

В. Б. Тарасов

МГТУ им. Н.Э. Баумана, Москва, Россия
e-mail: tarasov@rk9.bmstu.ru

ОТ СПЕЦИФИКАЦИИ КОГНИТОНОВ И ИНЖЕНЕРИИ ИНТЕНЦИЙ К ОБОБЩЕННОЙ АРХИТЕКТУРЕ ДЕЯТЕЛЬНОСТИ АГЕНТОВ

Введение

В заключительной части своей книги «Фантазия или наука» [1], опубликованной в 1982 г., Д. А. Поспелов выдвинул обширную программу исследований в области искусственного интеллекта (ИИ). Эти исследования должны быть нацелены на «изучение *психики человека* с целью ее *имитации* в технических системах, решающих определенный набор практических задач, традиционно считающихся интеллектуальными». Следует напомнить, что в психологии термин «психика» охватывает как *сознание*, так и сферу *бессознательного*.

Еще раньше, в 1967 г., Д. А. Поспелов по аналогии с бионикой ввел термин *психоника*, означающий науку, предмет которой — исследование психической деятельности человека и животных с целью использования результатов этих исследований при проектировании технических устройств [2]. В качестве трех главных направлений исследования в психонике указывались: 1) построение в машине модели внешней среды; 2) моделирование поведения животных и деятельности чело-

века, включая мотивационные и эмоциональные факторы; 3) создание искусственной личности. По сути, уже на отечественных семинарах по психонике конца 1960-х гг., проходивших в МЭИ, докладывались и обсуждались первые попытки моделирования психики в русле изучения и имитации известной психологической триады «интеллект — чувства — воля». Эти идеи и результаты намного опередили свое время.

Хотя слово *психоника* не прижилось в среде специалистов по кибернетике, информатике и ИИ, биологически и психологически инспирированные подходы, модели и архитектуры стали «визитной карточкой» исследований в области создания передовых интеллектуальных технологий, в особенности, агентно-ориентированных технологий.

Разработка спецификаций и построение архитектур автономных агентов сильно зависят от интерпретации агентов. К настоящему времени известны, по меньшей мере, четыре интерпретации искусственных агентов: искусственный организм, активный объект, помощник пользователя, *виртуальный деятель*. В то время как модели и архитектуры агентов, соответствующие первым трем интерпретациям, достаточно хорошо разработаны и нашли практическую реализацию, теоретические и практические вопросы *инженерии деятельности* (или *инженерии действий* — эти термины берутся по аналогии с инженерией знаний) все еще остаются малоизученными.

В нашей работе проанализированы существующие когнитивные архитектуры интеллектуальных систем и автономных агентов и констатировано отсутствие теоретически обоснованных обобщенных архитектур их деятельности. В этой связи изложены элементы *общей теории деятельности* агентов, раскрыты принципы и механизмы *самоорганизации* деятельности. Предварительно систематизированы и описаны базовые *интенциональные* характеристики агентов как пусковые механизмы их деятельности. Рассмотрены разнородные когнитивные единицы сознания и деятельности — *когнитоны*, позволяющие связать между собой подходы когнитивного моделирования с основами самоорганизации деятельности.

Главный результат работы — развитие и формализация обобщенной архитектуры деятельности СинАрД, исходные идеи, связи и структуры которой были предложены в нашей монографии [3].

1. Искусственное сознание и когнитивные архитектуры

Как известно, с самого начала работ по ИИ в нем возникли две программы исследований — прагматическая и имитационная. Сторонники первой исходили из того, что при разработке искусственных интеллектуальных систем необязательно опираться на известные принципы и механизмы работы естественного интеллекта (ЕИ), т. е. эти области можно рассматривать как совершенно разные и не связанные между собой: ЕИ \neq ИИ. Напротив, адепты второй программы настаивали на необходимости использования в исследованиях по ИИ всей доступной информации о человеческом мозге с целью имитации его структуры и функционирования, что можно схематически изобразить как ЕИ \Rightarrow ИИ.

На начальных этапах развития интеллектуальных систем основные результаты были достигнуты в рамках первой программы, но в XXI в. акцент исследований смещается в сторону второй, имитационной, программ.

Аналогично обстоят дела с проблемой *искусственного сознания* (Artificial Consciousness) [4]. Сегодня на первом этапе развития этого направления его исследователи абстрагируются от воззрений тех психологов, которые заявляют о недизъюнктивности психики, сознания и мышления [5], невозможности их построения из отдельных частей (как обычно строят технические системы). Соответственно, *искусственное (машинное) сознание* пока развивается как научная область, связанная с проблемами памяти, обучения, опыта, понимания, а также с моделированием поведения, управляемого сознанием [6] и разработкой биологически (психологически) инспирированных когнитивных архитектур (ВІСА) [7].

