

К построению модульной модели распределенного интеллекта

^{1,2}Ю.Л. Словохотов <slov@phys.chem.msu.ru>

³И.С. Неретин <ivan@neretin.ru>

¹Московский государственный университет имени М.В. Ломоносова,
119991, Россия, Москва, Ленинские горы, д. 1

²Институт элементоорганических соединений РАН,
119991, Россия, Москва, ул. Вавилова, д. 28

³Rock Flow Dynamics, 117418, Россия, Москва, ул. Профсоюзная, д. 25А

Аннотация. Описание и моделирование динамики мультиагентных социальных систем методами, заимствованными из статистической физики «неживых» многочастичных систем, не отражает принципиальную особенность совокупности взаимодействующих автономных агентов: способность воспринимать, обрабатывать и использовать внешнюю информацию. Распределенный интеллект социальных систем следует непосредственно учитывать в их экспериментальных и теоретических исследованиях. В работе предложена «модульная» модель интеллектуальной деятельности, включающая производство новой информации и пригодная для описания как индивидуального, так и распределенного интеллекта, перечислены возможные области ее использования. «Количественную оценку» эффективности распределенного интеллекта иллюстрирует компьютерная модель искусственной социальной системы; обсуждаются полученные результаты.

Ключевые слова: мультиагентные социальные системы; распределенный интеллект; моделирование интеллектуальной деятельности

DOI: 10.15514/ISPRAS-2018-30(3)-23

Для цитирования: Словохотов Ю.Л., Неретин И.С. К построению модульной модели распределенного интеллекта. Труды ИСП РАН, том 30, вып. 3, 2018 г., стр. 341-362. DOI: 10.15514/ISPRAS-2018-30(3)-23

1. Введение

Изучению и прогнозированию динамики реально существующих мультиагентных социальных систем – таких как пешеходные потоки [1], объекты управления [2], участники политического противоборства [3], избирательные кампании [4] и многих других – посвящена обширная литература. Основным формальным инструментом анализа служит теория

игр, призванная типологизировать поведение агентов¹ и установить возможные равновесия в системе. Структуру совокупности взаимосвязанных агентов исследуют методами теории графов, представляя достаточно большие системы в виде сложных сетей [5]. Это позволяет выделить характерные структурные элементы (концентраторы, клики, сообщества, деревья), определить количественные характеристики сети (диаметр, распределение порядков вершин, степень кластеризации и др.), а также моделировать динамику различных процессов на сетевых субструктурах [6–8].

При большом количестве положительных результатов и новых эмпирических наблюдений, в теоретико-игровых и теоретико-графовых исследованиях мультиагентных систем до сих пор не выработано их общепринятой модели. Моделирование мультиагентных процессов методами теории игр сталкивается с необходимостью учитывать как многоуровневую рефлексивность агентов, обладающих интеллектом [9], так и «ограниченно рациональное» поведение людей в реальных условиях [10]. Это усложняет модели, делает их стохастическими и зависящими от эмпирических параметров, снижая предсказательную способность. Даже для систем с простыми стратегиями агентов, которые удается выразить квази-динамическими уравнениями, расчетный краткосрочный прогноз требует больших вычислительных мощностей (транспортный поток на шоссе [11]) либо оказывается невозможным за пределами нескольких идеализированных режимов (биржа [12]). С другой стороны, результаты моделирования социальных процессов на сетях во многих случаях показывают отсутствие сильного влияния структуры субструктуры (которое неявно постулируется в большинстве «сетевых» работ). Так, основные качественные результаты математической социологии, полученные во 2-й половине XX века на квадратных решетках, изображавших систему социальных связей [13, 14], в целом воспроизводятся и на сложных сетях; различия в основном состоят в замедлении процесса и неоднородных «промежуточных» распределениях состояний агентов [14, 15].

С формальной точки зрения, большинство современных мультиагентных моделей построены по аналогии с моделями многочастичных физических систем [14]. Подобно взаимодействиям в ансамбле «неживых» частиц, взаимодействия агентов основаны на балансе количественно или качественно оцениваемых факторов (выигрышей, намерений, интересов и т.д.), а динамика каждого агента направляется стремлением к максимуму энергоподобной функции полезности и «размывается» стохастическим шумом, нередко называемым температурой. Усложнения агентных моделей, по сравнению с их физическими прототипами, сводятся к использованию нечетко определенных характеристик (начиная с «полезности»), неоднозначных зависимостей

¹ В рамках этой работы понятия «индивид», «индивидуум» и «агент» применительно к единичному «актору» социальной системы используются как синонимы.

(рефлексии) и сложной дискретной конфигурации взаимодействий (сети). Предсказуемый результат состоит в сильном увеличении объема расчетов и неопределенности прогноза в таких системах, где динамика процесса не задается самой сетевой структурой (как в эпидемиологии, энергетических и транспортных сетях, некоторых других областях).

Имеющиеся результаты мультиагентного моделирования позволяют предположить, что многочастичные модели, заимствованные из статистической термодинамики и физики стохастических процессов, не вполне адекватны фундаментальным особенностям социальных систем и процессов в таких системах. (Утверждение о несводимости социума к физике является общим местом в работах гуманитарного направления). Интригующим дополнительным обстоятельством служат эмпирические аналогии коллективного поведения людей и динамики, наблюдаемой для систем индивидов со значительно меньшими интеллектуальными возможностями (муравейник, термитник [16]), а также для формаций программируемых автоматов [17]. Таким образом, наличие индивидуального сознания у людей не является принципиальным отличием социальной мультиагентной системы от физической системы «неживых» частиц (молекул и атомов при сколь угодно детальном описании их состояний и взаимодействий): сходную динамику могут проявлять социальные системы, сильно различающиеся по уровню интеллекта у их агентов.

Мы предполагаем, что главной особенностью, отличающей систему взаимодействующих агентов от ансамбля частиц физической среды, является распределенный интеллект: способность мультиагентной системы воспринимать внешнюю информацию, обрабатывать и использовать ее в своей коллективной динамике. Наличие распределенного интеллекта следует непосредственно учитывать как в эмпирическом описании социальных мультиагентных систем, так и в их моделировании. Далее будут рассмотрены некоторые возможные подходы к исследованиям в этом направлении.

2. Определения

Мультиагентная социальная система (МСС) – это динамическая совокупность автономных агентов, которые воспринимают информацию и взаимодействуют с внешней средой и другими агентами в ходе собственной деятельности (*динамики*). Взаимодействующие биологические существа одного вида составляют социальную систему в узком смысле слова. «Неживые» программируемые агенты (роботы, беспилотные аппараты и т.д.) при наличии взаимодействия между ними образуют искусственную МСС. Индивидуальные агенты составляют МСС 1-го уровня (сообщество живых организмов, автомобили на шоссе и т.д.). Взаимодействующие мультиагентные системы (экономические субъекты, научные школы, политические «акторы») сами могут выступать в роли агентов в социальной системе более высокого уровня.

