

УДК 512.6, 517.9

**ИНТЕЛЛЕКТУАЛЬНОЕ КОГНИТИВНОЕ УПРАВЛЕНИЕ РОБОТИЗИРОВАННЫМИ  
СОЦИОТЕХНИЧЕСКИМИ СИСТЕМАМИ.****Ч. 1: РОБОТОТЕХНИЧЕСКИЕ СИСТЕМЫ И МОДЕЛИ ВЗАИМОДЕЙСТВИЯ  
«ЧЕЛОВЕК – РОБОТ» В «ИНДУСТРИЯ 4.0»****Тятюшкина Ольга Юрьевна<sup>1</sup>, Ульянов Сергей Викторович<sup>2</sup>**

<sup>1</sup>Кандидат технических наук, доцент;  
Государственный университет «Дубна»;  
141980, Московская обл., г. Дубна, ул. Университетская, 19;  
e-mail: tyatyushkina@mail.ru.

<sup>2</sup>Доктор физико-математических наук, профессор;  
Государственный университет «Дубна»;  
141980, Московская обл., г. Дубна, ул. Университетская, 19;  
Ведущий научный сотрудник;  
Объединенный институт ядерных исследований;  
141980, Московская обл., г. Дубна, ул. Жолио-Кюри, д. 6;  
e-mail: ulyanovsv@mail.ru.

С момента запуска международного проекта «Индустрия 4.0» в 2011 году до сегодняшнего дня его влияние ощущалось практически во всех секторах рынка разработок новейших информационных технологий. Многие индустриально развитые страны начали свою индивидуальную политику, чтобы поддержать данную технологическую революцию. Однако развитие данной технологии находится в начале переходного периода. В этой статье описываются некоторые особенности применения интеллектуальных когнитивных робототехнических комплексов в социотехнических системах на основе концепции проекта «Индустрии 4.0» и объяснительного искусственного интеллекта (XAI).

**Ключевые слова:** роботизированные социотехнические системы, когнитивные роботы, интеллектуальное управление, человеко-машинные когнитивные интерфейсы, проект «Индустрия 4.0».

**Для цитирования:**

Тятюшкина О.Ю., Ульянов С.В. Интеллектуальное когнитивное управление роботизированными социотехническими системами. Ч. 1: Робототехнические системы и модели взаимодействия «человек – робот» в «Индустрия 4.0» // Системный анализ в науке и образовании: сетевое научное издание. 2021. № 3. С. 44–101. URL : <http://sanse.ru/download/447>.

**INTELLIGENT COGNITIVE CONTROL OF ROBOTIC SOCIOTECHNICAL SYSTEMS.  
PT. 1: ROBOTIC SYSTEMS AND «HUMAN BEING – ROBOT» INTERACTIVE MODELS  
IN PROJECT «INDUSTRY 4.0»****Tyatyushkina Olga Yu.<sup>1</sup>, Ulyanov Sergey V.<sup>2</sup>**

<sup>1</sup>PhD in Engineering sciences, associate professor;  
Dubna State University,  
Institute of the system analysis and management;  
19 Universitetskaya Str., Dubna, Moscow region, 141980, Russia;  
e-mail: tyatyushkina@mail.ru.

<sup>2</sup>Grand PhD in Physical and Mathematical Sciences, professor;  
Dubna State University;  
19 Universitetskaya Str., Dubna, Moscow region, 141980, Russia;  
Leading Researcher of LIT JINR;  
Joint Institute for Nuclear Research;  
6 Joliot-Curie Str., Dubna, Moscow region, 141980, Russia;  
e-mail: ulyanovsv@mail.ru.

*Since the launch of the international project "Industry 4.0" in 2011 to the present day, its influence has been felt in almost all sectors of the market for the development of the latest information technologies. Many industrialized countries have started their own individual policies to support this technological revolution. However, the development of this technology is at the beginning of a transition period. This article describes some of the features of the application of intelligent cognitive robotic complexes in socio-technical systems based on the concept of the "Industry 4.0" project and explanatory artificial intelligence (XAI).*

**Keywords:** robotic socio-technical systems, cognitive robots, intelligent control, human-machine cognitive interfaces, the Industry 4.0 project.

#### For citation:

Tyatyushkina O., Ulyanov S. Intelligent cognitive control of robotic sociotechnical systems. Pt. 1: Robotic systems and «Human being – robot» interactive models in project «industry 4.0», 2021;(3):44–101(In Russ). Available from: <http://sanse.ru/download/447>.

### Введение. Интеллектуальная когнитивная робототехника в социально-технических системах: модели роботов

На основе анализа полученных отзывов и материалов от многочисленных коллективов разработчиков-исследователей в области робототехники онлайн-группа экспертов составила окончательный список 30 наиболее важных тем и направлений исследований. Далее они были сгруппированы в 10 основных проблем (рис. 1), которые могут иметь серьезные прорывы, значимые исследования и / или социально-экономические последствия в следующие 5-10 лет [1].

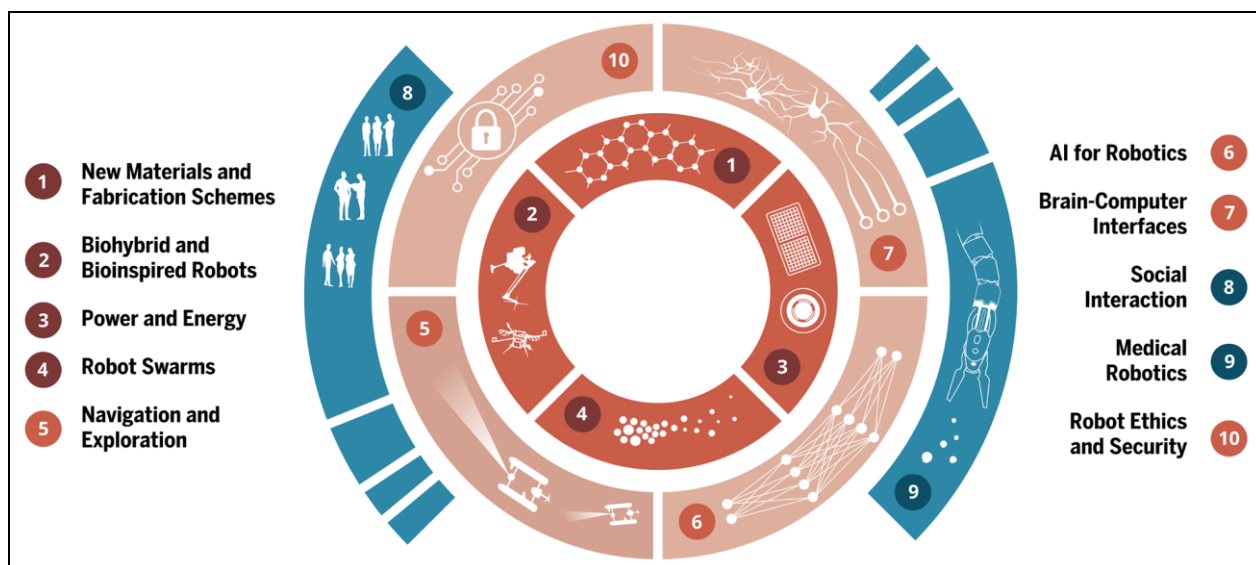


Рис. 1. Десять основных задач научной робототехники

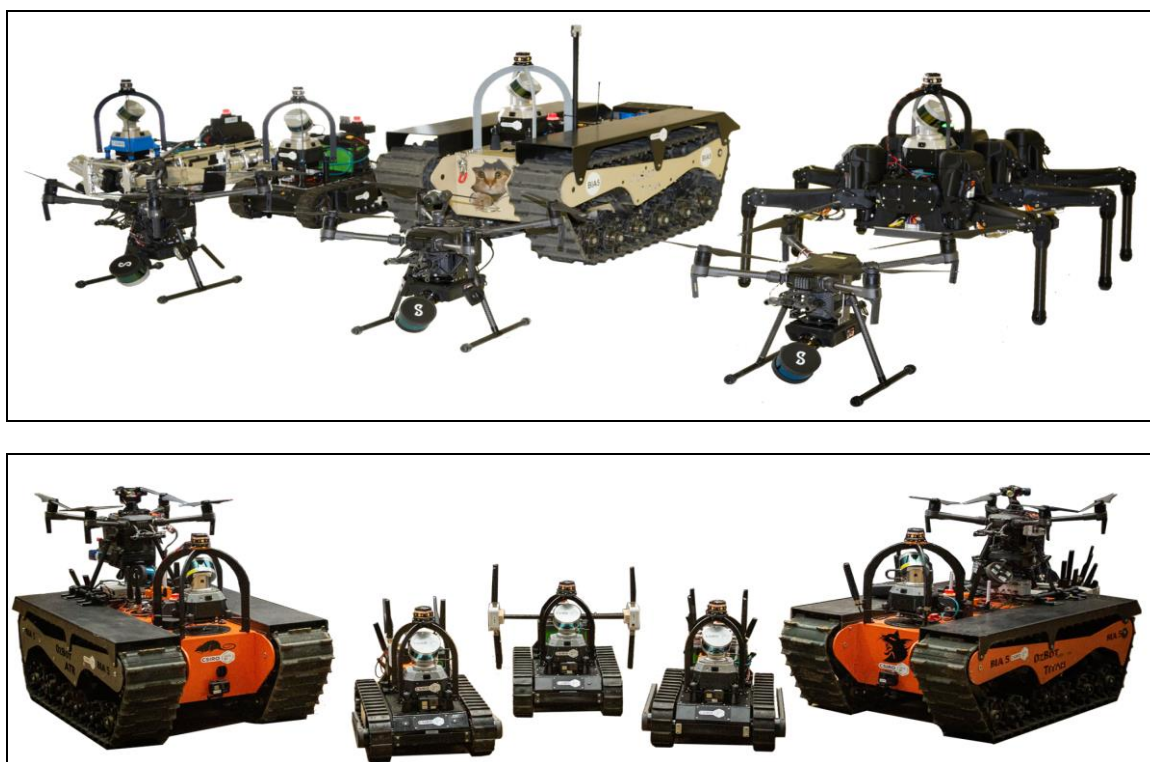
При этом отмечен достигнутый прогресс во многих аспектах робототехники – от микромашин для биомедицины до крупномасштабных систем для робототехники и от роботов для космического пространства до тех, кто занимается глубоководными исследованиями. Наблюдается эволюция «мягких» роботов и то, как новые материалы и схемы изготовления привели к деформируемым приводам, которые являются совместимыми, универсальными и самовосстанавливающимися. Существует множество примеров биоинспирированных конструкций, от прыгающего робота с регулируемой мощностью, ловкость и мощность которого приближены к таковым как у галаго (животное с самой высокой способностью к вертикальному прыжку), до биомиметической роботизированной платформы для изучения специализации полета летучих мышей и биороботический липкий диск для подводных автостопов, вдохновленных реморами-присосками. Отмечено 10-летие операционной системы роботов (ROS), промежуточного программного обеспечения для робототехники с открытым исходным кодом, которое создает большие успехи в реализации своей миссии по обеспечению энергией мировых роботов, от задач космических роботов до автономного вождения, промышленной сборки и хирургии. Сфера робототехники обширна и охватывает многие базовые и смежные технологические

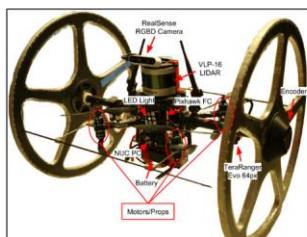
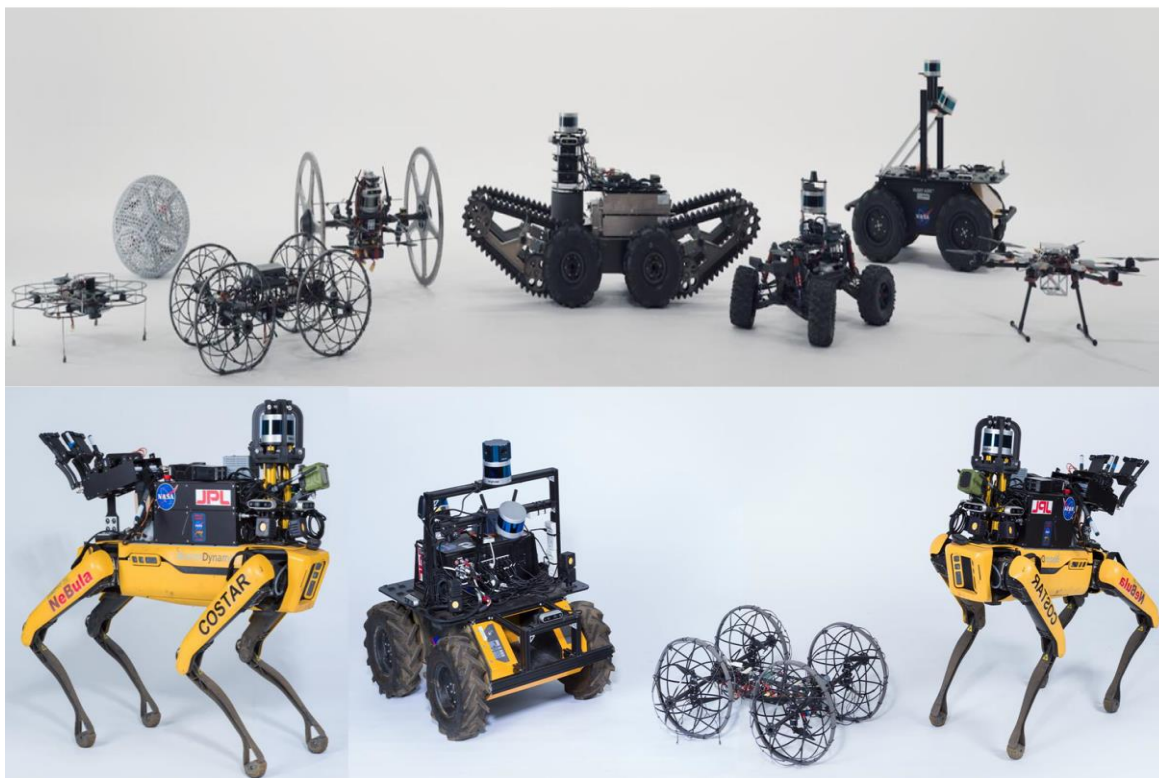
области. Выявление этих проблем было сложной задачей, и есть много не перечисленных подтем, которые одинаково важны для будущего развития.

Таким образом, приведенный выше на рис. 1 список не является исчерпывающим. Из 10 основных проблем, перечисленных здесь, первые семь представляют собой базовые технологии, которые оказывают более широкое влияние на все области применения робототехники. Для решения следующих двух задач в число прикладных областей развития включили социальную робототехнику и медицинскую робототехнику, чтобы подчеркнуть существенное воздействие на общество и здоровье, которое они принесут. Наконец, последняя проблема связана с ответственными инновациями и с тем, как следует тщательно учитывать этику и безопасность при дальнейшем развитии технологии.

В данной работе рассматривается сложившаяся ситуация с разработкой новых видов аппаратной реализации сложных видов роботизированных социотехнических систем, взаимодействующих с человеком-оператором и антропоморфных (автономных и коллективов) роботов с наделенными новыми когнитивными способностями качествами (такими как эмоционального когнитивного управления, обмена информацией и способностью оценивать ситуацию поведения внешней среды и мн. др.)

На рис. 2 приведены примеры [1 – 5] разрабатываемых гетерогенных роботизированных комплексов в рамках проектов типа *DARPA* и мн. др.