В соответствии с ГОСТ Р ИСО/МЭК 15288, 2008, *архитектурой системы* называется принципиальная организация системы, воплощенная в ее элементах, их взаимоотношениях друг с другом и со средой, а также принципы, направляющие ее проектирование и эволюцию. В свою очередь, программная архитектура — это организационная структура, в которой идентифицированы компоненты, их интерфейсы и концепция взаимодействия между ними.

Архитектура автономного агента тесно связана с методологией его построения. По организации взаимодействия структурных уровней различают горизонтальные и вертикальные архитектуры. В целом архитектуры программных агентов можно разделить на два больших класса [3]: 1) архитектуры, основанные на знаниях; 2) архитектуры, характеризующие реактивное поведение. Известны также гибридные архитектуры, сочетающие системы рассуждений на знаниях, с чисто поведенческими, реактивными механизмами, например InterRaP и Turing Machine.

Наряду с «чисто агентскими» архитектурами следует обратить внимание и на когнитивные архитектуры, первоначально не связанные с искусственными агентами, но легко расширяемые в этом русле.

Когнитивная архитектура есть вычислительная модель, в которой представлены как функциональные компоненты интеллектуальной системы (причем не только знания), так и взаимодействия между ними. В частности, она определяет базовую инфраструктуру интеллектуального агента, отображая такие его свойства, которые постоянны во времени и в различных областях применения. Наиболее известные примеры когнитивных архитектур: EPIC, Soar (рис. 1), Prodigy, ACT-R, ASMO, 4CAPS, Cougaar, ICARUS, Apex, FORR, Spaun, QuBIC, VisNet, и др.

Как правило, когнитивная архитектура включает кратковременную и долговременную память (рис. 1), где хранится информация о мнениях, знаниях и целях агента, а также модули обучения [8]. Одна из главных целей разработки когнитив-

ных архитектур заключается в обобщении результатов когнитивных наук и их реализации к компьютерной модели. Архитектурное проектирование интеллектуальных агентов и МАС [9; 10] предполагает итеративную разработку прототипа открытой когнитивной архитектуры с его последующими модификациями и расширениями. Например, в версии Soar 9 дополнительно присутствуют модули образных представлений и формирования эмоций. По сути, на базе референсного решения создается целое семейство все более общих и многоаспектных архитектур.

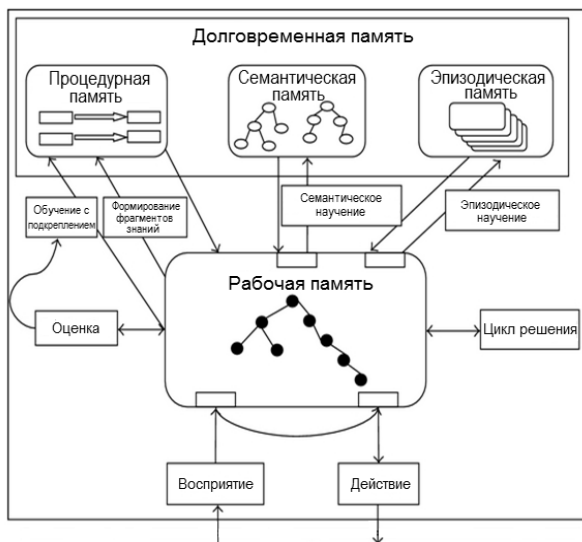


Рис. 1. Архитектура Soar

Обсуждение в апреле 2016 г. нашей лекции [11] на международной школе FIERCES on BICA показало, что в отличие от когнитивных архитектур до сих пор не существует общепринятой архитектуры самоорганизации деятельности агента, а также архитектур, реализующих принцип «единства сознания

и деятельности». В настоящей работе сделана попытка продвинуться в этом направлении путем соединения идей инженерии когнитонов, интенциональной самоорганизации и билатерального регулирования в архитектуре деятельности.

2. Инженерия когнитонов

В философии и ИИ сторонники *ментализма* считают, что любые попытки объяснить психическую деятельность, психические свойства, состояния и процессы на основе только физических, физиологических и нейрофизиологических подходов являются некорректным и недопустимым упрощением существа дел. В соответствии с познавательными установками ментализма, для моделирования и понимания действий и поступков агентов требуется обращение к ментальным (т. е. описывающим ненаблюдаемый внутренний мир) категориям.