Каждая МСС (1) обладает структурой, (2) осуществляет некоторую суммарную («системную») деятельность, (3) в ее ходе воспринимает, обрабатывает и использует информацию. Структуру МСС определяют воздействия окружающей среды на агентов («внешнее поле») и взаимодействия между агентами (внутренняя структура, в общем случае – взвешенный ориентированный граф, зависящий от внешнего поля и от времени). Динамика МСС определяется изменениями внешней среды, собственной динамикой агентов и эволюцией внутренней структуры во времени.

Системная динамика МСС складывается из индивидуальных динамик агентов, однако ее результаты не сводятся к сумме индивидуальных результатов агентов (перемещение стада животных из засушливой местности в благоприятную, биржевой крах, автоматизация производства и т.д.). Поскольку агенты-индивиды в МСС 1-го уровня воспринимают информацию и преследуют определенный набор целей, социальным системам как «акторам» можно приписать *системное целеполагание*. Объективной целью МСС на заданном интервале времени является оптимальный для системы результат ее динамики. Разнообразные МСС с вариативной динамикой агентов, не предполагающей сложной интеллектуальной деятельности, воспроизводимо и гибко преследуют системные цели (пчелиный рой, муравейник, «невидимая рука рынка»). Одной из объективных целей социальной системы в большинстве случаев является сохранение либо увеличение количества агентов и поддержка их функционирования: *самосохранение системы*.

Восприятие, обработка и использование информации социальной системой являются результатом восприятия информации ее агентами, взаимодействий между ними и целеполагания агентов. (Так, формация автоматов, движущихся в заданном направлении по алгоритму «следовать за соседом» и «избегать столкновений», огибает препятствия). Таким образом, социальные системы любого уровня не только преследуют определенные цели, но также, в процессе их достижения, анализируют информацию – как внешнюю (сила стимула или угрозы извне), так и внутреннюю (состояния агентов и их соседей в зоне восприятия). Иными словами, динамика МСС существенно включает *распределенный интеллект*: восприятие, анализ и использование информации социальной системой как единым «актором» и, возможно, как агентом в МСС более высокого уровня.

3. Моделирование интеллекта: состояние проблемы

Мы полагаем, что описание, типология и моделирование социальных систем не могут быть корректными без анализа распределенного интеллекта, которым обладают различные типы МСС. В этом случае особую важность приобретает формальный анализ *индивидуального интеллекта* – прежде всего

человеческого мышления – и степень переносимости его результатов на «коллективный разум».

Аналізу умственной деятельности людей и элементов мышления животных посвящена большая область *когнитивных наук* – «междисциплинарных исследований познания, понимаемого как совокупность процессов приобретения, хранения, преобразования и использования знаний живыми и искусственными системами» [18, 19]. В этой области получен огромный фактический материал по психологическим механизмам мышления, структуре сознания и нейрофизиологическим процессам («коррелятам»), сопровождающим работу мозга у человека и животных. К когнитивным наукам примыкает область конструирования и исследования *искусственного интеллекта* [17, 20]. Одним из наиболее известных достижением здесь являются искусственные нейронные сети (ИНС), успешно используемые в аппроксимации функций, распознавании образов, финансовом прогнозировании и поиске многопараметровых корреляций. Важное место в исследованиях ИНС, до настоящего времени остающихся перспективным, но не вполне объясненным техническим средством, занимают изучение механизма их функционирования и осуществляемой ими переработки информации [21].

Несмотря на многочисленные положительные результаты когнитивных исследований и компьютерного моделирования познавательных процессов, общепринятой формальной модели интеллектуальной деятельности до настоящего времени также не существует. В частности, детально разработанный алгоритм «обучения» нейронных сетей позволяет ИНС распознавать лишь такие виды внешнего воздействия, которые были заданы в ходе обучения. За пределами теории остается главное содержание мышления: производство новой, ранее не известной информации.

В то же время в психологии мышления имеется общепринятая описательная схема умственной деятельности, приводящей к открытию новых знаний [22], и выработаны эмпирические рекомендации («эвристики»), повышающие вероятность ее успеха – такие, как «мозговой шторм» [23]. Рядом авторов – в частности, в работах Д.С.Чернавского [24] – был предложен «генетический» механизм отбора *полезной информации*, который закрепляет новое знание, генерируемое в случайном процессе. В исследованиях динамики головного мозга методами электроэнцефалографии (ЭЭГ) и функциональной магнитно-резонансной томографии (фМРТ) экспериментально установлены локализация информации в определенных участках коры [25], «хранение» человеческих знаний и эмоций в различных областях мозга [26].

Характерными особенностями творческой умственной деятельности, установленными в психологии мышления [22, 27], являются:

1. обучение как необходимое условие мышления; желательность возможно более широкой «базы знаний» для плодотворного мышления;

2. спонтанное возникновение правильного решения («озарение», или *инсайт*) у людей и животных, поставленных в «проблемную ситуацию»;
3. повышение вероятности инсайта при активизации случайного перебора вариантов ответа и ослаблении ограничений на варьирование компонентов «проблемной ситуации».

Перечисленные положения использует, в частности, эвристическая схема стимулирования изобретательской деятельности (ТРИЗ). В ее рамках проблемная ситуация названа «изобретательской ситуацией», а изобретение (инсайт) состоит во «взгляде с новой стороны» на компоненты проблемной ситуации и в изменении логической связи между ними [23].

4. Модульная модель интеллекта

Термин «модульность» (modularity) весьма распространен в описании структуры интеллекта и его функций (познания), однако в когнитивных науках он используется скорее на вербальном уровне [28]. Мы предлагаем «модульную» модель индивидуального интеллекта, которая воспроизводит выработку новой информации в процессе мышления и может быть распространена на описание распределенного интеллекта МСС. Наша модель не использует математических соотношений теории информации, ограничиваясь «наивным» представлением об информации как имеющемся (или возникающем) человеческом знании. Ее основой служат перечисленные выше характеристики интеллектуальной творческой деятельности (п.п. 1 – 3) и «угадывание» (а не строгий логический вывод) правильного решения в проблемной ситуации.

Суть модели заключается в блочном (модульном) характере схематичного «опосредованного отражения», или *внутренней репрезентации* [29], вырабатываемой сознанием в ответ на внешнее воздействие, и в генетическом алгоритме подбора модулей, составляющих репрезентацию. Мы предполагаем, что внешнее воздействие на сознание вызывает в нем «отпечаток», интерпретируемый (т.е. воспринимаемый) путем сравнения с некоторой схемой, ранее выработанной для сходных «отпечатков» и вызываемой из памяти. Предполагается, что эта схема (далее – «образ-схема») состоит из небольшого числа знаковых модулей-«иероглифов», библиотека которых составляется в процессе обучения и хранится в памяти (рис. 1). Разные «отпечатки» интерпретируются разными комбинациями модулей. «Единицами хранения» в памяти являются не образы и не аппроксимирующие их схемы: в памяти хранится библиотека модулей и содержатся «ключи» (наборы идентификаторов модулей), позволяющие вызвать ранее построенные из них схемы (рис. 2).

