

СИСТЕМО-ОРГАНИЗАЦИОННЫЙ ПОДХОД В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ

В.Б.Тарасов

Abstract. Classical AI paradigm, based on an individualistic approach to the nature of intelligence, supposes the definition and modeling of specific intelligence aspects (e.g. knowledge processing). Hence it has some principal limitations in the scope of building intelligent organizations. Before discussing two main tendencies in modern AI - integration and decentralization - various historical schools in AI are mentioned, and basic postulates of classical (logical) school are considered. Some principles of systemic approach in AI are formulated, and as a result, systemic approach is confronted versus logical approach in AI. So a new systemic-organizational paradigm in AI is introduced, and some levels of AI studies together with appropriate basic entities are shown. Among basic units in modern AI the concept of agent is of primary concern. Different definitions of agents are discussed, and the classification of agents is given. A general activity model is suggested as a kernel for selecting agent's architectures.

Начиная с исторического Дартмутского семинара, положившего начало самой научной дисциплины "искусственный интеллект" (ИИ), первые 30-35 лет исследований в этой области были проникнуты духом скорее аналитических, чем синтетических подходов, и опирались скорее на индивидуалистические, чем коллективистские представления о природе интеллекта.

С одной стороны, такие аналитические, разрозненные подходы, основанные на выделении некоторого частного свойства или аспекта понятия "интеллект" и развитии соответствующих формальных моделей, привели к возникновению различных, расходящихся направлений в ИИ. Например, сведение понятия "интеллект" к обработке знаний (символов) и решению задач обусловило развитие и долгое безраздельное господство когнитивистских, и в особенности, логических моделей в ИИ. Определение интеллекта как способности к обучению или адаптации в изменчивой, динамической среде вызвало появление бионического направления, представленного как коннекционистскими моделями - нейронными сетями, так и алгоритмами эволюционного моделирования (в частности, генетическими алгоритмами). В свою очередь, рассмотрение восприятия как основы построения внутренней модели внешнего мира привело к становлению машинного восприятия, обработки изображений и анализа сцен, а трактовка интеллекта как способности к коммуникации и пониманию фраз на естественном языке - к созданию естественно-языковых систем и интерфейсов и средств машинного перевода [1]. Таким образом, осуществлялось противопоставление и происходило неуклонное расхождение отдельных школ ИИ, например, логической и нейробионической школ, инженерии знаний и распознавания образов, и т.п. (рис.1).

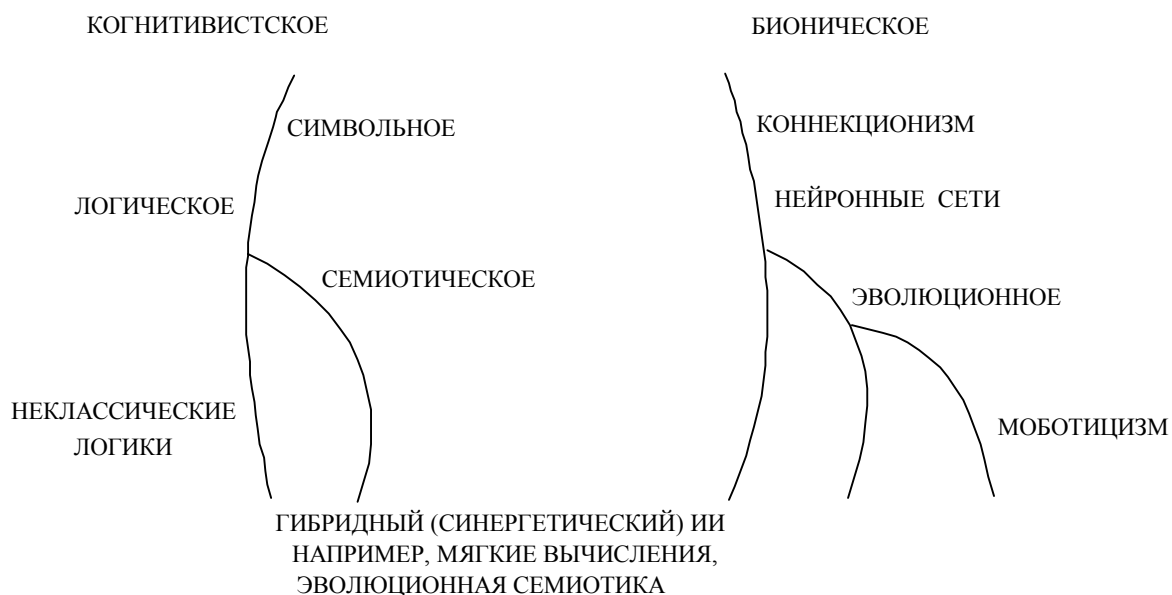


Рис.1. Основные направления в ИИ (1-й этап - расхождение, 2-й этап - схождение)

С другой стороны, с самого начала исследований по ИИ возобладал индивидуалистический подход к разработке программ, согласно которому они рассматривались как своего рода одинокие роденовские мыслители, способные при загрузке достаточного объема экспертных знаний самостоятельно решать задачи. Именно из этой идеи индивидуализма интеллектуальных процессов, трактовки интеллекта как характеристики отдельной личности возникли такие знаменитые в свое время проблемы как вопрос, может ли машина мыслить, вызвавший к жизни всевозможные варианты сопоставления функциональных возможностей человека и ЭВМ, а также известный тест Тьюринга, в котором машине было предписано имитировать поведение отдельного человека (и она считалось интеллектуальной, если ее было невозможно отличить от этого человека в процессе диалога), и т.п. Такой "солипсический" подход повлек за собой развитие последовательных, сосредоточенных, централизованных интеллектуальных систем, в которых управление было связано с процессами, выполняемыми один за другим, а для решения задачи оказывалось достаточно знаний, циркулирующих в одной экспертной системе. Фактически принимался постулат о "всеведущем эксперте" - некоем воплощении "господа-бога в данной проблемной области".