Один из современных вариантов ментализма — *когнитивизм*, развивающий познавательные трактовки интеллекта. Истоки когнитивных течений восходят к пониманию роли знаний как одного из ведущих факторов, определяющих эффективность действий агента.

При создании когнитивных агентов и проектировании их взаимодействий центральное место занимает *теория когнитонов* (см. [3; 12; 13]). В ней выделяются и исследуются когнитивно-регулятивные, когнитивно-коммуникативные, когнитивно-ресурсные аспекты деятельности агентов.

Когнитонами называются разнородные когнитивные единицы, характеризующие разные стороны сознания — познание, деятельность, общение. Когнитон — это общий термин для синтеза деятельности агентов с позиций ее внутренних описаний [13]. Он был введен с целью единого представления семейств ментальных единиц, которые лежат в основе формирования динамических структур агентов, например, BDI-структуры или WILL-архитектуры. Классификация когнитонов дана на рисунке 2.

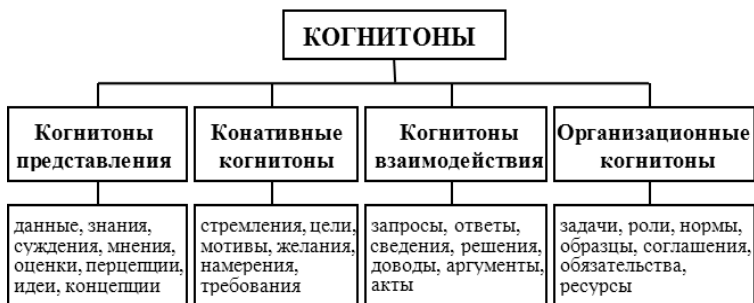


Рис. 2. Классификация когнитонов

Соответственно, возникает проблема *инженерии когнитонов* как ядра разработки когнитивных агентов, проектирования их деятельности, а также организации их взаимодействия и обеспечения взаимопонимания (подобно тому, как инженерия знаний определяет разработку классических интеллектуальных систем). Инженерия когнитонов расширяет обычные подходы инженерии знаний, видоизменяя и адаптируя их к проблемам формализации интенциональных и организационных свойств агентов (намерений, требований, обязательств, и т. д.).

Одной из важнейших характеристик когнитивных агентов является грануляция когнитонов, установление между ними отображений «один-ко-многим» (например, «когнитивный интервал — уровень регуляции»), что определяет возможность их анализа на различных уровнях абстракции-детализации [14]. При этом гранулярные структуры образуются как из однородных когнитонов, так и с использованием разнородных когнитонов. Примером гранулярной структуры, состоящей из когнитонов представления, служит структура «данные — информация — знания» (переход от данных к знаниям есть преобразование мелкозернистых гранул в крупнозернистые гранулы), а классической гранулярной конативной структурой выступает вектор «мотив-цель». В свою очередь, известными

вариантами гранулярных структур из неоднородных когнитонов являются BDI-модель [15; 16] (структура «мнение — желание — намерение») и структура «цель — ресурсы — восприятие — действие» ЦРВД [17]. Еще одним примером служит горизонтальная итеративная структура «потребность — мотив — план — знания» (см. ниже общую архитектуру деятельности агента).

Подробнее опишем конативную (интенциональную) основу агентов. Термин *конативный* (от лат. *conatus* — попытка, усилие, стремление) относится к мотивационным и волевым процессам регуляции деятельности агентов, обеспечивая связь знаний и эмоций с поведением. Слово *интенции* (интенциональные свойства агента) объединяет все виды побуждений агента к деятельности (действиям). Далее будем использовать эти два слова как синонимы, не проводя имеющихся легких различий между ними.

3. Интенции как причины деятельности агентов

Еще Ф. Brentano и Э. Гуссерль подчеркивали, что *интенциональность* есть фундаментальное отличие между физическими и психическими системами. Все психические явления (и только они) имеют интенциональную основу. У Д. Деннета одним из главных атрибутов сознания выступает интенциональная установка. Принятие интенциональной установки заключается в рассмотрении агента, действующего в соответствии с интенциональным образцом. По мнению Д. Деннета, созданные человеком артефакты также могут быть носителями интенциональности [18].

Соответственно все агенты отличаются от обычных объектов наличием побуждений. *Побуждение* — это любой внутренний источник деятельности агента, например внутренний стимул, «запускающий» деятельность. Побуждения индивидуальных агентов связаны с некоторой формой удовлетворения базовых потребностей, которые выступают как стремление к

определенному состоянию (гедонистический принцип). *Стремление* — это первичное побуждение, эмоциональное переживание потребности, относящейся к некоторому объекту. Бессознательное стремление выражается в форме *влечения*, а осознанная потребность выступает в форме *желания*, когда фиксируются не только объекты, но и пути достижения потребности.