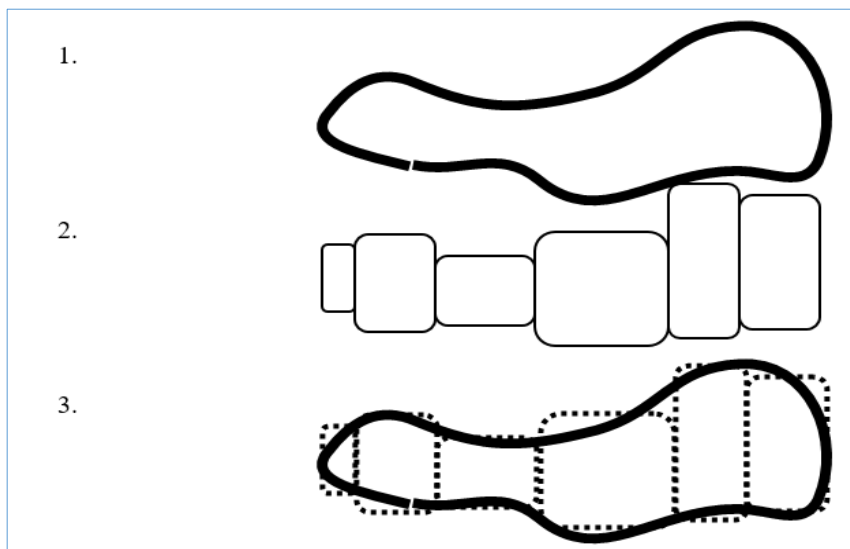


Рис. 1. Компоненты восприятия сознанием внешних стимулов: (1) «отпечаток» внешнего воздействия, (2) его интерпретация («образ-схема»), (3) сопоставление «отпечатка» и «образа-схемы»

Fig. 1. Components of perception of external stimuli: (1) «imprint» of the external action, (2) its interpretation («image-scheme»), (3) comparison of the «image-scheme» with the «imprint»

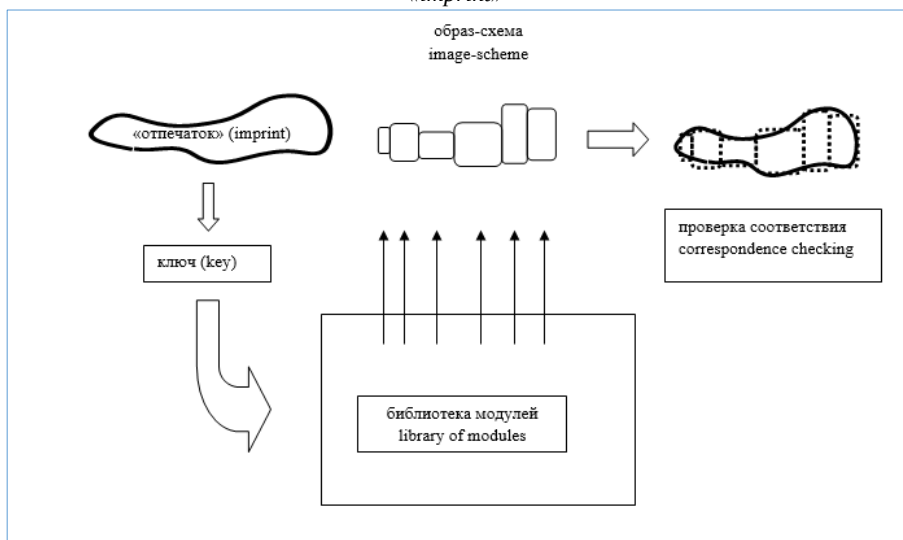


Рис. 2. Алгоритм восприятия: построение «образа-схемы» из библиотеки модулей

Fig. 2. Building of «image-scheme» from the library of modules

По аналогии со знаками иероглифической письменности, модули, составляющие схему, в процессе их приобретения (обучения), вероятно, характеризуют определенные признаки внешних воздействий, но используются далее как универсальные компоненты «образов-схем»: комбинаций модулей, индивидуальных для каждого «отпечатка». Соответствие вызванной модульной схемы «отпечатку» внешнего воздействия проверяется в мозге по некоторому быстрому «автоматическому», не вполне установленному механизму [26].

При хорошем соответствии происходит узнавание. Плохое соответствие вызывает неудовлетворенность (в терминах ТРИЗ – «обостренное противоречие»). Стремление преодолеть неудовлетворенность стимулирует построение новой схемы, лучше соответствующей «отпечатку», из блоков, имеющихся в библиотеке. Создание новой модульной схемы и запоминание ее ключа соответствует производству новой информации в ходе умственной деятельности.

Число модулей, составляющих «образ-схему» (m), можно оценить сверху числом параметров или факторов, которые человек в состоянии воспринимать параллельно (« 7 ± 2 »). Разумной оценкой общего количества N модулей в библиотеке (безусловно, определяемого продолжительностью и глубиной обучения) может служить количество иероглифов, которыми оперирует грамотный носитель японского или китайского языка (до нескольких тысяч). Оценки $m\sim 10$ и $N\sim 1000$ дают практически бесконечное число возможных сочетаний модулей в схеме $\binom{m}{N}\sim 1023$. Вместе с тем идентификатор схемы, позволяющий вызвать ее модули из памяти в заданном порядке, может весьма экономно состоять из «номеров» или иных признаков модулей в библиотеке – для чего в предельном случае достаточно одного нейрона, локализованного в коре головного мозга, на один модуль.

Предлагаемая модель на качественном уровне отражает главное содержание интеллектуальной деятельности: «придумывание нового» в форме подбора новой комбинации модулей, которая аппроксимирует воспринимаемый образ лучше всех комбинаций, хранящихся в памяти. Из невозможности, при принятых оценках чисел m и N , построить образ-схему прямым перебором модулей следует достаточно правдоподобный алгоритм творческой интеллектуальной деятельности:

- 1) поиск наилучшего образа-схемы для «отпечатка» внешнего воздействия по имеющимся в памяти ключам;
- 2) выявление в найденной схеме модуля (или модулей) с наихудшим соответствием «отпечатку»;
- 3) замена «плохого» модуля на другие модули, хранящиеся в памяти, ограниченным случайным перебором;
- 4) нахождение модуля, улучшающего соответствие схемы «отпечатку»;
- 5) запоминание новой комбинации модулей: введение в память ее ключа (рис. 3).