В целом, классические модели ИИ проникнуты рационализмом, поскольку в них интеллект связывается главным образом с рациональными методами решения задач с использованием эвристических процедур [2]. Подобные трактовки интеллекта носят весьма ограничительный и статичный характер, поскольку в них игнорируются реальности возникновения и развития интеллекта человека, связанные с коммуникативными и регулятивными процессами. В самом деле, естественный интеллект не может появиться вне общества, и 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.

Таким образом, можно утверждать, что возникшая с самого начала несистемная, асоциальная, индивидуалистическая парадигма развития ИИ имеет принципиальные ограничения в плане моделирования естественных интеллектуальных процессов. При этом базовая концепция интеллектуальных систем как систем, основанных на знаниях, неявно предполагает универсальность глобального централизованного подхода, согласно которому для решения проблемы возможно собрать в единой базе и эффективно использовать знания различной природы, поступающие из различных источников.

ОСНОВНЫЕ ПОСТУЛАТЫ КЛАССИЧЕСКОЙ ШКОЛЫ ИИ: ПРЕДЕЛЫ ПРИМЕНИМОСТИ И НЕОБХОДИМОСТЬ ПЕРЕСМОТРА

Следуя идеям работ [1-5], сформулируем набор основных постулатов, характеризующих классические интеллектуальные (экспертные) системы первого поколения, т.е. логическую школу ИИ. Это:

- 1) **постулат замкнутости** (closed world assumption), определяющий изучение интеллектуальных систем как изолированных, без учета процессов коммуникации с внешней средой и, в частности, с другими интеллектуальными системами;
- 2) **постулат централизации**, согласно которому интеллектуальная система рассматривается как централизованно управляемая;
- 3) **постулат представления (эпистемология ИИ)**, в соответствии с которым одна из главных задач работ по ИИ состоит в разработке методов внутреннего представления внешнего мира;
- 4) **постулат знаниецентризма** в ИИ, согласно которому ведущую роль в ИИ играет инженерия знаний, в частности, автоматическое формирование понятий и рассуждений, ("рассуждения и планы первичны, действие вторично" [4]), т.е. важнейшее место в моделировании познания занимает логический вывод;
- 5) **постулат бестелесности ИИ**, утверждающий возможность отрыва интеллекта от тела, например, изучения когнитивных процессов и знаний, абстрагируясь от особенностей регуляции действий и перцептивно-моторных процессов;
- 6) **постулат относительной независимости познания и эволюции** в ИИ, согласно которому когнитивные процессы можно рассматривать вне и независимо от обучения и эволюции;
- 7) **постулат соответствия познания и языка** в ИИ: кинематика познания представима в пропозициональных терминах, т.е. траекторию состояний знаний можно описать лингвистически в терминах естественного языка.

В последнее время в связи с развитием тенденций интеграции и децентрализации в ИИ эти казавшиеся ранее незыблемыми постулаты требуют пересмотра. Какие именно постулаты здесь следует пересмотреть зависит от того, какие именно модели, относящиеся к различным направлениям ИИ, объединяются, и на какой основе. Так, например, коннекционизм, будучи связанным с нейропсихологией, настаивает на важнейшем значении обучения в исследовании познания, а также проповедует идею возникновения символической системы формальных нейронов [6]. Полностью противоположную позицию по отношению к логическому подходу занимают приверженцы так называемого моботицизма (roboticism) - своеобразной апологии оперативного взаимодействия без внутреннего представления в рамках ИИ -, отрицающие и централизованное управление, и роль формирования понятий и рассуждений, и рациональную кинематику, и гипотезу "бестелесного интеллекта", и отделимость знания от обучения (см. работу "Интеллект без представления" Р.Брукса [7]). Похожую, хотя и менее радикальную позицию, занимают сторонники теории

сообщества разума (society of minds) М.Мински [8], распределенного ИИ и децентрализованного ИИ [9-11], антиформалистского подхода Р.Шенка [12] и др.

На наш взгляд, суть происходящей ныне смены парадигмы в ИИ, можно выразить, вспомнив название знаменитой монографии И.Пригожина "От существующего к возникающему" [13] (исследование и моделирование процессов интеллектогенеза) и призыв Д.Боброва [14] рассмотреть пути преодоления "isolation assumptions" - предположений, справедливых только для изолированных систем ИИ. Необходимо ввести "размерность взаимодействия", обеспечивающую переход от индивидуальных интеллектуальных систем к группам и сообществам совместно (параллельно) работающих интеллектуальных систем. Некоторые принципы такого перехода сформулированы в [15].

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В ИИ

Двумя наиболее важными тенденциями развития искусственного интеллекта в 90-е годы стали **интеграция и децентрализация**. Во-первых, еще в 80-е годы экспертные системы стали объединяться с традиционными информационными технологиями, в результате чего появились гибридные экспертные системы [16-18], что повлекло за собой развитие ряда **эмпирических и теоретических способов интеграции**, включая метамодели и архитектуру доски объявлений (имеющую важное значение и в контексте распределенного искусственного интеллекта). В таких интегрированных системах должны поддерживаться не только различные модели знаний, но и различные типы рассуждений, и реализовываться как инженерия знаний, так и обработка образов [19].

