

# РАСПРЕДЕЛЕННЫЙ ИНТЕЛЛЕКТ МУЛЬТИАГЕНТНЫХ СИСТЕМ.

## Ч. 1. Основные характеристики и простейшие формы

Ю.Л. Словохотов, Д.А. Новиков

**Аннотация.** Рассмотрены признаки и эмпирические характеристики распределенного интеллекта (РИ) как способности коллективного агента воспринимать, обрабатывать и использовать информацию для достижения собственных целей. В первой части обзора обсуждаются наиболее существенные признаки «протоинтеллектуальных» и интеллектуальных систем, особенности РИ, индивидуального интеллекта человека и искусственного интеллекта (ИИ). Подчеркнута неразрывная связь РИ организационных и социальных систем с индивидуальным человеческим интеллектом. На примерах простейших форм роевого интеллекта перечислены факторы, определяющие эффективность РИ мультиагентной системы, включая структуру взаимодействий между агентами, коллективное целеполагание, запись, свертывание и обработку внешней информации, стандартные образы внешних воздействий. Их совокупностью определяется способность мультиагентной системы к целенаправленным действиям, выходящим далеко за пределы возможностей отдельных составляющих ее агентов. Во второй части обзора будут рассмотрены формы коллективного интеллекта в «человеческих» социальных системах и дана общая классификация известных видов РИ.

**Ключевые слова:** мультиагентные системы, распределенный интеллект, роевой интеллект, организационные системы, коллективный интеллект.

### ВВЕДЕНИЕ

Изучение интеллектуальной деятельности человека, а также ее математическое и техническое моделирование – магистральные направления когнитивной психологии, информационных технологий, кибернетики, робототехники и ряда других наук [1]. На основе огромного объема данных в этих областях найдены ключевые признаки, общие для интеллекта человека (англ. *human intelligence*, HI) и различных компьютерных реализаций искусственного интеллекта (ИИ; англ. *artificial intelligence*, AI). К данному направлению примыкают исследования и моделирование кооперативной динамики в разнообразных мультиагентных системах: биологических, социальных, экономических и организационных, а также в группах автономных технических устройств. Процессы обработки и использования информации в таких системах, обычно относимые к проявлениям *распределенного интеллекта* (РИ; англ. *distributed intelligence*, DI), составляют предмет настоящего обзора.

Коллективная обработка информации и ее использование организационными системами (ОС) общеизвестны [1]. Примерами могут служить конкуренция фирм, боевые действия армейских подразделений, участие политических партий в избирательной кампании и другие процессы. Но повседневные проявления распределенного интеллекта в мультиагентных системах значительно шире. Они включают в том числе разделение движения пешеходов на встречные потоки (что увеличивает пропускную способность тротуаров и тоннелей), замедление автомобильного транспорта на перегруженных улицах (где потери времени от возникновения пробок перевешиваются соображениями безопасности [2, 3]), обвальные продажи активов на бирже, свидетельствующие о наступлении кризиса, и многое другое. В отличие от стандартных механизмов управления ОС, «консенсус» участников в таких системах устанавливается децентрализованно и без общих дискуссий: главной основой их согласованного поведения являются взаимная связанность и структура межагентных взаимодействий.

Многочисленные проявления РИ, не включающие осознанной обработки информации индивидами<sup>1</sup>, хорошо известны для общественных насекомых [4, 5], движения стай птиц и рыб [6, 7], пешеходных потоков [2, 8], формаций беспилотных аппаратов [8, 9]. Сюда же относятся такие формы деятельности людей, как рынок и биржа, где стремление участников к максимуму прибыли преобразуется в коллективную оценку стоимости товаров [10]. Существенно, что РИ мультиагентных систем самой разной природы не сводится к стандартным схемам индивидуального или коллективного принятия решений либо иерархического или сетевого управления: он представляет собой самостоятельный и недостаточно изученный аспект кооперативной динамики [11].

В настоящем обзоре рассмотрены основные известные на сегодня виды РИ в системах взаимосвязанных *агентов*, включая биологические, технические, социальные и организационные. Не углубляясь в огромную область моделирования интеллекта, перечислим основные виды РИ и рассмотрим их аналогии с интеллектуальным поведением человека и животных<sup>2</sup>, а также с некоторыми видами ИИ.

Из-за большого объема материала обзор разбит на две части. В представленной здесь первой части обсуждаются общие признаки всех видов интеллекта, кратко рассмотрены примеры их частичной реализации в «протоинтеллектуальных» системах автоматического управления и в существующих моделях ИИ. Более подробно в этой части обзора рассматриваются простейшие формы *роевого интеллекта* в группах биологических особей (общественные насекомые, косяки рыб, стаи птиц) и в технических системах (роевой интеллект в робототехнике), включая имитацию РИ в агентных моделях и в современных алгоритмах компьютерных расчетов. Различным видам РИ в человеческих сообществах, интеграции ИИ и ИИ в «коллективный разум» и общей классификации всех видов интеллекта будет посвящена вторая часть обзора.

## 1. ЭМПИРИЧЕСКИЕ ХАРАКТЕРИСТИКИ ИНТЕЛЛЕКТА

В обширнейшей литературе, посвященной интеллектуальной деятельности человека, парадоксальным образом отсутствует общепринятое определение интеллекта. Так, например, в фундамен-

тальной монографии “The Cambridge Handbook of Intelligence” [12] различным подходам к описанию и объяснению этого явления посвящены все пять глав первой части. Феномен интеллекта исследуется в широком спектре дисциплин от технических наук, биологии и психологии до философии с центром тяжести в когнитивных науках. Большинство определений интеллекта во всех этих областях с разной степенью детализации дают его эмпирическое описание как способности использовать поступающую извне ограниченную информацию в постановке и решении нестандартных задач, способствующих адаптации индивидуума в изменяющейся внешней среде. Их основой, безусловно, служит интеллект человека, однако такой характеристике отвечает значительно более широкий круг объектов и систем, включая технические, биологические и социальные (см. выше).

Разные виды интеллекта и его носителей, обсуждаемые в современной литературе, представлены на рис. 1. Чтобы не загромождать эту далеко не полную схему, в ней не показаны логические связи однотипных проявлений интеллектуальной деятельности у индивидов и сообществ разной природы (роевой интеллект низших животных и «краудсорсинг» в человеческом обществе, компьютерные алгоритмы «природных» вычислений, «эмерджентный» ИИ мультиагентных систем и многое другое), которые активно обсуждаются в литературе и будут рассмотрены ниже. Единственное исключение сделано для многообразных реализаций коллективного интеллекта человека в различных социальных структурах (штриховая стрелка), составляющих главный предмет нашего обсуждения. В описании разных видов интеллекта используются различные формальные модели, общие признаки которых будут далее проанализированы. Процессы в социальных системах наиболее крупного масштаба, выделенных серым цветом на схеме, чаще всего обсуждаются в гуманитарных дисциплинах на описательном уровне [13], однако в них, безусловно, проявляются многие виды коллективного интеллекта, поэтому они также включены в классификацию.

Если проявления РИ в слабо структурированных биологических и человеческих сообществах вызывают большой академический интерес (см. книги [10, 11] и цитированную в них литературу), то «умными» техническими и организационными мультиагентными системами преимущественно занимаются науки об управлении [14, 15] и прикладные инженерные дисциплины [16, 17]. В последнее десятилетие в этой области приобрел популярность термин *emergent intelligence*, EI: «спонтанно возникающий», «рождающийся», также

<sup>1</sup> Ниже для всех случаев, где это не вызывает разночтений, понятия «индивид», «индивидуум» и «особь» используются как синонимы, предпочтительно заменяемые словом «агент».

<sup>2</sup> В биологической литературе чаще называемого когнитивными способностями животных.

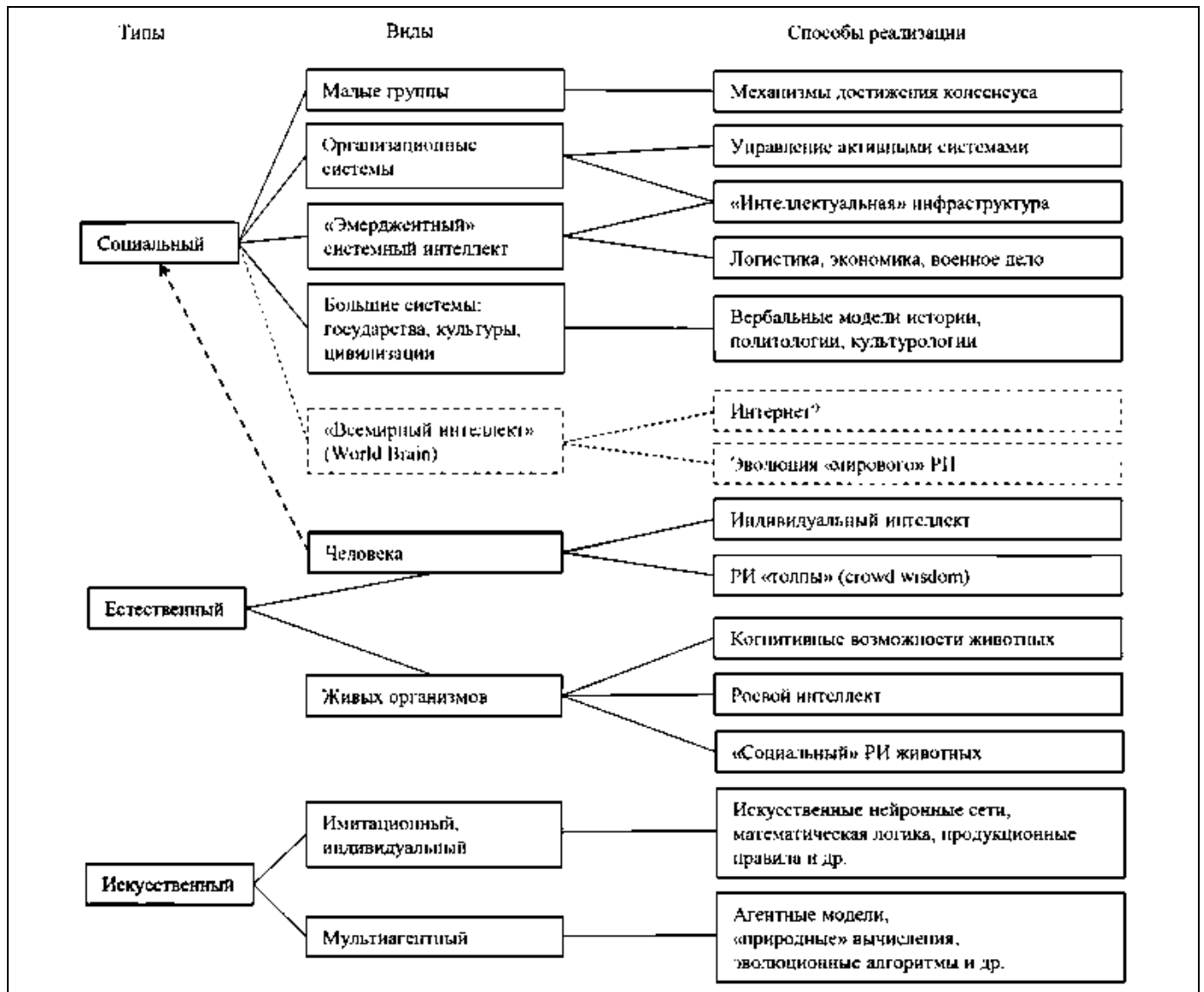


Рис. 1. Виды интеллекта и способы его реализации

*эмерджентный интеллект* в естественных и искусственных сложных системах. Так нередко обозначают гибкие мультиагентные системы планирования производственно-логистической деятельности, результаты которых очевидно превосходят индивидуальные возможности образующих их единиц. Этот термин особенно часто используют в описании РИ экономических, транспортных и других больших систем (см. работу [16] и др. книги из этой серии).

В приложениях также весьма распространены скорее рекламные словосочетания («интеллектуальный терминал», «интеллектуальная система» и т. д.) как характеристики передовой, многоцелевой и гибкой технической либо организационной схемы. Это особенно характерно для огромной области *робототехники* [18], где термин «интеллект» нередко используют как метафору далеко за

пределами его трактовки в психологии или когнитивистике. Ниже будет показано, что системы автоматического управления автономных технических устройств действительно проявляют некоторые существенные признаки *всех* видов интеллекта и в этом смысле могут называться «протоинтеллектуальными».