(a)



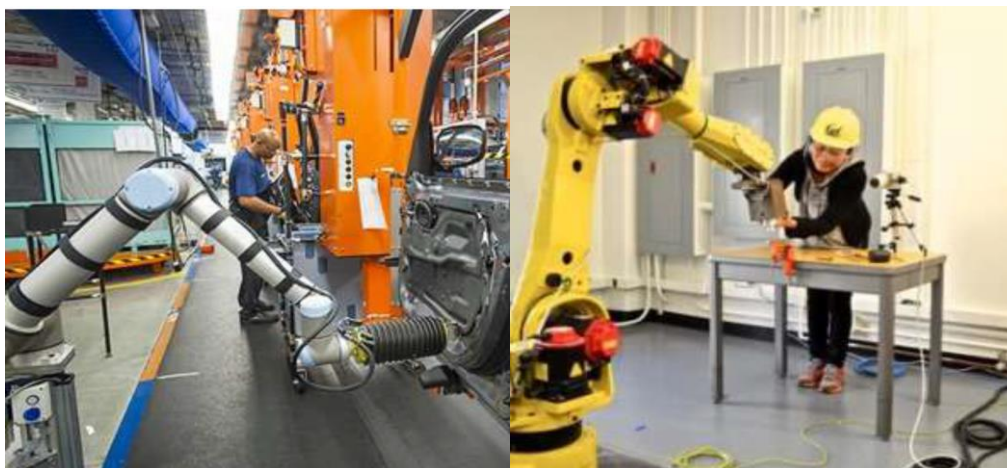
(b)

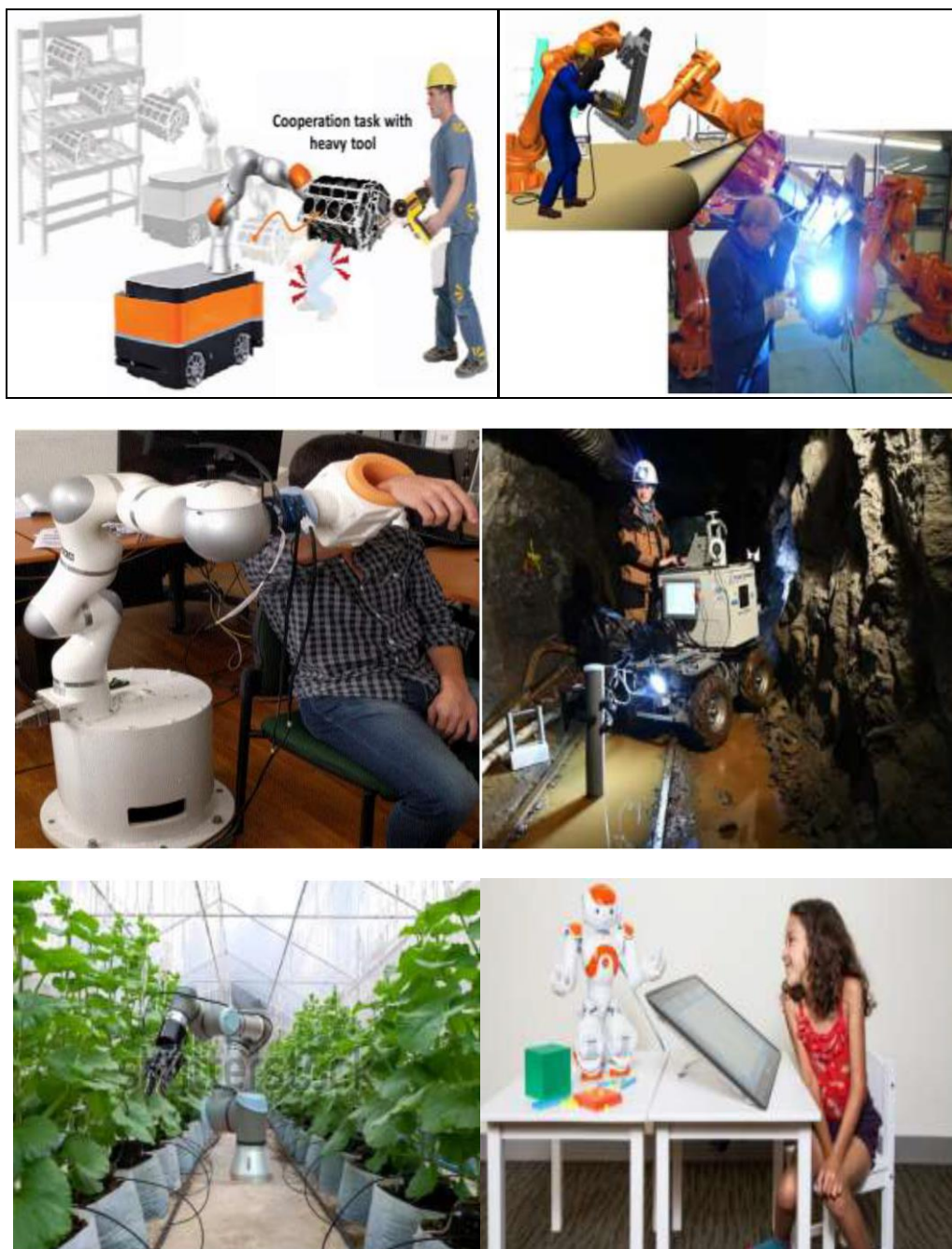


Рис. 2. Примеры платформ сложных гетерогенных роботизированных комплексов

Помимо традиционных требований типа управляемость, мобильность, устойчивость, надежность и др. в данных проектах настойчиво подчеркивается необходимость наделения таких роботизированных комплексов новыми, когнитивными способностями, основы которых изложены в данной работе.

Примеры, приведенные на рис. 2 и 3 показывают актуальность затрагиваемой темы в данной работе и применение разрабатываемых методов интеллектуального когнитивного управления позволяет повысить эффективность работы данных роботизированных систем, не меняя облика конструктивных технологических решений аппаратной реализации.





*Рис. 3. Примеры взаимодействия человека-оператора с автономными роботами*

Следует также отметить, что разрабатываемые обобщенные структуры систем управления (см. например, рис. 4) не содержат такого рода инструментария, и как отмечается самими авторами [3], требуется погружение в контур управления многоагентными системами когнитивных функций взаимодействующих роботов.



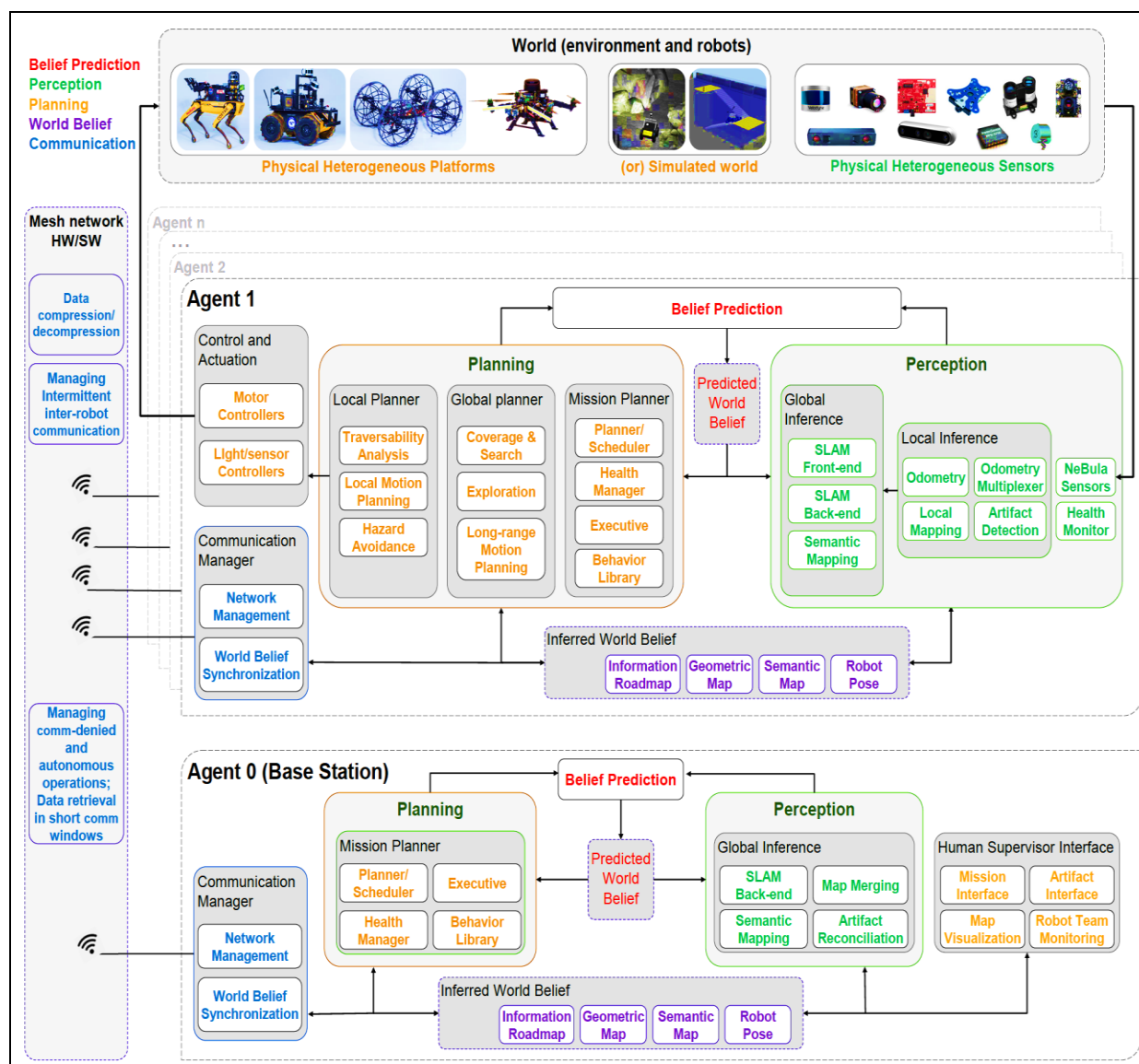


Рис. 4. Пример системы управления гетерогенным роботизированным комплексом

Рассмотрим некоторые особенности промышленного искусственного интеллекта (ИИ) во взаимосвязи с проектом «Индустрия 4.0».

## Особенности промышленного объяснительного искусственного интеллекта (ХИИ)

Искусственный интеллект (ИИ) считается «новым проводником» так называемой четвертой промышленной революции (Индустрия 4.0), которая несет в себе несколько сквозных (*end-to-end*) прорывных технологий и может развиваться без катастрофических прецедентов в истории человечества благодаря своей скорости, надежности и масштабам [6 – 8].

*Примечание.* По прогнозам [8], к 2032 году технологии ИИ добавят мировой экономике 15 трлн долларов США. Объем рынка ИИ вырастет на 75,54 миллиарда долларов в течение 2019-2023 годов. В течение прогнозируемого периода динамика роста рынка ускорится из-за устойчивого увеличения годового роста. Прогноз рынка принимает во внимание точку зрения конечных пользователей и географический ландшафт (Азиатско-Тихоокеанский регион, Европа, МЕА, Северная Америка и Южная Америка), как показано на рис. 5.

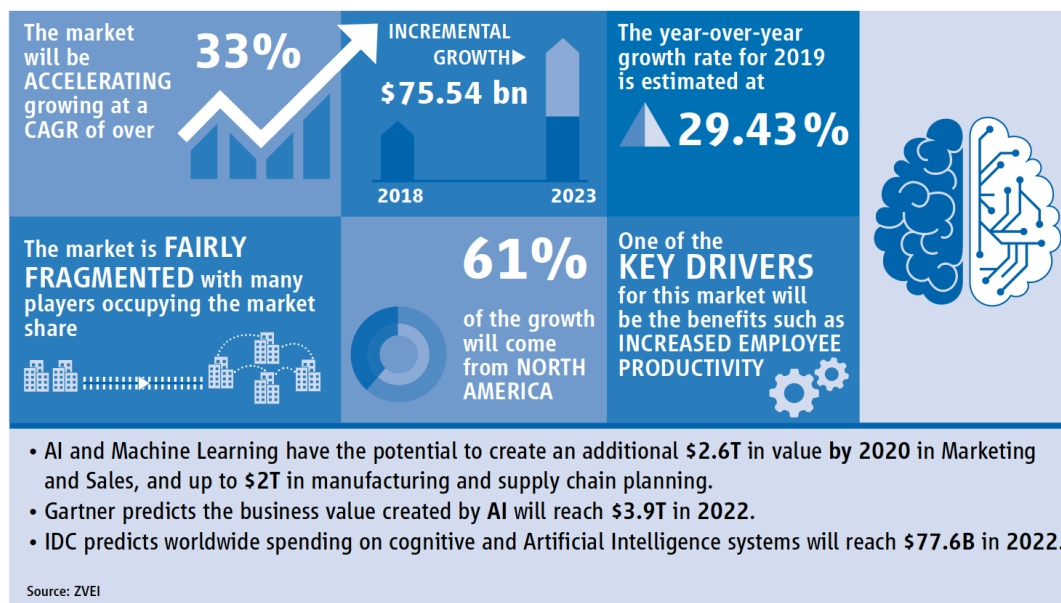


Рис. 5. Объем рынка ИИ

Ожидается, что доля ИИ на производственном рынке будет расти со среднегодовым темпом роста 39,7% с 2019 по 2027 год и достигнет 27 миллиардов долларов к 2027 году.

Организации активно внедряют решения и услуги на основе ИИ, чтобы изменить свои бизнес-операции и повысить прибыльность. Внедрение технологий ИИ, например, в обрабатывающей промышленности набирает обороты, чтобы создавать новые возможности и расширять операционные возможности за счет использования новых технологий, ускорения процессов и адаптации организаций к изменениям в будущем.

Есть два факта об Индустрии 4.0, которые можно привести с некоторой долей уверенности. Новый виток цифровой революции неизбежен, и эта революция все еще продолжается. Пока получаем множество преимуществ от этих новых технологий, проблемы и вызовы будут возникать. В краткосрочной перспективе, во-первых, небольшая автономия во многих текущих системах будет ограничивать движение промышленности к интеллектуальному производству. Во-вторых, нехватка полосы пропускания в большинстве текущих сетевых протоколов может стать узким местом, требующим десятилетнего улучшения. В-третьих, многие отрасли еще не обеспечили качество и целостность записываемых данных. Отсутствует стандартный подход к аннотациям сущностей данных. В-четвертых, моделирование и анализ сложных систем еще не удовлетворительны для практических целей. В-пятых, существует множество трудностей, связанных с корректировкой текущих производственных маршрутов, чтобы адаптировать большую динамическую реконфигурацию к индивидуализированным и индивидуализированным продуктам. Наконец, до сих пор неясно, как лучше всего инвестировать в разные сектора (например, малые и средние предприятия по сравнению с Fortune 500) и какова оптимальная поддержка со стороны правительств каждой страны [7].

*Примечание.* Как и в случае недостатков других технологий и ИИ (или, по крайней мере, ограниченных алгоритмов обучения, которые часто в настоящее время называют «искусственным интеллектом»), потенциальные выгоды от Индустрии 4.0, несомненно, огромны. Промышленные компании стремятся к более точному и качественному производству при сниженных эксплуатационных расходах; сокращение времени простоя благодаря профилактическому обслуживанию и анализу цепочки поставок; и меньше травм в производственных цехах благодаря более адаптируемому оборудованию. А за пределами завода другие отрасли могут выиграть от наличия «нейрона-подобной» распределенной системы датчиков, аналитики для обработки «озер» данных и своевременного реагирования на возникающие проблемы авиации, энергетики, логистики и многих других предприятий, которые полагаются на надежные, предсказуемые факты и также могут получить дополнительный импульс в своем развитии.



## **Взаимоотношение «Умное» производство и «Интеллектуальное» производство: роль технологий ИИ**

В своем будущем Индустрии 4.0 включает в себя интеллектуальные предприятия, использующие аддитивное производство, такое как 3D-печать с помощью селективного лазерного спекания и другие производственные системы, управляемые компьютером, смогут адаптивно производить детали по запросу прямо из цифровых конструкций. Датчики отслеживают необходимые компоненты и упорядочивают их на основе моделей спроса и других алгоритмических деревьев решений, выводя производство «точно в срок» на новый уровень оптимизации. Оптические датчики и системы, основанные на машинном обучении, контролируют качество компонентов с большей согласованностью и точностью, чем потенциально уставшие и скучающие люди в линейке продуктов. Промышленные роботы работают синхронно с людьми, выполняющими более деликатные задачи, или полностью заменяют их. Целые цепочки поставок могут меняться с появлением новых продуктов, изменениями в потреблении и экономических колебаниями. И машины могут подсказать людям, когда машины должны быть отремонтированы, прежде чем они даже сломаются, или подскажут людям, как лучше организовать линию, и все благодаря искусственному интеллекту, обрабатывающему огромные объемы данных, генерируемых производственным процессом [9, 10].