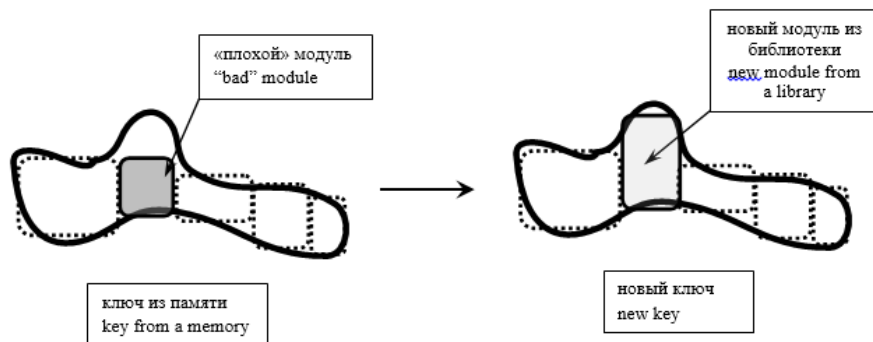


Рис. 3. Замена модуля в образе-схеме, улучшающая соответствие «отпечатку»
Fig. 3. Replacement of a module in the «image-scheme» improving its correspondence to «imprint»

Наша модель согласуется с такими известными условиями успешного решения творческих задач, как предварительное обучение, эрудиция (создание и расширение библиотеки), опыт распознавания проблемных ситуаций (алгоритм нахождения наилучшего образа в памяти), свободный ассоциативный поиск («ключи») и *умственные способности*, т.е. способность без ошибок обрабатывать большие объемы информации. В эмпирических описаниях творческой деятельности пункту (2) нашего алгоритма соответствует «проблемная», или «изобретательская» ситуация, п. (3) – состояние «нерешенной проблемы», которому сопутствуют процессы «расшатывания» образа (и снятие ограничений в поиске как одно из условий успеха), п. (4) – «озарение» (инсайт), а п. (5) – рождение новой информации. В бесструктурной схеме Д.С.Чернавского [24] аналогом перебора модулей является хаотический «перемешивающий слой», или странный аттрактор, в фазовом пространстве континуальной математической модели, а ценность новой информации, полученной случайным путем, определяется ее пригодностью для решения текущих задач. В нашей модели генетический подбор модулей контролируется сопоставлением получаемых схем с «отпечатком» воздействия; хорошее согласие предположительно вызывает биохимически контролируемое удовлетворение [26, 30].

Обсуждаемая модель согласуется с наличием локализованных нейронов, воспроизводимо активируемых в коре головного мозга при предъявлении определенного стимула («когнитивной специализацией нейронов в мозге человека» [25]). Локализованные нейроны в этом случае являются частью физического «субстрата» модуля, или воспроизводимо активируемого ансамбля нейронов [31], в составе «образа-схемы», возникающей при распознавании внешнего воздействия. В терминах, используемых при

обсуждении сетевых структур мозга, физическим носителем модуля является *ког*: «распределенная группа нейронов, сцепленная единым когнитивным опытом», т.е. кодирующая определенный аспект восприятия внешних воздействий [25]. Следует заметить, что нейроны, входящие в состав «паттерна активации» [31], могут быть локализованы в коре головного мозга без определенного алгоритма, ситуативно – подобно записи информации на магнитный диск. В этом случае тесная связь мышления с конкретными сетевыми структурами мозга (*коннектомом*) и сознания (*когнитомом*), постулируемая в данной области [25], подобно предполагаемой связи сетевой структуры с динамикой социальных систем, может не соответствовать действительности (см. разд. 1).

5. К моделированию распределенного интеллекта

Предлагаемая «модульная» модель индивидуального интеллекта естественно распространяется на описание распределенного интеллекта (РИ) мультиагентных социальных систем. Наличие интеллекта у МСС определяется способностью агентов воспринимать информацию, наличием целей у агентов и их взаимодействием (см. Введение). Уровень РИ соответственно зависит от когнитивных возможностей агентов, от преследуемых ими целей, от характера и структуры межагентных взаимодействий, а также от внешних условий.

При распаде структуры МСС (отсутствие ресурсов, изобилие ресурсов либо «война всех против всех») коллективная обработка информации не реализуется. Распределенный интеллект системы в общем случае возрастает с увеличением числа агентов и с усложнением структуры их взаимодействий.

Анализ и использование информации организационными системами достаточно высокого уровня, функционирующими в человеческом обществе, имеет очевидные аналогии с описанием индивидуального интеллекта по «модульной» схеме (табл. 1). В качестве «отпечатка» внешнего воздействия на систему (например, аварии на производстве) здесь выступает гетерогенная совокупность реакций системы как целого и индивидуальных реакций агентов. (Так, по должностным инструкциям, работники предприятия при аварии обязаны сообщить о ней руководству, однако серьезная авария может вызвать панику и потерю управления).

Составными единицами при интерпретации воздействия системой служат инструкции и должностные обязанности. Сильные внешние воздействия изменяют структуру системы (при аварии – эвакуация персонала, создание аварийного штаба и комиссии). «Образом-схемой» в данном случае является стандартный ответ системы (действие по инструкциям), корректировкой схемы – оценка успешности действий и испытание новых предложений (работа комиссии), заменой модуля (модулей) в «образе-схеме» – выбор оптимальных действий (выводы комиссии), запоминанием новой информации – дополнение и изменение инструкций.

Табл. 1. Сравнение компонентов индивидуальной интеллектуальной деятельности и распределенного интеллекта организационной системы
Table 1. A comparison of activity of individual human's intellect with a distributed intelligence of an enterprise

индивидуальный интеллект человека	распределенный интеллект предприятия
1-й этап восприятия: «отпечаток» внешнего воздействия	сумма системных и индивидуальных реакций работников на внешнее воздействие
библиотека модулей и «хранилище ключей»	рабочая информация, должностные инструкции
2-й этап восприятия: выбор модульной «схемы-образа»	оценки параметров воздействия и его предварительная характеристика; действия по инструкциям
поиск лучшей имеющейся приближенной схемы, оценка ее расхождений с образом	оценка успешности стандартных действий («создание комиссии»)
случайная замена модулей	корректировка действий; испытание новых предложений («работа комиссии»)
«генетический контроль»	оценка успешности скорректированных действий («выводы комиссии»)
запоминание новой схемы	дополнение и изменение инструкций

Представленный модельный пример иллюстрирует как сложность и неоднородность структурированной социальной системы, так и наличие нескольких уровней ее распределенного интеллекта (так, при потере управления персонал предприятия будет действовать группами, преследующими собственные цели). Существенно, что рациональность системной динамики увеличивается при ограничении набора возможных действий агентов и при наличии «библиотеки» стандартных реакций руководства предприятия и персонала.