В характерной статье [19], открывающей сборник материалов американского симпозиума 2006 г. по «общему», или «сильному» ИИ (*Artificial General Intelligence*)<sup>3</sup>, приводится 18 общенаучных и 35 психологических определений интеллекта, ис-

<sup>3</sup> Обсуждение «слабого» ИИ (Narrow AI), реально применяемого на практике, и пока гипотетического «сильного» ИИ (General AI), не уступающего интеллекту человека, не является предметом нашего обзора.

пользуемых в литературе наряду с 18-ю разными общепринятыми в том же смысле определениями ИИ. В Британской энциклопедии [20], также цитируемой авторами статьи [19], справедливо отмечено: «большинство споров в данной области происходит из попыток дать точное определение интеллекта». На этом фоне неудивительно стремление к тривиальным («интеллект как мера способности агента достигать своих целей в широко варьируемой внешней обстановке» [19], «искусственный интеллект: способность компьютера или управляемого компьютером робота выполнять действия, характерные для разумных существ» [20]) и даже несколько пародийным характеристикам («ввиду всей этой сложности, мы используем простое определение коллективного интеллекта: такие совместные действия группы индивидуумов, которые выглядят разумными» [21]) в самых серьезных источниках. В большинстве используемых формулировок интеллект фигурирует как интуитивно понятная сущность, избранным аспектам которой посвящено конкретное исследование. Подобные разночтения вносят неопределенность в обсуждение интеллекта как явления – одновременно создавая некоторую свободу выбора терминов при его анализе.

Часто упоминаемые, но отнюдь не все качества человеческого интеллекта в его наиболее распространенных определениях включают: когнитивные возможности, целеполагание, обучаемость (в том числе способность обучаться на опыте), адаптацию к новым ситуациям, продуктивное использование абстрактных понятий, а также концепций и образов, рефлексии, применение знаний для целенаправленного воздействия на внешнюю среду, обмен информацией, использование речи. Помимо аморфности ряда терминов (знания, концепции, речь), в разных контекстах, тоже имеющих много разных смыслов, за их пределами остаются такие виды интеллекта, где агенты не используют абстрактных понятий и речи («интеллектуальные» способности животных), не склонны к рефлексии (общественные насекомые) либо не имеют когнитивных способностей в обычном понимании этих слов (формации роботов и «роевая робототехника», см. книгу [18]). В этой ситуации наиболее разумно начать с максимально широкого определения «*протоинтеллекта*», которому удовлетворяет большинство объектов, в специальной литературе (иногда метафорически) называемых интеллектуальными. Обсуждение конкретных видов РИ с необходимостью будет сопровождаться детализацией понятий и сужением области их применимости.

Необходимым, хотя и не достаточным условием наличия интеллекта будем называть **способность автономного агента стремиться к определенному состоянию (цели), воспринимая, обрабатывая и используя внешнюю информацию для ее достижения**. Столь широкому определению удовлетворяют в том числе хемотаксис бактерий и фототропия зеленых растений (в литературе действительно относимые к прототипам интеллектуальной деятельности [7, 11]), а также действия роботов [18], которые за пределами робототехники обычно не считают интеллектуальными. Несмотря на расхождения с более точными из многочисленных (и нередко конфликтующих) описаний интеллекта, такое рабочее определение соответствует большинству его известных реализаций. Это облегчает основную задачу обзора: сопоставление различных объектов и систем, в литературе относимых к интеллектуальным, и выявление их наиболее общих свойств.

Динамику автономных устройств, не обладающих интеллектом, но гибко выполняющих поставленную им задачу в изменяющихся внешних условиях, описывает *теория автоматического управления* (ТАУ) [14, 15, 22–24]. В ее рамках эволюция объекта управления во времени направляется физическими динамическими факторами и управляющими воздействиями, которые вырабатывает система автоматического управления (САУ) на основе наблюдаемых параметров объекта (быть может, зашумленных). Параметры, характеризующие состояние объекта в заданный момент времени, вместе с оценками их неточностей рассчитываются из доступной системы информации по рекурсивным алгоритмам; их эволюция в *пространстве представлений* (англ. *belief space*) создает приближенный «образ» фактической динамики в *пространстве состояний*, точно не известных управляющей системе<sup>4</sup>.

<sup>4</sup> Объективные параметры агента в ТАУ представляет  $n$ -мерный вектор состояния  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ , а доступную управляющей системе информацию об этих параметрах –  $m$ -мерный вектор наблюдения  $\mathbf{y} = (y_1, y_2, \dots, y_m)$ . Динамику агента в дискретном времени отражает система линейных уравнений

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \omega_k \quad (1, a)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \theta_k, \quad (1, б)$$

где  $\mathbf{x}(t) \in X$ ,  $X \subset \mathbb{R}^n$  – *пространство состояний* (*state space*);  $\mathbf{u}(t) \in U$  – вектор управления;  $U \subset \mathbb{R}^p$  – пространство управляющих воздействий (*control space*);  $\mathbf{y}(t) \in Y$ ,  $Y \subset \mathbb{R}^m$  – пространство наблюдений ( $p, m \leq n$ );  $\mathbf{A} \in \mathbb{R}^{n \times n}$  – динамическая матрица;  $\mathbf{B} \in \mathbb{R}^{n \times p}$  – матрица управления;  $\mathbf{C} \in \mathbb{R}^{m \times n}$  – матрица наблюдений;  $\omega(t)$  (возмущение системы) и  $\theta(t)$  (шум) – случайные ве-

Таким образом, САУ неодушевленного технического устройства уже включает *отражение действительности*: феномен, свойственный всем общепризнанным видам интеллекта. Такое отражение со статистическим «размыванием» оценок состояния системы в пространстве представлений получило дальнейшее развитие в *нечетких системах управления* [26], где образы внешнего окружения автомата непосредственно включают *свертывание информации* – еще один неотъемлемый атрибут «настоящего» интеллекта. Практическая реализация обоих фундаментальных принципов применяется в алгоритмах картирования обстановки (англ. *Simultaneous Location & Mapping (SLAM)* и др. подходы, см. книгу [18]), применяемых в робототехнике. Иными словами, некоторые важные признаки деятельности интеллекта на эмпирическом уровне проявляются уже у «протоинтеллектуальных» автоматических систем.

Теория автоматического управления действительно представляет собой завораживающий пример успешного применения теории динамических систем для постановки и решения задач управления самыми разными техническими системами. Одной из причин ее триумфальных успехов в середине и в конце XX в. является, наверно, то, что очень многие механические системы адекватно описываются дифференциальными или разностными уравнениями невысоких порядков. Начиная с 1960-х гг. многие исследователи пытались транслировать результаты ТАУ на экономические, социальные и/или живые системы. Успехи здесь, к сожалению, достаточно скромны. Ведь, казалось бы, сказав «пусть управляемая система является линейной дискретной динамической системой» (см. формулы (1, а, б)), можно применять к ней весь богатый математический аппарат ТАУ. Проблема заключается в обосновании корректности пресловутого «пусть»: доказательстве того, что такое описание моделируемой экономической, социальной и/или живой системы адекватно, что система идентифицируема и т. д. Многие исследователи о подобных обоснованиях «забывают», некоторым они не удаются, но есть и исключения (к сожалению, немногочисленные и немассовые).

По совокупности разнородных определений, базовыми качествами «настоящего», или «сильно-

го», интеллекта можно считать одновременное наличие следующих его свойств:

- восприятие («отражение»), обработка, использование и передача информации;
- автономность, в том числе целенаправленное поведение;
- извлечение, накопление и анализ знаний; обучаемость;
- абстрагирование, обобщение, генерация новых знаний;
- сознательное автономное целеполагание, самосознание, творчество, эмоции, рефлексия.

Согласно вышеизложенному два первых признака проявляются уже на уровне автоматов, цели которых запрограммированы (заданы экзогенно). На возрастающих уровнях когнитивных возможностей агентов их интеллектуальные свойства модифицируются и расширяются. В дополнение к перечисленным базовым признакам возникают *автономное целеполагание* и активная обучаемость (начиная с животных), мысленное отражение внешнего мира и схемы действий в нем, в психологии называемое когнитивной картой [27] (также начиная с животных, но больше относящееся к человеку), сознательное конструирование образов действительности и рефлексия (в основном характерные для человека, но также и для некоторых аспектов ИИ [28]) и, далее, высшие формы интеллектуального творчества. В настоящем обзоре будет проиллюстрировано разное «наполнение» перечисленных качеств на разных уровнях интеллекта – со всеми оговорками на некоторую неопределенность этого фундаментального понятия. У большинства проявлений индивидуального интеллекта на эмпирическом уровне обнаруживаются аналогии с реализациями ИИ, а также с РИ систем взаимосвязанных агентов, которые могут обладать весьма разными интеллектуальными способностями.

В своем анализе мы пользуемся «наивным», или интуитивным, представлением об информации как упорядоченном отражении внешних воздействий на индивидуального агента либо на мультиагентную систему. Отражение реальности как характеристика интеллекта не поддается точному отделению от обучаемости и способности создавать новое знание<sup>5</sup>. Два первых качества «абстрактного» интеллекта лежат в основе большинства реализаций ИИ. Однако эмпирически очевидная способность «придумать новое» математическими моделями непосредственно не воспроизводится, несмотря на уже построенное формальное

личины; подстрочные индексы  $k$  и  $k+1$  соответствуют последовательным моментам времени  $(t, t + \Delta t)$ . Оценки состояния  $\{\hat{y}_k\}$  и их дисперсии  $\{P_k\}$  образуют *пространство представлений (belief space)*  $G = \{b_k\} = \{\hat{y}_k, P_k\}$ , где, например,

$$b_0 = (\hat{y}_0, P_0), b_{k+1} = K(b_k, \mathbf{u}_k, \mathbf{y}_{k+1}), \mathbf{u}_{k+1} = f(b_k)$$

и  $K$  – фильтр Калмана; в простейшем одномерном случае  $\hat{y}_{k+1} = K y_{k+1} + (1 - K)(\hat{y}_k + u_k)$  [25].

<sup>5</sup> Здесь и далее понятие *знания* используется также в «бытовом» интуитивном смысле.

описание творческой человеческой деятельности (см., например, работу [29] и цитированную в ней литературу).

Характерной чертой интеллекта людей и элементов когнитивной деятельности животных, связанной с решением нестандартных задач, является *инсайт* («озарение») – спонтанный приход к правильному ответу в *проблемной ситуации*. Спонтанное решение проблемы, позволяющее миновать некоторые этапы последовательного логического вывода при обязательной концентрации индивида на поиске решения – твердо установленный результат психологических экспериментов [30, 31] с многочисленными эвристиками для прикладных изобретательских целей [32].

Проблема *свертывания, или концентрирования, информации* в процессе интеллектуальной деятельности, несмотря на успешное применение «нечетких» эвристик, еще не имеет общего формального описания. Многочисленные варианты однотипных внешних воздействий в задаче *распознавания образов* образуют множества, не поддающиеся анализу прямым перебором. Так, например, три параметра монохроматического звука (что уже упрощает частотный спектр звучания музыкальных инструментов), различимые лицами с музыкальным образованием – высота (пять октав, т. е. 60 вариантов ноты), громкость (шесть ступеней от *pp* до *ff*) и длительность (шесть вариантов от целой до  $1/32$ ) – вместе с тремя формами извлечения звуков (стаккато, нонлегато, легато) составляют  $60 \times 6 \times 6 \times 3 = 6480$  комбинаций для одной ноты. В этом случае короткому фрагменту музыкального текста из 100 нот отвечает количество вариантов, много большее  $10^{100}$ , или числа *googol*. Свертывание информации здесь, как и в распознавании текстов и зрительных образов, достигается при обучении индивидуума, в результате которого множество комбинаторно возможных сочетаний кардинально сокращается и преобразуется в относительно небольшой набор блоков – например, мелодий, гармоний или слов естественного языка.

Индивидуальный интеллект человека составляет предмет большинства исследований в области когнитивных наук. Именно для него установлены перечисленные выше эмпирические характеристики, которые с середины XX в. служат основой для разработок искусственного интеллекта. В работах биологов были найдены многочисленные параллели когнитивных способностей животных с интеллектом человека, для некоторых видов включающие способность к абстракции и счету, использование языков-посредников, формирование образов действительности в сознании (со всеми оговорками об условности этого термина, применяемого не

к сознанию людей) и инсайт как основной механизм выработки решений в проблемной ситуации [33]. Сообщества животных также обладают социальной структурой (иногда весьма сложной [34]) и способны к кооперативным действиям. Тем не менее, РИ структурированных систем будет далее обсуждаться на примерах ОС в человеческом социуме – лучше изученных, актуальных и отличающихся значительно более разнообразным интеллектуальным поведением.