Проект «Индустрия 4.0» вызвал значительный интерес со стороны фирм, правительств и частных лиц как новая концепция будущих компьютерных, промышленных и социальных систем. Однако эту концепцию еще предстоит полностью изучить в научной литературе. Учитывая широкий охват темы, сделана попытка понять и прояснить Индустрию 4.0 путем анализа 660 журнальных статей и 3901 общих обзоров с помощью интеллектуального анализа текста и адаптивных алгоритмов машинного обучения.

Основываясь на результатах, в [11] выявлена 31 исследовательская и прикладная проблемы, связанных с Индустрией 4.0. Эти проблемы классифицируются и описываются в рамках пятиуровневой иерархии: 1) развитие инфраструктуры для подключения, 2) разработка ИИ для принятия решений на основе данных, 3) оптимизация системы и процессов, 4) промышленные инновации и 5) социальная сфера. Кроме того, предлагается структура для конвергенции в Индустрии 4.0, включающая шесть измерений: соединение, сбор, коммуникация, вычисление, управление и создание. Результаты исследования согласуются с существующими дискуссиями и дебатами по Индустрии 4.0 и дополняют их, что подтверждает полезность и эффективность подхода, основанного на данных, в этой работе для поддержки взглядов экспертов на Индустрию 4.0. Результаты проведенной работы в [11] помогают установить общую основу для понимания Индустрии 4.0 с разных точек зрения, обеспечивая дальнейшие исследования и разработки для промышленных инноваций и социального прогресса.

*Примечание.* Последние разработки в производственных процессах и автоматизации привели в результате к новой промышленной революции, получившей название «Индустрия 4.0». Индустрию 4.0 можно рассматривать как обширную область, которая включает в себя: управление данными, конкурентоспособность производства, производственные процессы и эффективность. Термин «Индустрия 4.0» включает в себя множество ключевых технологий, таких как киберфизические системы, Интернет вещей, ИИ, анализ больших данных и цифровых двойников, которые могут рассматриваться как основные составляющие автоматизированных и цифровых производственных сред. Устойчивое развитие можно рассматривать как основу бизнес-стратегии, которая выделена в повестке дня Организации Объединенных Наций (ООН) в области устойчивого развития на период до 2030 года и включает интеллектуальное производство, энергоэффективные здания и индустриализацию с низким уровнем воздействия. Технологии Индустрии 4.0 помогают добиться устойчивости деловой практики [11-13]. Однако в очень ограниченных исследованиях сообщалось об обширных обзорах по этим двум областям исследований. В [12] используется метод систематического обзора литературы, чтобы выявить текущий прогресс в исследованиях и будущий исследовательский потенциал технологий Индустрии 4.0 для достижения устойчивости производства. Подробно обсуждается роль и влияние различных технологий Индустрии 4.0 на устойчивость производства. Результаты этого исследования предоставляют новые объемы исследований и будущие направления исследований в различных областях исследований Индустрии 4.0, которые будут полезны для промышленности и академических кругов для достижения устойчивости производства с помощью технологий Индустрии 4.0.

Были обнаружены [13] различные структуры или архитектуры, связанные с «умным» производством (*SM* – *smart manufacturing*) и интеллектуальным производством (*IM* – *intelligent manufacturing*), в которых предлагалось систематическое внедрение и стандартизация технологий *SM* и *IM*:

- 1) Экосистема *SM* (*SME*) от *NIST*, США;
- 2) Модель эталонной архитектуры *Industrie 4.0* (*RAMI 4.0*), разработанная «Платформой стандартизации и эталонной архитектуры» *Industrie 4.0*, Германия;
- 3) Архитектура *IMS* (*IMSA*), Министерство промышленности и информационных технологий (*MIIT*) Китайской Народной Республики и Управление стандартизации Китая (*SAC*);
- 4) Архитектура интеллектуальных систем для производства (*ISAM*) – эталонная модель архитектуры для *IMS*, разработанная *NIST*, США.

Эти четыре репрезентативных каркаса или эталонные архитектуры показаны на рис. 6.

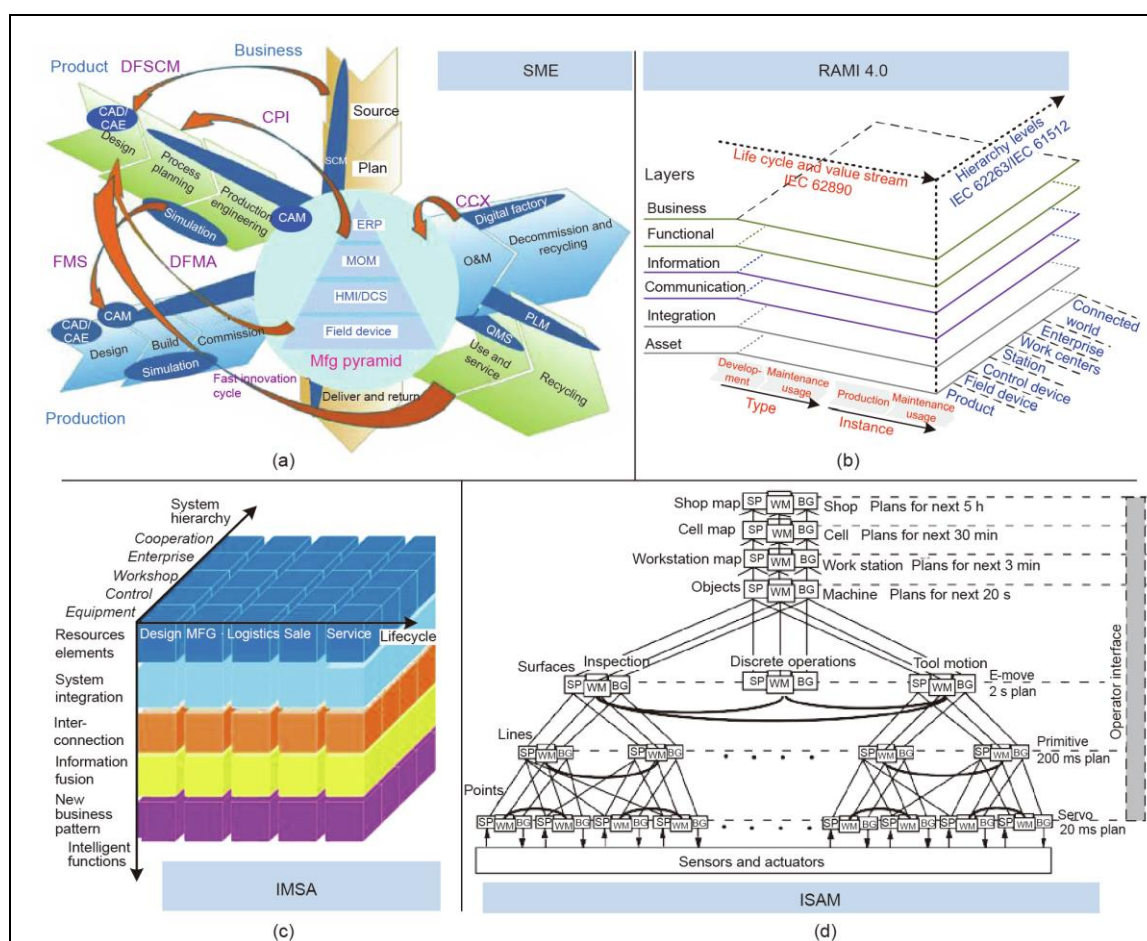


Рис. 6. Типичная структура / архитектура для *SM* (a) *SME*, (b) *RAMI 4.0* и для *IM* (c) *IMSA* и (d) *ISAM*

**Примечание.** На рис. 6 введены следующие обозначения: *CAE* – автоматизированная инженерия; *CAM* – автоматизированное производство; *CCX* – непрерывный ввод в эксплуатацию; *CPI* – постоянное совершенствование процесса; *DCS* – распределенные системы управления; *DFMA* – дизайн для изготовления и сборки; *DFSCM* – дизайн для управления цепочкой поставок; *ERP* – планирование ресурсов предприятия; *FMS* – гибкая производственная система; *HMI* – человеко-машинный интерфейс; *Mfg* – производство; *MOM* – управление производственными операциями; *O&M* – эксплуатация и обслуживание; *PLM* – управление жизненным циклом продукта; *QMS* – система менеджмента качества; *SCM* – управление цепочкой поставок; *BG* – формирование поведения; *SP* – сенсорная обработка; *WM* – моделирование мира.

Таким образом, *SM* / *IM* – это совместный продукт информационных технологий, технологий промышленного производства или операционных технологий (OT), человеческого интеллекта и творческих способностей, ведущих к быстрой эволюции производственной системы. Однако *SM* / *IM*

– это всего лишь один инструмент для достижения максимальной цели сокращения дефектов при повышении качества, улучшении производительности при снижении затрат, сокращении времени простоя за счет прогнозирования отказа до того, как он произойдет, минимизации потерь при повышении устойчивости и поддержании конкурентных преимуществ за счет понимания, накопления и применения знаний в предметной области – часть производственного процесса и систем.

*Примечание.* Быстрые технологические изменения, робототехника, большие данные, ИИ и связанные объекты открывают новые возможности для производителей. Но они также приносят свои вызовы обрабатывающей промышленности. Например, в Квебеке обрабатывающая промышленность находится под угрозой, что составляет почти 800 000 прямых и косвенных рабочих мест. Чтобы успешно перейти к Индустрии 4.0, производственная компания должна изучить новые требуемые навыки и потребность в квалифицированном персонале. Наиболее востребованные отраслевые навыки в 4.0 – это управление данными; безопасность данных; человеко-машинное взаимодействие; дизайн пользовательского интерфейса; разработка программного обеспечения; программирование; научные данные (*data science*); и аналитика данных.

Индустрия 4.0 ведет к эпохе цифровизации. Все становится цифровым: бизнес-модели, окружающая среда, производственные системы, машины, операторы, продукты и услуги. Все это взаимосвязано внутри цифровой сцены с соответствующим виртуальным представлением. В эпоху глобальных изменений гибкое мышление – ключевая компетенция, необходимая для обучения в среде с улучшенными технологиями. Благодаря этому интеллектуальный гибкий подход можно легко адаптировать для технологического усовершенствования исследования. Искусственный интеллект (ИИ), роботы, автоматизация и Интернет вещей (*IoT*) – четвертая промышленная революция – это термины, которые вошли в основное русло повседневных разговоров. Четвертая промышленная революция касается не только искусственного интеллекта, роботов или умных машин. Какими бы важными они ни были, Четвертая промышленная революция, в конце концов, касается не машин, а людей, как мы живем, учимся, зарабатываем и развиваем гибко бизнес-модели. По мере того, как Четвертая промышленная революция продолжает развиваться и разрушать существующие социальные и экономические системы, будущее труда будет все больше определяться использованием цифровых технологий не просто для вытеснения людей, а для расширения человеческих способностей и опыта.

Вышеупомянутый метод выбран для изучения будущих проблем и возможностей ИИ в промышленной революции в настоящем и будущем. Подход рекомендуется на этот раз, чтобы сделать эту работу правильно. Три технологических мегатренда являются основными движущими силами этой трансформации в производстве: возможности подключения, интеллектуальные возможности и гибкая автоматизация. На передовых производственных площадках, которые охватили эти мегатенденции в масштабе, произошел скачок в производительности.

Прочно уходящие корнями в Интернет вещей (*IoT*) и производство с использованием киберфизических систем, прорывные парадигмы, такие как «Умная фабрика будущего» и «Индустрия 4.0», предполагают наукоемкие промышленные интеллектуальные среды, в которых интеллектуальные персонализированные продукты создаются с помощью интеллектуальных процессов и процедур. Четвертая промышленная революция будет основана на киберфизических системах, которые будут отслеживать, анализировать и автоматизировать бизнес-процессы, преобразовывая производственные и логистические процессы в интеллектуальные производственные среды, в которых возможности больших данных, облачные сервисы и интеллектуальные инструменты поддержки принятия решений с прогнозированием используются для повышения производительности и эффективности. Этот подход дает представление о последних разработках в этих областях и определяет актуальные исследовательские задачи и возможности для формирования будущего интеллектуальных производственных сред. Следующее поколение Индустрии 4.0 обещает большую гибкость в производстве, а также массовую настройку, лучшее качество и повышенную производительность.

Таким образом, это позволяет компаниям, чтобы справиться с проблемами производства, производить все более индивидуализированной продукции с коротким сроком вывода на рынок и более высоким качеством. Интеллектуальное производство играет важную роль в Индустрии 4.0.

На рис. 7 ниже показан подход к интеграции искусственного интеллекта с Индустрией 4.0.

Искусственный интеллект (ИИ) – это когнитивная наука, которая позволяет человеку исследовать множество интеллектуальных способов моделирования наших процессов восприятия и рассуждения. Промышленный ИИ – это систематическая дисциплина, позволяющая инженерам системно

разрабатывать и внедрять алгоритмы ИИ с постоянным и неизменным успехом. ИИ в настоящее время привлекает значительный интерес и внимание со стороны промышленности, исследователей, правительств, а также инвесторов, которые вкладывают рекордные суммы денег в разработку новых технологий и приложений машинного обучения.



Рис. 7. Адаптированный интеллектуальный и гибкий подход к разработке Индустрии 4.0 [9]

Все более сложные алгоритмы используются для поддержки человеческой деятельности не только при прогнозировании задач, но и при принятии реальных решений, влияющих на общество, бизнес и отдельных лиц. Будь то в производственном секторе, где роботы приспосабливают свое поведение для работы вместе с людьми, или в домашней среде, где холодильники заказывают продукты питания на основе предпочтений домовладельца. ИИ постоянно вторгается в области, ранее предназначенные для человеческих навыков, суждений или принимать решение. Новаторские производители уже используют ИИ для решения этих многих проблем. Вот ключевые способы, которыми «Индустрия 4.0», последние тенденции в умных предприятиях, использует автоматизацию, обмен данными и новые технологии.

Промышленный ИИ обычно относится к применению ИИ в промышленности. В отличие от общего ИИ, который является передовой исследовательской дисциплиной для создания компьютеризированных систем, которые выполняют задачи, требующие человеческого интеллекта, промышленный ИИ больше озабочен применением таких технологий для решения промышленных проблем, связанных с созданием ценности для клиентов, повышением производительности и открытием новых идей. Хотя в антиутопическом видении приложений ИИ интеллектуальные машины могут лишать людей рабочих мест и вызывать социальные и этические проблемы, промышленность в целом более позитивно относится к ИИ и считает эту трансформацию экономики неудержимой и ожидает огромных возможностей для бизнеса в этом процессе. В результате ведущие компании сейчас ищут прорывные технологии для следующего горизонта повышения производительности. Многие начинают экспериментировать с такими технологиями, как цифровая связь между машинами (промышленный Интернет вещей или *IIoT*), искусственный интеллект (AI), машинное обучение, расширенная автоматизация, робототехника и аддитивное производство. Ожидается, что влияние этого сдвига будет настолько преобразующим, что его обычно и называют четвертой промышленной революцией или Индустрией 4.0.

Эта новая волна технологий и инноваций предлагает компаниям возможности не только для повышения производительности и эффективности, но и для достижения стратегической ценности бизнеса за счет создания конкурентных преимуществ в том, как они управляют всей цепочкой создания стоимости «от производства к доставке». Характер и масштаб возможностей будет варьироваться от сектора к сектору и от компании к компании, в зависимости от таких факторов, как факторы стоимости, динамика рынка и операционная зрелость. Однако регулярно видим, как успешные преобразования с помощью технологий резко меняют отдельные факторы стоимости. Например, производитель аэрокосмической отрасли, имеющий репутацию производителя высококачественной продукции, но страдающий от высоких затрат на рабочую силу и медленного производства, внедрил рабочие ин-



струкции с дополненной реальностью для сложных сборок, чтобы уменьшить количество совершаемых ошибок.

Как отмечалось ранее, три технологических мегатренда являются основными движущими силами этой трансформации в производстве: возможности подключения, интеллектуальные возможности и гибкая автоматизация. На передовых производственных площадках, которые охватили эти мегатенденции в масштабе, произошел скачок в производительности. Четвертая промышленная революция знаменует собой серию социальных, политических, культурных и экономических потрясений, которые развернутся в 21 веке.

