

## **НОВОСТИ ИИ. – 1998. – №3.**

### **КРУГЛЫЙ СТОЛ «ПАРАДИГМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА»**

**Руководители: д.т.н. О.П.Кузнецов, к.т.н. В.Б.Тарасов, к.ф.-м.н. А.Н.Аверкин, д.т.н. В.Н.Вагин**

#### **О.П.Кузнецов**

Известный науковед Т.Кун в книге «Структура научных революций» подверг серьезной критике привычное представление о поступательном, кумулятивном, (т.е. монотонно накапливающим знанию) развитии наук и выдвинул циклическую схему их развития. Этот цикл начинается с формирования парадигмы – набора понятий, исходных положений и установок, принимаемых и широко распространяемых научным сообществом. Парадигма включает в себя не только строгие утверждения и закономерности, но и психологические факторы. После установления парадигмы начинается этап «нормального развития» науки – разворачивания широкого фронта работ, основанных на данной парадигме. В ходе этих работ постепенно накапливаются данные, необъяснимые в рамках существующей парадигмы. Вначале кажется, что это – естественные проблемы, которые со временем будут решены. Однако, если эти проблемы не решаются и продолжают накапливаться, то они начинают восприниматься как серьезные аномалии. Когда количество необъяснимых явлений превышает некий критический уровень, возникает кризис: сама парадигма перестает быть бесспорной. Начинается процесс ее пересмотра, который завершается установлением новой парадигмы. В истории науки наиболее известные примеры кризисов, вызвавших смену парадигм, – это кризис классической ньютоновой механики, приведший к возникновению теории относительности, а также парадоксы теории множеств, попытки преодоления которых привели к появлению новой математической дисциплины – оснований математики.

И физика, и математика являются примерами «вечных», классических наук, в которых объект неисчерпаем, а смена парадигмы происходила медленно. Как правило, на протяжении жизни нескольких поколений ученых. Но и в этих науках очень часто смена парадигмы воспринималась ученым как личная катастрофа. И очень многие весьма солидные ученые так и не могли до конца своей деятельности, отказавшись от старого, признать новую парадигму.

Информатика дала нам другой тип наук. Информатика – это конгломерат микронаук: теории помехоустойчивых кодов, теории автоматов, теории формальных языков и грамматик, и т.д. История одной такой микронауки – это история одной парадигмы. Исчерпывается парадигма, закрывается эта микронаука.

Эти науки сверкнули и умерли на протяжении жизни одного поколения ученых. Но умерли не в том смысле, что от них отказались. Просто они ушли в технологии, и то, что от них осталось, сводится к научному обслуживанию технологий. Такие науки стали «прикладными в квадрате».

В связи с этим мы наблюдаем феномен, который я бы назвал «дефундаментализацией информатики».

#### **В.Л.Стефанюк**

Дефундаментализация - это когда убирается финансирование РФФИ.

(Дружный смех в зале).

### **О.П.Кузнецов**

Все время подобные микронауки, составляющие тело информатики, уходят в технологии. Фундаментальных проблем в информатике как бы становится все меньше.

Были надежды на то, что фундаментальность информатики сохранится надолго. В частности, покойный А.П.Ершов (не только он один) пытался создать теорию программирования, но из этого мало что получилось. Программирование расплодилось на технологии. Таких примеров много.

И для многих из нас возникли драматические обстоятельства: для одних – по морально-психологическим причинам (особенно, для тех, кто привык формировать системы понятий и доказывать теоремы), а для других, в частности, для тех, кто пишет заявки в РФФИ – еще и по формально-бюрократическим причинам.

Получается, что надежды науки на сколько-нибудь продолжительную жизнь определяются ее способностью к самообновлению, т.е. способностью сменить внутри себя парадигму, оставаясь самой собой. Если наука на это способна, то она может жить.

Таким образом, для тех наук, которые я назвал «вечными», т.е. для естественных наук, смена парадигмы – это довольно медленный и естественный процесс эволюции. А для наших наук смена парадигмы – это просто главный способ выживания. Не сменишь парадигму – неминуемо уйдешь в технологии.

### **И.Б.Фоминых**

А разве это плохо? Это замечательно.

### **О,П,Кузнецов**

Я не говорю, что это плохо. Возникновение технологий – это замечательно. В конце концов, ради них мы и стараемся. Но, что при этом умирает наука – грустно. Хотя смотрю для кого. Я понимаю, что все это неоднозначно.

Так вот, такая наука, которая способна внутри себя сменить парадигму, оставаясь по сути той же наукой, в информатике есть. Это искусственный интеллект. Связано это с тем, что его объект столь же неисчерпаем, как объект физики или объект химии. Этот объект – интеллект, интеллектуальная деятельность человека.

Теперь, немного поговорим о парадигмах ИИ. Я позволю себе опираться (что мне сильно сократит изложение) на прекрасную лекцию В.К.Финна, прочитанную здесь в Пушкино, в которой он развернул перед нами двухтысячелетнюю картину теории рассуждений, обозначив в ней три крупных этапа. В наших терминах, эти три этапа связаны со сменой макропарадигм.

Кстати, я не сказал, что помимо макропарадигм, которые определяют всю науку, существуют еще парадигмы, соответствующие различным ее направлениям. Например, в физике были в свое время парадигмы флогистона, эфира и т.д. Такие микропарадигмы имеются всегда. Но сейчас речь идет о макропарадигме.

Итак, картина выглядела как некоторая триада. Вначале были установки психологического характера: считалось, что логика исследует формы мышления, хотя в результате возникла, прежде всего, аристотелева логика с ее силлогизмами. Потом был

этап логицизма, связанный с мощным развитием «чистой» математической логики и оснований математики. Наконец, на этапе, который протекает сейчас, логика понимается как некий «естественный процесс», а рассуждения – как «эмпирический феномен», т.е. происходит некий возврат к психологическому контексту, к реалиям интеллектуальной деятельности.

Эта триада характеризует филогенез двухтысячелетней науки о рассуждениях. И такая филогенетическая схема (как и в биологии) в какой-то степени, в модифицированном виде повторяется в онтогенезе конкретного экземпляра этой науки, т.е. в ИИ.

Действительно, что мы имеем в первый период развития ИИ? Мы имеем кристаллизацию его задач и первых подходов из романтических общекибернетических рассуждений (типа проблемы, «может ли машина мыслить», проблемы, которую обсуждали и такие великие люди как А.Тьюринг, Н.Винер и Дж.фон Нейман) Отсюда же возникла по-видимому исторически первая задача алгоритмизации плохо алгоритмизируемого в распознавании образов.

Второй период, на который как раз приходится все основные, реальные достижения ИИ, - это период, который я упрощенно назвал бы «компьютерной парадигмой». Есть еще известное выражение «компьютерная метафора». Данная метафора исходит из того, что мозг – это, грубо говоря, сложный компьютер, и поэтому его можно изучать путем моделирования на компьютерах. То, что я понимаю под «компьютерной парадигмой» есть в некотором смысле «компьютерная метафора наоборот». Сравним между собой работу мозга и современного компьютера. Мозг работает медленнее, в миллионы раз медленнее электроники. Он устает, подвержен эмоциям, ненадежен, забывает. Тогда, спрашивается, зачем нам изучать мозг, если электроника все делает быстрее и лучше. Мы просто будем брать те задачи, которые решает мозг и решать их по-своему, алгоритмизируя, разрабатывая те или иные формы представления информации и т.д.