Необходимо отметить, что индивидуальный человеческий интеллект не может сформироваться и существовать вне социальной среды, где он является неотъемлемой частью коллективного интеллекта. Также в современной интеллектуальной деятельности – как индивидуальной, так и коллективной – в возрастающей степени используются компьютеры, роботы, системы технического зрения и другие элементы ИИ. Таким образом, все существующие формы интеллекта неразрывно связаны и входят в коллективный РИ человеческого социума как его составные части. Технические «протоинтеллектуальные» системы и компьютерные реализации ИИ (тоже не существующие вне социума) вносят существенный вклад в структуру РИ современных социальных систем и по этой причине должны быть кратко охарактеризованы.

## 2. МОДЕЛИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Искусственный интеллект как научное направление в первом приближении подразделяют на вычислительный и логический. Его наиболее распространенные реализации, не исчерпывающие всех видов, представлены на рис. 2. Как указано в докладе Стэнфордского университета (США) Artificial Intelligence Index Report 2022, общее ежегодное число публикаций по тематике ИИ в мире увеличилось со 162 тыс. в 2010 г. до 334 тыс. в 2021 г. [35<sup>6</sup>]. Не пытаясь анализировать такой огромный объем информации, перечислим здесь некоторые направления моделирования интеллекта, наиболее близкие к задачам настоящего обзора.

### 2.1. Искусственные нейронные сети

Самое массовое (и модное) на сегодня техническое средство, позволяющее воспроизвести некоторые аспекты интеллектуальной деятельности –

<sup>6</sup> См. также Каспарьянц Д. Обзор доклада Стэнфордского университета «Индекс искусственного интеллекта 2022» [https://rdc.grfc.ru/2022/05/artificial\\_intelligence\\_index\\_report\\_2022/](https://rdc.grfc.ru/2022/05/artificial_intelligence_index_report_2022/) (дата обращения 15.02.2023).



Рис. 2. Основные средства математического и компьютерного моделирования интеллекта

искусственная нейронная сеть (ИНС) [36]. Функционирование ИНС включает *отражение* воспринимаемого «входа» (изображений, текстов, аудиосигналов и др.) и *свертывание* информации в цифровые массивы в процессе обработки – т. е. фундаментальные признаки интеллекта, реализуемые уже на уровне САУ и роботов (см. выше). Новым качеством ИИ здесь является *обучаемость* – получение корректного выхода на заведомо известных типах входящего сигнала<sup>7</sup>. Цель обработки ин-

формации искусственными нейронными сетями по-прежнему задается извне.

Одно из главных направлений исследований в области ИИ составляют задачи *глубокого обучения* (*deep learning*) ИНС и других компьютерных архитектур с последовательным свертыванием информации. Достигнутые результаты в этой области включают распознавание лиц и объектов по данным видеонаблюдений («компьютерное зрение»), оцифровку рукописей и устной речи, составление логически связанных текстов, машинный перевод, автоматическое создание программных кодов и многие другие достижения. Однако компьютерные

<sup>7</sup> В схеме ИНС, представляющей собой упрощенную имитацию сетевой структуры нейронов мозга, входящий сигнал (например, изображение с разверткой по пикселям) передается элементами входного слоя к связанным с ними элементам следующего (*скрытого*) слоя, которые, в свою очередь, в зависимости от строения ИНС, транслируют сигнал на выходной элемент либо на последующий промежуточный слой, а в *рекурсивных* сетях и на предыдущие слои. Активационные функции узлов, или *фильтры*, которые определяют их способность передавать входной сигнал следующему узлу, могут иметь разнообразную ступенчатую форму. Результат анализа выводится в форме *n*-мерного вектора  $(p_1, p_2, \dots, p_n)$ , компонентами которого служат вероятности  $\{p_i\}$  отнесения введенного массива информации к одному из *n* заданных типов ( $\sum p_i = 1$ ). При *обучении* ИНС, по аналогии с функционированием сети нейронов мозга, изменяют «проводимости» связей между узлами, добиваясь корректной идентификации заведомо известных типов входящего сигнала. Далее обучающая выборка массивов входной информации заменяется на анализируемый массив.

В активно развиваемых многослойных *сверточных* ИНС сигналы, входящие в *сверточный слой* искусственных нейронов и

выходящие из него в следующий слой, а также наборы функций фильтров представляются в виде упорядоченных массивов чисел, в области ИТ обычно называемых тензорами. Так, изображение размером  $N_1 \times N_2$  пикселей, имеющее *C* цветовых каналов, задается массивом **X** размера  $N_1^{(1)} \times N_2^{(1)} \times C$ , компонентами которого  $\{X_{ijk}\}$  служат нормированные интенсивности всех цветовых каналов для каждого пикселя, а набор *S* фильтров с размерами  $m \times m$  и тем же числом цветовых каналов – массивом размера  $m \times m \times C \times S$  с компонентами  $\{F_{ijkp}\}$ . Результаты преобразования входного массива данных

$$Y(p, q, s) = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^C F(i, j, k, s) \times X(p+i-1, q+j-1, k)$$

(индексы компонент для упрощения записи приведены в скобках) образуют тензор **Y** размера  $(N_1 - m + 1) \times (N_2 - m + 1) \times S$  с меньшим числом компонент, который служит входным массивом информации для следующего слоя.

средства моделирования ИНС до сих пор не имеют адекватного общенаучного описания («общей теории ИИ»), а свернутые массивы информации на промежуточных слоях уязвимы к малым помехам, включая целенаправленные атаки [37]. В этом смысле нейронные сети повторяют судьбу многих фундаментальных технических достижений человечества (паровая машина, электронные схемы, летательные аппараты и др.), строгая количественная теория которых была разработана лишь после их продолжительного применения. Вопросы строения и функций современных ИНС обсуждаются в обзоре [38].

## 2.2. ИИ на основе математической логики

Особое направление когнитивных наук посвящено моделированию интеллекта методами математической логики [39]. В этом подходе внешняя информация воспринимается компьютерной программой в форме *высказываний*, или логических утверждений, которые могут быть истинными или ложными. Высказыванием устанавливаются связи между «сущностями» (объектами, так или иначе отражающими реальность) по правилам естественного языка. Логические операции позволяют производить *исчисление высказываний*, образующих бесконечное счетное множество, определять их истинность или ложность на основе заданного набора аксиом. Этим открывается путь к выводу новых содержательных высказываний на основе истинных высказываний – в частности, к «автоматическому» доказательству математических теорем. Важнейшим компонентом данного подхода является *логическое программирование*.

Детальный обзор средств реализации ИИ, в том числе различных видов логики, представлен в книге [40]. Методы математической логики внедрены в многообразные компьютерные средства, облегчающие пользователям решение интеллектуальных задач при обработке больших слабо упорядоченных массивов информации в *базах знаний* (как структурированных наборах установленных фактов), информационных, экспертных и иных «интеллектуальных» системах. Подходы, разработанные в этой области, включают *объектно-реляционные модели* (англ. *Entity-Relationship*, ER) и концептуальный дизайн [41], *семантические сети* [42] и другие методы обработки слабо структурированных множеств, а также *нечеткую логику* (см. Part B в книге [40]) и современное перспек-

тивное направление *анализа формальных понятий* (АФП; англ. *Formal Concept Analysis*, FCA [43]).

Реализации автономного обучаемого ИИ, решающего творческие задачи полностью на основе математической логики, в литературе не описаны. Более всего логические подходы применяются в прикладных задачах формирования и использования знаний на основе компьютеров и математических алгоритмов, в том числе в сочетании с ИНС (см. обзор [38]). Помимо распознавания образов, автоматического доказательства теорем, методов порождения знаний [44] и создания программ сверхчеловеческого уровня в стратегических играх (шахматы, го)<sup>8</sup>, важными задачами здесь являются представление, извлечение и обработка информации, выраженной на естественном языке (англ. *Natural Language Processing*, NLP). Наибольшие успехи достигнуты в создании комбинированных интеллектуальных систем «человек + компьютер», где реализуется диалог пользователя с программой (ввод и вывод информации) на естественном языке с простой графикой, но эта перспективная область не является предметом настоящего обзора. Концептуальные проблемы искусственного интеллекта обсуждаются в книге [45].

Не оспаривая достоинств использования математической логики в «интеллектуальных» системах, отметим, что такие системы не могут генерировать новую нетривиальную информацию иначе как в форме последовательного вывода из заданного набора аксиом и объема знаний. В отличие от ИНС, «логические» реализации ИИ способны создавать новые знания, проявляя еще один фундаментальный признак интеллекта (см. выше), но математический вывод не воспроизводит инсайта – существеннейшего механизма когнитивной деятельности людей. Кроме того, логический ИИ, в отличие от ИНС, не содержит схем прямого свертывания информации и, аналогично ИНС, имеет лишь заданное извне целеполагание. (Конструирование онтологий и их использование в выявлении знаний [1, 43, 44], а также разработка самообучающихся компьютерных программ, потенциально

<sup>8</sup> По справедливому замечанию О.П. Кузнецова, сам факт равных состязаний гроссмейстера с компьютером, в миллионы раз превосходящим человека по своим вычислительным возможностям, свидетельствует о принципиальных преимуществах системы обработки информации в человеческом мозге. На формализацию деятельности человеческого интеллекта нацелены многие современные работы в области ИИ (см. книгу [45]).

способные ликвидировать наиболее существенные отличия ИИ от интеллекта человека, пока остаются областью интенсивного поиска.)

Другим фактором, ограничивающим применение «автоматической» логики, является неполнота и нередко противоречивость данных в базах знаний – обстоятельство, легко преодолеваемое человеческим мышлением (см. книгу [45]). Наиболее крупные достижения здесь тоже связаны с использованием логических и семантических соотношений в комбинации «человек + компьютер», где на каждом последующем шаге решение принимает человек: интеллектуальные системы поддержки принятия решений (СППР) на основе подбора связанных терминов в построении производственных, экономических и логистических схем (когнитивных карт в их математическом понимании [46]). На этой основе разрабатываются гибридные архитектуры систем человеческого интеллекта, поддержанного ИИ (см. часть Human-Computer Interaction в книге [21]), также не обсуждаемые в нашей работе.

Принципиальной и еще малоисследованной особенностью интеллекта человека является *прерывистый* и асинхронный тип решения задач, отличный от существующих континуальных методов его компьютерного моделирования. Эта особенность проявляется как алгоритм *прерывистого управления* (*intermittent control*) в широком спектре деятельности людей: управлении транспортными средствами, работе операторов, стратегиях биржевых игроков и многих других видах поведения [47]. Для всех них характерно непрерывное наблюдение обстановки в сочетании с эпизодическими действиями, корректирующими ее динамику; в ряде случаев такая комбинация похожа на стандартное непрерывное управление, но не эквивалентна ему. Прерывистое управление, по видимому, является эволюционно отобранным средством экономии ресурсов мозга: оно также наблюдается у животных и, вероятно, тесно связано с явлением инсайта. Его элементы проявляются в РИ биологических и социальных систем.

### 3. ПРОСТЕЙШИЕ ФОРМЫ РАСПРЕДЕЛЕННОГО ИНТЕЛЛЕКТА

Как можно было видеть выше, фундаментальные качества индивидуального человеческого интеллекта – оперирование образами реальности, запись и свертывание внешней информации, целепо-

лагание, обучаемость и «придумывание нового» – частично проявляются уже в «протоинтеллектуальных» системах автоматического управления, а в более полном виде – в компьютерных реализациях ИИ. Ниже будет показано, что различные виды распределенного интеллекта в мультиагентных системах проявляют аналогичные свойства. Механизмы их реализации в разных видах РИ также обнаруживают большое сходство. Это позволяет предположить, что «интеллектуальный» характер системы определяется некоторым фиксированным и общим набором условий. В соответствии с большинством определений, признаком распределенного (для групп людей – *коллективного*) интеллекта будем считать *свойство эмерджентности* – способность системы взаимодействующих агентов справляться с задачами, превышающими возможности отдельных агентов [21].