Опираясь на широкую доступность цифровых технологий, которые стали результатом Третьей промышленной, или цифровой, революции, Четвертая промышленная революция будет в значительной степени обусловлена конвергенцией цифровых, биологических и физических инноваций. Адаптация и инновации жизненно важны для общества и для промышленности сегодня. ИИ обладает огромным потенциалом, чтобы помочь инженерам и исследователям в подходах, а также смягчить проблемы, вызванные жесткой конкуренцией, постоянным изменением предпочтений потребителей в отношении продуктов и бюджетными ограничениями. Производители постоянно ищут конкурентное преимущество, позволяя им оставаться на плаву в многолюдной отрасли, подверженной нестабильным условиям, например, экономическому спаду. Технологии на базе ИИ упрощают переговоры в режиме реального времени и взаимодействие с регулирующими организациями. Это, в свою очередь, способствует сокращению циклов разработки и ускорению вывода продукта на рынок.

«Умное» производство (*SM – smart manufacturing*) появляется как новая версия интеллектуального производства (*IM*), отражающая масштабы и влияние интеллектуальных технологий, таких как Интернет вещей, облачные вычисления, киберфизические системы и большие данные, на Индустрию 4.0. Глобализация и конкурентоспособность заставляют компании переосмысливать и обновлять свои производственные процессы в соответствии с так называемой парадигмой Индустрии 4.0. Он представляет собой интеграцию инструментов, которые уже использовались в прошлом (большие данные, облако, робот, 3D-печать, моделирование и т.д.), которые теперь подключены к глобальной сети путем передачи цифровых данных. Внедрение этой новой парадигмы представляет собой огромное изменение для компаний, которые сталкиваются с большими инвестициями. Чтобы воспользоваться возможностями интеллектуальной революции, компании должны иметь предпосылки, необходимые для того, чтобы противостоять изменениям, создаваемым интеллектуальной системой.

Кроме того, новые сотрудники, которые сталкиваются с миром труда в проекте «Индустрия 4.0», должны обладать новыми навыками в области автоматизации, оцифровки и информационных технологий, не забывая при этом о социальных навыках.

Умное производство имеет решающее значение для повышения качества обрабатывающей промышленности. В интеллектуальном производстве существует тенденция включать в анализ безопасности процессов различные виды информационных технологий нового поколения. В настоящее время экологически чистое производство сталкивается с серьезными препятствиями, связанными с управлением безопасностью, из-за использования большого количества опасных химикатов, что приводит к пространственной однородности химических промышленных процессов и все более строгим нормам безопасности и охраны окружающей среды. Новые информационные технологии, такие как ИИ, весьма перспективны в качестве средства преодоления этих трудностей. Основываясь на интеллектуальных и гибких методах ИИ и сложных отношениях безопасности в перерабатывающей промышленности, выявляются и обсуждаются несколько технических проблем, связанных с безопасностью процессов: получение знаний с использованием скудных этикеток для обеспечения безопасности процесса; обоснование безопасности процесса на основе знаний; точное слияние разнородных данных из разных источников; и эффективное обучение динамической оценке рисков и принятию решений. В этом контексте также обсуждаются текущие и будущие работы.

Использование «Индустрии 4.0» позволяет собирать и анализировать данные на разных машинах, обеспечивая быстрые, более гибкие и более эффективные процессы для производства товаров более высокого качества с минимально возможными затратами. Эта производственная революция, несомненно, повысит производительность, изменит экономику, будет способствовать промышленному росту и изменит профиль рабочей силы, в конечном итоге изменив конкурентоспособность компаний и регионов.

С развитием робототехники важность исследований в области взаимодействия между человеком и роботом (*HRC – human – robot collaboration*) стала расти. Чтобы роботы могли автономно взаимодействовать с людьми, им необходимо активное принятие решений с учетом человеческих партнеров. Тем не менее, современные исследования в HRC часто предполагают разделение «лидер – последователь», при котором взаимодействие возглавляет один агент. Это вызвано отсутствием надежного представления о человеке и окружающей среде, позволяющего принимать автономные решения. Эту проблему можно преодолеть с помощью воплощенного подхода к *HRC*, который вдохновлен психологическими исследованиями взаимодействия человека и человека (*HHI – human – human interaction*). Мы рассматриваем нейробиологические и психологические открытия сенсомоторных паттернов, управляющих *HHI*, и рассматриваем их в контексте робототехники. Кроме того, рассматриваем результаты, достигнутые сообществом роботов в направлении воплощенного HRC. Мы фокусируемся на механизмах, которые необходимы для активного физического взаимодействия человека и робота. Наконец, мы обсуждаем сходства и различия в двух областях исследования, которые определяют направления будущих исследований.

### Взаимодействие человека и робота (*HRI – human – robot interaction*)

В последние десятилетия взаимодействие человека и робота (HRI) стало важной областью робототехники. Поскольку у роботов и людей растет число рабочих мест, разработка безопасных и интуитивно понятных систем имеет большое значение. Физическая близость взаимодействующих партнеров создает проблему для интеграции действий человека в процесс принятия решений роботом. Обычно этой проблемы удастся избежать, следуя принципу «ведущий – ведомый», то есть создавая совместимую роботизированную систему, которая реагирует на действия человека и просто помогает человеку в достижении цели. Однако многие сценарии взаимодействия в повседневной жизни требуют активно сотрудничающих партнеров, которые могут автономно участвовать в общей задаче. Можем увидеть различия в уровне взаимозависимости и общих представлений, как показано на рис. 8, а. Взаимодействие между людьми и роботами можно охарактеризовать как обучение, сотрудничество или сотрудничество. На рис. 8, б изображены зависимости робота и человека в целенаправленном взаимодействии с течением времени для каждой из этих трех категорий

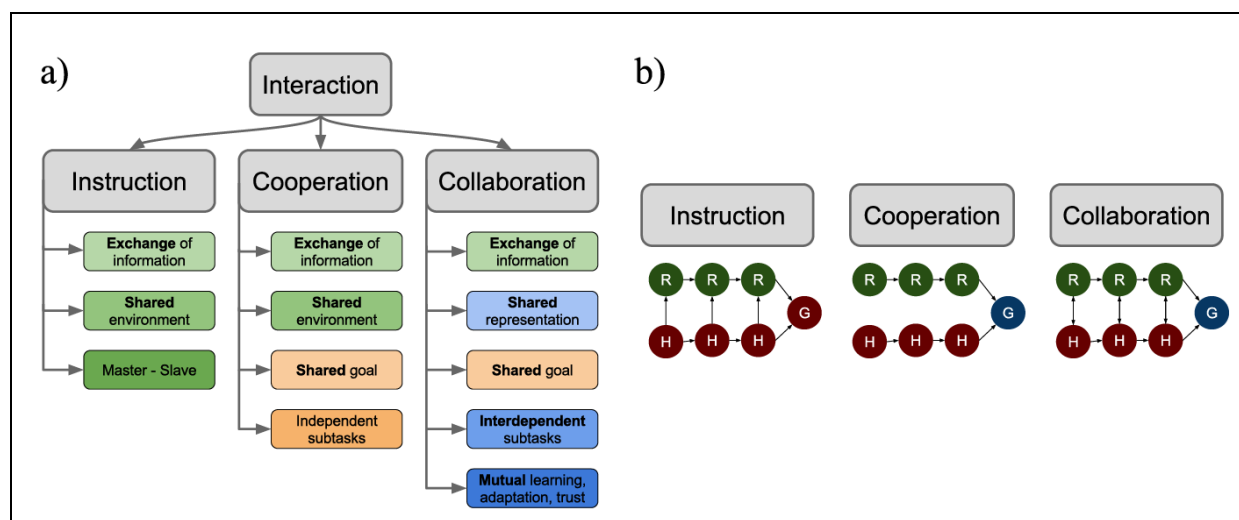


Рис. 8. (а) Три типа взаимодействия между людьми и роботами, степень взаимозависимости возрастает. (б) Взаимодействие между роботом (R, зеленый) и человеком (H, красный) для достижения человеческой (G, красный) или общей (G, синий) цели с течением времени. [Стрелки указывают зависимости] [14]

Исследования *HRI* включают проблемы, связанные с тем, как «понять и формируются взаимодействия между одним или несколькими людьми и одним или несколькими роботами». Разные среды и задачи требуют, чтобы робот демонстрировал разные уровни автономии, адаптации, способности к обучению и соответствия. В разработках стремятся решать проблемы, которые необходимо для обеспечения физического взаимодействия человека и робота. Робот не рассматривается в режиме «хозяин – раб», автономное поведение в социальном взаимодействии может быть достигнуто с помо-

стью сенсомоторных непредвиденных обстоятельств. Представление *socSMCs* в *HRC* обсуждается, как люди могут участвовать в невербальном общении и совместных действиях, полагаясь на постоянный обмен сенсомоторной информацией. Используя эти результаты создаются алгоритмы проверочных *SMC* и синхронизирующих *SMC* в *HRC* соответственно и эти механизмы могут привести к активному *HRC* и созданию методов оценки эффективности в задачах *HRC*.

## Аффорданс-исчисление и имитационное обучение

Робот может получить карту действий из собственного подхода позировать человеческому телу, произвольно выполняя действия и подражая человеку. Корреляции между этими двумя сущностями моделируются с помощью нейронных сетей. Считают, что подражание взрослым требует меньше усилий, чем подражание детям. Кроме того, детям с диагнозом расстройство аутистического спектра сложнее, чем обычно развивающимся детям. Таким образом, имитация человеческих поз зависит не только от отображения, но и от структуры демонстрации. По сравнению с подходами на основе корреляции, описанными выше, другой подход заключается в использовании примитивов динамического движения (*DMP*). *DMP* помещают проблему управления в структуру динамических систем в сочетании с непараметрической регрессией. Их можно использовать либо для изучения поведения на демонстрации человека, либо для приобретения примитивов интерактивного движения в совместной задаче

Подражание на высоком уровне часто ассоциируется с эксплуатацией таинств. А обряды относятся к ключевым концепциям воплощенного познания. Вместо кодирования объектов по их типу, например стул есть стул, они кодируются ассоциациями с теми действиями, которые они совершают. Стул, на котором можно сесть, просунуть под стол, залезть, чтобы добраться до верха полки и т. д. Поведение, основанное на предписаниях, следует за возможностями действий, которые предлагает окружающая среда. Эта концепция была принята в области робототехники, поскольку знания и представление мира сводятся к взаимодействию и исследованию окружающей среды. Несколько подходов к обучению навыкам обучения включали активное исследование случайностей действия-восприятия-эффекта, в основном в визуальной области, но также учитывалась тактильная обратная связь. Поскольку аффорданс-исчисление является общепринятым понятием, люди и роботы разделяют большинство данное исчисление. Таким образом, после изучения эти свойства можно использовать для имитации действий человека или для прогнозирования будущих действий. С другой стороны, предметы можно получить, наблюдая за манипуляциями человека.

В *HRC*, особенно во время физического взаимодействия, необходимо общее понимание общего рабочего пространства. Концепция приказов предлагает инструмент для установления этого понимания и позволяет командам людей и роботов работать для достижения общей цели. Подражание в этом контексте – это больше, чем обучение путем демонстрации, это естественный способ приобретения новых *SMC*. Необходимо изучать не только объектно-ориентированные действия, но также необходимо понимать социально важные действия и их интерпретации, такие как взгляд, мимика или рукопожатие. Однако эти сигналы зависят как от отправителя, так и от получателя, т.е. должны быть отнесены к категории синхронизирующих *SMC*, которые обсуждаются ниже.

Автономное, совместное поведение требует взаимодействия между различными механизмами, которые были представлены выше. Для достижения настоящего сотрудничества, а не инструктивного или кооперативного взаимодействия, важно, чтобы робот вел себя естественно для человека, чтобы могла происходить взаимная адаптация и обучение, чтобы направлять команду человек-робот к общим целям. Поскольку подавляющее большинство исследований в *HRC* сосредоточено либо на роботизированной реализации, либо на человеческом опыте взаимодействия, мы не знаем об исследованиях, в которых пытались объединить общее представление, прогнозирование и сигнализацию для достижения истинного сотрудничества.

*Пример: проблема автономного размещения плюшевого мишки на ранее невидимом стуле для сидения.* Для достижения этой цели опишем новый метод [15], позволяющий роботам представить сидящую позу медведя, физически моделируя виртуального гуманоидного агента, сидящего на стуле. Разработана роботизированная система использует систему *SE* для планирования движений робота-гуманоида, идущего к креслу, и движения всего тела, чтобы установить медведя, соответственно. Кроме того, чтобы справиться со случаями, когда стул не находится в доступной позе для размещения медведя, вводится структура взаимодействия человека и робота (*HRI*), в которой человек следует языковым инструкциям, данным роботом, чтобы повернуть стул и помочь сделать кресло доступным.

Метод реализуется с помощью робота-руки и робота-гуманоида. Откалибрована предложенная система с 3 стульями и тщательно протестирована на 12 ранее невидимых стульях как в доступных, так и в недоступных позах. Результаты показывают, что метод позволяет роботу автономно поставить плюшевого мишку на 12 невидимых стульях с очень высокой вероятностью успеха. Также было показано, что структура HRI очень эффективна в изменении доступности стула.

На рис. 9 показаны результаты эксперимента.

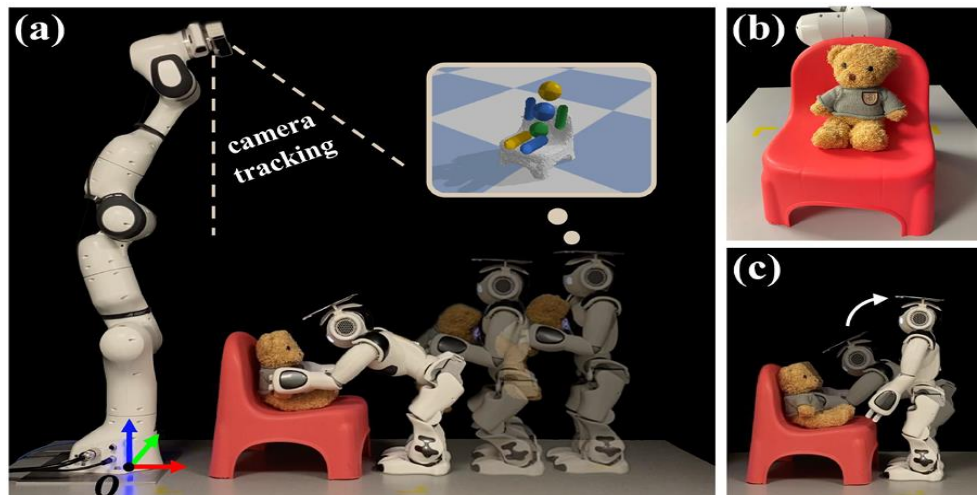


Рис. 9. (a) Робот представляет собой сидячую позу для ранее невидимого стула и усаживает медведя на стул. Рука робота с установленной камерой глубины сканирует кресло и отслеживает кресло и робота-гуманоида. Рамка в нижнем левом углу указывает базовую рамку манипулятора робота. (b) Результат посадки медведя на стул. (c) Гуманоидный робот поднимает руки и поднимает тело вверх после того, как посадил медведя на стул

Представим структуру взаимодействия человека и робота (HRI) с визуальной обратной связью с замкнутым контуром, чтобы робот мог проинструктировать человека повернуть стул и сделать стул доступным для размещения медведя.

На рис. 10 показан конвейер реализации метода, и его подробности.