Тесную связь «интеллектуальности» организационной системы с формализацией ее структуры и действий иллюстрирует табл. 2. Рациональность индивидуального поведения людей также возрастает с сокращением возможностей выбора (переход улицы по светофору, покупка или продажа акций на бирже по фиксированной цене) и становится «ограниченной» при выборе из многих возможностей (переход улицы без светофора, произвольная цена акций на бирже).

Табл. 2. Формальные условия функционирования разных организационных систем
Table 2. Formal conditions of activity for different organizing systems

«идеальная комиссия»	МИТИНГ	«коллективный идиот»
общие знания по специальности	общие намерения	ничего общего
целевой отбор участников по критерию квалификации	случайный отбор участников по близости	свободный вход

	настроенный	
сильное управление (председатель с решающим голосом)	слабое управление	нет управления
формализованный обмен информацией и мнениями, исключение эмоций	неформальный обмен мнениями и эмоциями	случайный обмен эмоциями
количественное сравнение значимости мнений (голосование)	декларации мнений (призывы)	нет формулируемых мнений
подчинение меньшинства большинству	неподчинение меньшинства большинству	нет большинства
обязательность исполнения решений	необязательность исполнения решений	нет решений

Общеизвестные схемы коллективного поведения «живых» социальных систем с более примитивными агентами (улей, муравейник, гнездо ос) также неплохо согласуются с основными этапами «модульного» алгоритма интеллектуальной деятельности [16]. Как и в организационной системе, состоящей из людей, в поведении насекомых трудно разделить формирование «отпечатка» (фиксацию системой внешнего воздействия) и стандартную реакцию на него («образ-схему»); примером может служить атака муравьев или ос на чужеродный объект вблизи гнезда. (Разделение воздействия, его отражения в сознании и репрезентаций образа в психологии также является непростой задачей [19]). «Танцы» пчел в улье и «тропинки» муравьев, которые можно рассматривать как модули, направляющие движения особей, изменяются при изменении внешних условий. Перемещение муравьиного гнезда из неблагоприятной зоны в благоприятную, восстановление поврежденного муравейника и многие другие действия коллективных насекомых разбиваются на несколько стандартных воспроизводимых режимов (возбуждение, несогласованные действия, согласованные действия, «успокоение» при достигнутой цели) – этот перечень легко продолжить.

6. «Измерение» распределенного интеллекта модельной системы

Компьютерное моделирование динамики МСС неявно включает оценку уровня распределенного интеллекта модельной системы (например, транспортного потока [11] – по скорости движения, пропускной способности дороги, предсказываемому числу аварий и т. д.) и максимизацию этого уровня. Мерой РИ в таких расчетах является близость достигаемого результата к объективной цели системы – в случае транспортного потока к максимально быстрому безаварийному движению. Для компьютерной иллюстрации этих положений мы использовали анимационную модель, в которой агенты с одинаковым радиусом r_0 перемещались по коридору с препятствиями, блокирующими сквозное движение (рис. 4). Состояние агента

задавали пять числовых параметров ($R(n)$, k_1 , k_2 , k_3 , k_4), где $R = nr_0$ – радиус восприятия агентом его окружения, а остальные $0 \leq k_i \leq 0.1$ – параметры движения агента:

k_1 – ускорение вправо: приращение скорости на каждом шаге дискретного времени как доля максимальной скорости v_0 (целеполагание агента),

k_2 – торможение перед препятствием в радиусе восприятия,

k_4 и k_3 – соответственно следование за другими агентами в радиусе восприятия и коррекция собственной скорости по скорости их движения.

Столкновение агента с препятствиями и другими агентами происходили как абсолютно упругие с соответствующими изменениями направления движения и скоростей. Более детальное описание схемы расчета будет представлено в отдельной публикации.

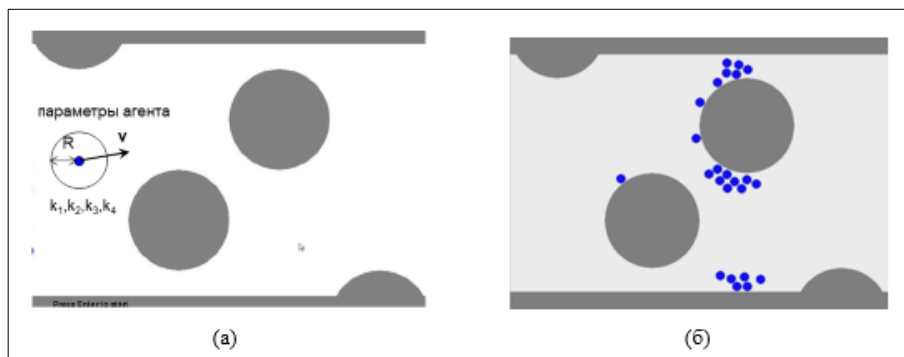


Рис.4. (а) Схема «коридора» с препятствиями и параметры агента, (б) движение в «обученной» системе

Fig. 4. (a) Pathway with obstacles and agent's parameters; (b) movement of agents in a «trained» system

В первом цикле расчета 24 агента одинакового радиуса r_0 с фиксированным радиусом восприятия $R=3r_0$ и параметрами движения $(0.01, 0, 0, 0)$, случайным образом размещенные в полосе шириной $5r_0$ перед «стартовой» линией, перемещались по коридору слева направо до «финиша», т.е. прохождения полной длины коридора. Направления движения и скорости агентов корректировались на каждом шаге дискретного времени в соответствии с параметрами $\{k_i\}$ по объектам (препятствиям и другим агентам) в радиусе восприятия. В последующих циклах параметры движения 12 «медленных» агентов, преодолевших коридор на предыдущем цикле позже 12 «быстрых», заменяли параметрами 12 «быстрых» агентов со случайными вариациями $\{k_i = \pm 0.001\}$ (генетический отбор). Количественной оценкой РИ системы служило среднее время t прохождения коридора агентом в каждом цикле.

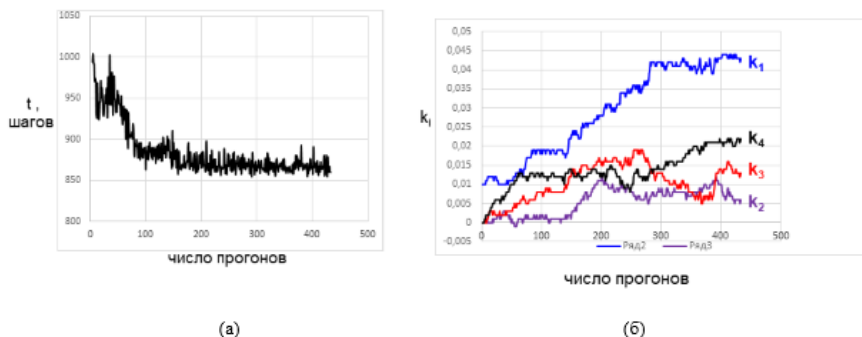


Рис.5. (а) Изменение среднего времени прохода t при обучении» системы, (б) изменения коэффициентов движения