На мой взгляд, именно эта парадигма, основные моменты которой – символическое представление информации и ее обработка с помощью логических и комбинаторных процедур – соответствует концепциям логицизма в триаде В.К.Финна.

Сейчас начинается третий период, предполагающий возврат к реалиям интеллектуальной деятельности. И если уж «чистые логицисты» выдвигают тезис о рассуждении как эмпирическом феномене, значит, действительно, возврат начинается.

Начало этого третьего периода по схеме Т.Куна означает кризис второго. Кризис второго, «компьютерного периода» (хотя, безусловно, он себя еще далеко не исчерпал) в научном плане выражается в постоянном столкновении с вычислительной сложностью, в трудностях самой формализации многих интеллектуальных процессов. Кроме того, кризис выражается в том, что происходит дефундаментализация ИИ. В рамках компьютерной парадигмы второго периода у нас остается все меньше фундаментальных задач, мы все дальше уходим в технологии, и все больше людей занимается ими. Это, конечно, хорошо, но терять фундаментальность в ИИ нам безусловно не хотелось бы.

С одной стороны, возврат к реалиям интеллектуальной деятельности означает наш возврат к первоисточнику – человеческому мозгу и его деятельности, что призвано обеспечить долгую продолжительность жизни нашей науки как фундаментальной. Но с другой стороны, стимул этого возврата и степень удаленности некомпьютерной парадигмы от компьютерной можно трактовать по-разному. В конце концов и неклассические логики, и немонотонные рассуждения тоже возникли из стремления к психологической адекватности, но все же они не выходят за рамки второго периода,

т.е. остаются под «прессингом» компьютерной парадигмы. И боюсь (хотя это конечно спорное заявление), что многим замечательным достижениям на этом пути будет суждено «остаться на бумаге» из-за неизбежных вычислительных сложностей, которые при их реализации должны возникнуть. Вспомним хотя бы судьбу проекта ЭВМ пятого поколения.

И вообще вычислительная сложность грозит нам полным тупиком, если не призвать на помощь все тот же первоисточник – мозг, которым еще недавно пренебрегали ввиду его медлительности и якобы ненадежности. И в этом плане (т.е. в связи с возвратом к психологическим реалиям) я хочу высказать достаточно радикальную точку зрения, на всеобщность которой я совершенно не претендую, а просто предлагаю ее здесь обсудить.

На мой взгляд, надо, наконец, основательно и систематически обратиться к тому огромному материалу, который наработан в когнитивной психологии. Не только потому, что ее спектр познавательных процедур намного шире того набора процедур, с которым мы пока оперируем в ИИ. Но прежде всего, потому, что существуют задачи, которые мозг решает (и, возможно, всегда будет решать) быстрее, чем компьютер.

Вот О.И.Ларичев в своем пленарном докладе произнес ключевое слово «тайна», и в формулировке моей позиции, которую я постараюсь далее пояснить, это слово также присутствует. Для меня важно понять «тайну быстроты медленного компьютера». Почему вообще, в принципе, есть задачи, которые «медленный мозг» решает очень быстро? В каком смысле мозг будет решать некоторые типы задач всегда быстрее компьютера? Всегда - не в смысле числа секунд. Уже было правильно отмечено, что с фундаментальной точки зрения обстоятельства поражения Гарри Каспарова от программы Deep Blue нам не интересны. Нам понятно, что, грубо говоря, такие программы делаются по принципу «сила (вычислительная мощность) есть, ума не надо».

А вот важнейшая особенность человеческого мозга заключается в том, что он решает подобные задачи совершенно по-другому. Как? Пока мы точно этого не знаем. Но очень естественно предположить, что он их решает во-первых с очень малой глубиной, т.е. за очень малое число шагов. Надо сказать, что на это обстоятельство обращал внимание еще Дж. фон Нейман. В одной из его статей в самом конце есть раздел с названием «Язык мозга не есть язык математики». Он устроен совсем по-другому, и одна из главных особенностей «быстрых» интеллектуальных процессов» - их малая глубина. Вторая особенность – очень высокая степень параллелизма, причем настолько высокая, что она, по-видимому, недостижима при использовании дискретной вычислительной техники. И связана с наличием несимвольных представлений.

Поскольку я уж сказал о несимвольных представлениях, то замечу, что ничего пугающего (для адептов традиционных взглядов на ИИ) в этих представлениях нет. Известно, что у человека есть по крайней мере две системы представления, хранения и обработки знаний. Это образная и символично-логическая системы. Как раз образная система и есть система несимвольных (субсимвольных) представлений.

Психологи уже давно отказались от примитивного представления образов как отражений реальности «один к одному». Образы необязательно имеют сенсорную природу, они могут формироваться и на основе знаков (слов). Существенно, что именно человеческий мозг обрабатывает такие несимвольные представления быстро, в то время как обработка несимвольных представлений, например, графических изображений, компьютером – всегда задача медленная. При пошаговом решении подобных задач всегда требуется очень большое число шагов.

При постановке задачи адекватного моделирования таких «быстрых» интеллектуальных процессов с методологической точки зрения наше положение гораздо выгоднее, чем у психологов, потому что наши критерии адекватности намного слабее. Психологи строят свои модели, стремясь доказать, что мозг работает именно так и не иначе. А доказывать это очень трудно. Все их гипотезы по этому поводу не слишком достоверны. Почти нет гипотез, которые разделялись бы всем психологическим сообществом, всегда находится группа, которая имеет свое мнение.

Более того, на мой взгляд, стопроцентной уверенности, абсолютно достоверного результата здесь не будет никогда. Серьезные интеллектуальные процессы надо изучать на целостном мозге (они очень сильно распределены), а при исследовании таких структур на нейронном уровне, какая бы тонкая аппаратура ни была, неизбежно влияние характеристик приборов на экспериментальные данные («наблюдающие приборы будут с неизбежностью изменять и даже разрушать исследуемые объекты»), т.е. будет работать принцип неопределенности, аналогичный соотношению неопределенностей В.Гейзенберга в физике.

Для построения эффективных структур, работающих с малой глубиной и очень высоким параллелизмом, конечные реализации должны выполняться на нейронных сетях. Примером возможной реализации «быстрых» интеллектуальных процессов на нейронных структурах является модель псевдооптической нейронной сети, о которой я неоднократно рассказывал, и которая опубликована в журналах «Теория и системы управления» (1995, №5), «Автоматика и телемеханика» (1996, №12).

В заключение, я хочу просто назвать две задачи, которые было интересно рассмотреть в рамках новой «некомпьютерной» парадигмы. Одна из них – это представление понятий. Хорошо известно классическое представление понятий в виде набора признаков, родо-видовых структур и т.д. Это традиция, идущая еще от аристотелевой логики. Аналогичные модели есть и в психологии. Однако, существует и другая психологическая модель, которая в терминах компьютерной парадигмы формализуется плохо, по крайней мере, не в виде «быстрого» процесса. Эта модель опирается на концепцию прототипа – типичного представителя для набора характерных признаков или репрезентативных примеров. Всем хорошо известно, что человек гораздо лучше и быстрее понимает, когда ему объясняешь на примерах. Создается впечатление, что задачу адекватного моделирования прототипического представления понятий гораздо проще решить на основе новой парадигмы в ИИ.