Под *мультиагентной социальной системой* (МСС) будем понимать динамическую совокупность автономных агентов, которые взаимодействуют с внешней средой и друг с другом. Каждый индивидуальный агент воспринимает и обрабатывает информацию, поступающую из окружающей среды, и использует ее для достижения некоторой *цели* – оптимального результата своей эволюции во времени. Взаимодействующие биологические существа одного вида (муравьи, пчелы, приматы, люди) образуют социальную систему в узком смысле слова. «Неживые» программируемые агенты (роботы, беспилотные аппараты и др., см. работы [7, 8]) при наличии взаимодействия между ними образуют *искусственную* МСС.

Целеполагание агентов в большинстве систем следует из их биологической или социальной природы, но для системы взаимодействующих технических устройств (или, например, армейского подразделения) цель задается извне. Индивидуальные агенты составляют МСС *1-го уровня*. Взаимодействующие МСС (экономические субъекты, организационные системы, политические партии и т. д.) сами могут выступать в роли агентов в социальных системах более высокого уровня. Общая классификация систем, образованных при взаимодействиях людей, и производимой ими деятельности представлена в книге [48].

Моделированию динамики МСС на основе теории игр [28, 49], а также методами, перенесенными из статистической физики ([7, 9, 50, 51], см. рис. 3), посвящена обширная литература. Динамику ряда систем (транспортные потоки [52], некото-

рые экономические процессы [53], биржи [54], влияние [50, 55] и распространение мнений [56] в социальной среде) на коротких горизонтах прогнозирования удастся предсказывать с помощью математических моделей. В таких моделях обработку и использование информации агентами обычно учитывают косвенно в форме квазифизических потенциалов отталкивания движущихся агентов от препятствий и их притяжения к целям [51, 52], стремления к максимуму полезности [53, 55], кластеризации единомышленников и игнорирования пользователей социальной сети с противоположным мнением [56] и других изначально заданных факторов, направляющих коллективную динамику.



Рис. 3. Описание мультиагентной системы методами междисциплинарной физики

Однако в общем случае описание социума осложняют нечетко определенные параметры, неоднозначные зависимости, многоуровневая рефлексия [28] и, с другой стороны, ограниченно рациональное поведение агентов [57, 58], а также ряд других обстоятельств, предпочтительно рассматриваемых на вербальном уровне в гуманитарных

науках. Это затрудняет расчеты, которые формально становятся возможными после введения квазифизических «сил». В целом агентные модели, основанные на балансе количественно либо качественно оцениваемых факторов (выигрышей, влияния, интересов) и поиске максимума энергоподобной целевой функции при воздействии случайных помех (шума), обычно дают лишь качественное описание социальных систем (см. рис. 3).

Мы полагаем, что фундаментальной характеристикой МСС, адекватно не отраженной в существующих моделях, является именно *распределенный интеллект* – т. е. способность взаимосвязанных агентов кооперативно воспринимать, обрабатывать и использовать информацию [59]. Динамика системы не сводится к простой сумме эволюций ее агентов; в частности, она включает системное целеполагание. *Объективной целью* МСС на заданном интервале времени является достижение состояния, оптимального для системы как целого. Формальный критерий оптимальности (утилитарный, эгалитарный либо иной [53], оптимум Парето или равновесие Нэша и т. д.) часто зависит от предпочтений исследователей. Тем не менее, на эмпирическом уровне трудно отрицать, что разнообразные системы, в том числе не предполагающие сложной интеллектуальной деятельности агентов (пчелиный рой, муравейник, рынок) воспроизводимо и гибко преследуют системные цели<sup>9</sup>. Существенно, что наличие индивидуального сознания у людей не является необходимым условием существования РИ: системы, сильно различающиеся по уровню когнитивных возможностей агентов, могут проявлять сходную динамику.

Распределенный интеллект, направляющий эволюцию системы, следует непосредственно учитывать в описании социальных процессов. Многообразные виды РИ, включая распределенный ИИ технических систем, схемы взаимодействия «человек – компьютер», биологические МСС и разнообразные аспекты человеческой деятельности обсуждаются в монографии [21]. Задача настоящего обзора – проанализировать условия проявлений РИ в динамике мультиагентных систем (в том числе состоящих из агентов с малыми либо нулевыми когнитивными возможностями) и на этой основе

<sup>9</sup> В большинстве случаев в объективные цели входит самосохранение системы: сохранение либо увеличение числа агентов и поддержка их функционирования.

сформулировать общие принципы функционирования «интеллектуального агента».

### 3.1. Роевой интеллект

Многие автономные агенты, формирующие систему, могут самостоятельно перемещаться («моторные», «самодвижущиеся» либо «живые» частицы). При превышении некоторого порогового количества таких частиц в единице объема уже у жгутиковых бактерий возникают корреляции перемещений. Они приводят к изменению режима коллективного движения частиц от неупорядоченного («рой») к вихревому, или «бактериальной турбулентности» [60]. Данный термин, возникший в литературе в начале 2000-х гг., не относится к турбулентному движению жидкости, реализуемому при достаточно высоком числе Рейнольдса  $Re \sim 10^3$ , поскольку размерам и скоростям жгутиковых бактерий (измеряемым соответственно в мкм и десятках мкм/с) отвечают значения  $Re \sim 10^{-3} - 10^{-4}$ . В обычном макроскопическом масштабе расстояний и скоростей этому соответствует перемещение безмассовых частиц в очень вязкой среде [61]. Однако именно благодаря гидродинамическому взаимодействию бактерии разворачиваются и увлекаются за соседними движущимися частицами, образуя «бактериальные вихри» (рис. 4).

Корреляция перемещений при достаточно высокой плотности агентов и низком уровне шума

реализуется во всех системах «самодвижущихся» частиц [7]. Она также наблюдается в «неживых» системах из анизотропных коллоидных частиц и в гибридных материалах – лиофильных жидких кристаллах с добавленными к ним живыми бактериями (см. статью [61]).

У биологических МСС особей, способных к активному восприятию информации (насекомых, рыб и птиц), существуют «библиотеки» разных режимов кооперативной динамики, активируемых внешними воздействиями (например, взлет стаи при возникновении опасности). Переключения системы с одного режима на другой в ответ на изменения внешних условий служат средством приспособления, способствующим выживанию вида [7, 11]. Это явление, в когнитивных науках называемое *роевым интеллектом* (*swarm intelligence*), имеет многочисленные аналогии среди других типов РИ мультиагентных систем [11].

Начиная с 1990-х гг. различные виды кооперативного движения «живых частиц» воспроизводились в простых мультиагентных моделях (см. обзор [7] и цитированную в нем литературу). В часто используемой компьютерной модели И. Кузена и соавторов [6] коллективная динамика системы агентов определяется соотношением радиусов области отталкивания от ближайших соседей  $R_1$ , области корреляции перемещений  $R_2$  и области притяжения к центру группы  $R_3$ . В зависимости от плотности агентов, их скоростей и уровня шума

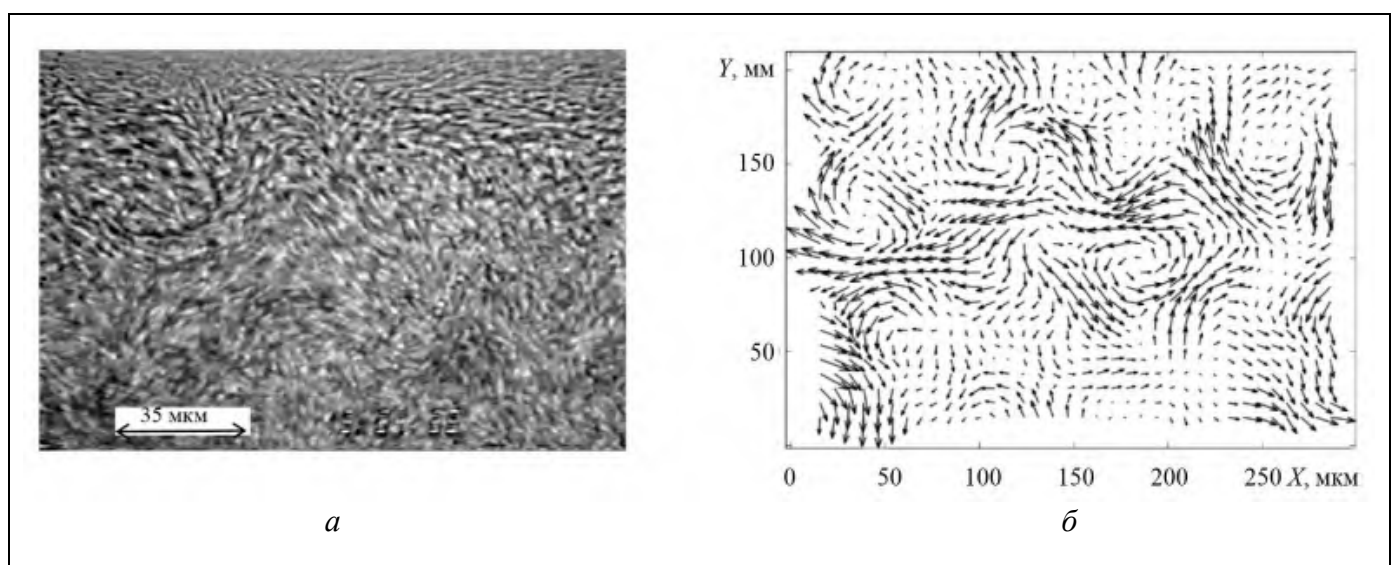


Рис. 4. «Вихри» *Bacillus subtilis*: а – фото, б – распределение скоростей [60]

реализуются режимы беспорядочных смещений внутри группы (собственно *рой*), вихревого движения (*тор*) и согласованного перемещения агентов (*стая*). При изменении радиуса корреляции движения наблюдался гистерезис (рис. 5). В реальной динамике косяка рыб (*fish school*) на корреляции движения влияет гидродинамическое взаимодействие с соседними особями [62]. В больших стаях птиц, где вихревые структуры не образуются, богатую коллективную динамику создает механизм *распределенных переменных лидеров*, также воспроизведенный в агентных моделях: при достаточном числе особей, одинаково изменяющих направление полета, по ним корректирует свое движение вся стая (рис. 6).

Коллективное движение биологических МСС всегда включает элементы хаотического поведения, особенно сильно выраженные у насекомых. Неупорядоченное расположение и случайные перемещения агентов являются приспособительным фактором, облегчающим распознавание опасности и источников питания. В частности, благодаря хаотическому круговому обзору и корреляциям сме-

щений рой «живых частиц» осуществляет коллективное наблюдение далеко за пределами возможностей составляющих его особей [4–6, 11]. В моделировании динамики косяка рыб при атаке хищника (вызывавшего реакцию избегания), несмотря на полную параметризованность агентов, были воспроизводимо зафиксированы сложные изменения структур для разных стратегий атаки (рис. 7).

Случайные изменения направления и скорости также характерны для птиц в стае; при одинаковых отклонениях нескольких особей от общего курса они могут стать временными лидерами, влияющими на общую траекторию полета [63]. Этим достигается гибкая реакция системы в изменчивой окружающей среде, которая не сводится к комбинациям запрограммированных действий. Механизмы проявления роевого интеллекта – корреляция перемещений, флуктуации положений и скоростей, инициирование коллективной динамики «локальным большинством» однотипно движущихся особей – не имеют общего теоретического описания, однако успешно воспроизводятся в простых агентных моделях.