Во-вторых, пошел более фундаментальный процесс **интеграции различных атрибутов** (или если угодно, различных "рабочих определений") интеллекта в интересах взаимной компенсации недостатков и объединения преимуществ разнородных моделей и, как следствие, получения синергетических (нелинейных эффектов) [20] (рис.1). Примерами могут служить интегрированные нейро-логические [21], нейро-нечеткие модели [22, 23] и "мягкие вычисления" по Л.Заде [24], нейро-оптические модели О.П.Кузнецова [15], интегрированные модели совмещенной разработки [25], варианты интеграции нейронных сетей и сетей Петри [26] и пр. Так в рамках "мягких вычислений" три аспекта интеллекта - управление неопределенностью, обучение и адаптация в процессе эволюции - объединяются путем представления нечетких продукционных моделей в обучаемой нейронной сети, оптимизация которой происходит с помощью генетических алгоритмов. Модели совмещенной разработки требуют построения модальных нечетких и немонотонных нечетких логик, а нейро-оптические модели предполагают объединение коннекционизма с волновой теорией.

Вторая тенденция - распределение или децентрализация ИИ - связана в первую очередь с исследованием процессов коммуникации, кооперации и координации. При этом распределенные системы ИИ могут иметь единый орган управления, а в децентрализованных системах управление происходит только за счет локальных взаимодействий [9-11, 27-29]. В обоих этих случаях интеллектуальные процессы рассматриваются в контексте коллективного поведения, а центральным объектом исследования становится группа или сообщество интеллектуальных систем.

Все вышеуказанное обосновывает целесообразность развития нового, системно-организационного подхода в ИИ, основы которого излагаются в настоящей работе (см. также [30]).

ОСНОВЫ СИСТЕМНО-ОРГАНИЗАЦИОННОГО ПОДХОДА В ИИ

Системно-организационный подход к исследованию интеллекта включает в себя следующие главные принципы:

- (1) **Принцип исследования интеллекта в иерархии взаимодействующих систем**, что означает необходимость рассмотрения **микро-, мезо- и макроуровней интеллектуальной организации** и целесообразность изучения **метаинтеллектуальных процедур**. Эти уровни соответствуют триаде "субъект - группа - организация". Метаинтеллектуальные процедуры и системы определяют, например, нормы взаимоотношений агентов в многоагентных системах (МАС). Их введение в ИИ представляется столь же естественным, как и определение понятий метазнаний [31, 32], метамodelей [33] и метаэкспертных систем [34]. Основным классификационным атрибутом метаинтеллектуальной системы является способность предсказать, организовать и управлять коллективным поведением и групповой динамикой агентов в МАС, что означает поддержку рекурсивных и рефлексивных процессов.
- (2) **Принцип учета коллективной природы интеллекта** (исследование генезиса и развития интеллектуального поведения в процессах коммуникации и совместной деятельности), что предполагает обращение к семиотическим аспектам интеллекта.
- (3) **Принцип определения рекурсивных связей между интеллектом и деятельностью**, согласно которому интеллект агента выступает как подсистема управления деятельностью, позволяющая ему организовать свои действия или действия другого агента. В то же время, интеллект имеет коммуникативную природу и формируется в процессах взаимодействия (коммуникации) агента с другими агентами, а потребность в коммуникации связана с реализацией целенаправленной деятельности. Подобная рекурсия предполагает, в частности, **анализ интеллекта как регулятора деятельности**, что означает "непосредственное вовлечение интеллекта в действия" (благодаря построению общей модели деятельности, определения актуальных функций интеллекта и условий для их выполнения, и пр.).
- (4) **Принцип приоритета координационных (горизонтальных) связей** над вертикальными в интеллектуальных системах и организациях. Иными словами, предполагается невозможность решения сложных задач отдельными агентами, опирающимися на локальные модели (и в то же время, невозможность построения ими полных или глобальных моделей внешнего мира). Следствием этого является преобладание кооперативных и коалиционных стратегий над чисто конкурентными стратегиями в интеллектуальном поведении.
- (5) **Принцип дополненности различных моделей интеллекта** (аналогичный принципу Н.Бора), согласно которому невозможно отразить в одной модели многомерный характер понятия интеллект; для этого требуется построение семейства взаимодополняющих моделей.
- (6) **Принцип выделения системных единиц интеллекта**. В основу системного синтеза распределенного (децентрализованного) интеллекта целесообразно положить формирование функционально-структурной единицы (ФСЕ) как "универсального строительного блока" или "клетки" МАС. Системные единицы следует отличать от элементов: структурный элемент - это простейшая, неделимая часть системы, которая обычно не сохраняет свойства системы как целого, тогда

как важнейшим требованием к функционально-структурной единице является сохранение важнейших свойств организации всей системы.

Основные требования к системной единице интеллекта таковы: а) **единица** должна быть **целостным функционально-структурным образованием**, связанной организационно-управленческой структурой; б) единица должна быть **способной к развитию** (и саморазвитию), т.е. должна обладать порождающими свойствами и возможностями для ее трансформации; в) единица должна быть **неоднородной**, т.е. отображать в некотором смысле противоположные функции, структуры и свойства интеллекта;

В соответствии с принципом исследования интеллекта в иерархии систем, **базовыми единицами описания на микроуровне** являются **объекты, акторы и агенты**. Классические экспертные системы, базы знаний и подсистемы рассуждения также относятся к этому уровню представления характеристик индивидуального интеллекта. В целом традиционная проблематика ИИ будет обозначаться термином **"искусственный интеллект-1"**.

На мезоуровне мы исходим из **моделей взаимодействия** между отдельными единицами и феноменов коллективного интеллекта, поэтому здесь основными объектами исследования оказываются **группы единиц** - различные интегрированные и многоагентные системы (**"искусственный интеллект-2"**). Так например МАС зародились на пересечении теории систем и распределенного искусственного интеллекта. С одной стороны, речь идет об открытых, активных, развивающихся системах, в которых главное внимание уделяется процессам взаимодействия агентов как причинам возникновения систем с новыми качествами. С другой стороны, достаточно часто МАС строятся по принципам распределенного интеллекта как объединение отдельных интеллектуальных систем, обладающих своими базами знаний и средствами рассуждений.

