом устанавливается в такое критическое значение, при котором существует опасность непредсказуемых лавин непредсказуемого размера.

Чрезвычайно важно для живых эволюционирующих систем найти механизмы, которые устанавливают параметры в такие значения, при которых

- 1) эволюция наиболее быстра и эффективна,
- 2) жизненные формы стремятся к усложнению,
- 3) метасистемные переходы поощряются.

Параметры, о которых идет речь, — это параметры, определяющие алгоритм эволюции: как гены шифруют “алгоритм жизни”, по какому правилу происходит рекомбинация и мутация, что определяет взаимное желание двух особей

соединиться. Кому-то может показаться странным, что алгоритм эволюции может рассматриваться как варьируемый параметр – ни природа, ни, тем более, мы не подвластны над законами эволюции. Но если подняться на ступень выше, то можно представить себе, что алгоритмов эволюции было несколько, а то, что мы с вами наблюдаем — это плоды того алгоритма, который оказался наиболее успешным в конструировании “жизни”. Плодов других алгоритмов эволюции просто нет. Таким образом “жизнь” — это, в некотором смысле, коллекция удивительных вещей, если уж человек в неё попал, то ждите, что на соседних полках находятся ничуть не меньшие чудеса и разные “невероятные штучки”. Вспоминается персонаж из книги братьев Стругацких, который притягивал к себе маловероятные события. Все человечество и планета Земля является именно таким персонажем в жизни вселенной.

Конечно, несколько нечестно утверждать, что сам факт существования сложных живых систем и есть та обратная связь, которая устанавливает параметры системы в нужные критические значения, поощряющие жизнь. Но, с другой стороны, на фоне признания того, что квантовую механику и общую теорию относительности нельзя рассматривать как теории, полно описывающие вселенную, а также на фоне отсутствия физической теории, объясняющей феномен сознания, можно делать самые разные предположения. Например, можно высказать предположение, что в фазовом пространстве вселенной состояний, соответствующих активной “жизни”, значительно больше. Это может быть следствием некоторого неоткрытого пока закона, из-за которого объём кусочков фазового пространства тем больше, чем более непредсказуемы и сложны траектории, идущие из этого кусочка. Либо можно высказывать предположения в духе Пенроуза, что в гамильтониане вселенной из ПТКГ присутствует член, измеряющий жизненную активность и поощряющий рост коэффициентов при тех собственных функций, где жизнь наиболее активна и разнообразна.

Но отвлечемся от необоснованных предположений и вернёмся к определению ?? сложности реализации с заданной глубиной. Другое объяснение феноменальной эффективности эволюции может быть основано на том, что эволюция происходит в условиях ограничения сложности реализации алгоритма жизни, которую можно связывать с размером описания генотипа. Когда увеличение кода, описывающего “алгоритм жизни”, стоит дорого, автоматически начинают выживать только те особи, у которых алгоритм жизни сложен, а код, его реализующий, мал. Это как раз и означает, что поощряется reuse – вторичное использование, то есть формирование фундаментальных, часто используемых на разных уровнях функций. Это также означает, что эволюцией поощряются использование функций друг другом, многократная рекурсия, увеличение глубины функций, а значит, и метасистемные переходы.

8 Заключение

Следует признать, что многие из приведенных здесь рассуждений носят философский оттенок. Поэтому хотелось бы еще раз выделить две конструктивных мысли:

- Тест “Универсальный Игрок” есть первый хорошо формализованный тест на способность искусственных систем осуществлять метасистемные переходы. И в отличие от “Универсального Учёного” он поддается штурму и является тем шагом, который уже пара делать в теории алгоритмов.
- Моделирование эволюции алгоритмов с явным введением макро-языка мышления является актуальной задачей. Исследование возможностей эволюционирующих систем, основанных на идее “геном-алгоритм”, может дать много интересных результатов. Данное здесь определение “глубины функции” позволяет объективно сравнивать эволюционирующие способности различных моделей и говорить о числе метасистемных переходов, осуществленных во время эволюции. Необходимо исследовать, как влияет на эволюционирующую способность систем введение платы за размер описания “генома-алгоритма”.

Автор выражает благодарность Никите Винокурову и Георгию Геннадьевичу Малинецкому за интересные беседы и полезные замечания.

Список литературы

- [1] Турчин В. Ф., Феномен науки: Кибернетический подход к эволюции, М.: Наука, 295 с., 1993, <http://pespmc1.vub.ac.be/POSBOOK.html>.
- [2] Turchin V. F., A Dialogue on Metasystem Transition, in Heylighen F., Joslyn C., Turchin V. F. (eds.): The Quantum of Evolution. Toward a theory of metasystem transitions, (Gordon and Breach Science Publishers, New York) (special issue, Vol. 45:1-4, of “World Futures: the journal of general evolution”), pp. 5–58, 1995.
- [3] Тьюринг А., Может ли машина мыслить? (С приложением статьи Аж. фон Неймана “Общая и логическая теория автоматов”. Перевод и примечания Ю.В. Данилова). М.: ГИФМЛ, 1960.
- [4] Roger P., The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics, рус. перевод: Роджер Пенроуз, Новый ум короля, Едиториал УРСС, 384 с., 2003.
- [5] Редько В. Г., Эволюционная биокибернетика, Вестник Российской Академии Наук, 67, №9. с. 800-803, 1997.