У простейших реактивных агентов интенциональная основа сводится к наличию предпочтительных состояний, к которым они стремятся (например, трофические основы интенций). Здесь несомненный интерес представляет описание поведения на базе экстремальных принципов. В свою очередь, интеллектуальные агенты способны формировать не только желания, но и намерения, для чего служат механизмы памяти, обработки информации и принятия решений.

Намерение означает *волю действовать*. Оно предполагает стремление агента совершить действие в соответствии с целью и намеченным планом, когда во внимание принимаются наличные ресурсы, а также внешние требования и ограничения. Так, П. Коэн и Г. Левек назвали свою основополагающую статью «Намерение есть выбор + обязательство» [19]. Агент, в силу ограниченности имеющихся у него ресурсов, должен выбирать какое-либо желание, связанное с достижением цели, что означает принятие обязательства в той или иной форме. Намерение уже подразумевает выработку частичного плана действий, которому следует агент [16].

Наличие у любого агента хотя бы в зачаточном виде *встроенных источников мотивации, механизмов образования предпочтений* или формирования *целей* обеспечивает ему принципиально новый уровень автономии по сравнению с объектом. В отличие от объекта программный агент может принять на себя определенные обязательства или, наоборот, отказаться от предлагаемого задания, мотивируя это отсутствием компетентности, занятостью другой задачей и т. п. В то же время программный агент способен выполнять такие дей-

ствия, как порождение, подавление и замена других программных агентов, активизация функций (как своих, так и других агентов), активизация сценариев деятельности, запоминание текущего состояния других агентов и пр.

Итак, отличительная особенность синтеза когнитивных агентов — это необходимость построения и реализации специальной подсистемы представления ментальных состояний, включающей хотя бы некоторое подмножество из множества *интенциональных характеристик*.

Известны различные классификации интенций. Так, по временному критерию различают интенции настоящего времени и интенции, устремленные в будущее. Последние связаны с планированием действий по достижению желаемого будущего состояния в некотором мире и могут порождать производные интенции. Наоборот, интенция настоящего времени обеспечивает выполнение конкретного текущего действия.

Согласно Г. Киссу, среди интенциональных характеристик выделяются: а) *когнитивные* (мнения, знания, свидетельства), *конативные* (намерения, обязательства, планы) и *аффективные* (цели, желания, предпочтения). Здесь когнитивные отношения связаны с эпистемическими структурами, конативные — с управлением и действием, а аффективные — с динамикой поведения агента [20]. И. Шоэм и Р. Казинс [21] подразделяют интенции на *информационные* (мнения, знания), *мотивационные* (цели, желания, намерения, планы), *социальные* (обязательства, разрешения).

По Р. Братману [16], *знания и мнения* агента в многоагентной системе понимаются как социальные установки, которые ориентированы на ценности и предписывают определенные варианты поведения. В то же время *желания, намерения* и пр. представляют собой факторы мотивации, направляющие поведение агента. В отличие от мнений, которые могут быть частично истинными или ложными, желания или намерения обычно рассматриваются в контексте их удовлетворения.

Пример BDI-архитектуры агента представлен на рисунке 3.

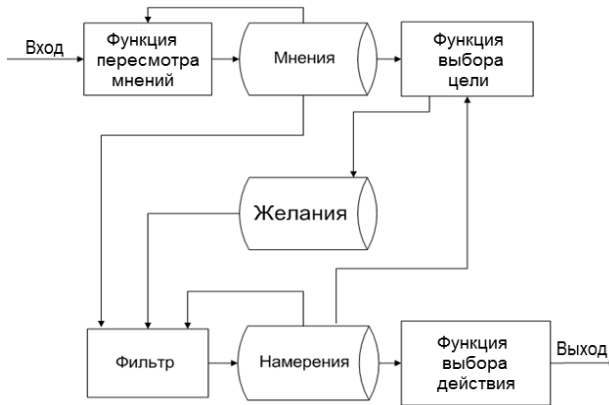


Рис. 3. Пример BDI-архитектуры когнитивного агента

4. Особенности самоорганизации и моделирования деятельности и действий агентов

До последнего времени (а именно до широкого распространения искусственных агентов и многоагентных систем в информатике и ИИ) современный деятельностный подход был вне поля зрения специалистов по искусственному интеллекту. В американском научном сообществе это отчасти объяснялось долгой популярностью воззрений бихевиоризма и необихевиоризма. В то же время одним из главных достижений отечественной психологии стало создание в последней четверти XX в. *общей теории деятельности* (см.[22; 23]), в которой понятия сознания и психики в целом используются как «средства объяснения».