Наконец, **на макроуровне**, центральное место занимает изучение **сообществ (популяций) многоагентных систем** и моделирование **социальных законов**, определяющих существование и эволюцию таких сообществ, понимаемых как **"искусственные"** (виртуальные) или, в ряде случаев, как **интеллектуальные организации**. Под **виртуальной организацией** понимается искусственно образованная в компьютере, не существующая в реальном физическом пространстве метасистема, включающая множество МАС. Виртуальная организация представляет собой сетевую, компьютерно интегрированную организационную структуру, состоящую из как правило неоднородных компонентов, расположенных в различных местах. В свою очередь, **интеллектуальная организация** определяется как метаинтеллектуальная МАС, возникающая в процессах коммуникации, происходящих как внутри организации, так и между отдельными организациями, и обеспечивающих **организационный симбиоз**, которая функционирует в соответствии с принципами максимальной адаптации к динамической среде, самореорганизации в интересах самосохранения и саморегуляции деятельности благодаря сохранению и накоплению знаний и опыта, прогнозированию потребностей и стратегическому планированию, разделенной мотивации и коллективному обучению. В частности, интеллектуальные организации характеризуются процессами интенсивной переработки знаний, связанными с преобразованием интеллектуальных ресурсов в преимущества над конкурентами. Такой организационный уровень исследований в ИИ (**"искусственный интеллект-3"**) значительно перекрывается с проблематикой искусственной жизни в широком смысле этого термина [5, 35-37].

Таким образом, в русле системно-организационного подхода к интеллекту радикально изменяется точка зрения на предмет и объект ИИ. В рамках классической парадигмы объектом ИИ как научно-технической дисциплины являются сложные человеко-компьютерные системы, обеспечивающие инженериию (приобретение, представление, порождение, пополнение, поддержку и передачу) знаний, а его предметом выступает разработка методологии, теории, методов, моделей, программно-аппаратных средств и прикладных компьютерных систем, позволяющих решать интеллектуальные, т.е. основанные на знаниях (правилах) задачи или выполнять интеллектуальные функции, ранее считавшиеся прерогативой человека-эксперта.

Согласно новой парадигме, **основным объектом исследований в ИИ** является **группа** (или **сообщество**), вообще говоря, неоднородных, взаимодействующих **агентов**, а основное содержание разработок связано либо с **синтезом индивидуальных свойств и поведения агентов**, исходя из заданной **групповой динамики**, определяемой отношениями кооперации и конкуренции, конфликта и сотрудничества, субординации и координации и пр. (**нисходящее проектирование**), либо с **построением организационных структур из агентов** на базе **анализа основных функций организации**, определения состава агентов и их ролей (**восходящее проектирование**). Здесь предметом исследования являются: а) конечное множество агентов; б) ситуации взаимодействия агентов, определяющие возникновение и структуру МАС (тип сообщества агентов) в зависимости от таких параметров, как совместимость целей, наличие взаимных обязательств и ответственности, ограниченность ресурсов, недостаток индивидуального опыта, и пр.); в) множество факторов, обеспечивающих установление локальных пространственных и временных отношений между агентами (непосредственное общение или телекоммуникации, совпадение целей или столкновение интересов, и пр.); г) множество ресурсов взаимодействия, обеспечивающих поддержание некоторых отношений между агентами и их трансформацию. д) множество действий (совместных действий) агентов и пр.

Итак, принимаемая здесь базовая единица исследования - агент в организации - предполагает расширение и смещение психологического базиса ИИ от когнитивной психологии в психологию деятельности (см. работы А.Н.Леонтьева [38], Б.Ф.Ломова [39], А.М.Волкова [40] и др.) и социальную психологию [41-42] и, вообще, в социокультурную область (см. культурно-историческую теорию развития психики Л.С.Выготского [43]). Сегодня акцент ставится на моделировании не только и не столько когнитивных, сколько регулятивных и коммуникативных процессов, и не столько структуры, сколько эволюции интеллекта (см. теорию интеллекта Ж.Пиаже [44]). Ниже будет проведен анализ известных определений понятия "агент", рассматриваемого как средство анализа сложных систем, и построена общая архитектура деятельности агента.

АНАЛИЗ ОПРЕДЕЛЕНИЙ ПОНЯТИЯ "АГЕНТ"

Если рассмотреть некую абстрактную непрерывную полярную шкалу типа "пассивный - активный" и расположить на ней в виде интервалов базовые единицы ИИ, такие как **фрейм** (объект) и **агент**, то интуитивно ясно, что агент должен быть расположен правее, ближе к активному полюсу и может в частном случае интерпретироваться как "активный объект", а агентно-ориентированное программирование [28] представляет собой естественное развитие и специальный случай объектно-ориентированного программирования.

Различные определения понятия “агент” подразделяются на “слабые” и “сильные” [45, 46], программистские и антропоморфные. Вначале остановимся на некоторых наиболее слабых определениях, а затем постепенно перейдем к более сильным. Так в [47] утверждается, что агент есть любая сущность, воспринимающая среду посредством сенсоров и действующая на нее посредством эффекторов. Таким образом, вычлняются три исходных, агентообразующих фактора - среда, восприятие, действие. Это достаточно тривиальное определение не позволяет проводить различий между агентом и произвольной программой, и здесь необходимо ввести некоторые специальные ограничения на среду, восприятие и действие. В близком по смыслу определении Б.Хэйес-Рот [48] подчеркивается, что интеллектуальные агенты непрерывно выполняют три следующие функции: 1) восприятие динамики среды; 2) действия, изменяющие среду; 3) рассуждения в целях интерпретации наблюдаемых явлений, решения задач, вывода заключений и определения действий. Кроме того, она специально оговаривает динамичность среды и подчиненность рассуждения действию: агенты рассуждают для того чтобы выбрать действия.