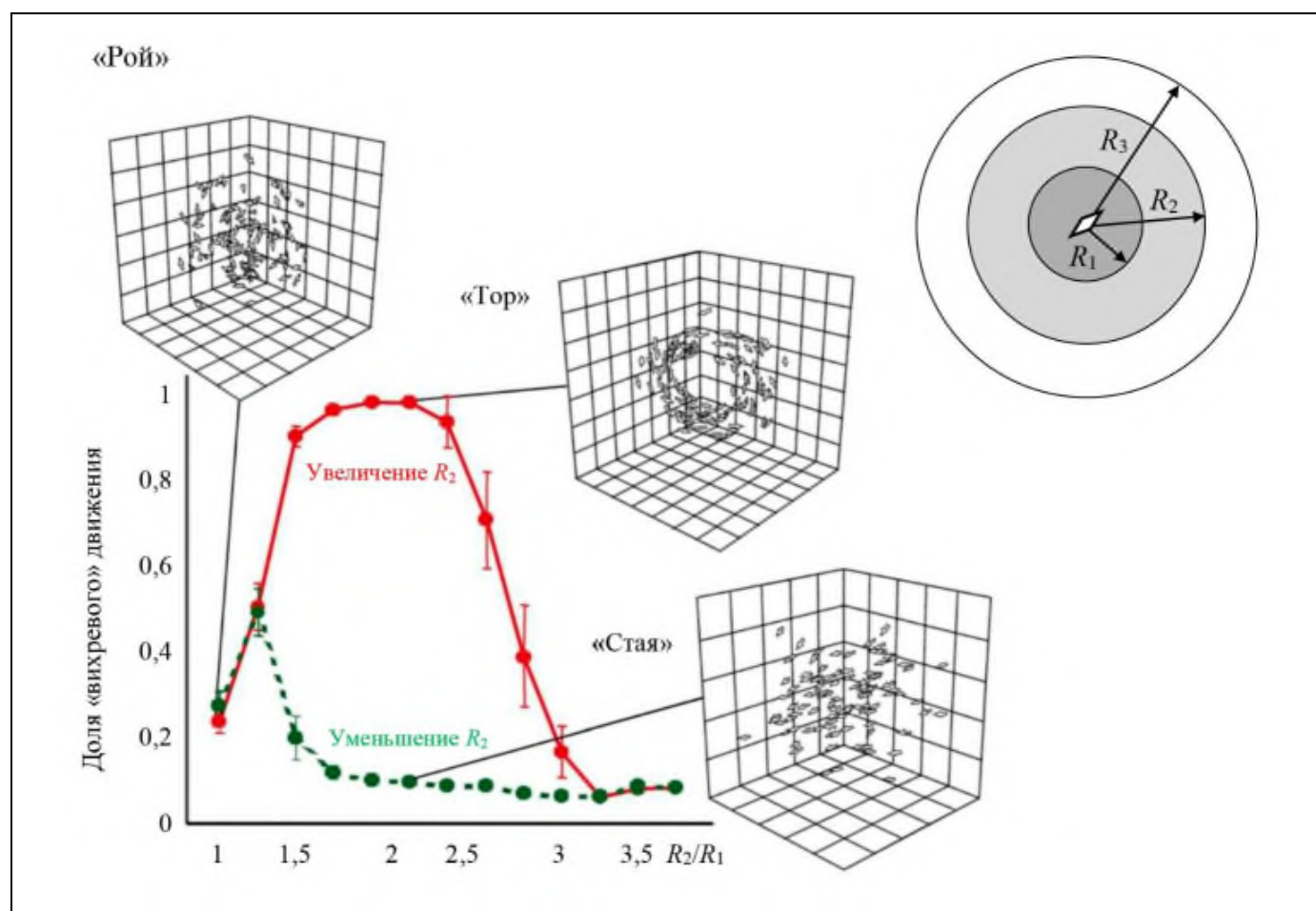


Рис. 5. Режимы коллективного перемещения агентов в зависимости от соотношения радиусов отталкивания  $R_1$  и корреляции движения  $R_2$  [6].

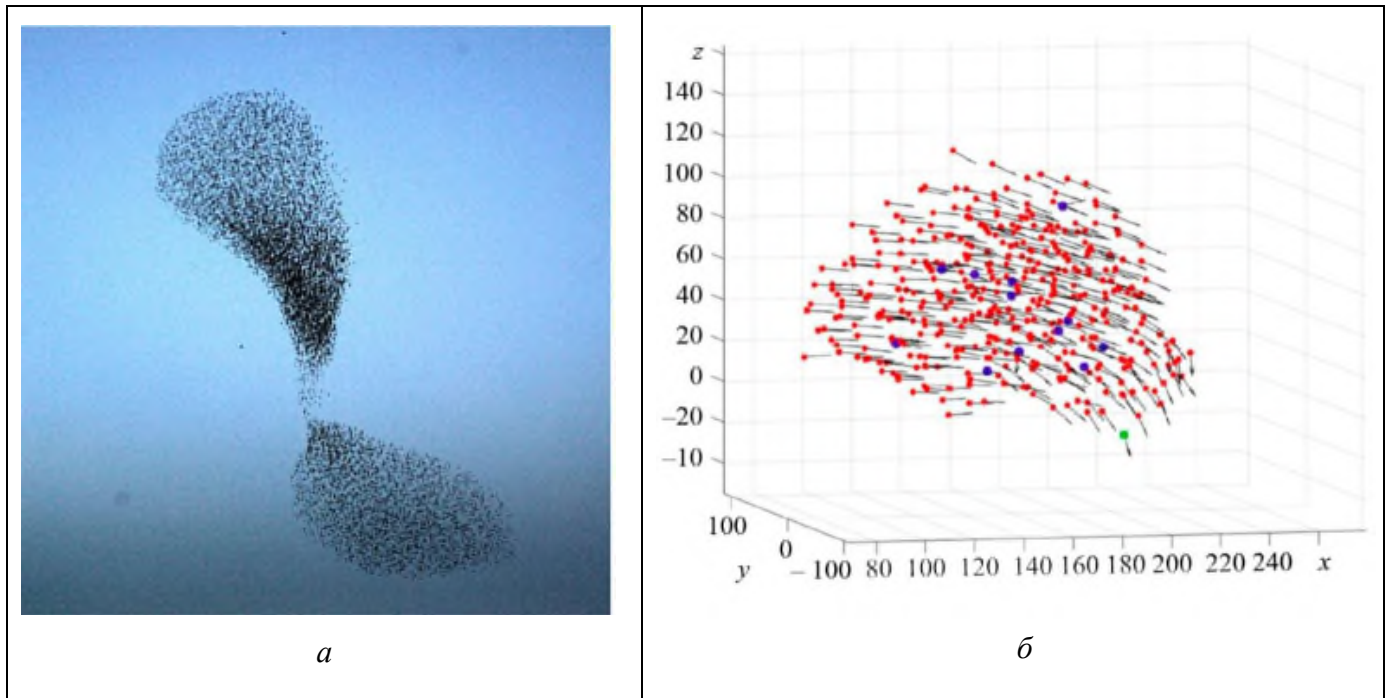


Рис. 6. Коллективное перемещение птиц: *a* – стая скворцов в воздухе, *б* – перемещения агентов в модели, синий цвет – временные лидеры [63]

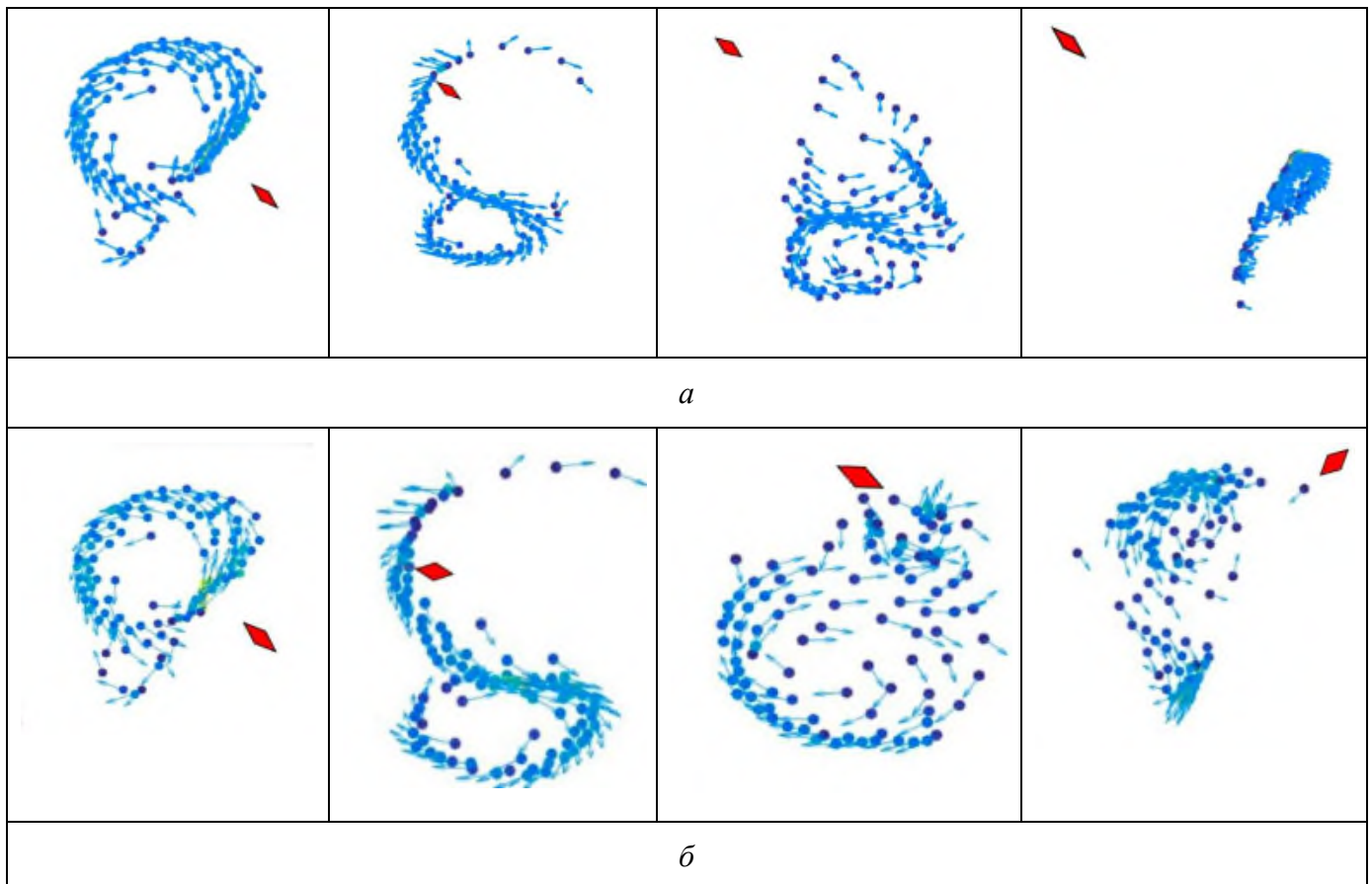


Рис. 7. Перестройка «тора» движущихся агентов при разных стратегиях «хищника» (закрашенный ромб): *a* – атака по прямой линии, *б* – атака на ближайшую особь (по изложенному в работе [62])

Случайные изменения направления и скорости также характерны для птиц в стае; при одинаковых отклонениях нескольких особей от общего курса они могут стать временными лидерами, влияющими на общую траекторию полета [63]. Этим достигается гибкая реакция системы в изменчивой окружающей среде, которая не сводится к комбинациям запрограммированных действий. Механизмы проявления роевого интеллекта – корреляция перемещений, флуктуации положений и скоростей, инициирование коллективной динамики «локальным большинством» однотипно движущихся особей – не имеют общего теоретического описания, однако успешно воспроизводятся в простых агентных моделях. Особенностью общественных насекомых является «запись» информации на окружающем ландшафте с использованием *феромонов*: химических аттрактантов, распознаваемых другими особями внутри системы<sup>10</sup>. Муравьиные тропинки, маркированные феромонами, по механизму положительной обратной связи выделяют оптимальные пути к источникам пищи и другим объектам жизнедеятельности муравейника (рис. 8, а) [7, 11]<sup>11</sup>. Сложная система химических регуляторов поведения насекомых формирует иерархическую систему муравейников, термитников, пчелиных ульев [4]. Хаотические действия особей при строительстве и заботе о потомстве благодаря взаимодействиям внутри системы синхронизируются, порождая циклы (рис. 8, б). Кооперативные действия в максимуме циклов – еще один отобранный эволюцией приспособительный механизм концентрации усилий колонии, способствующий выживанию в неблагоприятной среде [6].

### 3.2. Роевой интеллект в робототехнике

Кооперативные эффекты и динамические структуры также возникают в искусственных МСС из автономных взаимосвязанных аппаратов, способных воспринимать положения соседних устройств и избегать столкновений при перемещении к заданной цели [8, 9]. Группы роботов могут решать общую, поставленную извне задачу с коллективной обработкой информации по механизму

<sup>10</sup> В более простом примере атака осами человека, приблизившегося к гнезду, также инициируется их химическими выделениями, повышающими агрессивность [4].

<sup>11</sup> По аналогичному механизму (с заменой обоняния зрением) возникают пешеходные тропинки на снегу и на траве (см. работу [64]).