- [6] Редько В. Г., Проблема происхождения интеллекта и эволюционная биокibernетика, журнал “Высшая нервная деятельность”, 48, №2, стр. 358-369, 1998,
<http://www.keldysh.ru/pages/BioCyber/>
- [7] Бурцев М. С., Модель эволюционного возникновения целенаправленного адаптивного поведения. 2. Исследование развития иерархии целей, препринт ИПМ РАН, № 69, 2002.
- [8] Подлазов А. В., Новые математические модели, методы и характеристики в теории самоорганизованной критичности, Диссертация на соискание степени к.ф.-м.н. по специальности 05.13.18 – “Математическое моделирование, численные методы и комплексы программ”. – М., 2001.
- [9] Малинецкий Г. Г., Нейроромантизм или компьютерные вариации. Знание-сила №8, 1994.
- [10] Малинецкий Г. Г., От мозга к компьютеру и обратно. Знание-сила №9, 1994.
- [11] Малинецкий Г. Г., Потапов А. Б., Джокеры и русла, или поиски третьей парадигмы. Знание-сила №3, 1998.
- [12] Малинецкий Г. Г., Хаос. Структуры. Вычислительный эксперимент. Введение в нелинейную динамику. М., Эдиториал УРСС, 3-е издание, 2001.
- [13] Нейронные сети: общие положения, парадигмы,
<http://www.91.ru/Education/Books/Neural Net/Neural Teory/title.htm>
- [14] Koza J. R., Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge, Massachusetts, 1992.
- [15] Aske Plaat, Jonathan Schaeffer, Wim Pijls, and Arie de Bruin: Best-First Fixed-Depth Minimax Algorithms, In: Artificial Intelligence, Volume 87(1-2), pp. 255-293, 1996,
<http://www.cs.vu.nl/ aske/minimax.html>
- [16] Lambek J., On the calculus of syntactic types, in: R. Jakobson (Ed.), Structure of language and its mathematical aspects, American Mathematical Society, pp. 166–178, 1961.
- [17] Chomsky N., Language and Mind, New York: Harcourt, Brace & World, 1972.
- [18] Angluin D., Inductive inference of formal languages from positive data, Information and Control 45, pp. 117–135, 1980.
- [19] Smith J. M., Evolution and the theory of games, Cambridge University Press, 1982.
- [20] Hinton G. E., Nowlan S. J., How learning can guide evolution. In Complex Systems, 1, pp. 495–502, 1987.

- [21] Pinker S., Bloom P., Natural Language and Natural Selection, Behavioral and Brain Sciences, 13, pp. 707–784, 1990.
- [22] Kanazawa M., Learnable Classes of Categorical Grammars, Ph.D. thesis, Stanford University, 1994.
- [23] J. van Benthem and A. ter Meulen (Eds.), Handbook of Logic and Language. Elsevier Science, 1997.
- [24] Steels L., The synthetic modeling of language origins, Evolution of Communication, 1(1):1–34, 1997
- [25] Steels L., Kaplan F., Stochasticity as a source of innovation in language games. In C. Adami, R. Belew, H. Kitano, and C. Taylor, editors, Proceedings of Artificial Life VI, Los Angeles, June. MIT Press, 1998.
- [26] Batali J., The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In T. Briscoe, editor, Linguistic evolution through language acquisition: formal and computational models. Cambridge University Press, 2000.
- [27] Zuidema W. H., Westermann G., On the Relevance of Language Evolution Models for Cognitive Science, AI-MEMO 01-03, Vrije Universiteit Brussel, 2001.
- [28] Kirby S., Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. IEEE Transactions on Evolutionary Computation, 5(2):102–110, 2001.
- [29] Christiansen M. H., Kirby S. (Eds.), Language Evolution: The States of the Art. Oxford University Press, 2003, <http://www.ling.ed.ac.uk/~simon/0-19-924484-7.pdf>
- [30] Kirby S. (Ed.), Language Evolution and Computation, ESSLLI workshop proceedings, 2003.
- [31] Kirby S., Smith K., Brighton H., From UG to Universals: linguistic adaptation through iterated learning. Studies in Language, 28(3), 2004, <http://www.ling.ed.ac.uk/kenny/publications/ksb.pdf>
- [32] Komarova N., Niyogi P., Nowak M., The evolutionary dynamics of grammar acquisition, Journal of Theoretical Biology, 209: 43-59, 2001.
- [33] Komarova N. L., Hauser M. D., Building the tower of babble. Trends in Cognitive Sciences, 5(10):412–413, 2001
- [34] Komarova N. L., Nowak M., “Language, Learning and Evolution”, in [?] M. H. Christiansen and S. Kirby (Eds.), Language Evolution: The States of the Art, Oxford University Press, 2003.
- [35] Floréncio C. C., Consistent Identification in the Limit of the Class k -valued is NP-hard, in: LACL, 2002.

- [36] Groote P., Lamarche F., Classical non-associative Lambek calculus, *Studia Logica* 71.1 (2), 2002.
- [37] Foret A., Y. Le Nir, Lambek rigid grammars are not learnable from strings, in: COLING'2002, 19th International Conference on Computational Linguistics, Taipei, Taiwan, 2002.
- [38] Béchet D., Foret A., k-valued non-associative Lambek grammars are learnable from function-argument structures, in: Proceedings of the 10th Workshop on Logic, Language, Information and Computation (WoLLIC'2003), volume 85, Electronic Notes in Theoretical Computer Science, 2003.
- [39] Béchet D., Foret A., k-valued non-associative Lambek categorial grammars are not learnable from strings, in: ACL (Ed.), Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL 2003), pp. 351–358, 2003
- [40] G. de Pauw, A corpus-based natural language grammar optimization approach using agent-based evolutionary computing. In Ann Nowe et al (eds), Proceedings of the annual machine learning conference of Belgium and The Netherlands, pp. 30–37, Brussels, 2004.
- [41] Iwahashi N., Language acquisition by robots – Towards New Paradigm of Language Processing, *Journal of Japanese Society for Artificial Intelligence*, 18(1):49–58, 2003.