Вторая задача связана с формализацией загадочного понятия целостности. Есть много разных определений целостности. Например, в известной многим из нас статье Г.Р.Голицына и И.Б.Фоминых (см. «Новости ИИ».-1996.-№4) в качестве одного из определений целостности принимается «возможность восстановить целое по частям». Мне гораздо больше нравится другая идея, которая выражена в одной, не очень известной статье Губермана («Системные исследования» М., 1984), где представление о целостности связывается с понятием интерпретации: «отличительной особенностью целостного представления является то, что его элемент нельзя проинтерпретировать, не проинтерпретировав целого». Так если вырвать фрагмент изображения, то он может быть проинтерпретирован совершенно не так, неправильно. На этот счет, есть классические примеры, скажем, известная картина С.Дали «Невольничий рынок с исчезающим бюстом Вольтера», в которой можно увидеть две разные картины (невольничий рынок и бюст Вольтера) и, в результате, нельзя понять, что означает часть, если не поймешь картину целиком.

**Э.В.Попов**

Когда я давал рецензию на книгу Т.А.Гавриловой по приобретению знаний, я написал ей массу замечаний, одно из которых сводилось к известной фразе «страшно далеки вы от народа» (делающего реально работающие прикладные системы). Но вывод был таков, что лучшей книги еще никто не написал, и издавать надо безусловно. Так и здесь мне кажется, что в разговоре о парадигмах ИИ не надо противопоставлять научное знание и технологии. Любая парадигма зависит от достигнутого обществом уровня технологий. И настоящая наука не может полностью «раствориться, исчезнуть» в технологиях.

### **В.Б.Тарасов**

В том, что говорил Олег Петрович, мне очень понравились рассуждения о резком ускорении чередования парадигм в науках нового типа, таких как информатика и ИИ. Фактически он предложил еще один важнейший, дополнительный признак, отличающий науки об искусственном от естественных наук – скорость изменения парадигмы. В своей книге «Науки об искусственном», опубликованной у нас более четверти века назад, Г.Саймон по-видимому впервые отметил существенное различие между естественными (в широком смысле) науками, занимающимися анализом и описанием природных (естественных) явлений и объектов, и науками об искусственном, рассматривающими проблемы создания, синтеза новых искусственных объектов, обладающих желаемыми свойствами (эти свойства обычно связаны с имитацией поведения или выполнением функций соответствующего естественного объекта). Для обоснования этого различия он прибег к трем антиномиям: природа - человек, анализ - синтез, дескриптивная модель - нормативная модель. При этом главные тезисы Г.Саймона сводились к следующему.

1. Всякая естественная наука – это совокупность знаний о некотором классе вещей, объектов и явлений природы. Здесь вопрос «что изучать?» задает сама природа, тогда как в науках об искусственном, так же как и в ряде инженерных наук предмет исследования гораздо сильнее зависит от позиции исследователя (индивидуального или группового субъекта, научного сообщества). Его во-многом определяют предыдущие научные и технологические достижения общества.

2. Две вышеупомянутых разновидности наук отличаются по ведущему методу исследования, т.е. по ответу на вопрос «как изучать?». Классические науки в первую очередь применяют анализ, в то время как главный метод, находящийся в арсенале наук об искусственном – это синтез, конструирование.

3. Естественные науки и науки об искусственном различаются еще и по тому, какие модели – дескриптивные или нормативные - преимущественно используются. По сравнению с естественными, искусственные объекты рассматриваются не столько в описательных терминах, сколько на основе категории долженствования (что должен обеспечивать искусственный объект).

Теперь сюда можно еще добавить и предложенный О.П.Кузнецовым признак: быстрота смены парадигмы в науках об искусственном гораздо выше, чем для естественных наук.

Мне кажется интересным поговорить о том, как, в каких направлениях меняется парадигма ИИ. Ныне существующую парадигму в ИИ на мой взгляд можно охарактеризовать как несистемную и асоциальную. Для объяснения этой позиции, обратимся к истории ИИ.

Как появились основные направления, ведущие школы в ИИ? С позиций здравого смысла, бралось определение понятия интеллект через какой-то признак, какую-то человеческую способность. Например, интеллект понимался как способность решения задач на основе манипуляций над символами. Так зародилось символьное (логическое) направление в ИИ, обусловленное компьютерной (по терминологии О.П.Кузнецова) парадигмой. Или еще одно определение чисто на уровне здравого смысла: интеллект –

это прежде всего способность к обучению. И возникла коннекционистская парадигма, где ведущим направлением являются нейронные сети. Еще одно определение, утверждающее, что интеллект – это способность оперативной адаптации к динамической, быстро меняющейся, сложной среде. И стала развиваться эволюционистская парадигма в ИИ, представленная, например, школой генетических алгоритмов. Из аналогичных трактовок понятия «интеллект» по отдельному признаку (группе признаков) вышли такие классические направления ИИ как машинное восприятие и анализ сцен, машинный перевод и пр. Таким образом, в методологии ИИ на ранних этапах его развития всегда центральное место занимали редукционистский подход и принцип изучения целостной системы по ее частям (и их характеристикам).

Более того, с первых шагов ИИ как самостоятельной научной дисциплины в ней во главу угла был поставлен когнитивистский подход с акцентом на изучение и компьютерное представление процессов познания, когнитивной функции интеллекта. Соответственно, методы и результаты, полученные в когнитивной психологии, стали своеобразным психологическим базисом ИИ. Между тем другие важнейшие функции интеллекта – регулятивная, коммуникативная, ресурсная – оказались незаслуженно забытыми. Практически не рассматривались вопросы моделирования генезиса интеллекта, который тесно связан с процессами общения. Психологические исследования показали, что естественный интеллект не может появиться вне общества, и что человек не может интеллектуально развиваться без выполнения определенной деятельности в обществе себе подобных. Без адекватной внешней среды его когнитивное развитие оказывается сильно заторможенным, и если субъект не имел с раннего детства надлежащего социокультурного окружения, даже простое обучение членораздельной речи становится практически невозможным. На первых порах этот факт не нашел никакого отражения в работах по ИИ.

Неслучайно, в течение длительного времени в ИИ доминировал индивидуалистический подход к разработке программ, согласно которому они рассматривались как своего рода одинокие роденовские мыслители, способные при загрузке достаточного объема экспертных знаний самостоятельно решать те или иные задачи. Именно из этой идеи индивидуализма интеллектуальных процессов, трактовки интеллекта как характеристики отдельной личности возникли такие популярные в свое время проблемы, как вопрос о том, может ли машина мыслить, вызвавший к жизни многообразные варианты сравнения когнитивных возможностей человека и ЭВМ. В знаменитом тесте Тьюринга машине было предписано имитировать поведение отдельного человека, но не группы или микросоциума.

Таким образом, возникшая с самого начала в ИИ эгоцентрическая, несистемная, асоциальная парадигма имеет принципиальные ограничения в плане моделирования естественных интеллектуальных процессов. Трудно не согласиться с Д.Бобровым, который писал, что в целях повышения эффективности интеллектуальных технологий следует отказаться от «isolation assumptions» - предположений, справедливых только для изолированных систем ИИ. Необходимо ввести «размерность взаимодействия», обеспечивающую переход к групповому и организационному уровням интеллектуализации искусственных систем, которые связаны с развитием распределенного искусственного интеллекта и построением многоагентных систем и их сообществ.