Fig. 5. (a) Evolution of average passing time t during 'learning' (b) evolution of movement parameters

В ходе эволюции искусственной МСС среднее время прохода t за первые 150-200 циклов уменьшалось от 1040–1050 до 880–870 шагов и далее оставалось приблизительно постоянным (рис. 5 а). Параметры движения агентов при этом продолжали изменяться (рис. 5 б). Характер движения в результате эволюции изменялся от хаотического дрейфа при многочисленных столкновениях с препятствиями и между собой в «необученной» системе к согласованному движению группами, огибающими препятствия, в «обученной» системе (рис. 4 б). Время прохождения коридора агентами с фиксированными параметрами движения ($k_1, 0.01, 0.01, 0.01$) при $k_1 \leq 0.05$ уменьшалось с увеличением радиуса восприятия, проходя через минимум (рис. 6). (Дальнейшее увеличение k_1 приводило к достижению всеми агентами максимальной скорости v_0 за первые 10-15 шагов дискретного времени и сопровождалось увеличением t с его небольшим линейным ростом при увеличении R). Таким образом, расчеты наглядно продемонстрировали способность модельной МСС из искусственных агентов воспринимать информацию и оптимизировать динамику в ходе «обучения» (т.е. генетического отбора «быстрых» агентов).

7. Обсуждение

Проявлениям РИ в социальных системах, включая различные аспекты коллективной человеческой деятельности, посвящена обширная литература (см. [1, 10, 30, 32, 33]). Однако в большинстве современных публикаций интеллектуальная деятельность МСС трактуется весьма ограниченно: как обмен информацией между агентами [1, 33] или как выработка общего мнения

в совокупности людей («crowd wisdom» [34]). За пределами обсуждения, аналогично теориям индивидуального интеллекта, оказывается способность МСС не только воспринимать и обрабатывать информацию, но и генерировать новую информацию – например, в форме муравьиных «тропинок», ведущих к новому, не использованному ранее источнику пищи. При этом «обучаемая» система может состоять из агентов с нулевыми когнитивными возможностями – как в построенной нами модели движения по коридору, где динамику каждого агента задавали пять числовых параметров.

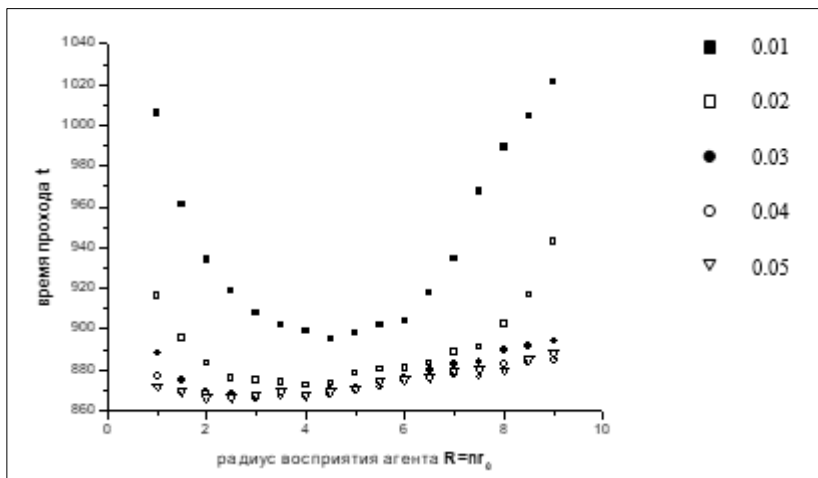


Рис.6. Зависимость времени прохода t от радиуса восприятия R при разных коэффициентах k_1

Fig. 6. Passing time t vs. perception radius R at different values of movement parameter k_1

Организационные системы, состоящие из людей, также способны не только к обмену индивидуальными мнениями, приводящему к «консенсусу» (что воспроизводится большинством социологических моделей, см. [14, 15]), но и к производству новой информации «коллективным разумом». РИ таких систем отличается от человеческого сознания как принципиально меньшим числом «узлов» (т.е. людей) в сетевых структурах, так и способностью этих узлов к глубокой переработке информации. Текущая информация о состоянии, генерируемая в системе, направляет ее коллективную динамику в переменных внешних условиях и фиксируется в виде административных «образов-схем» – т.е. распоряжений, правил и инструкций. При этом «глубина» РИ явно возрастает с повышением формализации связей в структуре МСС. Так, в модельных примерах таблицы 2, строгая регламентация «идеальной комиссии» (левая колонка) является эвристикой, стимулирующей ее содержательную работу, тогда как полностью деформализованная МСС (правая колонка таблицы), независимо от интеллектуального уровня

составляющих ее людей, вряд ли способна принять какое-либо общее решение.

«Модульный» подход представляется продуктивным также для конструирования искусственного интеллекта (ИИ) в вычислительной компьютерной среде. Известно, что при «обучении» многослойных нейронных сетей векторные реплики распознаваемых объектов в промежуточных слоях ИНС преобразуются к модульной структуре [21]. В общем случае распознаваемому объекту может сопоставляться набор модулей, фиксирующих определенные качества объекта – что открывает путь к созданию самообучающихся программ. В отличие от индивидуального человеческого интеллекта, число модулей в «образе-схеме», при хороших вычислительных мощностях, может быть достаточно большим, а перебор «плохих» модулей – быстрым и эффективным. Параметры «модульного» ИИ, таким образом, могут существенно превысить любые человеческие возможности, что указывает на вероятные риски при его реализации.

8. Перспективные направления исследований

Исследования распределенного интеллекта социальных систем будут проводиться в нескольких основных направлениях.

- (1) Разработка модульной модели индивидуального интеллекта, пока сформулированной на качественном уровне, наполнение ее количественным содержанием (структура модулей, ключей и библиотек, модульный состав «образов-схем», механизмы сравнения «отпечатка» с модульной схемой, замены модулей в схеме, обновления модулей в библиотеке и т.д.).
- (2) Эмпирические оценки уровня распределенного интеллекта для социальных систем разной природы (организаций, партий, политических течений и т.д.). Поиск корреляций «уровня интеллекта» со структурой социальной системы.
- (3) Формальное описание процессов восприятия и переработки информации индивидуальным сознанием и «коллективным разумом». Исследование математических свойств *информационного пространства*, объектами которого служат модули восприятия и «образы-схемы».
- (4) Модификация существующих моделей коллективного поведения (динамика пешеходных и автомобильных потоков, распространение мнений в онлайн-социальных сетях, политические кампании и др.) с непосредственным учетом распределенного интеллекта у системы взаимодействующих агентов. Сопоставление результатов моделирования с эмпирическими данными.