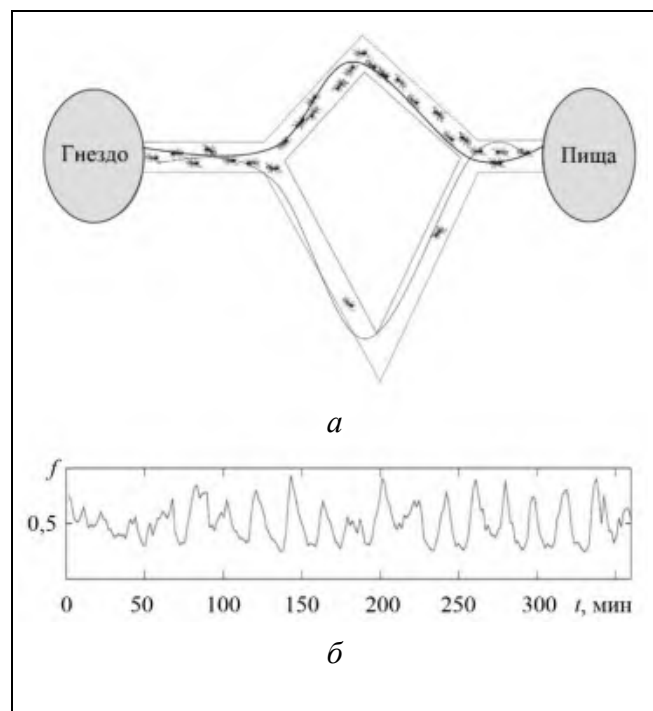


Рис. 8. Коллективная динамика общественных насекомых: а – выбор муравьями короткого пути к источнику пищи [11]; б – циклы активности ( $f$  – доля активных особей) при строительстве гнезда [6]

распределенного (а не централизованного) управления. «Роевому интеллекту» систем из относительно простых и недорогих взаимосвязанных технических устройств (*swarm robotics*), во многих отношениях имитирующему РИ общественных насекомых, доступен ряд практически важных операций:

- образование геометрической формы (*строя*) и коллективное движение;
- совместный сбор и рассредоточение по области пространства без потери связи;
- разделение функций;
- поиск, транспортировка и совместное перемещение объектов;
- коллективное позиционирование и картография.

В формациях роботов достигается «робастность» коллективных действий, т. е. взаимозаменяемость устройств и нечувствительность системы к поломкам некоторого числа единиц [18, 65]. Характерным примером практических приложений может служить коллективная навигация группы летающих дронов [9]. В этом направлении робототехники уже возникла терминология «социального сообщества роботов» [11, 66], хотя его современные задачи пока в основном сводятся к оптимизации совместного движения.

### 3.3. Имитация роевого интеллекта в компьютерных алгоритмах

Идея роевого интеллекта используется в компьютерных алгоритмах поиска и оптимизации, где возросшие вычислительные мощности изменили как технику, так и стратегию расчетов. В последние десятилетия в области информационных технологий был разработан ряд *алгоритмов, подсказанных Природой* (англ. *Nature-Inspired Metaheuristics*, NIMs), в которых имитация «интеллектуального роя» позволяет оптимизировать решение трудных вычислительных задач<sup>12</sup> [40, 67, 68]. Круг

<sup>12</sup> Так, в *муравьином* алгоритме [67], впервые предложенном в 1992 г. для поиска оптимального пути на сложных графах (*задача коммивояжера*), вероятность перехода изображающей точки между  $i$ -й и  $j$ -й вершинами

$$P_{ij}(t) = \frac{\tau_{ij}(t)^\alpha d_{ij}(t)^{-\beta}}{\sum \tau_{ij}(t)^\alpha d_{ij}(t)^{-\beta}}$$

определяется длиной пути  $d_{ij}$  между вершинами и «уровнем феромона»  $\tau_{ij}$  на этом пути (где  $\alpha$  и  $\beta$  – эмпирические параметры, а сумма в знаменателе отвечает всем возможным маршрутам из  $i$  в  $j$ ). В общем случае множитель  $d_{ij}^{-\beta}$  заменяется *эвристической ценностью*  $C_{ij}^{(k)}$ , которая вычисляется алгоритмически для каждого частного решения (маршрута). «Уровень феромона» уменьшается на каждом шаге дискретного времени  $t$  и увеличивается при прохождении маршрута новыми агентами:

$$\tau_{ij}(t+1) = (1-\varepsilon)\tau_{ij}(t) \sum_k \frac{\gamma}{C_{ij}^{(k)}}.$$

Здесь  $\varepsilon$  – «коэффициент испарения»;  $\gamma$  – ценность наилучшего решения; суммирование проводится по всем найденным маршрутам  $i \rightarrow j$ .

Имитируя положительную обратную связь привлекательности реальных муравьиных тропинок и числа прошедших по ним насекомых, муравьиный алгоритм позволяет экономно находить субоптимальные решения задач, в которых затраты вычислительных ресурсов в случае прямого перебора вариантов возрастают факториально. В альтернативном *пчелином* алгоритме каждый виртуальный агент проводит случайный поиск лучшего решения (например, максимума целевой функции  $U(p_1, p_2, \dots, p_n)$ ) по связной траектории в пространстве параметров  $\{p_i\}$ , учитывая общее знание наилучшей точки  $\mathbf{p}^*(t)$  на каждом шаге времени – «информационную панель» (*black-board*) [67]. В общей схеме роевых алгоритмов поиска оптимума положение и скорость виртуальной частицы-агента в  $n$ -мерном пространстве параметров в момент времени  $t + \Delta t$

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t)\Delta t, \\ v_i(t+\Delta t) &= \omega v_i(t) + a_1 r_1 [l_i(t) - x_i(t)]\Delta t + \\ &+ a_2 r_2 [g(t) - x_i(t)]\Delta t, \end{aligned} \quad (2)$$

рассчитываются исходя из положения и скорости в предыдущий момент времени с учетом положения локального оптимума  $l_i(t)$  в радиусе обзора агента и глобального оптимума  $g(t)$  по данным всех агентов с эмпирическими параметрами  $\omega$  (инерция),  $R$  (радиус обзора),  $a_1$  и  $a_2$  (коэффициенты ускорения к оптимумам) и случайными числами  $r_1, r_2 \in [0, 1]$ . В данной модификации классического метода молекулярной динамики множители  $a_1 r_1 [l_i(t) - x_i(t)]$  и  $a_2 r_2 [g(t) - x_i(t)]$  при приращении времени  $\Delta t$  в формулах (2) называются соответственно *когнитивным фактором* и *социальным фактором* [69].

таких вычислительных алгоритмов значительно шире эвристик, имитирующих рой живых организмов (включая синхронизацию вспышек свечения светлячков). В частности, к области NIM относятся схемы, воспроизводящие социальное поведение высокоорганизованных животных (волков, львов, обезьян [70]), а также *генетические* и *культуральные* алгоритмы [68]. В последних типах расчетов происходят изменения популяции виртуальных агентов, наследующих адаптационные признаки с варьированием в результате случайных «мутаций» (генетический алгоритм) либо «обучения» (культуральный алгоритм). Многообразие методов *естественных вычислений* рассмотрено в обзоре [71].

Эмерджентные эффекты, сопутствующие достижению оптимума целевой функции в расчетах по генетическим алгоритмам, позволяют по-новому взглянуть на содержание биологической эволюции. Так, например, в работе [59] целевая функция простой эволюционной модели – время прохождения группой агентов коридора с препятствиями – минимизировалась в последовательных циклах путем замены «медленных» 50 % агентов копиями «быстрых» с флуктуациями их параметров, которыми задавались ускорение при движении слева направо, отталкивание от препятствий и следование за соседними агентами в радиусе обзора. В результате обучения время прохождения коридора последним агентом в группе уменьшалось на 15–16 % (рис. 9, а), а их хаотический дрейф с многочисленными столкновениями сменялся согласованным движением в сгустках, не заданным в исходной модели (рис. 9, б). Подобные результаты указывают на аналогии процессов эволюции, в которых оптимизируется приспособленность биологических видов, с развитием их «системного РИ», давно обсуждаемого на качественном уровне [72].

Одним из направлений современных компьютерных исследований является искусственная эволюция роботов и их управляющих программ с произвольно задаваемыми приспособительными признаками [73]. Следует подчеркнуть, что алгоритмы NIMs являются эвристическими, т. е. не только не гарантируют нахождение глобального оптимума, но и, как правило, не позволяют получить оценок того, насколько далеко от последнего находится найденное ими решение.

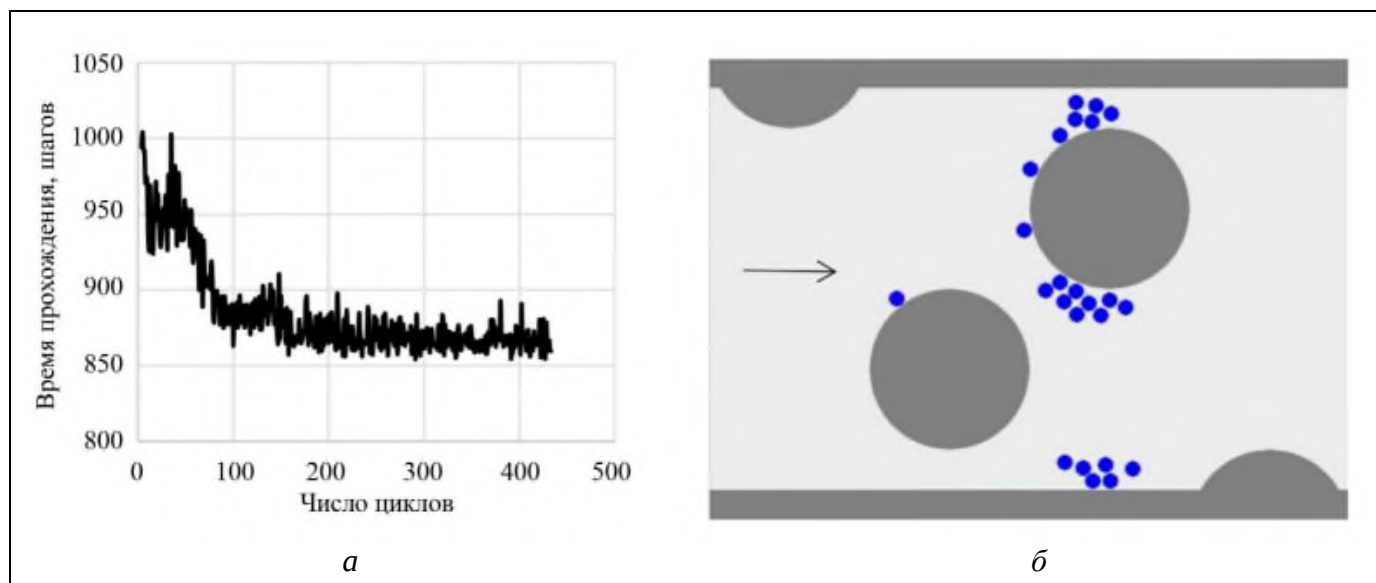


Рис. 9. Прохождение агентами коридора с препятствием в генетическом алгоритме отбора: а – время прохождения (число шагов) в последовательных циклах, б – движение агентов в «обученной» системе (стрелка – направление движения) [59]

## ЗАКЛЮЧЕНИЕ

Проведенный анализ позволяет утверждать, что ряд основных качеств интеллекта как фундаментального явления – автономное восприятие и обработка внешней информации, создание образов окружающей действительности, целеполагание, обучаемость и адаптация к изменяющимся внешним условиям – частично проявляется уже в неодушевленных «протоинтеллектуальных» системах автоматического управления, а также в разных версиях компьютерного искусственного интеллекта и в коллективной динамике систем, состоящих из агентов с малыми (общественные насекомые, рыбы, птицы) или нулевыми когнитивными возможностями (формации роботов, природоподобные компьютерные метаэвристики). Общим качеством таких систем является эмерджентная способность к восприятию, обработке и использованию внешней информации далеко за пределами возможностей индивидуальных агентов.