По мнению П.Маэс [49], автономные агенты - это компьютерные системы, функционирующие в сложной, динамической среде, способные ощущать и автономно действовать на эту среду и, таким образом, выполнять множество задач, для которых они предназначены. Иными словами, здесь предложены ограничения на среду агентов - "сложная, динамическая". Классификация различных сред агентов, служащая основой их типологии, изложена в [50].

Слабые определения смыкаются с чисто программистскими взглядами на природу агента [51-53]. Так в [51]. агенты понимаются как вычислительные единицы, поддерживающие локальные состояния и параллельные вычисления, а также способные в процессах коммуникации достигать состояния других агентов, и автоматически выполнять действия при удовлетворении некоторых условий среды. В свою очередь стратегия интеллектуальных агентов, изложенная в IBM White Paper [52], исходит из идеи “персонального ассистента”: интеллектуальный агент действует в интересах других лиц, которые передают ему определенные властные полномочия. Соответственно, интеллектуальные агенты понимаются как программные единицы, выполняющие некое множество операций для пользователя или другой программной системы, обладая определенной степенью автономности или независимости, и используя при этом знания или представления о целях или желаниях пользователя.

Большинство исследователей согласны с тем, что основными характеристиками агентов являются целенаправленность и автономность [50, 54, 55]. Так в [54] отмечено, что термин агент включает две ключевые способности: а) автономное выполнение некоторых действий; б) проведение целенаправленных проблемно-ориентированных рассуждений. Часто в качестве главных признаков агента берутся интеллектуальность и автономность, где интеллектуальность связывается с восприятием и рассуждениями, а автономность - с принятием решений и действием на среду [56] .

Более сильные определения отражают регулятивные, коммуникативные и поведенческие аспекты агента. Так в [57] указано, что программные агенты - это программы, которые имеют взаимные обязательства, определяемые в процессе диалога, ведут переговоры и координируют передачу информации. Здесь очевидно, что переговорный процесс предполагает наличие восприятия и действия, а диалог требует коммуникации. В [58] агентом называется единица модели, способная действовать на

себя и окружающую среду и общаться с другими агентами, причем ее поведение определяется восприятием и частичным представлением внешней среды, наличием знаний и опыта. Ее основная особенность определяется термином "живучий" (persistent) что отличает агентов от обычных подпрограмм: у агентов имеются свои собственные представления о том, как выполнять задачи, или свои программы действий.

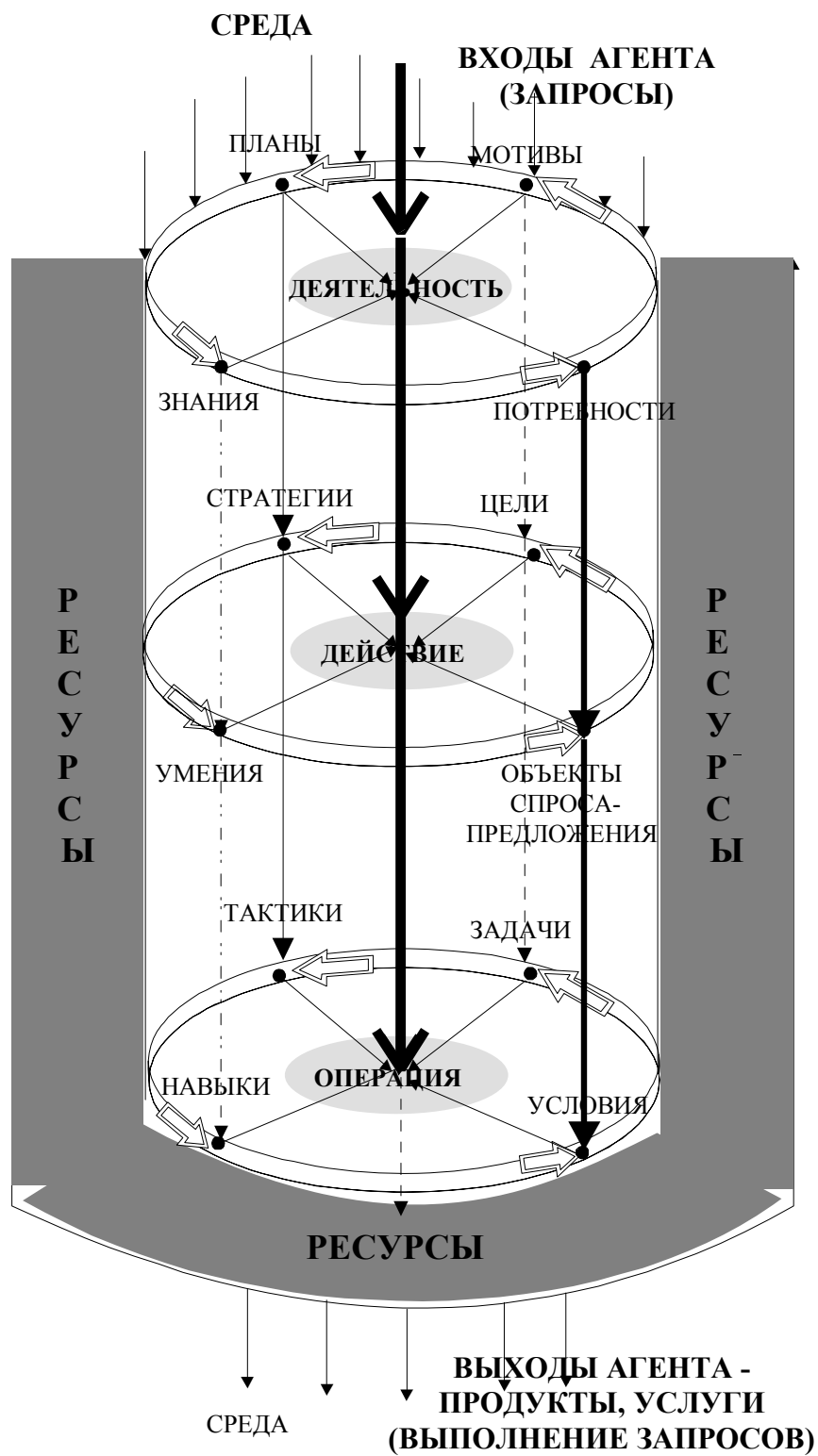
По-видимому первое в исследованиях по ИИ определение агента восходит к работам Д.А.Поспелова и В.Н.Пушкина [59] по теории гиромата. В гиромате - элементарной модели целесообразного поведения, способной адаптироваться к условиям решаемой задачи - уже присутствовали следующие модули: блок мотивации; блок селекции (рецепторы); блок построения внутренней модели внешней среды; блок выдвижения гипотез; блок модельного опыта; блок выработки решений; блок активного опыта; блок времени.