9. Заключение

Мы полагаем, что предложенный нами «модульный» алгоритм восприятия, обработки, использования и хранения информации применим к описанию индивидуального интеллекта человека и (в скорректированном виде) к моделированию распределенного интеллекта социальных систем. Оба вида информационной динамики реализуются в определенных сетевых структурах (см. [1, 8, 14, 17, 25]), взаимосвязь которых с уровнем интеллекта следует дополнительно изучать. Очевидными направлениями дальнейших исследований в данной области являются детализация и формализация модульной модели интеллекта, эмпирические оценки «глубины» распределенного интеллекта МСС и его непосредственный учет в моделировании мультиагентных систем.

Ю.Л.С. признателен доктору психологических наук М.В. Фаликман (НИУ ВШЭ) за предоставленную литературу по когнитивным исследованиям и плодотворное обсуждение «модульной» модели.

Список литературы

- [1] Moussaïd M, Helbing D, Theraulaz G, How simple rules determine pedestrian behavior and crowd disasters. *PNAS*, 108 (17), 2011, pp. 6884-6888
- [2] Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. 2-е издание. М.: Синтег, 2005, 138 с.
- [3] Galam S. *Sociophysics: a physicist's modeling of phycho-political phenomena*. Springer, 2012, 439 p.
- [4] Захаров А.В. Модели политической конкуренции: обзор литературы. *Экономика и математические методы*, том 45, вып. 1, 2009 г., стр. 110-128
- [5] Dorogovtsev S.N. *Lectures on Complex Networks*. Clarendon: Oxford, 2010, 134 p.
- [6] Newman M.E.J. The structure and functions of complex networks, *SIAM Review*, 45(2), 2003, pp. 167-225.
- [7] Берновский М.М., Кузюрин Н.Н. Случайные графы, модели и генераторы безмасштабных графов. *Труды ИСП РАН*, 2012, том 22, стр. 419-432. DOI: 10.15514/ISPRAS-2012-22-22.
- [8] Евин И.А. Введение в теорию сложных сетей. Компьютерные исследования и моделирование, том 2, вып. 2, 2010 г., стр. 121-141.
- [9] Новиков Д.А. Модели стратегической рефлексии. *Автоматика и телемеханика*, том 73, вып.1, 2012 г., стр. 1 -19
- [10] Kahneman D. Maps of Bounded Rationality: Psychology for Behavioral Economics *Amer. Econ. Rev.*, 2003, 93(5), pp. 1449-1475.
- [11] Гасников А.В. (ред.), Введение в математическое моделирование транспортных потоков. М.: МЦНМО, 2013, 428 с.
- [12] Адамчук А.Н., Есипов С.Е. Коллективно флуктуирующие активы при иналичии арбитражных возможностей и оценка платежных обязательств. *Усп. физ. наук*, том 167, вып. 12, 1997 г., стр. 1295-1306.
- [13] Schelling T. Dynamic models of segregation. *J. Math. Sociol.*, 1 (2), 1971, pp. 143-186
- [14] Castellano C., Fortunato S., Loreto V. Statistical physics of social dynamics. *Rev. Mod. Phys.* 81 (2), 2009, pp. 591-646.

- [15] Словохотов Ю.Л. Физика и социофизика. Проблемы управления, 2012, вып. 3, стр. 2–34.
- [16] Кипятков В.Е. Мир общественных насекомых, 3-е изд., М.: Либроком, 2009, 408 с.
- [17] Engelbtecht A.P. Fundamentals of computational swarm intelligence. N.-Y.: Wiley, 2005, 672 p.
- [18] Фаликман М.В. Основные подходы в когнитивной науке. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2017-02-09/presentation.pdf>. Дата обращения: 02.04.2018.
- [19] Величковский Б.М. Когнитивная наука: основы психологии познания. В 2-х т.. М.: Смысл; Академия, 2006.
- [20] Shaib-Draa B, Moulin B., Mandiau P., Millot P. Trends in distributed artificial intelligence. *Artific. Intelligence Rev.*, 6 (1), 1992, pp. 35-66.
- [21] Хайкин С. Нейронные сети. Полный курс. 2-е издание. Вильямс, 2016, 1104 с.
- [22] Петухов В.В. Психология мышления. Учебно-методическое пособие. М.: МГУ, 1987, 99 с.
- [23] Альтшуллер Г.С. Найти идею. Введение в ТРИЗ – теорию решения изобретательских задач. 4-е изд. М.: Альпина Паблишерз, 2011, 400 с.
- [24] Чернавский Д.С. Синергетика и информация: динамическая теория информации. 3-е изд. М.: Либроком, 2009, 304 с.
- [25] Анохин К.В. Когнитом: разум как физическая и математическая структура. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2016-09-27/presentation.pdf>. Дата обращения: 02.04.2018.
- [26] Шумский С.А. Моделирование работы мозга: состояние и перспективы. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2015-03-31/presentation.pdf>. Дата обращения: 02.04.2018.
- [27] Ohlsson S. Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking*. New York, NY: Harvester Wheatsheaf, 1992, pp. 1– 44.
- [28] Fodor J.A. *The Modularity of Mind*. MIT Press 1983, 142 p.
- [29] Курганский А.В. Понятие внутренней репрезентации в когнитивной нейронауке. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2017-02-28/presentation.pdf>. Дата обращения: 02.04.2018.
- [30] Dandurand F., Shultz T.R., Rivest F. Complex problem solving with reinforcement learning. *Proc. 6th IEEE Internat. Conf. on Development and Learning*, 2007, pp. 157-162.
- [31] Hebb D.O. *The Organization of Behavior: a Neuropsychological Theory*, Wiley, 1949, 335 p.
- [32] Mossaid M., Garnieer S., Theraulaz G., Helbing D. Collective information processing and pattern formation in swarms, flocks and crowds. *Topics Cogn. Sci.*, 1, 2009, pp 469-497.
- [33] Becker J., Brackbill D., Centola D. *PNAS*, 114, 2017, pp. E5070-E5076.

Toward construction of a modular model of distributed intelligence

^{1,2}Yu.L. Slovokhotov <slov@phys.chem.msu.ru>

³I.S. Neretin <ivan@neretin.ru>

¹Lomonosov Moscow State University,
GSP-1, 1 Leninskie Gory, Moscow, 119991, Russia

²Institute of Organoelement Compounds, Russian Academy of Sciences,
28 Vavilov St., Moscow, 119991, Russia

³Rock Flow Dynamics, 25A Profsoyuznaya St., Moscow, 117418, Russia

Abstract. Multi-agent social systems (MASS) in general are systems of autonomous interdependent agents each pursuing its own goals interacting with other agents and environment. Dynamics of MASS cannot be adequately modeled by the methods borrowed from statistical physics since these methods do not reflect the main feature of social systems, viz. their ability to percept, process and use the external information. This important quality of distributed (“swarm”) intelligence has to be directly taken into account in a correct theoretical description of social systems. However, discussion of distributed intelligence (DI) in the literature is mostly restricted to distributed tasks, information exchange and aggregated judgment – i.e. to ‘sum’ or ‘average’ of independent intellectual activities. This approach ignores empirically well-known option of a ‘collective insight’ in a group as a special demonstration of MASS’s DI. In this paper, a state of art in modeling social systems and studies of intelligence per se are briefly characterized, and a new modular model of intelligence is suggested. The model allows to reproduce the most important result of intellectual activity, i.e. creation of new information, which is not reflected in the contemporary theoretical schemes (e.g. neural networks). Using the “modular” approach, a correspondence between individual intelligence and DI of MASS is discussed, and prospective directions for future studies are suggested. Efficiency of DI was estimated numerically by computer simulations of a simple system of agents with variable kinematic parameters {*k_i*}, moving through a pathway with obstacles. Selection of fast agents with ‘positive mutation’ of parameters gives ca. 20% reduction of average passing time after 200-300 cycles and creates a swarm movement where agents follow a leader and cooperatively avoid obstacles.