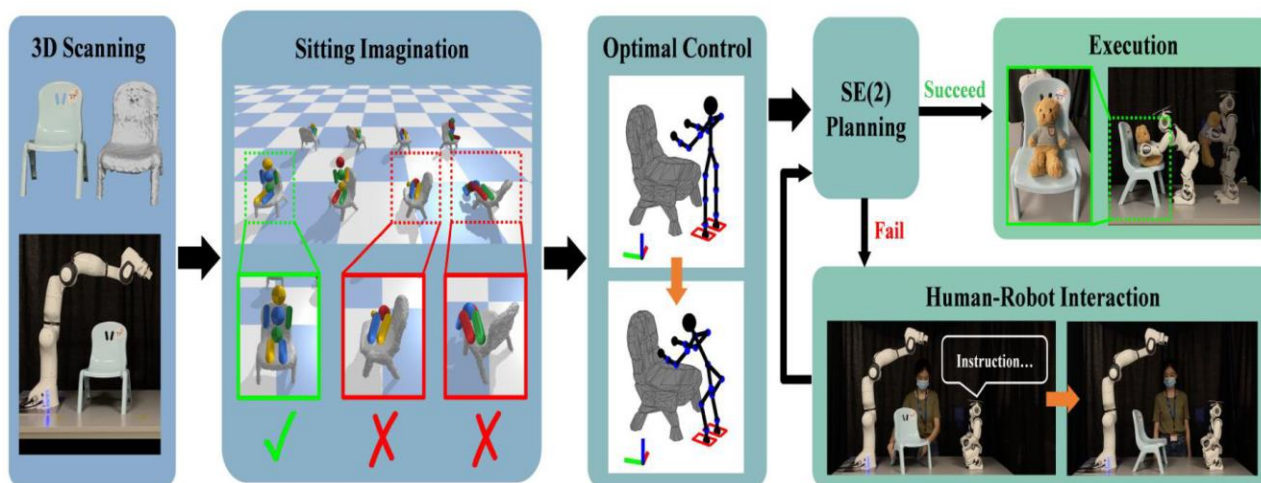


Рис. 10. Стул случайным образом размещается перед манипулятором робота для 3D-сканирования, которое воссоздает 3D-модель кресла. Три снимка в сидячем воображении – это результат трех сидячих испытаний. Только левый с контролем классифицируется как правильное сидение. Инструкция, которую дает робот при взаимодействии человека с роботом: «Пожалуйста, поверните кресло вокруг вертикальной оси на 90 градусов против часовой стрелки!»

Результаты показывают, что метод позволяет роботу автономно усадить медведя на 12 ранее невидимых стульях различной формы и внешнего вида в 72 испытаниях, включая как доступные, так и



недоступные позы стульев, с показателем успеха 94 – 4%. Кроме того, структура HRI оказалась очень эффективной. Он способен успешно изменить доступность стула для размещения медведя со 100% успешностью в 36 попытках с недоступными позами.

На рис. 11 приведено описание метода планирования движение робота.

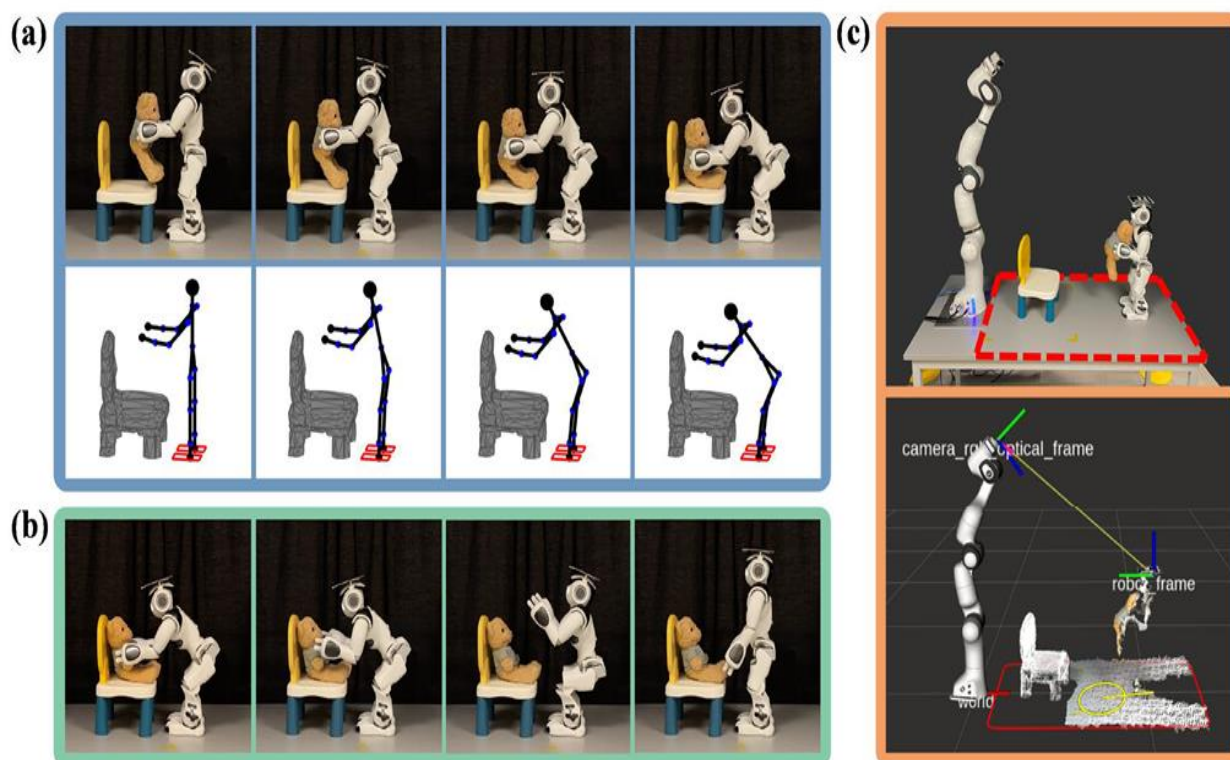


Рис. 11. Планирование движения. (а) Движение усаживания медведя на стул. (б) Уложив медведя, робот сначала отпускает медведя, разжимая его руки. Затем он поднимает руки и поднимает тело. (с) Настройка модуля планирования SE в реальной и виртуальной среде. Красные пунктирные линии на верхнем рисунке обозначают арену планирования в реальном мире. Желтый эллипсоид на нижнем рисунке показывает цель SE

Атрибуты объекта описывают потенциальные взаимодействия с объектом. Они играют важную роль в восприятии человеком объектов реального мира с огромным разнообразием внешнего вида. Возможности объекта очень связаны с функциональностью объекта. Например, стул может позволить на себе сидеть. Таким образом, он обладает возможностью сидения. Для роботов, работающих в неструктурированной среде и взаимодействующих с ранее невидимыми объектами, например персональными роботами, очень желательно понимание возможностей объектов. Он позволяет роботам получать доступ к невидимым объектам и понимать, как эффективно и разумно с ними взаимодействовать. Это дополнительно открывает возможность более естественного и интеллектуального взаимодействия человека и робота. Обсуждение аффордансов объекта также может помочь роботам раскрыть потенциал объекта для предоставления функциональности несмотря на то, что он не является типичным экземпляром класса объекта, с которым связана функциональность (например, импровизированный стул, собранный из книг и коробок на рис. 12 (d)). Самостоятельное определение аффордансов представляет собой более высокий уровень интеллекта роботов.



Рис. 12. Данные. (b) показывает захваченное изображение глубины, когда робот находится в позе захвата, показанной на (a). (c) показаны три позы захвата 3D-сканирования. (d) показывает все стулья в наборе данных. Среди стульев в испытательном наборе два из них, указанные в красном поле, не являются типичными стульями, но могут позволить себе сидячую возможность. Один из них - табурет; другой импровизирован, собирая книги и коробки

Метод реализован с помощью операционной системы роботов (ROS) на компьютере с процессором Intel Core i9-10920X с тактовой частотой 3,5 ГГц. Неоптимизированная реализация занимает около 70 с, 4 с и 29 с для 3D-сканирования (включая захват и реконструкцию модели), воображение сидя и планирование движения (включая планирование движения всего тела и SE) соответственно. Ходьба и постановка медведя занимают примерно 67 и 24 секунды соответственно. В серии экспериментов стул размещается так, чтобы он был доступен для сидения медведя. Направление сидения указывает на направление с отклонением в пределах 30 градусов. В этом случае HRI не нужны. Для каждого стула в тестовом наборе медведя помещали в 3 разные позы для тестирования, в результате чего в общей сложности было проведено 36 испытаний (рис. 13).

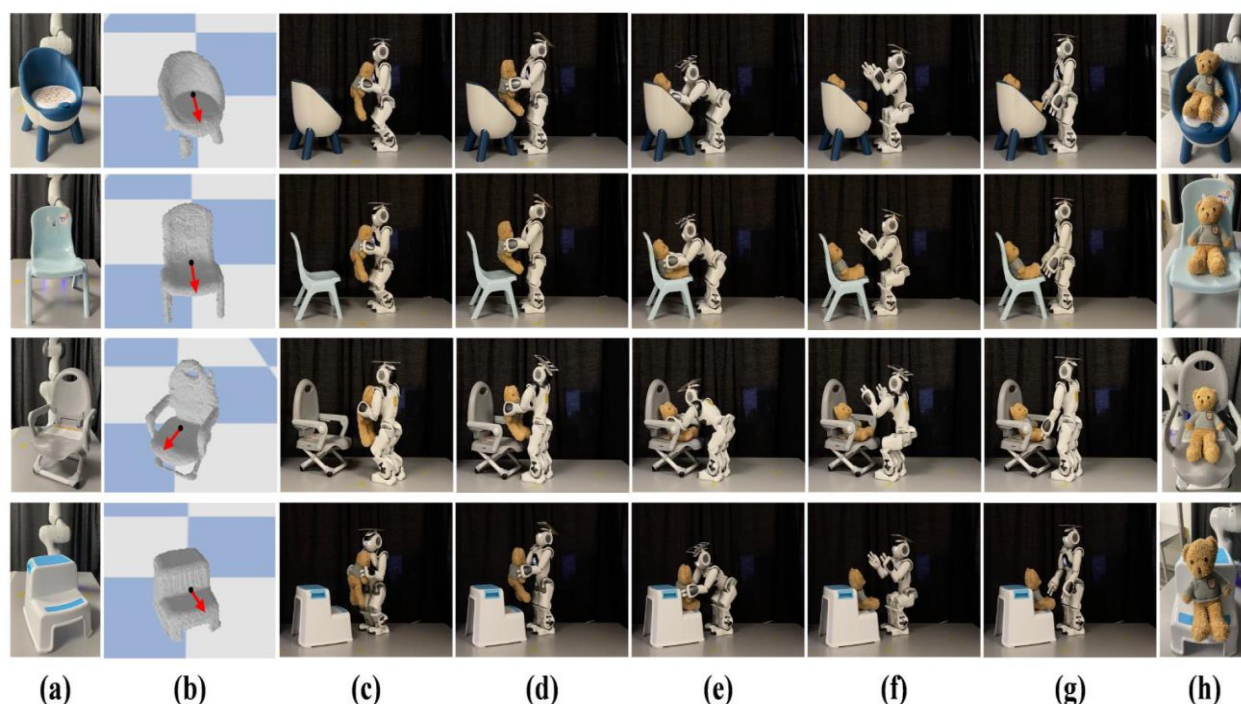


Рис. 13. Результат: доступный. (a) Снимок стула. (b) Воображаемая сидячая поза. (c) Робот подходит к стулу. (d) и (e) показывают начало и конец движения при посадке медведя на стул. (f) и (g) показывают движение извлечения рук и подъема тела в вертикальное положение после размещения медведя. (h) Результаты эксперимента



Есть много неопределенности в *HRI* (например, человек отвлекается или неправильно понимает инструкцию), что приведет к недоступности даже после взаимодействия. В этом наборе экспериментов проверялась надежность метода при устранении этих неопределенностей. Для каждого стула в тестовом наборе его помещали в недоступную позу, как в настройке «Недоступность + Подчинение человека» (всего 12 испытаний). Человек сознательно не подчинялся первой инструкции, данной инструкции, и подчинялся следующим инструкциям (рис. 14).

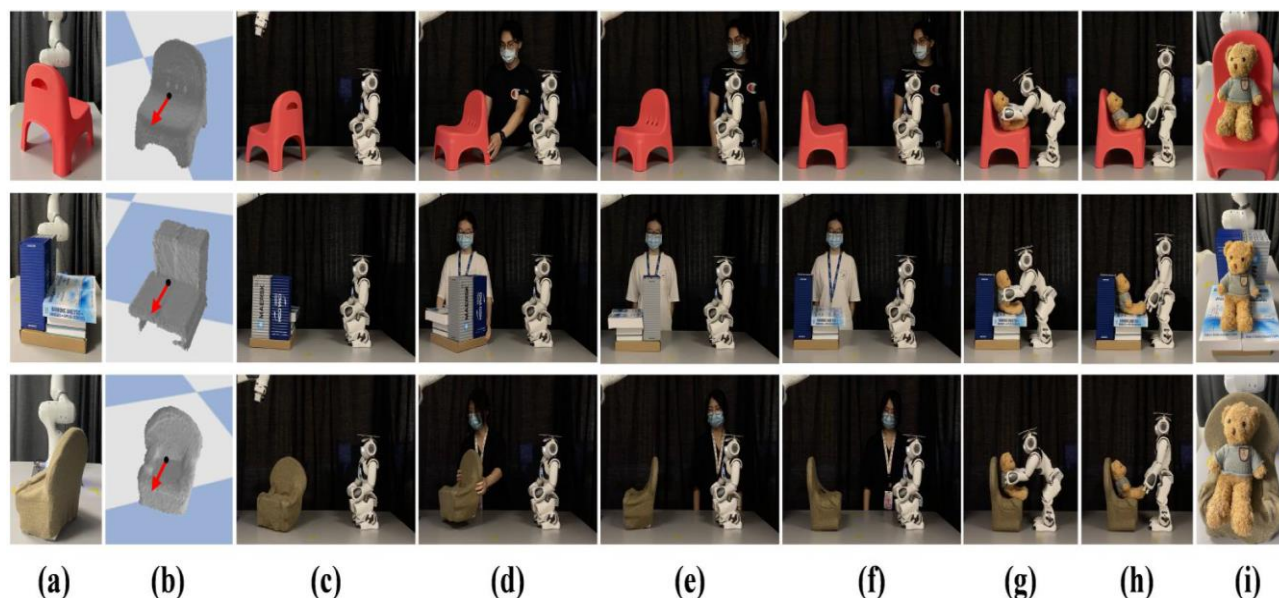


Рис. 14. Ситуация «Недоступно + Человеческое неповиновение». (а) Снимок стула. (б) Воображаемая сидячая поза. (с) До *HRI*. (д) Человек не подчиняется инструкциям в первой *HRI*. (е) После первой *HRI*. (ф) После последней *HRI*. (г) и (h) показывают начало и конец посадки медведя на стул. (и) Результаты эксперимента

В целом, можно достичь очень высокого уровня успеха 94: 4% на 12 невидимых стульях в тестовом наборе (68 успешных испытаний из всех 72 испытаний). Показатели успеха трех серий экспериментов примерно одинаковы. Стулья-ступенька приходится на все 4 случая неудач в 72 испытаниях. Два в ситуации «Доступные» и один в Недоступные + Человеческое подчинение» и «Недоступное + Человеческое неповиновение» соответственно. Во всех этих неудачных случаях воображение сидя не может найти сидячую позу, потому что глубина сиденья слишком мала. Успешное испытание шагового стула показано на рис. 13.

## Промышленный объяснительный искусственный интеллект (ХАИ) в Индустрии 4.0

Традиционная производственная парадигма крупносерийного производства не обеспечивает гибкости для удовлетворения требований отдельных клиентов. Ожидается, что новое поколение «умных» заводов будет поддерживать новые разнообразные и мелкосерийные индивидуальные режимы производства. Для этого ИИ обеспечивает производство с более высокой добавленной стоимостью за счет ускорения интеграции производственных и информационных коммуникационных технологий, включая вычисления, связь и управление. Характеристики индивидуализированной интеллектуальной фабрики включают в себя самовосприятие, оптимизацию операций, динамическую реконфигурацию и интеллектуальное принятие решений. Технологии ИИ позволят производственным системам воспринимать окружающую среду, адаптироваться к внешним потребностям и извлекать знания о процессах, включая бизнес-модели, такие как интеллектуальное производство, сетевое сотрудничество и модели расширенных услуг.

Необходима подробная информация об интеллектуальных производственных устройствах, интеллектуальном информационном взаимодействии и создании гибкой производственной линии. Рассмотрены современные технологии ИИ, которые могут быть использованы в «умных» производ-

ствах, например, машинное обучение, многоагентные системы, Интернет вещей, большие данные и облачные вычисления. Технологии на базе ИИ на индивидуальной умной фабрике проверяются на примере индивидуальной упаковки. Экспериментальные результаты продемонстрировали, что производство по индивидуальному заказу с использованием ИИ предлагает возможность более высокой производственной гибкости и эффективности.