В одном из наиболее популярных определений понятия "агент" у М.Вулдриджа и Н.Дженнингса [45] полагается, что агент - это программно или аппаратно реализованная система, обладающая следующими свойствами:

- автономность - способность функционировать без прямого вмешательства людей или компьютерных средств и при этом осуществлять самоконтроль над своими действиями и внутренними состояниями;
- общественное поведение (social ability), т.е. способность взаимодействия с другими агентами (а возможно, людьми), обмениваясь сообщениями с помощью языков коммуникации;
- реактивность - способность воспринимать состояние среды (физического мира, пользователя - через пользовательский интерфейс, совокупности других агентов, сети Internet, или сразу всех этих компонентов внешней среды);
- про-активность (pro-activity) - способность агентов не просто реагировать на стимулы, поступающие из среды, но и осуществлять целенаправленное поведение, проявляя инициативу.

В ряде сильных (антропоморфных) определений агенты наделяются еще рядом дополнительных, ментальных свойств. Как правило, особое внимание обращается на главное отличие агентов от объектов, связанное с наличием у агентов внутренних механизмов мотивации и целеполагания. Так по И.Шоэму [28], агент есть целостная единица, описание состояния которой включает такие психические (метаинтеллектуальные) компоненты как убеждения, способности, обязательства и решения. Иными словами, каждый агент имеет встроенные механизмы мотивации (формирования предпочтений).

На наш взгляд, для того, чтобы соотнести между собой упомянутые точки зрения на сущность агентов и построить их классификацию, следует рассмотреть общую схему деятельности агентов. Вариант такой схемы представлен на рис.2. Ядром этой общей модели является трехуровневая схема развертывания деятельности вида "деятельность-действие - операция". Здесь деятельность понимается как активная (асимметричная)



- ▶ ПОДСИСТЕМА ПРЕДСТАВЛЕНИЯ
- - - -▶ ИНТЕНЦИОНАЛЬНАЯ ПОДСИСТЕМА
- ▶ ПОДСИСТЕМА ПОВЕДЕНИЯ
- - - -▶ ПОДСИСТЕМА ОПЫТА

Рис.2. Общая схема деятельности агента

форма отношения агента к внешнему миру. В общем случае можно указать, что:
ДЕЯТЕЛЬНОСТЬ = F (СРЕДА, ПОТРЕБНОСТИ, МОТИВАЦИЯ, ПЛАНИРОВАНИЕ, ЗНАНИЯ);
ДЕЙСТВИЕ = G (ДЕЯТЕЛЬНОСТЬ, ОБЪЕКТЫ, ЦЕЛИ, СТРАТЕГИИ, УМЕНИЯ);
ОПЕРАЦИЯ = H (ДЕЙСТВИЕ, УСЛОВИЯ, ЗАДАЧИ, ТАКТИКИ, НАВЫКИ),
где F, G и H - отношения полиморфизма ("один-ко-многим").

В отличие от различных моделей деятельности типа "черного ящика", эта модель может рассматриваться как "прозрачный ящик", поскольку в ней раскрываются внутренние механизмы деятельности путем анализа подсистем представления (отражения внешней среды), опыта (самоотражения), поведения (внешней регуляции) и интенциональной подсистемы (саморегуляции). Генезис и эволюция деятельности изучаются на основе анализа как вертикальных отношений "мотивы - цели - задачи", "планы - стратегии - тактики", так и горизонтальных отношений "потребности - мотивы - планы - знания", "объекты спроса - цели - стратегии - умения", "условия - задачи - тактики - навыки".

Соответственно, понятие агента как деятеля можно определить в зависимости от выбранного (или технически достижимого) уровня антропоморфизма неким подмножеством из указанного набора характеристик деятельности. В общем случае, основываясь на схеме деятельности, изображенной на рис.2 и следуя [60], дадим следующее общее определение интеллектуального агента.

Определение 1. Под интеллектуальными агентами в информатике и ИИ понимаются любые физические или виртуальные единицы, способные:

- а) действовать на объекты в некоторой среде, на других агентов, а также на самих себя (**действие**);
- б) исходить из некоторых потребностей и образовывать свои собственные цели (**потребностно-целевая основа - причина действия**);
- в) общаться с другими агентами (**общение**);
- г) накапливать собственные ресурсы, обеспечивающие их автономию (**автономность**);
- д) воспринимать некую область среды (**восприятие с ограниченным разрешением**);
- е) строить частичное представление этой среды на основе ее восприятия, т.е. исходя из имеющихся перцептивных навыков и умений (**локальное представление среды**);
- ж) нести обязанности и предоставлять услуги (наличие **обязательств**);
- з) обучаться, адаптироваться к среде и эволюционировать (**эволюционный потенциал**);
- и) самоорганизоваться и самовоспроизводиться (**самосохранение**).