На мой взгляд, этап становления ИИ как эмпирической науки уже близок к завершению, и нам нужно перейти к рассмотрению интеллекта как сложной, открытой, целостной системы (см. например, статью «Системно-организационный подход в ИИ», опубликованную в журнале «Программные продукты и системы». – 1997. - №3).

В этой связи мне импонирует тезис о том, что без понимания целого нельзя постичь частей, прозвучавший в заключительной части выступления О.П.Кузнецова. Как это ни странно, в ИИ до сих пор по настоящему не развит системный подход, давно нашедший широкое применение во многих других науках, в том числе и в психологии. Я хотел бы обратить ваше внимание на два важнейших аспекта системного подхода в любой области знания: принцип исследования объекта в иерархии систем и принцип выделения системных единиц (в данном случае, единиц, сохраняющих главные качества интеллектуальных систем). Хотя конечно надо было бы перечислить и другие принципы, включая принцип дополнительности Н.Бора (один из центральных принципов при интеграции различных моделей и представлений в ИИ), принцип неопределенности (о котором упомянул Олег Петрович), принцип самоорганизации и саморазвития интеллекта, принцип учета его коллективной природы и пр.

При исследовании объекта в иерархии систем мы можем изначально выделить три уровня: микроуровень, мезоуровень (собственный уровень системы) и макроуровень. То, что было сделано на первой стадии развития ИИ (я имею в виду период до возникновения распределенного искусственного интеллекта и многоагентных систем), относится к микроуровню. На новом этапе развития ИИ все это ни в коем случае не отвергается, а включается в инструментарий реализации отдельных единиц микроуровня - объектов, акторов, агентов и пр. Все это можно условно обозначить термином «Искусственный интеллект -1».

И только теперь, с позиций системного подхода, мы начинаем понимать, что наряду с «ИИ-1» может существовать и «Искусственный интеллект-2», где главное содержание работ связано уже не с идеей инженерии знаний, а с моделированием взаимодействия агентов, учетом процессов коммуникации, кооперации, координации агентов. Собственно развитие глобальной сети Интернет подтолкнуло нас к этому выводу. Именно в сторону этих проблем, относящихся к мезоуровню исследований по ИИ, и смещается «центр тяжести» фундаментальных работ в этой области.

Наконец, макроуровень, «Искусственный интеллект-3», - это уровень не просто многоагентных систем, а сообществ многоагентных систем с соответствующими социальными законами, т.е. уровень интеллектуальных организаций, где основным объектом исследования сдвинется еще дальше от традиционной парадигмы инженерии знаний.

В связи с этим уже сейчас я хотел бы предложить обсудить тезис о принципиально новом объекте современного ИИ. Объектом ИИ в наши дни является не просто разработка систем, основанных на знаниях, а прежде всего моделирование жизнедеятельности автономного компьютерного агента, а также взаимодействия агентов в некоторой сложной динамической среде.

Это означает, что и психологический базис, о котором мы говорили, становится комплексным и включает не только когнитивную психологию, но и социальную психологию и психологию деятельности. Компьютерный агент есть не что иное как виртуальный деятель. По сути дела, это активный, автономный, целенаправленный, внутренне мотивированный объект, живущий и взаимодействующий с себе подобными в некоторой виртуальной среде. Поэтому такие основные понятия психологической теории деятельности как намерения, установки, мотивы, желания, цели, задачи отныне прочно занимают свое место в структуре исследований по ИИ.

Более того, при рассмотрении интеллекта как открытой системы речь уже идет о моделировании не только чисто интеллектуальных (в узком смысле), но и интенциональных, эмоциональных и волевых процессов. Античная психологическая

триада «интеллект – чувства – воля» возрождается и наполняется новым смыслом в эпоху виртуальных агентов.

В заключение, мне бы хотелось, чтобы кто-нибудь поддержал или опроверг мое мнение о том, что сейчас наиболее «горячие» точки ИИ связаны в первую очередь с агентно-ориентированным подходом, построением теории агентов и моделированием их взаимодействия. Формируется системная (системно-организационная и системно-эволюционная) парадигма в ИИ. Этот процесс смены парадигмы ИИ сегодня приобретает все большую актуальность и в связи с широким внедрением новых коммуникационных технологий на базе Internet/Intranet, и в связи с той революцией, которая разворачивается в бизнесе, о чем наверное сейчас будет говорить Эдуард Викторович Попов.

### **Э.В.Попов**

Я думаю, действительно, более молодому поколению ученых и исследователей по ИИ, по крайней мере, тем, кто моложе 50 лет (а таких людей и, тем более, молодых ученых до 35 лет, на нашем круглом столе, к сожалению, очень мало) нужна новая парадигма. Пока в наше движение не придут новые люди со свежими взглядами, мы будем напоминать закрытый клуб любителей джаза. И вымрем как динозавры. Будут думать, от чего они умерли? От финансовых потрясений в стране? От неумения приспособиться к рыночной экономике? Нет, от старости парадигмы.

### **И.Б.Фоминых**

Пока выяснилось, что мамонты умерли от отсутствия парадигмы.

(Дружный смех в зале)

### **Э.В.Попов**

Я не берусь судить, в чем и как будет выражена эта новая парадигма. Не знаю, кто ближе к истине - О.П.Кузнецов, говоривший о некомпьютерной парадигме (что касается общей тенденции, здесь я с ним согласен, но сам термин «некомпьютерная» в эпоху становления информационного общества вызывает у меня неприятие) или В.Б.Тарасов, кинувший клич о «собираании земель ИИ в единое государство» на основе системной парадигмы. Но в том, что реалии современной экономической жизни, включая такое новое и на шумевшее направление в менеджменте как реинжиниринг и усовершенствование предприятий, ставят новые требования к теории и практике ИИ, и традиционная инженерия знаний и ныне существующие инструментальные средства здесь недостаточны – в этом я уверен, это бесспорный факт.

### **А.Н.Аверкин**

Кратко резюмируя то, что сказали предыдущие ораторы, можно утверждать, что парадигма символьных вычислений устарела. Как ее изменить? Любую парадигму можно менять в две стороны – в сторону ее усложнения или, наоборот, в сторону упрощения. То, о чем говорил Валерий Борисович, - это усложнение. Вместо одного агента берутся группа или сообщество агентов – объект исследования становится сложнее.

### **В.Л.Стефанюк**

Всегда ли? Это коллективная ответственность, а с ней всегда проще.

(Смех в зале)

### **А.Н.Аверкин**

Возможно. Если мы упрощаем парадигму, то по-видимому можем опираться на какую-то часть ранее существовавшей. Олег Петрович считает, что плохо, когда наука, доходя до своего логического конца, создает технологии и переходит к обслуживанию этих технологий. Это означает гибель науки. Но в то же время это дает то, чего часто не имеют классические (фундаментальные) науки: приложение, внедрение, финансирование. И происходит переток научных сил на обслуживание технологий. В недрах ИИ можно указать по крайней мере две ветки, которые очень быстро прошли путь от теорий до технологий и сейчас уже идут на обратной волне, переключившись на обслуживание технологий. Это нечеткие системы и нейросети. Интеграция этих направлений в результате такой обратной волны называется «мягкими вычислениями» (soft computing), а в более широком контексте, «вычислительным интеллектом» (computational intelligence). Мягкие вычисления – это комбинация нечетких систем, нейросетей и генетических алгоритмов. Если в основе большинства классических моделей ИИ лежали вывод в символической логике и классические методы поиска, а в случае интегрированных систем, также традиционные вычисления, то в основу вычислительного интеллекта кладутся приближенные модели, в частности, приближенные рассуждения. В отличие от классических моделей, ориентированных на нахождение оптимальных решений (численных, точечных результатов по точным данным), мягкие вычисления направлены на поиск приближенных решений, т.е. решений, которые нас в какой-то степени удовлетворяют в данной ситуации. В подобных конструкциях любое изменение модели ведет к изменению практически всего объема накопленной информации.