### Индивидуальное производство (СМ) на основе искусственного интеллекта

ИИ охватывает теории, методы, технологии и приложения для увеличения человеческого интеллекта. Он включает в себя не только методы ИИ, такие как восприятие, машинное обучение (*ML – machine learning*), глубокое обучение (*DL – deep learning*), обучение с подкреплением и принятие решений, но и приложения с поддержкой ИИ, такие как компьютерное зрение, обработка естественного языка, интеллектуальные роботы и системы рекомендаций, как показано на рис. 15, а. ML превзошел традиционные статистические методы в таких задачах, как классификация, регрессия, кластеризация и извлечение правил. Типичные алгоритмы машинного обучения включают дерево решений, опорные векторные машины, регрессионный анализ, байесовские сети и глубокие нейронные сети.

Поскольку технологии ИИ продемонстрировали свой потенциал в таких областях, как проектирование изделий по индивидуальному заказу, изготовление изделий по индивидуальному заказу, управление производством, производственное обслуживание, управление клиентами, логистика, послепродажное обслуживание и анализ рынка, как показано на рис. 15, промышленное производство практики и исследователи приступили к их внедрению.

Метрики производительности системы, например гибкость, эффективность, масштабируемость и устойчивость, могут быть улучшены за счет внедрения технологий ИИ, таких как машинное обучение, графы знаний и взаимодействие человека с компьютером (*HCI*). Это особенно верно при обнаружении, взаимодействии, оптимизации ресурсов, эксплуатации и техническом обслуживании умной фабрики СМ. Поскольку парадигмы облачных вычислений, периферийных вычислений и локальных вычислений имеют свои уникальные сильные стороны и ограничения, их следует интегрировать, чтобы максимизировать их эффективность.

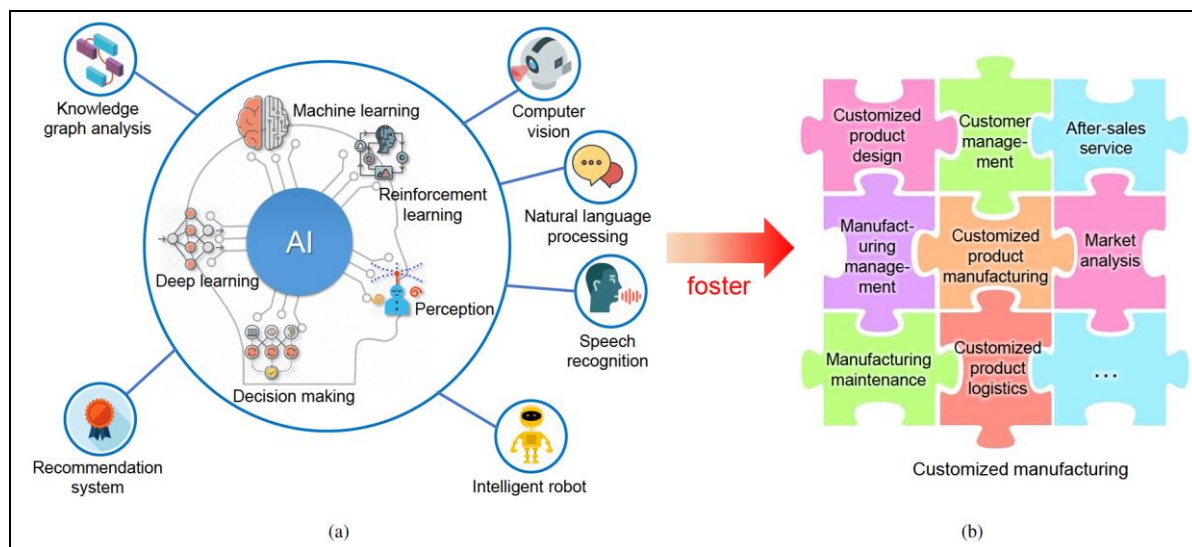


Рис. 15. ИИ и изготовление по индивидуальному заказу. (а) Технологии ИИ включают восприятие, машинное обучение, глубокое обучение, обучение с подкреплением и принятие решений, а также приложения с поддержкой ИИ, такие как компьютерное зрение, обработка естественного языка, интеллектуальные роботы и системы рекомендаций. (б) ИИ может способствовать индивидуальному производству в следующих аспектах: индивидуальный дизайн продукта, индивидуальное производство продукции, техническое обслуживание производства, управление клиентами, логистика, послепродажное обслуживание и анализ рынка

В то же время соответствующие алгоритмы ИИ должны быть переработаны, чтобы соответствовать соответствующей вычислительной парадигме. Облачный интеллект отвечает за всесторонний, не зависящий от времени анализ и решения, в то время как интеллектуальный анализ граничных и ло-



кальных узлов применим к контексту или средам с учетом времени. Интеллектуальные производственные системы включают интеллектуальное производство устройств, реализовывать интеллектуальное информационное взаимодействие и предоставлять интеллектуальные производственные услуги путем объединения технологий искусственного интеллекта.

Как показано на рис. 16, структура СМ на основе ИИ, которая включает интеллектуальные устройства, интеллектуальное взаимодействие, уровень ИИ и интеллектуальные сервисы.

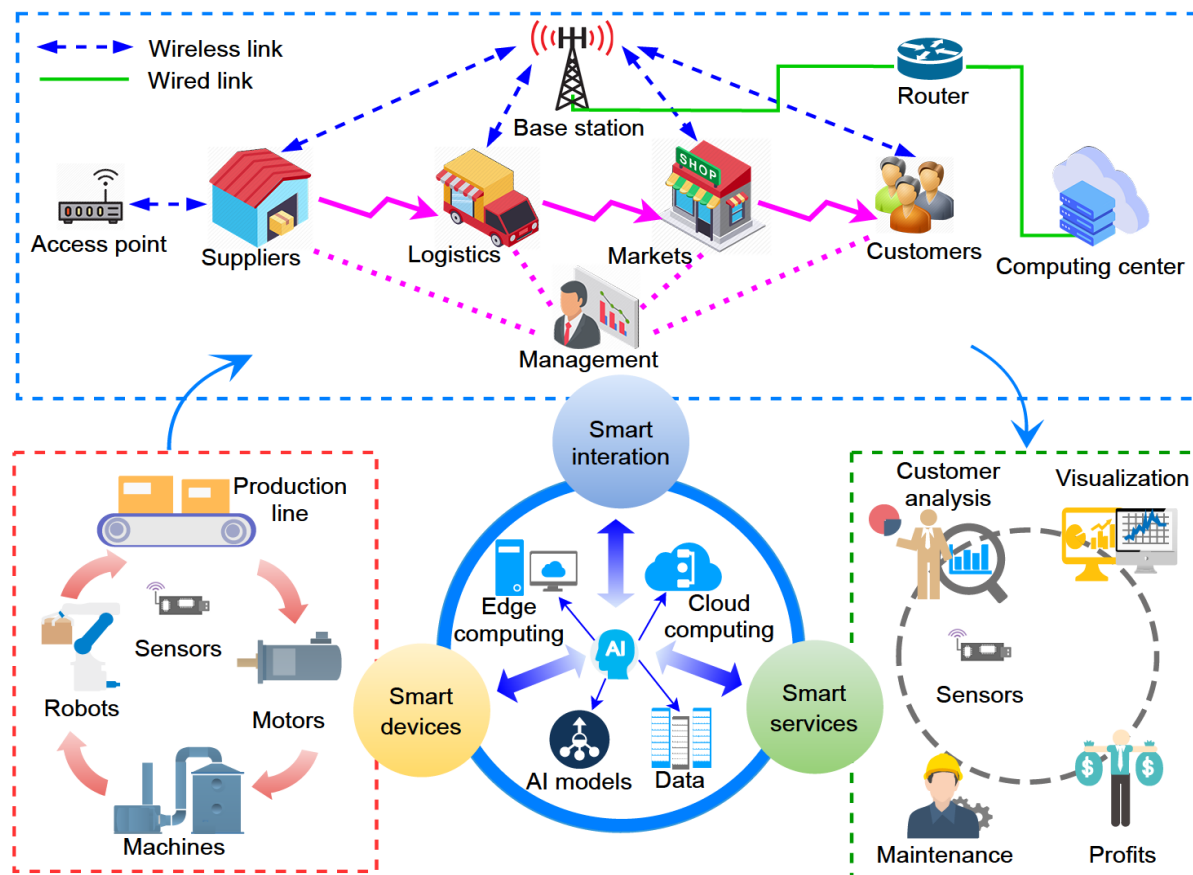


Рис. 16. Архитектура индивидуализированного производства с использованием ИИ включает интеллектуальные устройства, интеллектуальное взаимодействие, уровень ИИ и интеллектуальные услуги

Сотрудничество и взаимодействие между несколькими агентами необходимо для динамического создания групп сотрудничества для выполнения индивидуальных производственных задач. Несколько агентов с периферийными вычислениями предоставляют лучший вариант, чем одно устройство, для создания совместной операции для реализации СМ. Следовательно, комбинируя интеллектуальные агенты с помощью граничных вычислений и различные алгоритмы искусственного интеллекта, можно построить новую совместную операцию, как показано на рис. 17.

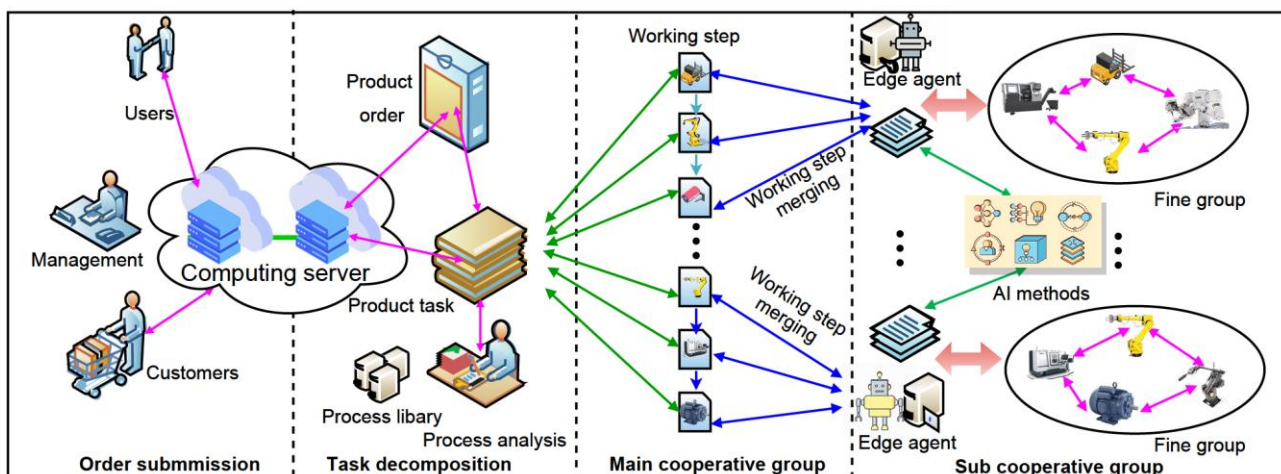
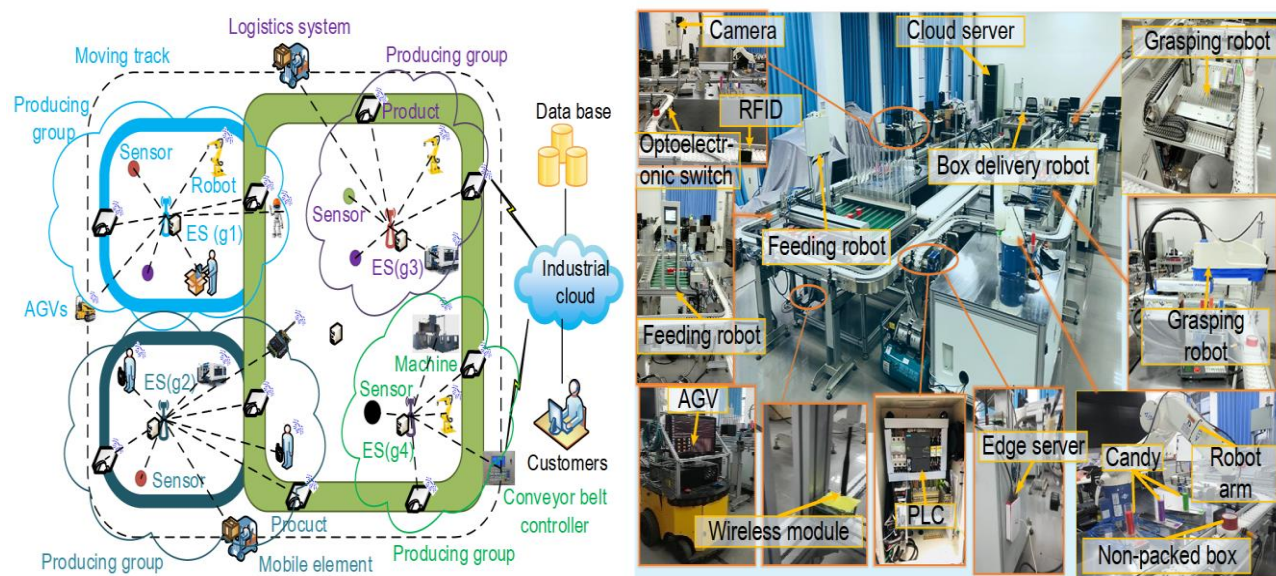


Рис. 17. Совместное использование нескольких агентов. Стратегию кооперативного взаимодействия нескольких агентов можно разделить на i) порядок представления, ii) декомпозицию задач, iii) кооперативную группу и iv) назначение

Прототипом каркаса СМ с использованием ИИ может быть рассмотрен пример, а именно производственной линии по производству упаковки для конфет.

Как показано на рис. 18 а, структура включает следующие компоненты: устройства СМ, промышленную сеть, конвейер и киберфизическую систему.



(а) Рамки для тематического исследования

(б) Сценарий внедрения СМ

Рис. 18. Структура интеллектуального производства индивидуальной упаковки конфет. (а) Структура включает устройства СМ, промышленную сеть, конвейер и киберфизическую систему. (б) Реализация линии по упаковке конфет по индивидуальному заказу состоит из различных устройств

Все компоненты связаны через промышленную сеть, т. е. через унифицированную архитектуру OPC (OPC UA) и службу распределения данных (DDS). Рис. 18, б иллюстрирует реализацию линии для упаковки конфет по индивидуальному заказу. Линия упаковки конфет в основном включает производственные станции и логистическую транспортную систему. В системе передачи логистики упаковочная коробка непрерывно перемещается ленточными конвейерами или грузовыми автомобилями. Производственные станции дискретно распределяются между магистралью и ответвлением, и RFID-метки используются для получения оперативной информации. Типы оборудования производственных участков включают подачу техники, захват конфет, доставку коробок, склад готовой продукции. Представленная система соответствует требованиям мелкосерийного производства. В част-

ности, упакованные конфеты соответствовали вкусу, аромату и цвету, заданным покупателями. Система включает четыре уровня, каждый из которых связан промышленным Интернетом вещей с различными функциями связи [16].

Производители робототехники лидируют в применении технологии ИИ, применяя аналитику данных на основе ИИ для повышения эффективности, качества продукции и безопасности сотрудников. Но они также сталкиваются с проблемами, связанными с более короткими сроками вывода на рынок, все более сложными продуктами и строгими правилами, и стандартами качества. Однако подавляющему большинству производственных компаний приходится преодолевать другие препятствия [17], мешающие цифровой трансформации и инициативам в области ИИ:

1. *Нехватка талантливых специалистов ИИ.* Опытных специалистов по обработке больших данных и ИИ мало. Для проектов ИИ требуется междисциплинарная команда специалистов по данным, инженеров машинного обучения, архитекторов программного обеспечения, аналитиков бизнес-аналитики и представителей малого и среднего бизнеса. Эта потребность особенно очевидна в производстве - секторе, который многие молодые специалисты по данным считают монотонным, повторяющимся и не стимулирующим. Эта проблема усугубляется тем, что производство, как ожидается, столкнется с серьезной нехваткой рабочей силы в течение следующих 10 лет, когда специалисты выйдут на пенсию. «AI Automation» и «AutoML 2.0» – важные технологии, которые могут устранить этот пробел в навыках и ускорить цифровую трансформацию в производстве.

2. *Технологическая инфраструктура и функциональная совместимость.* Производственные предприятия часто имеют множество машин, инструментов и производственных систем, в которых используются разные, а иногда и конкурирующие технологии, некоторые из которых могут работать на устаревшем программном обеспечении, несовместимом с остальной частью их системы. В отсутствие стандартов и общих рамок инженеры предприятий должны определить наилучший способ подключения своих машин и систем, а также какие датчики или преобразователи установить.