Keywords: multi-agent social systems; swarm intelligence; models of intelligence.

DOI: 10.15514/ISPRAS-2018-30(3)-23

For citation: Slovokhotov Yu.L., Neretin I.S. Toward construction of a modular model of distributed intelligence. *Trudy ISP RAN/Proc. ISP RAS*, vol. 30, issue 3, 2018, pp. 341-362 (in Russian). DOI: 10.15514/ISPRAS-2018-30(3)-23

References

- [1] Moussaïd M, Helbing D, Theraulaz G, How simple rules determine pedestrian behavior and crowd disasters *PNAS*, 108 (17), 2011, pp. 6884-6888
- [2] Goubko M.V., Novikov D.A. Game theory in control of organizing systems T. 2-nd ed.. Moscow, Sinteg, 2005, 138 p. (in Russian)

- [3] Galam S. Sociophysics: a physicist's modeling of psycho-political phenomena. Springer, 2012, 439 p.
- [4] Zakharov A.V. Models of political competition^ a review. *Economics and Mathematical Methods [Ekonomika i matematicheskiye metody]*, 45 (1), , 2009 pp. 110-128 (in Russian)
- [5] Dorogovtsev S.N. Lectures on Complex Networks. Clarendon: Oxford, 2010, 134 p.
- [6] Newman M.E.J., The structure and functions of complex networks, *SIAM Review*, 45(2), 2003, pp. 167-225.
- [7] Bernovskiy M.M., Kuzyurin N.N. Random graphs, models and generators of scale-free graphs. *Trudy ISP RAN/Proc. ISP RAS*, 2012, v. 22, pp. 419-432 (in Russian). DOI: 10.15514/ISPRAS-2012-22-22.
- [8] Yevin I.A. Introduction to a theory of complex networks. *Computer Research and Modeling [Kompyuternye issledovaniya i modelirovanie]*, 2 (2), 2010, pp. 121-141 (in Russian).
- [9] Novikov D.A. Models of strategic behavior. *Automation and Remote Control*, 73 (1), 2012, pp. 1 -19
- [10] Kahneman D. Maps of Bounded Rationality: Psychology for Behavioral Economics *Amer. Econ. Rev.*, 2003, 93(5), pp. 1449-1475.
- [11] Gasnikov A.V. (Ed.). Introduction into mathematic modeling of traffic flows. Moscow, MTsIMO, 2013, 428 p. (in Russian)
- [12] Adamchuk A.N., Esipov C.E. Collectively fluctuating assets in the presence of arbitrage opportunities, and option pricing. *Phys. Usp.*, vol. 40, 1997, pp. 1239–1248
- [13] Schelling T. Dynamic models of segregation. *J. Math. Sociol.*, 1 (2), 1971, pp. 143-186
- [14] Castellano C., Fortunato S., Loreto V. Statistical physics of social dynamics. *Rev. Mod. Phys.* 81 (2), 2009, pp. 591-646.
- [15] Slovokhotov Yu.L. Physics vs. Sociophysics. *Control Science [Problemy upravleniya]*, 2012 (3), pp. 2-34 (in Russian).
- [16] Kipyatkov V.E. World of social insects, 3rd Ed., Moscow, Librokom, 2009, 308 p.
- [17] Engelbrecht A.P. Fundamentals of computational swarm intelligence. N.-Y.: Wiley, 2005, 672 p.
- [18] Falikman M.V. The principal approaches in cognitive science <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2017-02-09/presentation.pdf> (in Russian). Accessed 02.04.2018.
- [19] Velichkovsky B.M. Cognitive science: foundation of cognition psychology. In 2 volumes. Moscow, Smysl, Akademiya, 2006 (in Russian).
- [20] Shaib-Draa B, Moulin B., Mandiau P., Millot P. Trends in distributed artificial intelligence. *Artific. Intelligence Rev.*, 6 (1), 1992, pp. 35-66.
- [21] Khaikin S. Neural networks. A comprehensive course. 2nd edition. Viljams, 2016, 1104 p. (in Russian)
- [22] Petukhov V.V. Psychology of thinking A textbook. Moscow, MGU., 1987, 99 p. (in Russian).
- [23] Altshuler G.S. To find the idea. Introduction into TRIZ: a theory of solving inventional problems. 4EP ed. Moscow, Alpina Publishers, 2011, 400 p. (in Russian)
- [24] Chernavsky D.S. Synergetics and information: dynamic theory of information. 3rd ed. Moscow, Librokom, 2009, 304 p. (in Russian)
- [25] Anokhin K.V. Cognitom: mind as a physical and mathematical structure. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2016-09-27/presentation.pdf> (in Russian). Accessed 02.04.2018.

- [26] Shumsky S.A. Modeling of brain activity: state of art and prospects. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2015-03-31/presentation.pdf> (in Russian). Accessed 02.04.2018.
- [27] Ohlsson S. Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking*, New York, NY: Harvester Wheatsheaf, 1992, pp. 1– 44.
- [28] Fodor J.A. *The Modularity of Mind*. MIT Press 1983, 142 p.
- [29] Kurgansky A.V. Internal representation in cognitive neuroscience. <http://www.soc-phys.chem.msu.ru/rus/prev/zas-2017-02-28/presentation.pdf> (in Russian). Accessed 02.04.2018.
- [30] Dandurand F., Shultz T.R., Rivest F. Complex problem solving with reinforcement learning Proc. 6th IEEE Internat. Conf. on Development and Learning, 2007, pp 157-162.
- [31] Hebb D.O. *The Organization of Behavior: a Neuropsychological Theory*, Wiley, 1949, 335 p.
- [32] Mossaid M., Garnieer S., Theraulaz G., Helbing D. Collective information processing and pattern formation in swarms, flocks and crowds. *Topics Cogn. Sci.*, 1, 2009, pp. 469-497.
- [33] Becker J., Brackbill D., Centola D. *PNAS*, 114, 2017, pp. E5070-E5076.