У истоков этой теории стояли Л. С. Выготский, А. Н. Леонтьев [24] и С. Л. Рубинштейн [25]; свое дальнейшее развитие она получила в работах Б. Г. Ананьева [26], Б. Ф. Ломова [27], А. М. Волкова [28], Г. В. Суходольского [29], Г. П. Щедровицкого [22] и др.

Под деятельностью искусственного агента будем понимать асимметричное взаимодействие (отношение) между агентом и объектами среды, которое возникает и развивается благодаря наличию у агента *интенциональных свойств*. Деятельность есть сложная, открытая, неаддитивная *единица «жизни»* [24] агента, что служит обоснованием применения *синергетической методологии* ее синтеза.

Как отмечает Г. П. Щедровицкий [22], отличительная особенность деятельности как системы заключается в ее неоднородности и полиструктурности. В русле современного синергетического подхода следует также указать на неравновесность, необратимость и нелинейность процессов в системе деятельности, причем центральное место здесь занимают проблемы моделирования процессов ее *возникновения и развития*.

Синтез деятельности агентов необходимо начинать с определения и исследования основных функций, компонентов и структур деятельности (см. [24; 27]). Следуя Г. В. Суходольскому [29], будем рассматривать онтологические, морфологические, гносеологические аксиологические и праксеологические структуры деятельности.

Онтология деятельности фиксирует ее существование в виде асимметричного отношения «агент — объект» или «агент — агент — объект». Здесь классический формализм — тройка **SAO** (модель «субъект — действие — объект», лежащая в основе универсального семантического кода В. В. Мартынова) или пятерка **saorc** (*спецификация создаваемого продукта в поисковом проектировании*: агент **s** действует **a** на объект **o** для получения нужного результата **r** при ограничениях **c**).

В отличие от деятельности в поведении реализуются связи «агент — агент» или «агент — объект — агент». Единицей деятельности выступает *действие*, а единицей поведения — *поступок* (табл.).

О различиях между деятельностью и поведением

Система	Деятельность	Поведение
Единица	Действие	Поступок

Морфология деятельности агентов характеризует ее состав и структуру. Так, деятельность состоит из набора действий, которые разбиваются на отдельные операции (рис. 4).

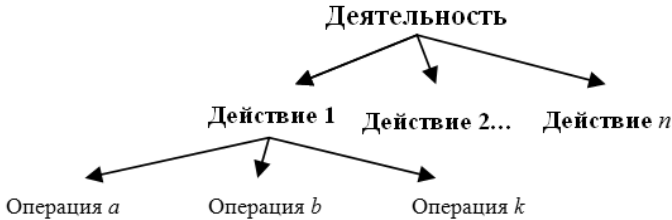


Рис. 4. Трехуровневая иерархическая структура деятельности

Действие — это единица деятельности, соотнесенная с некоторой целью. Согласно П. Жане, следует различать «первичное действие», или отношение между субъектом и объектом, и «вторичное действие», т. е. реакцию субъекта на свое собственное действие. В свою очередь, под операцией понимается элемент деятельности, который отвечает определенным условиям выполнения действия.

Аксиология деятельности агента объединяет ее потребности, результаты, ценности, оценки.

Наличие *потребности* — основа самоорганизации любой деятельности. Потребность как состояние нужды N в чем-либо есть функция рассогласования между наличной U_0 и желаемой ситуацией U_1 : $N = f(U_0, U_1)$.

Деятельность любого агента разворачивается как совокупность действий, направленных на удовлетворение некоторых потребностей. Обычно она предполагает внутреннее представление объекта (его интериоризацию), вынесение суждений и проведение рассуждений в интересах формирования действий, направленных на получение объективных результатов. Таким образом, деятельность выступает как процесс, связывающий агента и объект, в рамках которого осуществляются взаимные переходы между этими полюсами (переходы из внутреннего плана деятельности во внешний и наоборот).

Внешний план деятельности охватывает *потребности* вместе с *предметами* и *условиями* их удовлетворения, а внутренний план соответствует *интенциональной* (мотивационно-целевой) сфере (рис. 5).

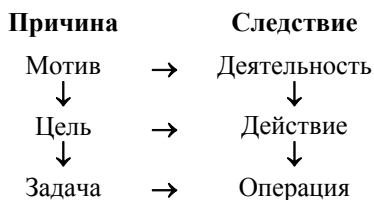


Рис. 5. Причинно-следственные взаимосвязи в структуре деятельности

Причиной деятельности выступает мотив как осознанная потребность, связанная с некоторым предметом. Мотив — это любое внутреннее представление некоторой объективной потребности. Всякая деятельность как процесс реализации мотива состоит из совокупности действий, направленных на достижение цели, которые распадаются на отдельные операции.