С одной стороны, в моделях мягких вычислений налицо определенный шаг назад по сравнению с традиционным ИИ: мы отказываемся от семантических интерпретаций и переходим на численный уровень. Но с другой стороны, выигрыш, который мы получаем за счет этого, часто позволяет перекрыть ряд слабых мест, присутствовавших в системах ИИ.

Так первые нечеткие системы в ИИ были ориентированы на извлечение и работу с символическими знаниями. Но за счет этого они не могли работать в реальном времени. Системы получались громоздкими, и ограничения по извлечению знаний не позволяли им работать в реальном времени, успевать за быстротекущим процессом. Теперь же нечеткие системы нового поколения и нейронные сети, опирающиеся на специальную элементную базу – нечеткие контроллеры и нейрочипы, - вполне позволяют работать в реальном времени.

Представление знаний в нейросетях и генетических алгоритмах – только числовое, в то время как в нечетких системах и классических системах ИИ – это смешанные числовые-символьные представления. Наконец, возможности оптимизации отсутствуют у всех этих компонентов, за исключением генетических алгоритмов.

Сочетание нечетких систем, нейросетей и генетических алгоритмов позволяет компенсировать недостатки всех этих моделей в отдельности. Оно позволяет решать те задачи, с которыми традиционный ИИ не всегда справляется, испытывая дефицит вычислительных ресурсов.

Теперь мне хотелось бы сделать некий прогноз. Представьте себе робота, который обладает инфракрасными датчиками, видеокамерой, манипуляторами, может

передвигаться по утилиту, принимать какие-то решения, обучаться. Как вы думаете, каков минимальный вес этого робота? Для современного уровня нанотехнологий.

### **Э.В.Попов**

Пятьдесят граммов.

### **А.В.Аверкин**

Я не ожидал такого уровня точности в оценке. На самом деле на выставке в Ганновере я видел семидесятиграммовый прототип подобного робота. Он целиком основан на технологиях вычислительного интеллекта. Это - нейрочипы, нечеткие контроллеры, генетические алгоритмы и средства гибридизации. Здесь нечеткая система играет роль мозга, нейросеть играет роль глаз, а генетические алгоритмы как бы оптимизируют работу этих глаз, не позволяя им попадать в локальные экстремумы и помогая в поиске глобального экстремума. Вышеназванная триада позволяет создавать системы, которые еще лет пятнадцать назад мы спокойно назвали бы системами ИИ. Но это не системы ИИ, поскольку они не основаны на символьных вычислениях.

Зато эти системы мы можем собрать в коллективы, как в случае с коллективом роботов, которые, например, вместе строят стену или решают какую-либо другую совместную задачу. Тогда мы переходим напрямую к проблематике, о которой говорил В.Б.Тарасов. Это проблема коммуникации и кооперации агентов в многоагентных системах, причем агенты обладают определенным интеллектом.

Таким образом, отказ от парадигмы символьных вычислений порождает две волны: вверх – в сторону усложнения исходных моделей, и вниз – в сторону их упрощения. Мне кажется, где-то на бесконечности эти волны сольются и породят то, что мы назовем ИИ XXI-го века.

### **В.Б.Тарасов**

Противопоставление идеологии искусственного интеллекта и вычислительного интеллекта, когда все, что было и есть в ИИ, отождествляется с символьной парадигмой, а другие представления о механизмах работы интеллекта, включая такие давно известные модели как нейронные сети, генетические алгоритмы, вероятностная логика, квантовые и колебательные модели интеллектуальных процессов, о которых еще в 70-е годы упоминал В.В.Чавчанидзе, выносятся за рамки дисциплины «искусственный интеллект», мне кажется надуманным. Мне ближе позиция, согласно которой на новом этапе исследований и разработок по ИИ, когда интеллект начинает пониматься как сложная, открытая, целостная система, как совокупность динамических структур, согласованно выполняющих определенные функции, одной из центральных становится задача интеграции результатов различных школ и направлений, построения гибридных, в том числе нейро-логических, моделей, которые более гибко и адекватно описывают феномен интеллектуального поведения в неопределенной среде. Да, формируются ранее невозможные, непривычные связи и гибриды. Но это новый этап работ по ИИ, а не что-то иное.

### **В.Л.Стефанюк**

Несколько замечаний по тому, что рассказывали все трое. Нет ли у вас ощущения, что когда речь пошла о нейронных сетях, это означает не смену парадигмы (исторически парадигма нейронных сетей была одной из первых, если не первой в ИИ), а скорее смену ситуации вокруг ИИ. Была кибернетика, в русле которой рассматривались и роботы, и проблемы управления с обратными связями. В ней были достигнуты серьезные практические результаты: построены автоматические регуляторы,

автопилоты и пр. Потом кибернетика исчерпала свои ресурсы как математически, так и технологически. Тогда и возникла необходимость в смене макропарадигмы, и появился ИИ. У меня был случай обсуждения с М.Минским этого вопроса, почему от кибернетики нужно было отказаться.

Итак, произошло изменение парадигмы. Но старая наука – кибернетика – не умерла. На самом деле не все ее ресурсы исчерпались. Немножко еще можно покрутить нейронные сети, как это сделал, например, Дж.Хопфилд. Та же ситуация с коллективами автоматов М.Л.Цетлина. Они возродились как прототип многоагентных систем, состоящих из простейших реактивных агентов. Мне кажется, что сегодня многие старые наработки вновь становятся актуальными. Мы думали, что они уже умерли, потому что пришла эпоха ИИ. Мы развиваем ИИ, а старая кибернетика, развиваясь и порождая новые идеи, нагоняет нас, наступая на пятки, и получается, что она делает почти то же самое. Наступает период конвергенции, слияния ранее расходившихся подходов.

В свое время был такой лысый француз Симон с горящими глазами, который когда только появились экспертные системы, не хотел признавать какой-либо их идейно-теоретической новизны. Он говорил: «А чем же мы всю жизнь занимались? Мы и делали всегда экспертные системы». То же самое сказал и Э.В.Попов, когда ему показали некоторые зарубежные системы: «У меня системы почище ваших, они давно работают с естественным языком, да и для более серьезных приложений».

Вот была ссылка на старую фундаментальную науку, а там ведь было то же самое. Два коротких примера. Возьмем физику. В свое время, после опытов Майкельсона большинство физиков отвергли идею эфира. А лет через двадцать выяснилось, что обработка результатов его опытов была ошибочной, и они отнюдь не показывают отсутствия эфира. Более того, до сих пор еще существуют ученые, которые, придерживаясь теории эфира, очень многое объясняют из явлений реального мира. В основном, это люди философского склада, но они составляют вполне заметную научную группу.

Один раз я слышал лекцию известного физика В.Брагинского, который занимается экспериментами в области современной теоретической физики. Он занимается гравитационными полями, ловит нейтрино. Все это очень абстрактные вещи, но пути решения задач – конкретные, чисто экспериментальные. Он поведал, что его очень интересовало, в какой момент теория Эйнштейна заняла свое теперешнее место в науке. И почему? Что произошло?