Распределенный интеллект мультиагентных систем во всех его известных формах обнаруживает общие черты, определяемые коллективной обработкой информации, не обязательно рефлекслируемой сознанием агентов. Простейшие роевые модификации РИ в биологических сообществах на фоне беспорядочного поведения особей демонстрируют отбор и «запоминание» полезной информации, системное целеполагание и стандартные реакции на внешние воздействия (режимы коллективного движения рыб и птиц, муравьиные

тропинки, циклы синхронизации динамики общественных насекомых). Элементы хаоса сами являются составной частью РИ, создавая круговой обзор стаи рыб, вариативность перемещения стаи птиц и другие действия, облегчающие коллективное выживание.

Возможности РИ определяются интенсивностью и структурой взаимодействий между агентами, их индивидуальными когнитивными способностями, а также балансом степени упорядоченности системы и случайного шума, который играет важную роль в динамике системы. Имитация роевого интеллекта в робототехнике и компьютерных «алгоритмах, подсказанных Природой», позволяет воспроизвести эмерджентные свойства РИ – в частности, в поиске субоптимальных решений сложных вычислительных задач. Успешное применение генетических и культуральных расчетных алгоритмов в ряде практических приложений открывает возможность для «когнитивной» интерпретации эволюции биологических видов, также обнаруживающей существенные признаки РИ. Обсуждение более сложных и эффективных форм коллективного интеллекта в социальных системах, состоящих из людей, будет представлено во второй части обзора.

**Благодарности.** Авторы выражают искреннюю признательность д-ру техн. наук, проф. О.П. Кузнецову, канд. техн. наук А.В. Макаренко и д-ру техн. наук, проф. П.О. Скобелеву за внимательное ознакомление с рукописью и ценные конструктивные замечания.

## ЛИТЕРАТУРА

1. Новиков Д.А. Кибернетика: навигатор. История кибернетики, современное состояние, перспективы развития. – М.: ЛЕНАНД, 2016. – 160 с. [Novikov, D.A. Kibernetika: navigator. Istoriya kibernetiki, sovremennoe sostoyanie, perspektivy razvitiya. – M.: LENAND, 2016. – 160 s. (In Russian)]
2. Cavens, D., Gloor, C., Illenberber, J., et al. Distributed intelligence in pedestrian simulations // In Waldau, N., Gattermann, P., Knoflachner, H., Schreckenberger, M. (Eds.) Pedestrian and Evacuation Dynamics 2005. – Berlin, Heidelberg: Springer, 2007. – P. 201–212.
3. Fu, X., Gao, H., Cai, H., et al. How to Improve Urban Intelligent Traffic? A Case Study Using Traffic Signal Timing Optimization Model Based on Swarm Intelligence Algorithm // Sensors. – 2021. – Vol. 21, no 2631.
4. Кипятков В.Е. Мир общественных насекомых. – Л.: Издательство ЛГУ, 1991. – 408 с. [Kipyatkov, V.E. Mir obshchestvennyh nasekomyh. – Leningrad: LGU, 1991. – 408 s. (In Russian)]
5. Chittka, L., Rossi, N. Social Cognition in Insects // Trends in Cognitive Sciences. – 2022. – Vol. 26, no. 7. – P. 578–592.
6. Couzin, I.D. Collective Cognition in Animal Groups // Trends in Cognitive Sciences. – 2009. – Vol. 13, no. 1. – P. 36–43.
7. Vicsek, T., Zafeiris, A. Collective Motion // Phys. Rep. – 2012. – Vol. 517. – P. 71–140.
8. Корепанов В.О. Модели рефлексивного группового поведения и управления. – М.: ИПУ РАН, 2011. – 127 с. [Korepanov, V.O. Modeli refleksivnogo gruppovogo povedeniya i upravleniya. – M.: ICS RAS, 2011. – 127 s. (In Russian)]
9. Vászárhelyi, G., Viragh, C., Somorjai, G., et al. Optimized Flocking of Autonomous Drones in Confined Environments // Science Robotics. – 2018. – Vol. 3, no. 20. – Art. no. eaat3536.
10. Шуровьески Дж. Мудрость толпы. Почему вместе мы умнее, чем поодиночке, и как коллективный разум формирует бизнес, экономику, общество и государство: Пер. с англ. – М.: ООО «И.Д. Вильямс», 2007. – 304 с. [Surowiecky, J. The Wisdom of Crowds. – N.-Y.: Anchor Books, 2005. – 336 p.]
11. Kennedy, J., Eberhart, R.C., Shi, Y. Swarm Intelligence. – San Francisco: Morgan Kaufman Publishers, 2001.
12. The Cambridge Handbook of Intelligence. 2<sup>nd</sup> Ed. Ed. by R.J. Sternberg. – Cambridge: Cambridge University Press, 2020.
13. Kornhaber, M.L. The Theory of Multiple Intelligences // In: The Cambridge Handbook of Intelligence. 2<sup>nd</sup> Ed. Ed. by R.J. Sternberg. – Cambridge: Cambridge University Press, 2020.
14. Поляк Б.Т., Хлебников М.В., Рапопорт Л.Б. Математическая теория автоматического управления: учебное пособие. – М.: ЛЕНАНД, 2019. – 500 с. [Polyak, B.T., Khlebnikov, M.V., Rapoport, L.B. Matematicheskaya teoriya avtomaticheskogo upravleniya: uchebnoye posobiye. – Moscow: LENAND, 2019 – 500 p. (In Russian)]
15. Теория управления (дополнительные главы): Учебное пособие / Под ред. Д.А. Новикова. – М.: ЛЕНАНД, 2019. – 552 с. [Teoriya upravleniya (dopolnitel'nye glavy). Uchebnoye posobiye. Ed. by D.A. Novikov. – M.: LENAND, 2019. – 552 s. (In Russian)]
16. Elhoseny, M., Yuan, X. (Eds) Studies in Distributed Intelligence (серия электронных книг). Artificial Intelligence Applications for Smart Societies. Recent Advances. – Cham: Springer, 2021. – 251 p.
17. Rzevski, G., Skobelev, P. Emergent Intelligence in Large Scale Multi-agent Systems // Intern. J. of Education and Inform. Technol. – 2004. – Vol. 1, iss. 2. – P. 64–71.
18. Springer Handbook of Robotics, 2<sup>nd</sup> Edition. Ed. by B. Siciliano, O. Khatib. – Berlin: Springer-Verlag, 2015. – 2227 p.
19. Legg, S., Hutter, M. A Collection of Definitions of Intelligence // In Adv. Artif. General Intelligence: Concepts, Architectures and Algorithms. Ed. by B. Goertzel, P. Wang. – Amsterdam: IOS Press, 2007. – P. 17–25.
20. Encyclopedia Britannica – URL: <https://www.britannica.com/> (дата обращения 15.02.2023). [Accessed February 15, 2023.]
21. Malone, T.W., Bernstein, M.S. Handbook of Collective Intelligence. – Cambridge: MIT Press, 2015. – 230 p.
22. Справочник по теории автоматического управления / Под ред. А.А. Красовского. – М.: Наука, 1987. – 712 с. [Spravochnik po teorii avtomaticheskogo upravleniya / Ed. By A.A. Krasovskij. – M.: Nauka, 1987. – 712 s. (In Russian)]
23. Поляков К.Ю. Основы теории автоматического управления. – СПб.: Изд-во СПбГМУ, 2012. – 234 с. [Polyakov, K.Yu. Osnovy teorii avtomaticheskogo upravleniya, – S.-Petersburg: SPbGMTU, 2012. – 234 s. (in Russian)]
24. The Control Handbook. Vol. I – III. / Ed. by W. Levine. – N.-Y.: CRC Press, 2010. – 3526 p.
25. Vasil, C.-I., Leahy, K., Cristofalo, E., et al. Control in Belief Space with Temporal Logic Specifications // Proc. IEEE Conference on Decision and Control. – Las Vegas, USA, 2016. – P. 7419–7424.
26. Кудин Ю.И., Пащенко Ф.Ф. Нечеткие системы управления / Теория управления (дополнительные главы): Учебное пособие. Под ред. Д.А. Новикова. – М.: ЛЕНАНД, 2019. – С. 400–425. [Kudin, Yu.I., Pashchenko, F.F. Nechetkiye sistemy upravleniya // In: Teoriya upravleniya (dopolnitel'nye glavy). Uchebnoya posobiye. Ed. by D.A. Novikov. – M.: LENAND, 2019. – S. 400–425. (In Russian)]
27. Толмен Э. Когнитивные карты у крыс и у человека / Хрестоматия по истории психологии. Под ред. Гальперина П.Я. и Ждан А.Н. – М.: МГУ, 1980. – С. 63–69. [Tolmen, E.C. Cognitive Maps in Rats and Men // The Psychological Review. – 1948. – Vol. 55, no. 4. – P. 189–208.]
28. Новиков Д.А., Чхартишвили А.Г. Рефлексия и управление: математические модели. 2-е изд. – М.: ЛЕНАНД, 2022. – 416 с. [Novikov, D.A., Chkhartishvili, A.G. Reflection and Control: Mathematical Modeling, 2<sup>nd</sup> Ed. – M.: LENAND, 2022 – 416 p. (In Russian)]
29. Белов М.В., Новиков Д.А. Структура креативной деятельности // Проблемы управления. – 2021. – № 5. – С. 20–33. [Belov, M.V., Novikov, D.A. The Structure of Creative Activity // Control Sciences. – 2021. – No. 5. – P. 17–28.]
30. Петухов В.В. Психология мышления. Учебно-методическое пособие. – М.: МГУ, 1987. – 99 с. [Petukhov, V.V. Psikhologiya myshleniya, Uchebno-metodicheskoye posobiye. – M.: MGU, 1987. – 99 p. (In Russian)]
31. Пономарев Я.А., Александров Ю.И., Ушаков Д.В. Психология творчества. Школа Я.А. Пономарева. – М.: Институт психологии РАН, 2006. – 624 с. [Ponomarev, Ya.A., Aleksandrov, Yu.I., Ushakov D.V. Psikhologiya tvorchestva. Shkola Ya.A. Ponomareva, – M.: Institute of Psychology RAS, 2006. – 624 s. (In Russian)]
32. Альтшуллер Г.С. Найти идею. Введение в ТРИЗ – теорию решения изобретательских задач. 4-е изд. – М.: Альпина Паблишерз, 2011. – 400 с. [Altshuler, G.S. Najti ideyu.