Внешняя наблюдаемая часть деятельности соответствует поведению. Структура организации поведения соответствует формированию планов, стратегий (тактик) и реакций.

Праксиология деятельности изучает функционирование, освоение и развитие деятельности в пространстве и во времени. С ней связана структура *опыта*, включающая знания, навыки и умения (как продукт координации знаний и навыков).

Наконец, *гносеология деятельности* подразумевает формирование когнитивной структуры, характеризующей познавательную активность агента, построение внутренней модели внешней среды, других агентов и самого себя.

5. Архитектура деятельности интенционального агента

В контексте развития теории агентов как искусственных деятелей центральной задачей становится разработка общей архитектуры деятельности интенциональных агентов. Здесь

обычная, принятая в информатике модель деятельности типа «черного ящика» (рис. 6) оказывается недостаточной. В интересах построения и обоснования архитектур агентов как деятелей необходимо провести анализ базовых принципов и механизмов самоорганизации деятельности.



Рис. 6. Представление деятельности как «черного ящика»

Деятельность агента есть сложная, отрывая, неоднородная, многофункциональная, полиструктурная, развивающаяся иерархическая система. В соответствии с подходом, описанным в [3; 28], в основу синергетической архитектуры деятельности СинАрД положены четыре взаимодействующие структуры — когнитивная, поведенческая, интенциональная, а также структура опыта.

Когнитивная структура характеризует познавательную активность агента, связанную с отражением объекта и условий деятельности, построением внутренней модели внешней среды.

Интенциональная структура охватывает основные внутренние механизмы деятельности агента, ее источники, причины, стимулы и ожидаемые результаты.

Структура организации поведения отвечает за регуляцию внешней части деятельности агента, прежде всего за выработку планов и стратегий поведения.

Структура опыта выражает динамику обучения агента, т. е. освоения и развития его деятельности.

Вариант общей архитектуры деятельности агента [3; 10; 11; 30] изображен на рисунке 7 в виде иерархической (многослой-

ной) семантической сети. Ядро этой архитектуры — трехуровневая схема развертывания деятельности: *деятельность — действие — операция* (по А. Н. Леонтьеву [24]).

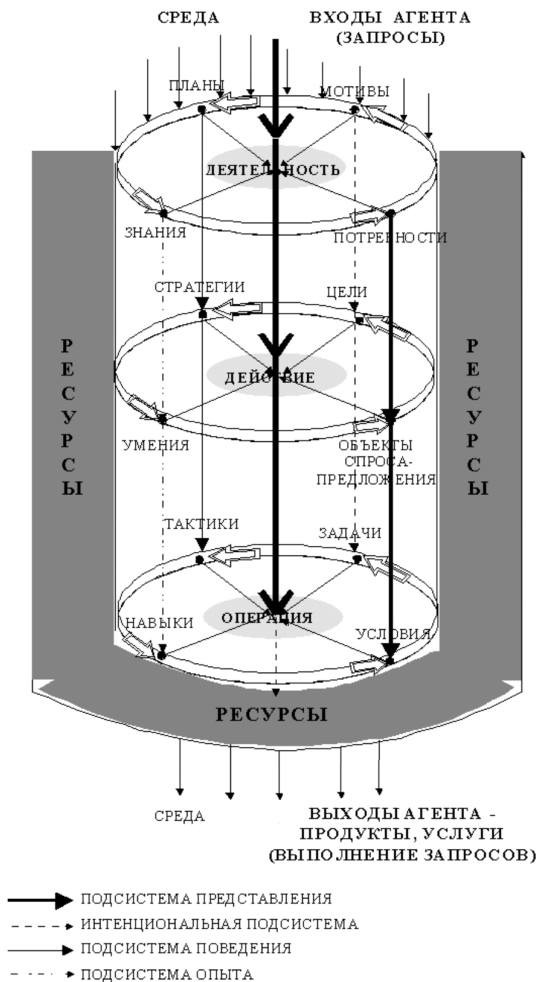


Рис. 7. Архитектура деятельности интенционального агента

В данной архитектуре на каждом уровне отражены как радиальные связи (от периферии к центру) между регулятивными (интенциональными, поведенческими) и морфологическими компонентами деятельности, так и кольцевые связи между базовыми функциями организации деятельности.