Известно утверждение А.Эйнштейна, что луч света приходит от звезды к нам на Землю и искривляется, потому что есть гравитационное поле, и т.д. Заметим, что и до А.Эйнштейна эти явления также находили весьма остроумные объяснения. Без всякой теории Эйнштейна. И эти объяснения были в книгах, учебниках, студенты учились по ним. Потом пришла теория Эйнштейна, которую вначале восприняли довольно плохо. Как и всякую другую новую теорию. Потом она стала публиковаться в учебниках как вторая теория, как вспомогательное объяснение. И Брагинский, просмотрев множество учебников, понял, что не было никаких объективных причин. Чисто психологический феномен. Просто старую теорию перестали указывать, а новая (в каком то смысле более полная) осталась. Но никакой научной революции с полной сменой парадигмы не было. Никто не показал, что старое было ошибочно, а новая теория все эти ошибки преодолевает.

**О.П.Кузнецов (В.Л.Стефанюку)**

То, что вы сейчас говорите, никак не противоречит вышесказанному. Эволюции и революции в науке сосуществуют и накладываются друг на друга.

### **Б.А.Кобринский**

Конечно и мозг не компьютер, и компьютер не мозг. Но в их сопоставлении, в компьютерной метафоре есть доля истины, поскольку частично компьютер использует некоторые процедуры, которые использует мозг, и наоборот, в мозгу имеются явления, которые были восприняты и исследовались в компьютерных терминах. И то что люди быстро решают сложные интеллектуальные задачи, например, задачи узнавания, в условиях высокой неопределенности (или недоопределенности) и нечеткости, происходит благодаря таким феноменам как интуиция или инсайт, Это неслучайно. Всегда остаются знания, которые мы не можем вербализовать или даже извлечь.

Вот Олег Петрович говорил о том, что есть символьная и есть аналоговая информация. Но и образ, и понятие-знак можно назвать одним термином – символ. И тогда можно с единых позиций трактовать информацию как символьную. На самом деле, это всего лишь вопрос классификации, придуманной нами. Может быть, даже говоря об образно-понятийных базах знаний, мы сможем в основе своей сохранить старую парадигму, подразумевая, что существуют символы и символы, т.е. символы разного порядка. Сюда же войдут и прототипы, т.е. некие целостные, нерасчленимые образы, работа с которыми является в основном прерогативой правополушарного мышления. Может такой путь окажется более продуктивным?

Конечно не надо отбрасывать всего остального. Все-таки человек работает в двух ипостасях: в системе рассуждений (аргументации) и системе образов. Тогда можно говорить о чередующихся включениях тех или иных механизмов символьного восприятия действительности. Они могут последовательно сменять друг друга. Например, сначала появляется образ, затем до какого-то момента идет аргументация. Опять происходит «всплеск восприятия» и потом на этапе рассуждений все уточняется. Все мы знаем замечательную историю с таблицей Менделеева, которая приснилась ему во сне. В поисках какого-то решения ученый бьется неделями, месяцами, годами, а потом оно спонтанно приходит. Сначала возникает нечто цельное, а вербализация приходит потом.

Мне кажется, что будущая парадигма ИИ может пойти по этому пути, став гибридной, о чем говорил, в частности, В.Б.Тарасов. Целесообразность гибридного подхода очевидна, вопрос главным образом состоит в том, как и что соединять. И в какой-то степени, как называть (хотя, безусловно это вещь вторичная). Тем не менее, может быть мы сами себя загоняем в ловушку, все время говоря, что есть разные типы понятий, мы хотим их «собрать в одну корзину», а они не ложатся. Если мы начнем их представлять как понятия одного порядка, но имеющие свои отличия, специальные свойства, то тогда у нас больше шансов их успешно объединить.

### **О.П.Кузнецов**

Одно локальное замечание, Борис Аркадьевич. Вы употребляете понятие символ как минимум в двух смыслах. Есть, например, понимание символа как знака в семиотике. Есть трактовка символа по К.Г.Юнгу. Когда же речь идет о символьных представлениях, то имеется в виду очень простое понимание символа как некоторого элементарного знака.

### **Г.С.Осипов**

В отличие от знака символ не имеет ни семантики, ни прагматики.

### **О.П.Кузнецов**

Всего-навсего элемент алфавита и не более того. Вот что имеется в виду под символом, когда говорят о символических представлениях. Поэтому лучше символичные и образные представления в ИИ называть разными именами. Да и психологи не поймут такого смешения.

### **Б.А.Кобринский**

Вы правы, Олег Петрович. Я использовал в данном случае понятие «символ» скорее в семиотическом смысле.

### **В.К.Финн**

Мне хотелось бы обратить внимание на некоторые специфические черты нашей науки. Первое – это знание, как таковое. Второе – это рассуждение, как таковое. И третье – это мышление, как таковое. В силу этих трех обстоятельств предмет этой нашей науки чрезвычайно сложен, и она, грубо говоря, не может загнуться. Если физика может найти элементарные частицы (а вдруг теория элементарных частиц когда-нибудь захлебнется, а вдруг электрон исчерпаем), то она в принципе может дойти до некоторого предела. Наука же, которая изучает знания, мышление и рассуждения не имеет пределов по той простой причине, что человек не в состоянии познать сам свое мышление (парадокс культуры). Как только он познал свое мышление до конца, он стал богом. А это теологически невозможно.

В этом смысле, ИИ – такая дисциплина, которая не имеет предела. И в отличие от ряда других дисциплин, например, гуманитарной науки социологии, или даже от ряда разделов физики, кризиса здесь быть не может. Хотя, как всегда в любой науке бывают Сцилла и Харибда, и между ними надо пройти. С одной стороны, предмет данной науки столь открыт, что в нем все время можно продвигаться, и это явление придает оптимизм. С другой стороны, имеется некоторая опасность, которую в плохом смысле слова я называю «инженерным мышлением».

Это отнюдь не означает некое высокомерное отношение со стороны математически мыслящих людей к инженерной деятельности как таковой. Наоборот, инженерная деятельность вызывает у меня всяческое уважение. Дело в том, что некоторые математики и физики действуют с позиций инженерного мышления. Под инженерным мышлением в плохом смысле слова я понимаю слепую веру в некий формальный аппарат.

### **Э.В.Попов**

Скорее, наоборот, инженеры в отличие от математиков более свободны в отношении с формальным аппаратом. Никакой-такой слепой веры у них нет. Этим больше грешат математики.

### **В.К.Финн**

Не знаю. Часто инженеры берут некий формальный аппарат и начинают где-то, не задумываясь, его применять. И убеждены, что комбинируя этот аппарат, наращивая его возможности, можно дойти до истины. Вот в это я не верю. Если мы возьмем нейросети, генетические алгоритмы, нечеткие множества, как-то их объединим и будем

вычислять, то это отнюдь не решение проблем, связанных со знанием, рассуждением и мышлением. Говоря словами Б.Окуджавы, «ах, это, братцы, о другом».

Теперь рассмотрим каждый из компонентов. Нечеткие множества. Мода гуляет по белу свету. Известно, как наказали Тома Сойера. Тетушка Полли наказала Тома Сойера, он стал красить забор, получая от этого удовольствие, и после этого все возжелали красить забор. То же самое происходит в науке. Один человек занялся нечеткими множествами, и пошла мода. Журналы. Конференции. Престиж.