- Vvedenie v TRIZ teoriyu resheniya izobretatel'skikh zadach. 4<sup>th</sup> ed. – М.: Alpina Publishers, 2011. – 400 s. (In Russian)]
33. Зорина З.А., Поletaева И.И. Зоопсихология. Элементарное мышление животных: учебное пособие. – М.: Аспект Пресс, 2002. – 320 с. [Zorina, Z.A., Poletayeva, I.I. Zoopsikhologiya. Elementarnoye myshleniye u zhivotnykh. – М.: Aspect Press, 2002. – 320 p. (In Russian)].
  34. Тинберген Н. Социальное поведение животных (пер. с англ.). – М.: Мир, 1993. – 152 с. [Tinbergen, N. Social Behaviour in Animals with Special Reference to Vertebrates. – London – New-York: Chapman & Hall, 1968. – 150 p.]
  35. Artificial Intelligence Index Report. – Stanford University Human-Centered Artificial Intelligence, 2022. – URL: <https://aiindex.stanford.edu/report/> (дата обращения 15.02.2023). [Accessed February 15, 2023.]
  36. Хайкин С. Нейронные сети. Полный курс. 2-е издание (пер. с англ.). – М.: Вильямс, 2016. – 1104 с. [Haykin, S. Neural Networks: A Comprehensive Foundation. 2<sup>nd</sup> Ed. – Singapore: Prentice Hall & Pearson Education, 1998. – 842 p.]
  37. Gu, T., Liu, K., Dolan-Gavitt, B., Garg, S. BadNets: Evaluating Backdooring Attacks on Deep Neural Networks // IEEE Access. – 2019. – Vol. 7. – P. 47230–47244.
  38. Макаренко А.В. Глубокие нейронные сети: зарождение, становление, современное состояние // Проблемы управления. – 2020. – № 2. – С. 3–19 [Makarenko, A.V. Deep Neural Networks: Origins, Development, Current Status // Control Sciences. – 2020. – No. 2. – P. 3–19. (In Russian)]
  39. Тей А., Грибомон П., Луи Ж. и др. Логический подход к искусственному интеллекту: от классической логики к логическому программированию (пер. с франц.). – М.: Мир, 1990. – 432 с. [Thayse, A., Gribomont, P., Louis, G., et al. Approche logique de l'intelligence artificielle. T. 1: De la logique classique a la programmation logique. – Paris: Bordas Editions, 1988. – 274 p. (In French)]
  40. Springer Handbook of Computational Intelligence. Ed. by J. Kacprzyk, W. Pedrych. – Berlin: Springer, 2015. – 1633 p.
  41. Batini, C., Ceri, S., Navathe, S.B. Conceptual Database Design: An Entity-Relationship Approach. – Redwood: Benjamin/Cummings, 1992. – 496 p.
  42. Lehmann, F. Semantic Networks in Artificial Intelligence. – New-York: Pergamon, 1992. – 768 p.
  43. Poelmans, J., Ignatov, D.I., Kuznetsov, S.O., Dedene, G. Formal Concept Analysis in Knowledge Processing: A Survey on Applications // Expert Systems with Applications. – 2013. – Vol. 40, no. 16. – P. 6538–6560.
  44. Васильев С.Н. Интерактивное порождение новых знаний на основе автоматических средств логического вывода. // Онтология проектирования. – 2023. – Т. 13, № 1. – С. 10–29. [Vassiliev, S.N. Interactive Generation of New Knowledge Based on Automatic Means of Logical Inference. Ontology of Designing. – 2023. – Vol. 13, no. 1. – P. 10–29. (In Russian)]
  45. Кузнецов О.П. Избранные труды. Автоматы, языки и искусственный интеллект. – М.: ИПУ РАН, 2016. – 464 с. [Kuznetsov, O.P. Izbrannye Trudy. Avtomaty, yazyki i iskusstvennyy intellect. – М.: ICS RAS, 2016. – 464 p. (In Russian)]
  46. Кулинич А.А. Семиотические когнитивные карты. Ч. 1. Когнитивный и семиотический подходы в информатике и управлении // Проблемы управления. – 2016. – № 1. – С. 2–10; Ч. 2. Основные определения и алгоритмы // Проблемы управления. – 2016. – № 2. – С. 24–40 [Kulinich, A.A. Semiotic Cognitive Maps. Part 1. Cognitive and Semiotic Approach in Informatics and Control Sciences // Control Sciences. – 2016. – No. 1. – P. 2–10; P. 2. The Basic Definitions and Algorithms // Control Sciences. – 2016. – No. 2. – P. 24–40. (In Russian)]
  47. Gawthrop, P., Loram, I., Lakie, M., Gollee, H. Intermittent Control: A Computational Theory of Human Control // Biological Cybernetics. – 2011. – Vol. 104. – P. 31–51.
  48. Белов М.В., Новиков Д.А. Модели деятельности (основы математической теории деятельности). – М.: ЛЕНАНД, 2021. – 216 с. [Belov, M.V., Novikov, D.A. Models of Activity (foundations of mathematical theory of activity). – М.: LENAND, 2021. – 216 p. (In Russian)]
  49. Новиков Д.А. Модели информационного противоборства в управлении толпой // Проблемы управления. – 2015. – № 3. – С. 29–39. [Novikov, D.A. Models of Informational Confrontation in Mob Control // Control Sciences. – 2015. – No. 3. – P. 29–39. (In Russian)]
  50. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и противоборства. 3-е изд. – М.: МЦНМО, 2018. – 224 с. [Gubanov, D.A., Novikov, D.A., Chkhartishvili, A.G. Social Networks: Models of Informational Influence, Control and Contest. 3<sup>rd</sup> Ed. – М.: MTsNMO, 2018. – 224 s. (In Russian)]
  51. Словохотов Ю.Л. Физика и социофизика // Проблемы управления. – 2012. – № 1. – С. 2–20; № 2. – С. 2–31; № 3. – С. 2–34. [Slovokhotov, Y.L. Physics vs. Sociophysics // Control Sciences. – 2012. – No. 1. – P. 2–20; No. 2. – P. 2–31; No. 3. – P. 2–34. (In Russian)]
  52. Гасников А.В. (ред.) Введение в математическое моделирование транспортных потоков. – М.: МЦНМО, 2013. – 429 с. [Gasnikov, A.V. (Ed.) Introduction to Mathematical Modeling of Traffic Flows. – М.: MTsNMO, 2013. – 429 s. (In Russian)]
  53. Романовский М.Ю., Романовский Ю.М. Математические начала эконофизики. – М.-Ижевск: Институт компьютерных исследований, 2019. – 360 с. [Romanovsky, M.Yu., Romanovsky, Yu.M. Mathematical foundations of econophysics. – Moscow-Izhevsk: Institut kompyuternykh issledovaniy, 2019. – 360 p. (In Russian)]
  54. Ширяев В.И. Финансовые рынки: нейронные сети, хаос и нелинейная динамика. 5-е изд., – М.: ЛИБРОКОМ, 2013. – 232 с. [Shiryayev, V.I. Financial Markets: Neural Networks, Chaos and Non-linear Dynamics. 5<sup>th</sup> Ed. – Moscow: LIBROKOM, 2013. – 232 p. (in Russian)].
  55. Барабанов И.Н., Новиков Д.А. Динамические модели управления возбуждением толпы в непрерывном времени // Управление большими системами. – 2016. – Вып. 67. – С. 71–86. [Barabanov, I.N., Novikov D.A. Continuous-Time Dynamic Models of Mob Excitation // Large-Scale Systems Control. – 2016. – No. 67. – P. 71–86. (In Russian)]
  56. Kozitsin, I.V. A General Framework to Link Theory and Empirics in Opinion Formation Models // Sci. Rep. – 2022. – Vol. 12, no. 5543. – P. 1–18.
  57. Spiegler, R. Bounded Rationality and Industrial Organization. – Oxford: Oxford University Press, 2014. – 240 p.
  58. Новиков Д.А. Ограниченная рациональность и управление // Математическая теория игр и ее приложения. – 2022. – Т. 14, вып. 1. – С. 49–84 [Novikov, D.A. Bounded Rationality and Control // Matematicheskaya teoriya igr i ee prilozheniya. – 2022. – Vol. 14, no. 1. – P. 49–84. (In Russian)]
  59. Slovkhotov, Y.L., Neretin, I.S. Towards Constructing a Modular Model of Distributed Intelligence // Programming & Computer Software. – 2018. – Vol. 44, no. 6. – P. 499–507.

60. Dombrowski, C., Cisneros, L., Chatkaew, S., et al. Self-Concentration and Large-Scale Coherence in Bacterial Dynamics // *Phys. Rev. Lett.* – 2004. – Vol. 93, no. 9. – P. 098103–1–098103–54.
61. Арансон И.С. Топологические дефекты в активных жидких кристаллах // *Усп. физ. наук*, 2013. – Т. 183, № 1. – С. 87–102. [Aranson, I.S. Aktivnaye kolloidy, *Physics-Uspekhi*, 31 Jan. – 2013. – Vol. 183, no. 1. – P. 87–102. (In Russian)]
62. Deng, J., Liu, D. Spontaneous Response of a Self-Organized Fish School to a Predator // *Bioinspiration & Biomimetics*. – 2021. – Vol. 16, no. 4. – Art. no. 046013.
63. Cristiani, E., Menci, M., Papi, M., Brafman, L. An All-leader Agent-Based Model for Turning and Flocking Birds // *J. Math. Biology*. – 2021. – Vol. 83. – Art. no. 45.
64. Helbing, D., Keltsch, J., Molnar, P. Modelling the Evolution of Human Trail Systems // *Nature*. – 1997. – Vol. 388, no. 3. – P. 47–50.
65. El Zoghby, N., Loscri, V., Natalizio, E., Cherfaoui, V. Robot Cooperation and Swarm Intelligence / In: *Wireless Sensor and Robot Networks: from Topology Control to Communication Aspects*. – Singapore: World Scientific Publishing Company, 2014. – P. 168–201.
66. Кулинич А.А., Карпов В.Э., Карпова И.П. Социальные сообщества роботов. – М.: ЛЕНАНД, 2019. – 352 с. [Kulinich, A.A., Karpov, V.E., Karpova, I.I. *Social Communities of Robots*. – М.: LENAND, 2019. – 352 p. (In Russian)]
67. МакКоннелл, Дж. Основы современных алгоритмов. – М.: Техносфера, 2004. – 368 с. [McConnell, J.J. *Analysis of Algorithms: An Active Learning Approach*. – Boston: Jones & Bartlett Publishers, 2001. – 297 p.]
68. *Nature-Inspired Metaheuristic Algorithms for Engineering Optimization Applications*. Ed. by S. Carbar, A. Toktas, D. Ustin / *Springer Tracts in Nature-Inspired Computing*. – Cham: Springer, 2021. – 404 p.
69. Winklerova, Z. Maturity of the Particle Swarm as a Metric for Measuring the Particle Swarm Intelligence // *Proc. 4<sup>th</sup> Int. Conf. Adv. Swarm Intel.* – Harbin, 2013. – P. 40–54.
70. Chakraborty, A., Kar, A.K. Swarm Intelligence: A Review of Algorithms / In: *Nature-Inspired Computing and Optimization: Theory and Applications*. Ed. by S. Patnaik, X.-S. Yang, K. Nakamatsu. – Cham: Springer, 2017. – P. 475–494.
71. Широкий А.А., Калашиников А.О. Применение методов естественных вычислений для управления рисками сложных систем // *Проблемы управления*. – 2021. – № 4. – С. 3–20. [Shiroky, A.A., Kalashnikov, A.O. *Natural Computing with Application to Risk Management in Complex Systems* // *Control Sciences*. – 2021. – No. 4. – P. 2–17].
72. Докинз Р. Эгоистичный ген (пер. с англ.). – М.: Corpus (АСТ), 1989. – 277 с. [Dawkins, R. *The Selfish Gene*. – Oxford: Oxford University Press, 1976. – 360 p.]
73. Alattas, R.J., Patel, S., Sobh, T.M. Evolutionary Modular Robotics: Survey and Analysis // *J. Intel. & Robotic Syst.* – 2019. – Vol. 95. – P. 815–828.

*Статья представлена к публикации членом редколлегии В.Е. Лепским.*

*Поступила в редакцию 06.03.2023,  
после доработки 11.05.2023.  
Принята к публикации 22.05.2023.*

**Словохотов Юрий Леонидович** – д-р хим. наук, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва; Московский государственный университет имени М.В. Ломоносова,  
✉ [yurislovo@yandex.ru](mailto:yurislovo@yandex.ru),  
ORCID iD: <https://orcid.org/0000-0002-6669-6210>

**Новиков Дмитрий Александрович** – академик РАН, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва,  
✉ [novikov@ipu.ru](mailto:novikov@ipu.ru),  
ORCID iD: <https://orcid.org/0000-0002-9314-3304>

© 2023 г. Словохотов Ю.Л., Новиков Д.А.



Эта статья доступна по лицензии Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная.

# DISTRIBUTED INTELLIGENCE OF MULTI-AGENT SYSTEMS.

## PART I: Basic Features and Simple Forms

Yu.L. Slovokhotov<sup>1,2</sup> and D.A. Novikov<sup>1</sup>

<sup>1</sup>Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

<sup>2</sup>Department of Materials Science, Moscow State University, Russia

✉ yurislovo@yandex.ru, ✉ novikov@ipu.ru

**Abstract.** This multi-part survey is devoted to the empirical characteristics and manifestations of intelligence as the capability of an autonomous agent to perceive, process, and use information in order to achieve its goal. Part I of the survey briefly describes the most significant attributes of “proto-intelligent” and intelligent systems, together with the main features of distributed intelligence (DI), individual human intelligence (HI), and artificial intelligence (AI). An inseparable connection is emphasized between the DI of human organizational and social systems and individual human intelligence. The simplest forms of “swarm” intelligence are considered as examples, and the key factors determining the effectiveness of DI in such multi-agent systems are presented, including the structure of interactions between agents, their collective goal-setting, external information recording, convolution, and processing, and the standard “images” of external influences. Their combination pushes the performance of a multi-agent system far beyond the capabilities of its individual agents. In part II of the survey, different forms of collective intelligence in human social systems will be analyzed and all known types of intelligence will be generally classified.

**Keywords:** multi-agent systems, distributed intelligence, swarm intelligence, organizational systems, collective intelligence.

**Acknowledgments.** We are grateful to Prof., Dr. Sci. (Eng.) O.P. Kuznetsov, Cand. Sci. (Eng.) A.V. Makarenko, and Prof., Dr. Sci. (Eng.) P.O. Skobelev for careful reading of the manuscript and helpful remarks.