3. *Качество данных.* Доступ к «чистым», информативно значимым и высококачественным данным имеет решающее значение для успеха инициатив в области ИИ, но может стать проблемой на производстве. Производственные данные часто субъективно предвзяты, устаревшие и полны ошибок, которые могут быть вызваны множеством факторов. Одним из примеров являются данные датчиков, собранные на производственном участке в экстремальных, суровых условиях эксплуатации, когда экстремальные температуры, шум и вибрация могут давать неточные данные. Исторически заводы строились с использованием многих проприетарных систем, которые не взаимодействуют друг с другом, где рабочие данные также могут быть распределены по нескольким базам данных в нескольких форматах, не подходящих для аналитики, требующих обширной предварительной обработки.

4. *Принятие решений в режиме реального времени.* Это становится все более важным в производственных приложениях, таких как мониторинг качества, соблюдение сроков поставки клиентам и т. д. Часто решения необходимо принимать немедленно - в течение нескольких секунд - для выявления проблемы, прежде чем она приведет к незапланированным отключениям, дефектам или проблемам безопасности. Для быстрого принятия решений требуются услуги потоковой аналитики и прогнозирования в реальном времени, которые позволяют производителям действовать незамедлительно и предотвращать нежелательные последствия.

5. *Пограничные развертывания.* Существует множество потенциальных вариантов использования периферийных вычислений в производстве, позволяющих производителям обрабатывать данные локально, фильтровать данные и сокращать объем данных, отправляемых на центральный сервер, либо на месте, либо в облаке. Кроме того, ключевая цель современного производства - иметь возможность использовать данные с нескольких машин, процессов и систем для адаптации производственного процесса в режиме реального времени. Этот точный мониторинг и управление производственными активами и процессами использует большие объемы данных, требует машинного обучения для определения наилучшего действия в результате анализа данных, а также требует периферийных вычислений. Возможность развертывания прогнозных моделей на периферийных устройствах, таких как машины, локальный шлюз или сервер, критически важна для работы приложений интеллектуального производства.

6. *Доверие и прозрачность.* Существенным препятствием для широкого внедрения ИИ является сложность технологии и недоверие производителей к ее возможностям. Люди без опыта в области науки о данных пытаются понять, как работает наука о данных и прогнозное моделирование, и не

верят в абстрактные алгоритмы, лежащие в основе технологии ИИ. Большая прозрачность предоставит информацию о процессе ИИ – используемых входных данных, какие алгоритмы были выбраны и как модель делала свои прогнозы.

В то время как большинство ИИ традиционно использует модели «черного ящика», новые подходы к науке о данных обеспечивают большую прозрачность всего конвейера ИИ. Это включает в себя понимание подробного процесса преобразования необработанных данных во входные данные машинного обучения (также известного как разработка функций) и того, как модель машинного обучения производит прогнозы, комбинируя сотни или даже больше функций. Предоставляя понимание того, как работают модели прогнозирования и обоснования прогнозов, производственные организации могут укрепить доверие к моделям и полученным в результате бизнес-идеям.

Хотя проблемы с внедрением ИИ все еще существуют, предоставление производителям возможности делать больше с меньшими затратами на автоматизацию ИИ – это верный способ ускорить цифровую трансформацию. Автоматизация ИИ помогает производственным компаниям сокращать расходы, повышать эффективность и решать новые проблемы.

Исходя из идеи, что множество связанных интеллектуальных технологий может объединить (и таким образом произвести революцию) производство [7], был введен термин «Индустрия 4.0», который включает в себе обещание свершения технологически новой промышленной революции, четвертой такой трансформации в истории промышленных революций производства (см. рис. 19).

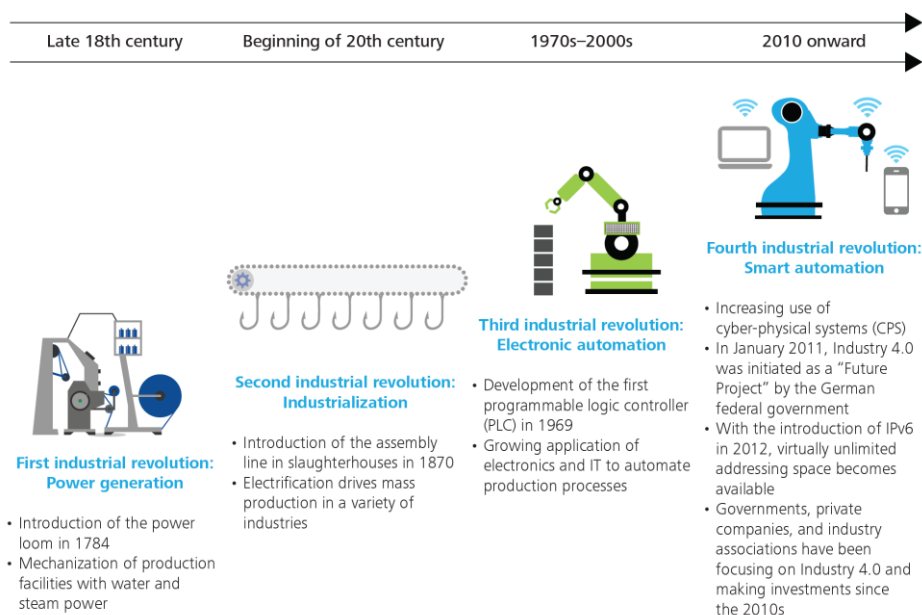


Рис. 19. История промышленных революций: эволюция отрасли с ключевыми событиями

**Примечание.** Определение Индустрии 4.0 было впервые представлено в 2011 году на выставке (Hannover Messe ярмарка), и была предметом рабочей группы Индустрии 4.0, созданной федеральным правительством Германии. *Germany Trade and Invest (GTAI)* определяет Индустрию 4.0 как: «Смена парадигмы стало возможным благодаря техническому прогрессу, который представляет собой обратную логику обычного производственного процесса. Проще говоря, это означает, что промышленное производственное оборудование не просто «обрабатывает» продукт, но и что продукт взаимодействует с оборудованием, чтобы точно сказать ему, что делать»

Чтобы прояснить свою концепцию Индустрии 4.0, *GTAI* использует концепцию киберфизических систем (CPS) – технологий, объединяющих цифровой и физический миры, обычно через датчики, прикрепленные к физическим устройствам, и сетевые технологии, которые собирают результирующие данные (см. рис. 20).



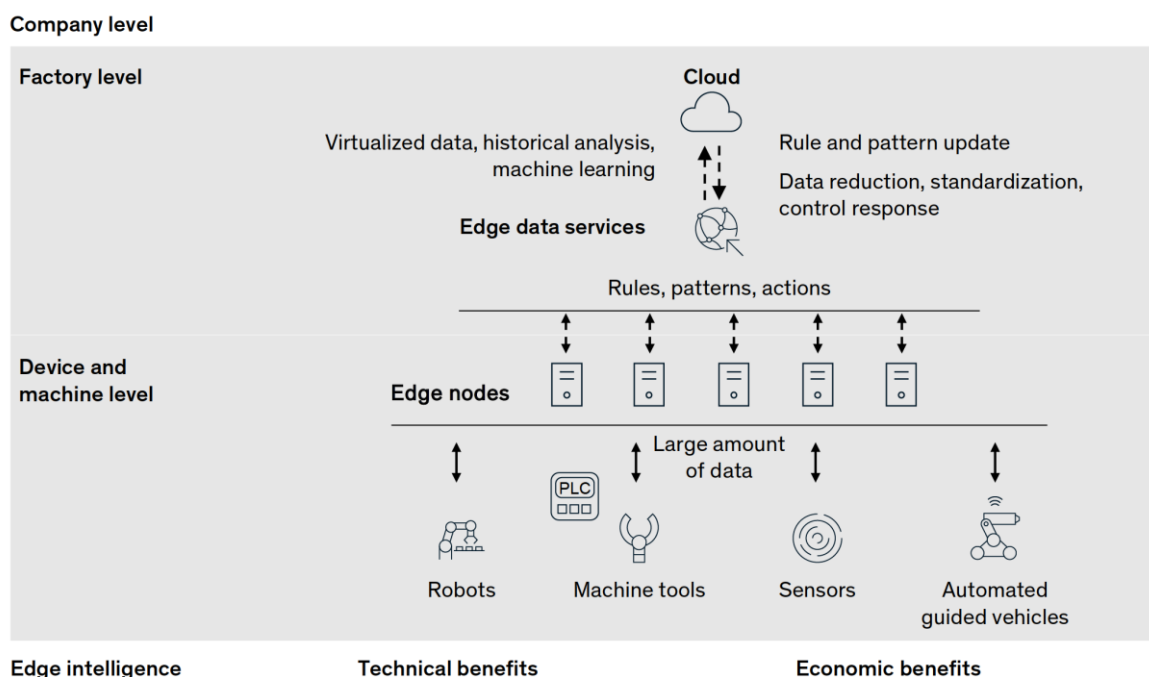
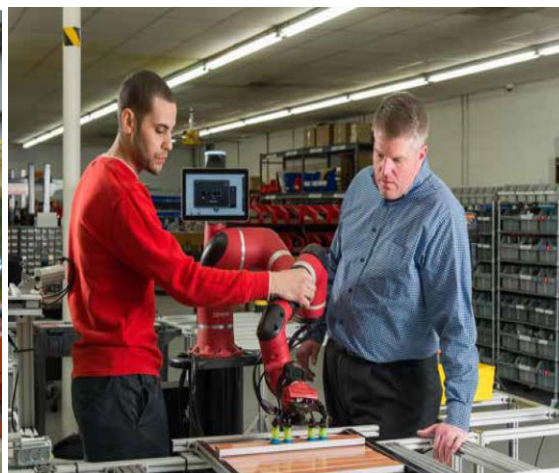
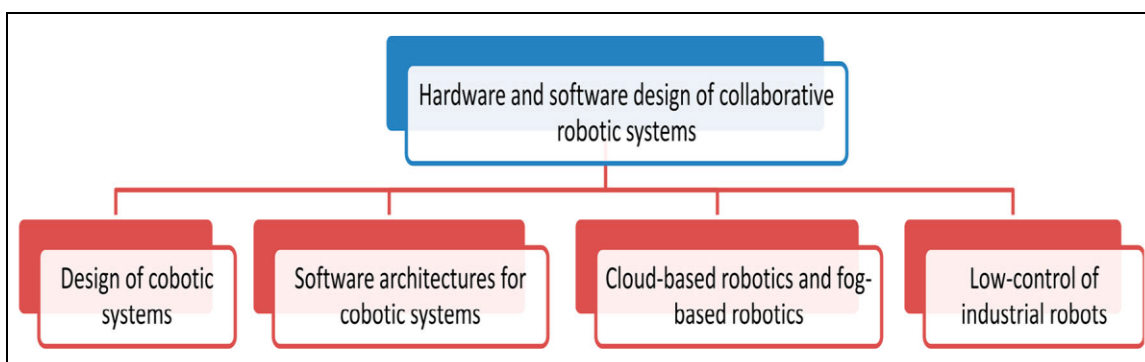
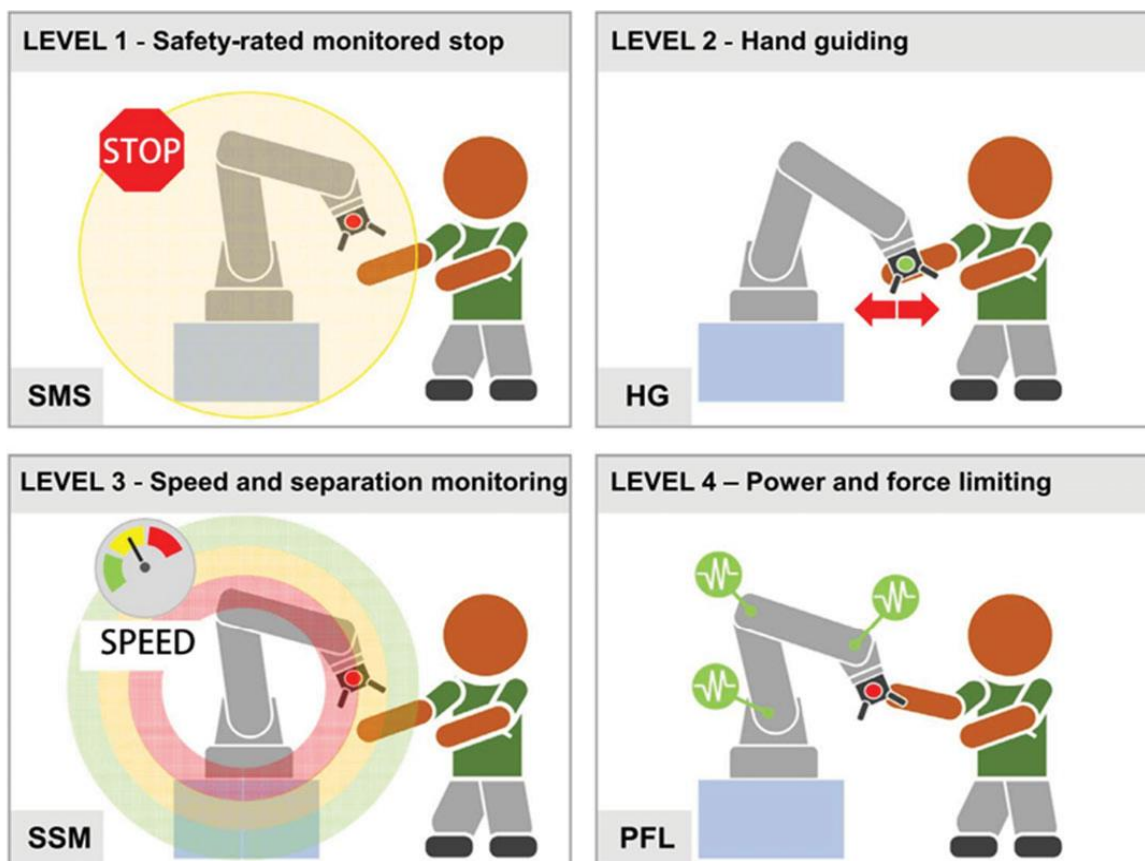


Рис. 20. Концепция построения модуля в проекте «Индустрия 4.0» компании GTAI. Эта концепция очень похожа на более часто упоминаемый Интернет Вещей (IoT)

Архитектура будущего *IIoT* использует преимущества периферийных вычислений с четким распределением задач на периферии по сравнению с облаком. Роботы играют важную роль в Индустрии 4.0. В зависимости от требований это может быть роботизированное вооружение, целая сборочная линия, марсоход, андроид или патрульные роботы на ногах. Многие уже можно увидеть в области химии, переработка мусора, фармацевтическое производство, производство продуктов питания и напитков и т.д. Вот несколько примеров. *KUKAiwa* – легкий робот для ответственных промышленных задач, разработанный *KUKA Robotics*. *Baxter* от *Rethink Robotics* – это интерактивный производственный робот для упаковки. *BioRob Arm* может быть использован в непосредственной близости и взаимодействия с людьми. Автоматизированное модульное оборудование предназначено для выполнения повторяющихся действий с высокой скоростью и точностью [18], а также для работы там, где человеческие возможности рабочих ограничены (см. примеры на рис. 21).