В общем случае можно указать, что:

ДЕЯТЕЛЬНОСТЬ = F (СРЕДА, ПОТРЕБНОСТИ, МОТИВЫ, ПЛАНЫ, ЗНАНИЯ);

ДЕЙСТВИЕ = G (ДЕЯТЕЛЬНОСТЬ, ОБЪЕКТЫ, ЦЕЛИ, СТРАТЕГИИ, УМЕНИЯ);

ОПЕРАЦИЯ = H (ДЕЙСТВИЕ, УСЛОВИЯ, ЗАДАЧИ, РЕАКЦИИ, НАВЫКИ),

где F , G и H — отношения полиморфизма («один-ко-многим»).

В отличие от представления деятельности как «черного ящика», эта трехуровневая и двухконтурная архитектура могут рассматриваться как «прозрачный ящик». В них раскрываются внутренние механизмы деятельности путем анализа *подсистем представления (моделирования внешней среды) и опыта, поведения (внешней регуляции) и интенциональной подсистемы (саморегуляции)*.

Генезис и эволюция деятельности здесь изучаются на основе как вертикальных отношений *мотивы — цели — задачи, планы — стратегии — реакции, знания — умения — навыки*, так и горизонтальных отношений *потребности — мотивы — планы — знания, объекты удовлетворения потребностей — цели — стратегии — умения, условия — задачи — реакции — навыки*, определяющих контуры самоорганизации (см. схему билатерального регулирования у Б. Г. Ананьева [26]).

Соответственно, понятие агента как деятеля можно определить в зависимости от выбранного (или технически достижимого) уровня антропоморфизма неким подмножеством из указанного набора характеристик деятельности. При разработке реактивных агентов со стимульно-реактивным поведением можно ограничиться моделированием характеристик нижнего уровня и их связей с функциями среднего уровня. Наоборот, при создании интеллектуальных агентов, которые должны

иметь достаточно развитые внутренние модели внешнего мира, средства проведения рассуждений, а также модули организации действий, необходимо подробно исследовать отношения между понятиями верхнего уровня, а также связи со средним уровнем.

Интеллектуальную среду (Ambient Intelligence) можно рассматривать как искусственного метаагента, который взаимодействует с физико-технической средой с целью обеспечения комфортных условий деятельности естественных агентов, а также в интересах проведения мониторинга технических объектов, объединенных средствами «интернета вещей», подготовки и поддержки принятия управленческих решений. Когнитивно-регулятивная координация метаагента при его взаимодействии с физико-технической средой описывается парой «когнитивный интервал — уровень регуляции».

Заключение

Перспективы архитектуры деятельности агента СинАрД связаны с ее валидацией и применением в различных областях. Так, в [31] модифицированная архитектура СинАрД используется для создания виртуальной бизнес-среды развития инноваций. При этом она рассмотрена вместе с внутренней подсистемой имитационного моделирования. Предложен вариант реализации имитационного аппарата агентов на основе системно-динамических моделей в многоагентной среде.

В этом контексте также следует развивать гибридные интеллектуальные системы с самоорганизацией [32], обеспечивающие гибкие переходы от кооперативных к конкурентным стратегиям деятельности и обратно.

Работа выполнена при поддержке РФФ (проект № 16-11-00018) и РФФИ (гранты № 14-07-00846 и № 15-07-05623).

Список литературы

1. *Поспелов Д. А.* Фантазия или наука: на пути к искусственному интеллекту. М. : Наука, 1982.