Выделение различных ведущих уровней деятельности агента может служить основанием для их классификации (см. [50, 60, 61]).

Работа посвящена вопросам разработки и применения методологии системного подхода в ИИ. Системное мышление предполагает интеграцию существующих направлений ИИ и порождение новых, неклассических парадигм. Предлагаемый системно-организационный подход к моделированию интеллектуальных процессов и систем расширяет традиционную сферу исследований в ИИ и соответствует новой агентно-ориентированной парадигме, согласно которой изучение деятельности и взаимодействия агентов в МАС и интеллектуальных организациях становится главным содержанием ИИ в канун XXI-го века. В этом подходе основной акцент ставится на построении архитектуры агента с учетом рекурсивных связей интеллекта и деятельности, включая коммуникативные и регулятивные процессы, а также на исследовании генезиса и эволюции МАС и интеллектуальных организаций.

Дальнейшие ведущиеся работы связаны с построением базовых архитектур для интеллектуальных организаций (см. [30]).

ЛИТЕРАТУРА

1. Поспелов Д.А. История искусственного интеллекта до середины 80-х годов// Новости искусственного интеллекта. - 1994. - № 4. - С.70-90.
2. Winograd T., Flores F.. Understanding computers and cognition: a new foundation for design. - Norwood: Ablex, 1986.
3. Kirsh D.. Foundations of AI: the big issues// Artificial Intelligence. - 1991. - Vol.47, № 1-3. - P.3-30.
4. Непейвода Н.Н. Логический подход как альтернатива системному в математическом описании систем// Экспертные системы: состояние и перспективы. - М.: Наука, 1989. - С.20-30.
5. Тарасов В.Б. От искусственного интеллекта к искусственной жизни: новые направления в науках об искусственном// Новости искусственного интеллекта. - 1995. - № 4. - С.93-117.
6. Pollack J.B. Connectionism: past, present and future// Artificial Intelligence Review. - 1989. Vol.3, № 1. - С.3-22.
7. Brooks R. Intelligence without representation// Artificial Intelligence. - 1991. - Vol.47, №1-3. - P.139-159.
8. Minsky M. The society of mind. - N.Y.: Simon and Shuster, 1986.
9. Readings in Distributed Artificial Intelligence/ Ed.by A.H.Bond and L.Gasser. - N.Y.: Morgan Kaufmann, 1988.
10. Durfee E. Coordination of Distributed Problem Solvers. - Amsterdam: Kluwer Academic Press, 1988.
11. Hewitt C. Open information systems semantics for distributed artificial intelligence// Artificial Intelligence. 1991. - Vol.47, №1-3. - P.79-106.
12. Шенк Р. Обработка концептуальной информации: Пер. с англ. - М.: Энергия, 1980.
13. Пригожин И. От существующего к возникающему: Пер. с англ. - М.: Наука, 1985.
14. Bobrow D. Dimensions of interaction: AAAI-90 Presidential Address// AI Magazine. - 1991. -Vol.12, № 3. - P.64-80.
15. Кузнецов О.П. Неклассические парадигмы в искусственном интеллекте//Теория и системы управления - 1995. - № 5. - С.3-23.
16. Kitzmiller C.T., Kowalik J.S. Symbolic and numerical computing in knowledge-based systems// Coupling Symbolic and Numerical Computing in Expert Systems/ Ed. by J.S.Kowalik. - Amsterdam: Elsevier Science Publishers, 1986. - P.3-17.
17. Поспелов Г.С. Искусственный интеллект - основа новой информационной технологии. - М.: Наука, 1988.
18. Знания - диалог - решение. Ч.1.Теоретические основы и методы. -Киев: НПО "Системотехника", 1990
19. Тарасов В.Б., Желтов С.Ю., Степанов А.А. Нечеткие модели в обработке изображений: обзор зарубежных достижений// Новости искусственного интеллекта. - 1993. - № 3. - С.40-64.
20. Fanabashi M., Maeda A., Morooka Y., Mori K. Fuzzy and neural hybrid expert systems: synergetic AI// IEEE Expert. 1995 August. - P.32-40.
21. Keller J.M., Yager R.R., Tahani H. Neural network implementation of fuzzy logic// Fuzzy Sets and Systems. - 1992. - Vol.45, № 1. - P.1-12.
22. Kosko B. Neural Networks and Fuzzy Systems. - Englewood Cliffs: Prentice Hall, 1992.
23. Захаров В.Н., Ульянов С.В. Нечеткие модели интеллектуальных промышленных регуляторов и систем управления. IV.Имитационное моделирование// Известия Академии наук: Техническая кибернетика. - 1994.- № 5. - С.168-210.
24. Zadeh L. Fuzzy logic, neural network and soft computing// Communications of the ACM. - 1994.- Vol.37, № 3. - P.77-84.
25. Tarassov V., Perfilyev S., Deneux D. Structural-technological solutions in concurrent design// Proc. of International Conference "Concurrent Engineering Europe'97: Building Tomorrow's Virtual Enterprise" (Erlangen, Germany, April 16-18 1997)/ Ed. by K.Reger. -Budapest: Simulation Councils Inc.,1997. - P.10-18.
26. Ahson S. Petri net models of fuzzy neural networks// IEEE Transactions on Systems, Man and Cybernetics. - 1995. - Vol.25, № 6. - P.926-933.
27. Decentralized Artificial Intelligence// Ed. by Y.Demazeau, J.-P. Muller. - Amsterdam: Elsevier North-Holland, 1990.
28. Shoham Y. Agent oriented programming// Artificial Intelligence. - 1993. - Vol.60, № 1. - P.51-92.
29. Варшавский В.А., Поспелов Д.А. Оркестр играет без дирижера. Размышления об эволюции некоторых технических систем и управлении ими. - М.: Наука, 1984.
30. Tarassov V.B. Artificial meta-intelligence: a key to enterprise reengineering// Proc.of the Second Joint Conference on Knowledge-Based Software Engineering (JCKBSE'96, Sozopol, Bulgaria, September 21-22, 1996). - Sofia: BAIA, 1996. - P.15-24.