Что же это такое? Если задуматься по поводу нечетких множеств, то мы поймем, что во-первых, это логики Лукасевича. И нечего мудрости разводиться по этому поводу. А про логики Лукасевича есть некоторое точное знание; в некотором смысле, это арифметические конструкции.

Теперь зададимся вопросом, нечеткость и неопределенность могут быть выражены чисто арифметически. Конечно, нет. Поэтому, совершенно очевидно, что при всей моде и некоторых практических успехах (все-таки это – формальный аппарат, он что-то дает) мы никак не можем считать, что решением всех наших проблем являются нечеткие множества. Ибо в их природе лежит неадекватная трактовка неопределенности.

Что касается генетических алгоритмов, то при всей их красоте и остроумии, я сохраняю определенный скепсис. Я убежден, что для науки очень важен принцип осмысленности. Принцип осмысленности математических преобразований, который к сожалению очень редко соблюдается, был сформулирован Л.Д.Ландау. Его идея заключалась в том, что физически осмысленной является такая теория, в которой каждый шаг математических преобразований имеет интерпретацию в наблюдаемых терминах. Вот такая строгая физическая осмысленность. Совершенно очевидно, что большая часть современной науки не имеет физической осмысленности по Ландау. Поэтому имеется некий разрыв между аппаратом и его интерпретируемостью. В некоторых случаях, как следствие гениальности создателей математического аппарата, эта интерпретируемость наблюдается. Но когда люди начинают сознательно при решении задачи искать прежде всего подходящий аппарат, то разрыв между интерпретируемостью и средствами решения увеличивается. Отсюда вытекает некий важный принцип для любой молодой науки: всякая молодая наука должна искать прежде всего проблемы, которые она решает, и адекватные средства решения этих проблем. И очень часто адекватные средства решения этих проблем отсутствуют как готовый формальный аппарат. И горе той науке, которая в первую очередь начинает подыскивать годный, готовый формальный аппарат.

Разумеется, это не означает, что не может вовсе быть случаев, когда такая адекватность присутствует. Булева алгебра и релейные контактные схемы. Динамические аналогии, когда выяснилось, что механические, электрические и прочие колебания описываются одинаковыми дифференциальными уравнениями. Но эти случаи редки и поэтому не могут быть методологическим принципом.

В связи с этим, очевидно, что искусственный интеллект как самостоятельная наука должен создавать свои методы. И прежде всего должен уважать свои методы, а не искать спасение от дяди, который что-то там придумал на стороне.

Мое убеждение состоит в следующем. Надо различать две цели создания интеллектуальных систем. Первая цель – имитировать человека. Вторая цель – усилить способности человеческих рассуждений при решении задач. У этих целей есть область пересечения, но это разные цели. Совершенно очевидно, что цель «усиление

рассуждений» открывает неограниченные возможности для создания аппарата, позволяющего решать проблему.

Всякая наука должна очертить круг своих проблем. Интеллектуальные системы – это системы с решателями задач. Решатель задач – это рассуждатель + вычислитель + интеллектуальный интерфейс, т.е. система, поддерживающая общение. В зависимости от богатства этих средств интеллектуальной системы мы будем получать те или иные возможности как для имитации, так и для усиления человеческих рассуждений.

В связи с этим, следует сформулировать чрезвычайно важный принцип, который направлен против применения готового формального аппарата в ИИ. Искусственный интеллект – наука совершенно экспериментальная. Эксперимент при создании интеллектуальных систем является важнейшим принципом. В этом смысле и логика является наукой экспериментальной, раз она связала свою судьбу с ИИ. Мы создаем фигуры рассуждений в связи с решением проблем и настраиваемся на эти проблемы. При этом мы ничем не ограничены, кроме математической культуры и культуры естествоиспытателя. Конечно, здесь могут быть разные средства. Так в последнее время в инструментарий математической логики были включены такие системы, какие раньше никогда не использовались, например, немонотонные логики и рассуждения.

Когда мы говорим о решении задач с помощью компьютеров, то для систем ИИ чрезвычайно важным признаком оказывается некоторая автономность. В самом деле, некоторые почтенные программные системы, которые решают сложные задачи, к системам ИИ относить не следует. Есть некоторые специальные черты, которыми характеризуются системы ИИ, я их выше перечислил. Эти черты также связаны с изменчивостью, приспособляемостью. Антропоморфные черты. Если они все в совокупности присутствуют, хотя бы в некоторой степени, то этого заведомо достаточно, чтобы система считалась интеллектуальной.

И последнее соображение. То, что сейчас происходит в мире исследований по ИИ, можно назвать предварительной стадией создания рассуждающих или «думающих» систем. Мы с Дмитрием Александровичем Поспеловым много беседовали по этому поводу. До сих пор мы не имеем средств представления знаний, которые были бы антропоморфными. Под антропоморфными средствами представления знаний я понимаю такие средства, в которых бы рождались утверждения из введенных туда понятий. До последнего времени все системы, которые у нас были, относились либо к сугубо декларативным, либо к процедуральным, либо к декларативно-процедуральным. Порождать представление знаний так, чтобы из некоторой развитой понятийной среды синтезировались утверждения – до этого мы еще не дошли. Есть, правда, замечательные работы Г.С.Плесневича, но это пока предыстория вопроса. Это еще далеко не полное решение данной проблемы, да и автор на это не претендует. Тем не менее, эти работы можно считать первыми ласточками на пути к синтезу утверждений в определенной понятийной среде. Когда мы действительно сможем от идей, рождающихся как предпонятийные конструкции, переходить к сочетанию этих идей и порождению утверждений, тогда аппарат систем ИИ станет более антропоморфным.

В заключение, отмечу следующий парадокс. Основные области применения ИИ - такие дисциплины, которые менее всего подвержены математизации. Чем меньше дисциплина подвержена математизации, тем она интереснее для ИИ. Для нас как специалистов по интеллектуальным системам чрезвычайно важны слабо формализованные области, в которых понятия плохо связаны. Вторжение ИИ в эти области поднимает уровень их формализации и экспериментально оправдывает ИИ.

Мы еще не расправились с возможностями рассуждающего субъекта. Когда мы их хотя бы аппроксимируем, тогда мы сможем говорить о коллективах субъектов. Пока

же, при всей привлекательности многоагентных систем, они все-таки не являются центральными для развития основного аппарата ИИ. Его главная проблема – это глубокая антропоморфность. Создание коллектива – это вторичная проблема.

### **Э.В.Попов**

Все, что сказал Виктор Константинович, очень интересно. Он выразил самую суть классической логической парадигмы в ИИ. Искусственный интеллект – это системы, работающие на знаниях и включающие логический решатель задач. Знания в логике включает понятия, суждения и рассуждения. Автоматическая генерация суждений и рассуждений в некоторой понятийной среде – вот главная перспективная цель. И никакие расширения предмета или переосмысления его основного содержания искусственному интеллекту не нужны, поскольку они грозят «размыть логический фундамент» ИИ, нарушить жесткий принцип физической (я бы сказал, логической) осмысленности науки. Очень понятная и симпатичная мне позиция, которую я все же не разделяю.