2. *Поспелов Д. А.* О задачах психоники // Вопросы бионики. 1967. С. 294—297.
3. *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М. : Эдиториал УРСС, 2002.
4. *Reggia J.* The Rise of Machine Consciousness: Studying Consciousness with Computational Models // *Neural Networks*. 2013. Vol. 44. P. 112—131.
5. *Брушлинский А.В.* Мышление и прогнозирование. М. : Мысль, 1979.
6. *Осинов Г.С., Панов А.И., Чудова Н.В.* Управление поведением как функция сознания // Теория и системы управления. 2014. №4. С. 49—62.
7. *Samsonovich A.V.* On a Roadmap for the BICA Challenge// *Biologically Inspired Cognitive Architectures*. 2011. Vol. 1. P. 100—107.
8. *Laird J.E., Congdon C.B.* The Soar User's Manual: Version 9.1. University of Michigan, 2009.
9. *Weyns D.* Architecture-Based Design of Multi-Agent Systems. Heidelberg : Springer-Verlag, 2010.
10. *Tarassov V.B.* Building Activity Architectures for Multiagent Systems — On the Way to Intelligent Organizations // Proceedings of the International Workshop «Distributed Artificial Intelligence and Multiagent Systems» (DAIMAS'97, St. Petersburg, Russia, June 15—18, 1997). SPb. : SPIIRAS Ed., 1997. P. 348—351.
11. *Tarassov V.B.* Ambient Intelligence: Concepts, Models, Architectures // Lecture at the First International Early Research Enhancement School on Biologically Inspired Cognitive Architectures (FIERCES on BICA, Moscow, April 21—24, 2016).
12. *Ferber J.* Les systèmes multi-agents. Vers une intelligence collective. Paris : InterEditions, 1995.
13. *Тарасов В.Б., Борисов А.В.* Логическое моделирование когнитивных и коммуникативных характеристик агентов: единый подход // Труды 10-й национальной конференции по ИИ (КИИ-2006, Обнинск, 25—28 сентября 2006 г.). М. : Физматлит, 2006. Т. 3. С. 916—928.
14. *Тарасов В.Б.* Грануляция информации — основа когнитивных процессов и предпосылка создания интеллектуальных систем новых поколений // Подходы к моделированию мышления / под ред. В.Г. Редько. М. : ЛЕНАНД, 2014. С. 219—261.

15. *Georgeff M.P., Rao A.S.* BDI Agents: From Theory to Practice// Proceedings of the First International Conference on Multi-Agent Systems (ICMAS'95, San Francisco, 1995)/ Ed. by V. Lesser. San Francisco : AAI Press/The MIT Press. P. 312—319.

16. *Bratman M.* Faces of Intention. Cambridge : Cambridge University Press, 1999.

17. *Dyundyukov V.S., Tarassov V.B.* Goal-Resource Networks and Their Application to Agents Communication and Co-Ordination in Virtual Enterprise // Proceedings of the 7th IFAC Conference on Manufacturing Modelling, Management and Control (MIM'2013, Saint-Petersburg, Russia, June 19—21, 2013). SPb. : IFAC Proceedings Volumes, 2013. P. 347—352.

18. *Деннет Д.С.* Виды психики: на пути к пониманию сознания / пер. с англ. М. : Идея-Пресс, 2004.

19. *Cohen P., Levesque H.* Intention Is Choice + Commitment // Journal of Artificial Intelligence. 1990. Vol. 42. P. 213—261.

20. *Kiss G.* Variable Coupling of Agents to Their Environment: Combining Situated and Symbolic Automata // Decentralized Artificial Intelligence 3 / ed. by E. Werner, Y. Demazeau. Amsterdam: Elsevier / North-Holland, 1992. P.231—248.

21. *Shoham Y., Cousins S.B.* Logics of Mental Attitudes in AI // Foundations of Knowledge Representation and Reasoning / ed. by G. Lakemeyer, B. Nebel. Heidelberg : Springer-Verlag, 1994. P. 296—309.

22. *Щедровицкий Г.П.* Исходные представления и категориальные средства теории деятельности // Избранные труды. М. : ШКП, 1995.

23. *Bedny G., Meister D.* The Russian Theory of Activity: Current Applications to Design and Learning. N. Y. : Psychology Press, 2014.

24. *Леонтьев А.Н.* Деятельность. Сознание. Личность. М. : Смысл ; Академия, 2005.

25. *Рубинштейн С.Л.* Избранные философско-психологические труды: основы онтологии, логики и психологии. М. : Наука, 1997.

26. *Ананьев Б.Г.* О проблемах современного человекознания. 2-е изд. СПб. : Питер, 2001.

27. *Ломов Б.Ф.* Психическая регуляция деятельности. Избранные труды. М. : Изд-во Института психологии РАН, 2006.

28. *Волков А.М., Микадзе Ю.В., Солнцева Г.В.* Деятельность: структура и регуляция. М. : Изд-во МГУ, 1987.

29. *Суходольский Г.В.* Введение в математико-психологическую теорию деятельности. СПб. : Изд-во СПбУ, 1998.

30. *Тарасов В. Б.* От инженерии знаний к самоорганизации деятельности: новая парадигма в искусственном интеллекте // Научная сессия МИФИ-2004 : сб. науч. тр. (Москва, 28 января 2004 г.). М. : МИФИ, 2004. Т. 3: Интеллектуальные системы и технологии. С. 22—23.

31. *Маслобоев А. В.* Гибридная архитектура интеллектуального агента с имитационным аппаратом // Вестник МГТУ. 2009. Т. 12, № 1. С. 113—124.

32. *Колесников А. В., Кириков И. А., Листопад С. В.* Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. М. : ИПИ РАН, 2014.