31. Lenat D. The measure of heuristics III// *Artificial Intelligence*. - 1983. - Vol.19, N° 3. - P.131-142.
32. Pitra J. *Metaconnaissances, futur de l'intelligence artificielle*. - Paris: Hermes, 1990.
33. Kiriya T., Yamamoto F., Tomiyama T. et al. *Metamodel: an intelligent modelling framework for intelligent CAD*// *Artificial Intelligence in Design*/ Ed. by G. S. Gero. - Southampton: Computational Mechanics, 1989. - P.429-449.
34. Стефанюк В.Л. Некоторые аспекты теории экспертных систем// *Известия АН СССР: Техническая кибернетика*. - 1987. - N° 2. - С.85-91.
35. *Artificial Life*/ Ed.by C.G.Langton. - Redwood-City: Addison Wesley, 1989.- 655 p.
36. *Artificial Life V*/ Ed.by C.G.Langton and T.Shimohara. - Cambridge MA: MIT Press, 1996.
37. Heudin J.-C. *La vie artificielle*. - Paris: Hermes, 1994.
38. Леонтьев А.Н. *Деятельность, сознание, личность*. - М.: ИПЛ, 1975.
39. Ломов Б.Ф. *Методологические и теоретические проблемы психологии*. - М.: Наука, 1984.
40. Волков А.М., Микадзе Ю.В., Солнцева Г.В. *Деятельность: структура и регуляция*. - М.: Наука, 1987.
41. Петровский А.В. *Личность, деятельность, коллектив*. - М.: Наука, 1982.
42. *Промышленная социальная психология*// Под ред. Е.С.Кузьмина, А.Л. Свенцицкого. - Л.: ЛГУ, 1982.
43. Выготский Л.С. *Собрание сочинений: в 6-ти томах. Т.1. Вопросы теории и истории психологии*. - М.: Педагогика, 1982.
44. Пиаже Ж. *Избранные психологические труды: Пер. с фр. и англ.* - М.: МПА, 1994.
45. Wooldridge M., Jennings N. *Agent theories, architectures and languages: a survey*// *Intelligent Agents. Proc. of ECAI-94 Workshop on Agent Theories, Architectures and Languages (Amsterdam, The Netherlands, August 8-9, 1994)*/ Ed. by M.Wooldridge, N.Jennings. - Amsterdam: Springer Verlag, 1995. - P. 1-22).
46. Городецкий В.И., Грушинский М.С., Хабалов А.В. *Многоагентные системы (обзор)* // *Новости искусственного интеллекта*. - 1997. - N° 1.
47. Russell S.J., Norvig P. *Artificial Intelligence: a Modern Approach*. - Englewood Cliffs NJ: Prentice Hall, 1995.
48. Hayes-Roth B. *An architecture for adaptive intelligent systems*// *Artificial Intelligence*. - 1995. - Vol.72. - P.329-365.
49. Maes P. *Artificial life meets entertainment: life like autonomous agents*// *Communications of the ACM*. - 1995. - Vol.38, N° 11. - P.108-114.
50. Поспелов Д.А. *От коллектива автоматов к мультиагентным системам*// *Труды Международного семинара "Распределенный искусственный интеллект и многоагентные системы" (DAIMAS'97, Санкт-Петербург, Россия, 15-18 июня 1997)*. - С.319-325.
51. Городецкий В.И. *Многоагентные системы: современное состояние исследований и перспективы применения*// *Новости искусственного интеллекта*. - 1996. - N° 1. - С.44-59.
52. The IBM Agent (<http://activist.gpl.ibm.com/White Paper/ ptc2.htm>)
53. Petrie C. *What is an agent?*// *Intelligent Agents III. Proc. of ECAI-96 Workshop on Agent Theories, Architectures and Languages (ALAL, Budapest, The Hungary, August 12-13, 1996)*/ Ed. by J.-P.Muller, M.Wooldridge, N.Jennings. - Berlin: Springer Verlag, 1996. - P.41-43.
54. Virdhagriswaran S. *White Paper, Crystaliz Inc.* <http://www.crystaliz.com/logicware/mubot.html>.
55. Franklin S., Graesser A. *Is it an agent, or just a program?*// *Intelligent Agents III. Proc. of ECAI-96 Workshop on Agent Theories, Architectures and Languages (ALAL, Budapest, The Hungary, August 12-13, 1996)*/ Ed. by J.-P.Muller, M.Wooldridge, N.Jennings. - Berlin: Springer Verlag, 1996. - P.21-35.
56. Chaib-Draa B., Moulin M., Mandiau R., Millot P. *Trends in distributed artificial intelligence*// *Artificial Intelligence Review*. - 1992. - Vol.6.
57. Coen M., MIT AI Lab, <http://www/ai/mit/edu/people/sodabot/slideshow/total/P001.html>.
58. Smith D.C., Cypher A. and Spohrer J. *KidSim: Programming agents without a programming language*// *Communications of the ACM*. - 1994. - Vol.37, N° 7. - P.55-67..
59. Поспелов Д.А., Пушкин В.Н. *Мышление и автоматы*. - М.: Сов.Радио, 1972.
60. Ferber J. *Les systemes multi-agents. Vers une intelligence collective*. Paris: InterEditions, 1995.
61. Braspenning P. *Plant-like, animal-like and humanoid agents and corresponding multi-agent systems*// *Proc. of the International Workshop on Distributed Artificial Intelligence and Multi-Agent Systems (St.Petersburg, Russia, June 15-18, 1997)*. - P. 64-77.