### **А.Н.Аверкин**

Мне представляется, что есть явное противоречие между тезисом об основных сферах приложения ИИ как трудноформализуемых областях и неприятием «инженерного мышления», которое часто связано с развитием и приспособлением некоего математического аппарата к описанию конкретной проблемной области, как правило открытой и характеризующейся неполнотой, неточностью, нечеткостью располагаемой информации. Задолго до Л.Заде были известны трехзначная логика Лукасевича и k-значная логика Поста, взвешенные графы, а также способы построения распределений, предложенные в математической статистике, но именно он, предложив простую и наглядную формальную интерпретацию неточности и многозначности описаний на естественных языках, выдвинул целостную программу построения прикладной математической дисциплины нетрадиционного типа, опирающейся на нечеткие множества и теорию возможности. Эта программа основывалась на сформулированном Л.Заде принципе несовместимости: «по мере возрастания сложности исследуемых нами систем наша способность делать точные и в то же время значимые суждения об их поведении уменьшается, вплоть до некоторого порога, за которым точность и осмысленность (релевантность) становятся практически несовместимыми характеристиками». То, что идеи Л.Заде привлекли внимание целой армии прикладников к ранее не востребуемым, элитарным областям математики, хотя «чистота и невинность» последней пострадали, это безусловное благо. И конечно появление в 80-е – 90-е годы специализированных нечетких аппаратных средств, а также большого количества коммерческих нечетких систем в самых различных областях, их широкое использование в промышленности, на транспорте, в бытовой технике, служат наилучшим доказательством целесообразности и жизнеспособности данного направления.

Я признаю, что в рамках этого течения было и есть немало «пены», немало работ и результатов, подверженных влиянию моды и не удовлетворяющих критериям строгости, принятым в классической математике. Однако помимо интересных приложений, здесь имеются и глубокие математические результаты, особенно в нечетких алгебрах и нечетких топологиях. И конечно существующий спектр нечетких и лингвистических логик, когда расширения по сравнению с классическими логиками касаются и набора значений истинности (переход к континуальной или лингвистически заданной истинности), и интерпретации операторов конъюнкции, дизъюнкции, отрицания и импликации (можно, в частности, использовать различные семейства операторов конъюнкции и дизъюнкции на базе аксиоматики треугольных

норм и конорм, или строить гибкие параметризованные операторы, настраивая их под конкретную задачу), и существа базовых аксиом, нельзя свести к логикам Лукасевича.

### **В.Б.Тарасов**

Я согласен с Виктором Константиновичем в том, что, скажем мягкие вычисления, предполагающие построение гибридных систем из нечетких регуляторов, нейронных сетей и генетических алгоритмов, - это отнюдь не универсальная схема интеграции знаний, отнюдь не панацея от всех бед. Ее эффективность безусловно зависит от типа проблемной области. Эта схема лучше подходит для задач управления динамическими системами или технологическими процессами, но слабо приспособлена для решения задач проектирования или обучения. Например, для задач моделирования процесса проектирования прежде всего нужны гибридные немонотонные нечеткие и модальные нечеткие логики (кстати, такими логическими гибридами стали заниматься многие почтенные логики, например, Д.Габбэй). Так что мягкие вычисления – это лишь пример одной из многих возможных схем гибридизации. Уже известны и другие примеры, в частности, варианты интеграции нейронных сетей и сетей Петри, нейро-оптические модели О.П.Кузнецова, и т.д.

Что же касается предложения сначала полнее отразить особенности рассуждений индивидуальных субъектов, в частности, создать системы порождения утверждений в определенной среде понятий, а затем уже выходить на групповой уровень, то такая позиция представляется не слишком конструктивной. И с точки зрения психологических предпосылок организации рассуждений (поскольку, данные ряда психологов, например, А.В.Брушлинского и О.К.Тихомирова, говорят о том, что протекание процессов мышления и рассуждений часто зависит от условий коммуникативного взаимодействия думающих субъектов). Сопоставление сходства и различия механизмов индивидуальных и групповых рассуждений – очень важная и интересная задача для специалистов по ИИ. И конечно в современных условиях не следует настаивать на индивидуалистической парадигме в контексте практических приложений ИИ, где удельный вес неуклонно сдвигается в сторону распределенных систем и сетевых организаций.

### **В.Н.Поляков**

Мне кажется, что ИИ – сначала моделирующая наука и лишь потом экспериментальная. Ведь мы экспериментируем на образцах, на прототипах будущих систем, словом, на моделях.

Касаясь соотношения известных аппаратов и новых аппаратов, отмечу, что в свое время в физике этот вопрос был решен очень просто. Экспериментальная физика сперва пытается применить известный аппарат, а когда невозможно объяснить некий важный феномен, теоретическая физика выдумывает что-то новое ради его объяснения.

Что касается противоречия парадигм, я его не вижу. Парадигмы должны взаимодействовать, поскольку каждая парадигма нацелена на решение своего класса задач. А круг этих задач чрезвычайно широк и постоянно меняется.

Я думаю, что когда В.Б.Тарасов говорил о признаках, характеризующих интеллект, им была пропущена очень важная особенность интеллекта – обработка естественного языка. Сегодня уже говорилось, что около 15% узлов в Интернете, занимающихся проблемами ИИ, относятся к естественному языку.

Мне кажется, что сейчас речь идет не столько о смене парадигмы, сколько о смене целей ИИ. Что является источником знаний? Основным источником знаний являются

книги. Классик говорил: «любите книгу и придете к знаниям». Понимание текста есть прямой путь к знаниям. Большинство знаний приходит из текстов. И научить без текста нельзя.

**Э.В.Попов**

Можно.

**Г.С.Осипов**

Я хотел бы еще раз сформулировать принципиальный вопрос «что есть искусственный интеллект»? Это – наука об устройстве мозга или это наука о том, как усилить человеческие возможности в решении сложных задач (именно в плохоструктурированных областях, о чем говорил Виктор Константинович)? Если берется первый ответ, то (по отношению к основным достижениям классического ИИ) это действительно «песня о другом». Если это второе, то следует признать, что все средства хороши, которые помогают усилить интеллектуальные возможности человека, решающего сложные задачи в плохоструктурированных предметных областях. И пусть расцветают сто цветов: каждая парадигма, каждая дисциплина, которая помогает решать эту задачу, должна выполнять свою роль. Если мы говорим о перцепции или о восприятии гештальтов, то чем плохи нейронные сети. Как раз они здесь очень уместны. А перцепция, восприятие – первый шаг к интеллекту. Дальше, если мы говорим о рассуждениях, то разумеется надо развивать символическую парадигму. Но кто мешает нам строить и изучать гибриды? Как взаимодействуют, скажем, генетические алгоритмы или нейронные сети с базами знаний и решателями? Это достаточно интересная проблема и сейчас, мне кажется, она выходит на передний план в связи с активными попытками создания прототипов интеллектуальных систем нового поколения.

Я не сторонник чисто инженерного подхода к ИИ, но также считаю, что это – экспериментальная наука. И полигоном для испытания его концепций и методик должно быть решение практических задач, которые в свою очередь могут дать толчок развитию новых подходов. А с другой стороны, в какой-то момент надо продемонстрировать возможности науки. Любая наука должна удовлетворять некоторые социальные заказы. Поэтому я считаю, что неотъемлемой частью ИИ являются и наука о рассуждениях, знаниях и мышлении, и исследование способов восприятия, быстрой обработки образной информации, и анализ проблем взаимодействия и кооперативного поведения интеллектуальных систем.