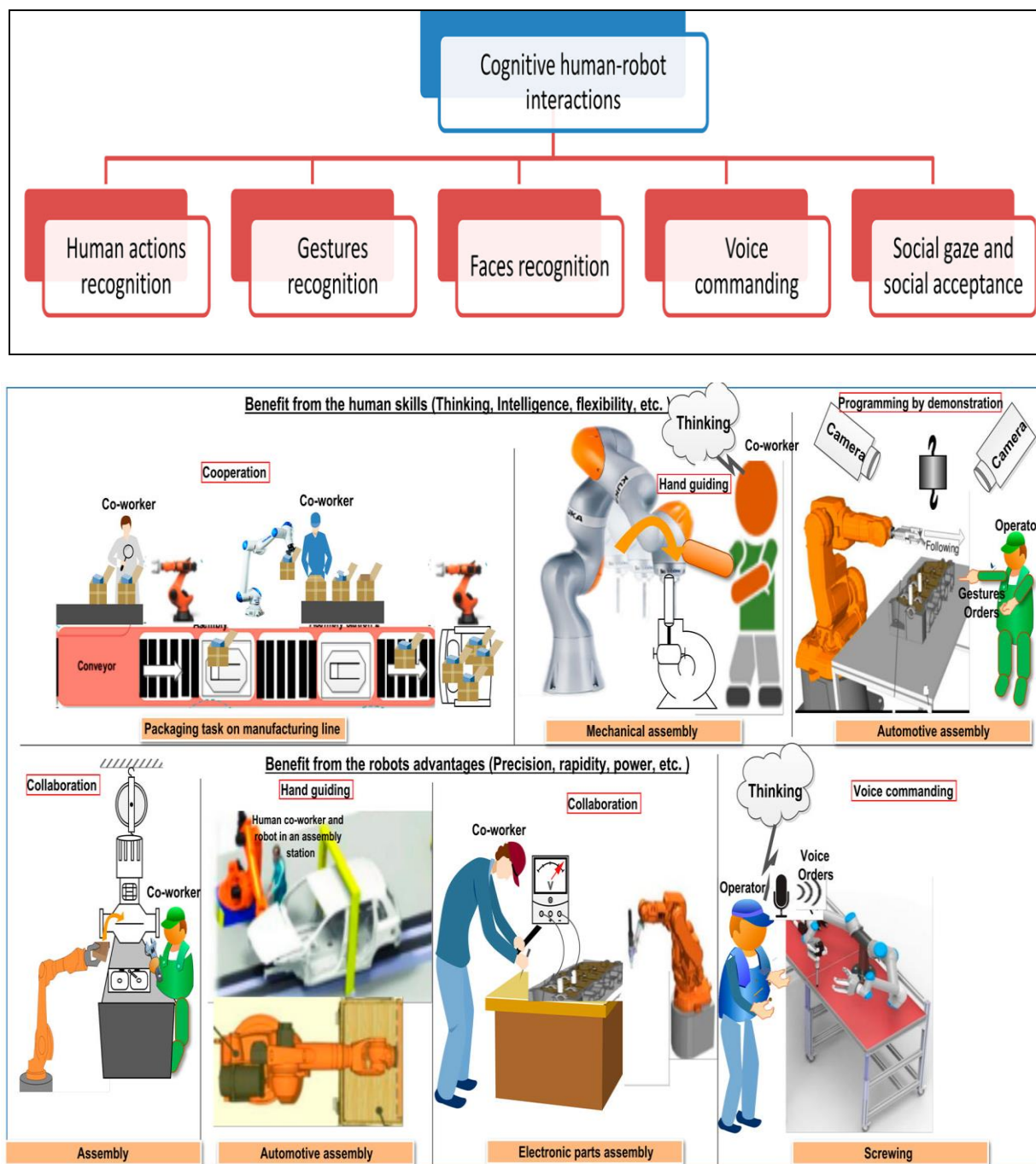


Рис. 21. Примеры роботизированных модулей

Роботизированные социотехнические системы оснащаются роботами, способными совершать сложные действия, реализуя движения нелинейной механики.

На рис. 22 – 26 приведены примеры подобного рода движений с заданной точностью [19 – 25].

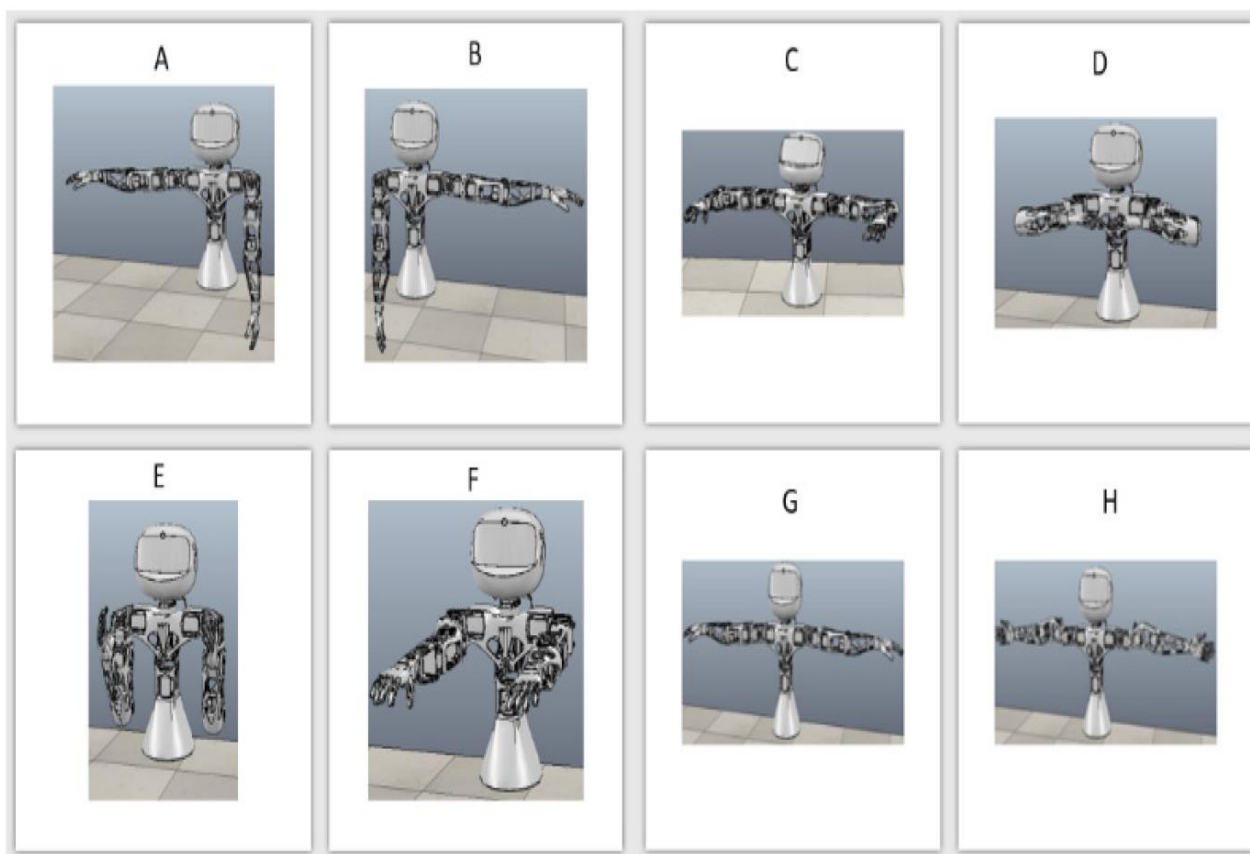
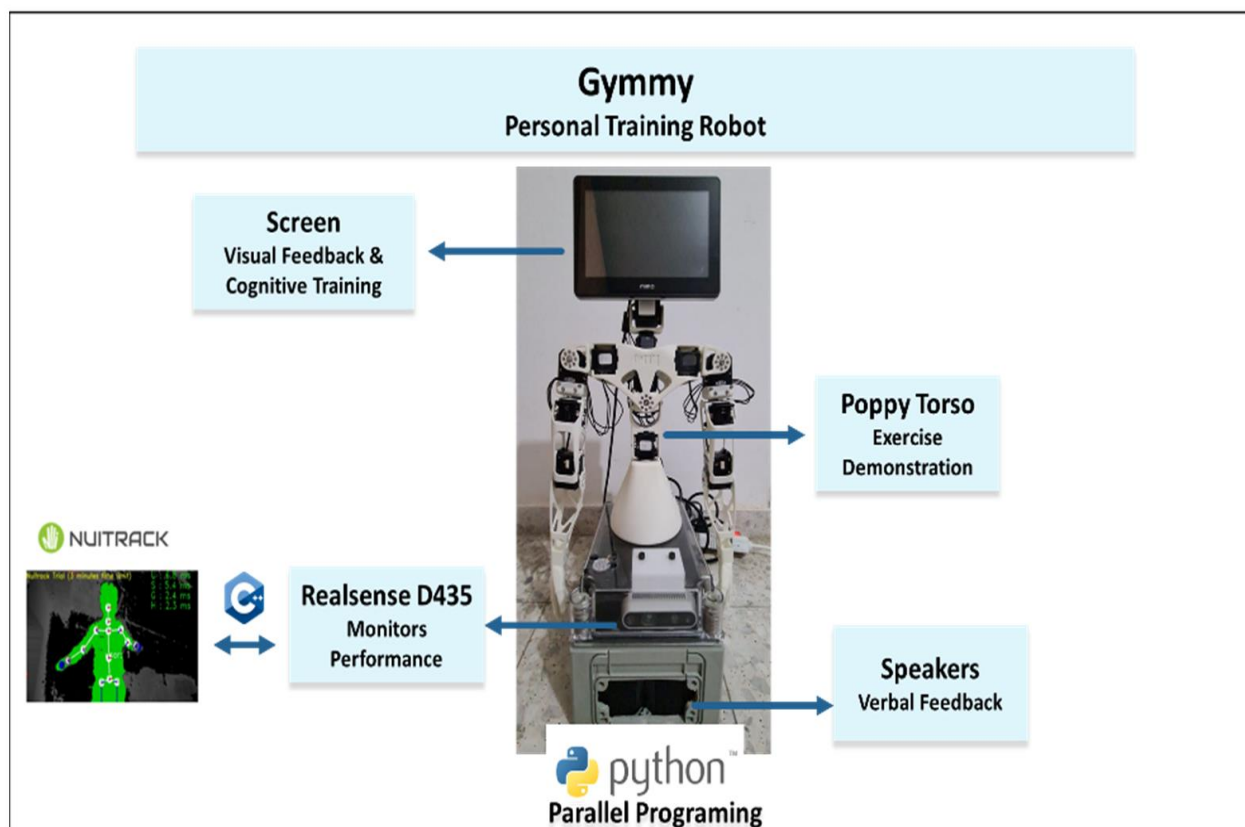


Рис. 22. Робот сервисного обслуживания пожилых людей “Гутту” [19, 20]



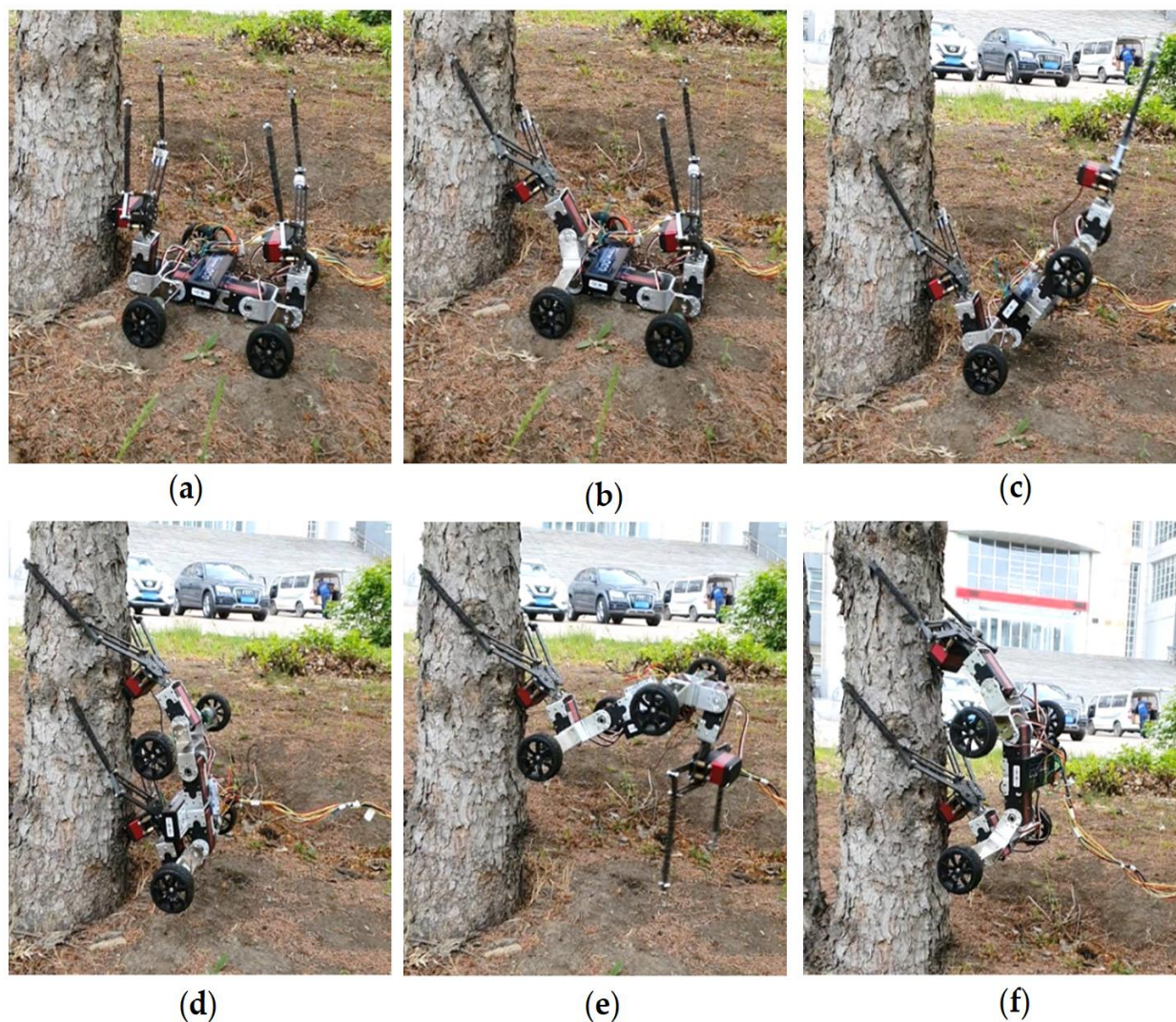
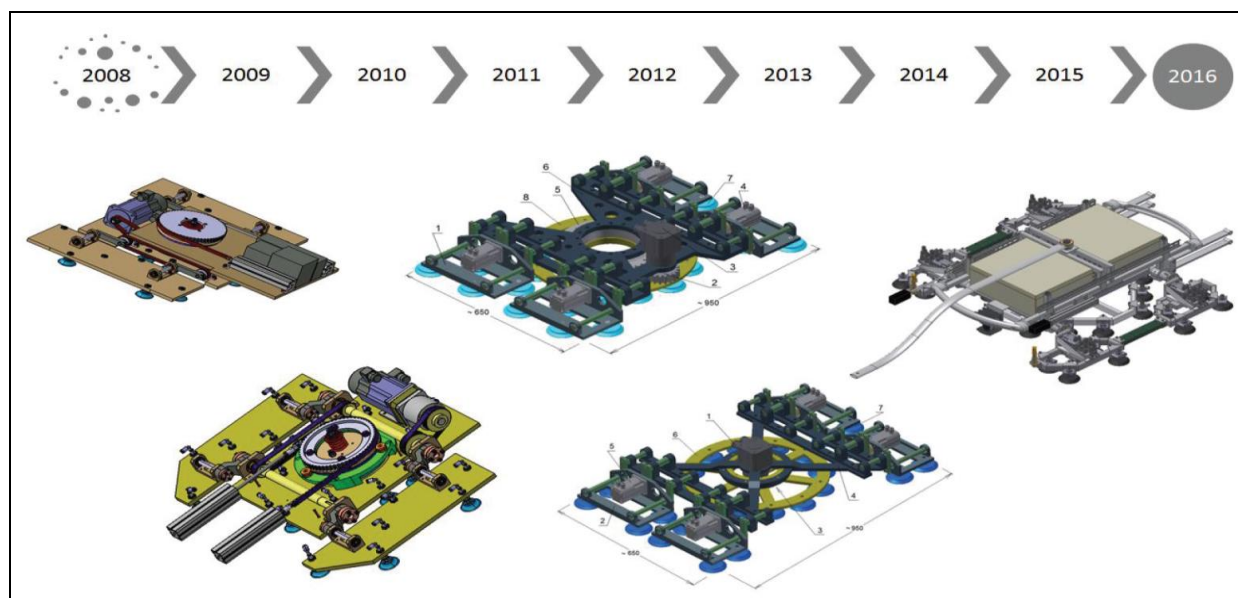


Рис. 23. Робот вертикального перемещения с переменной структурой – Альпинистский эксперимент с группой контрольных параметров. (a) Подход к столбу; (b) Захват штанги; (c) Отрыв от земли. (d) Второй захват штанги; (e) Действие переворачивания; (f) Робот завершает переворот; (g) повторный захват шеста [21]





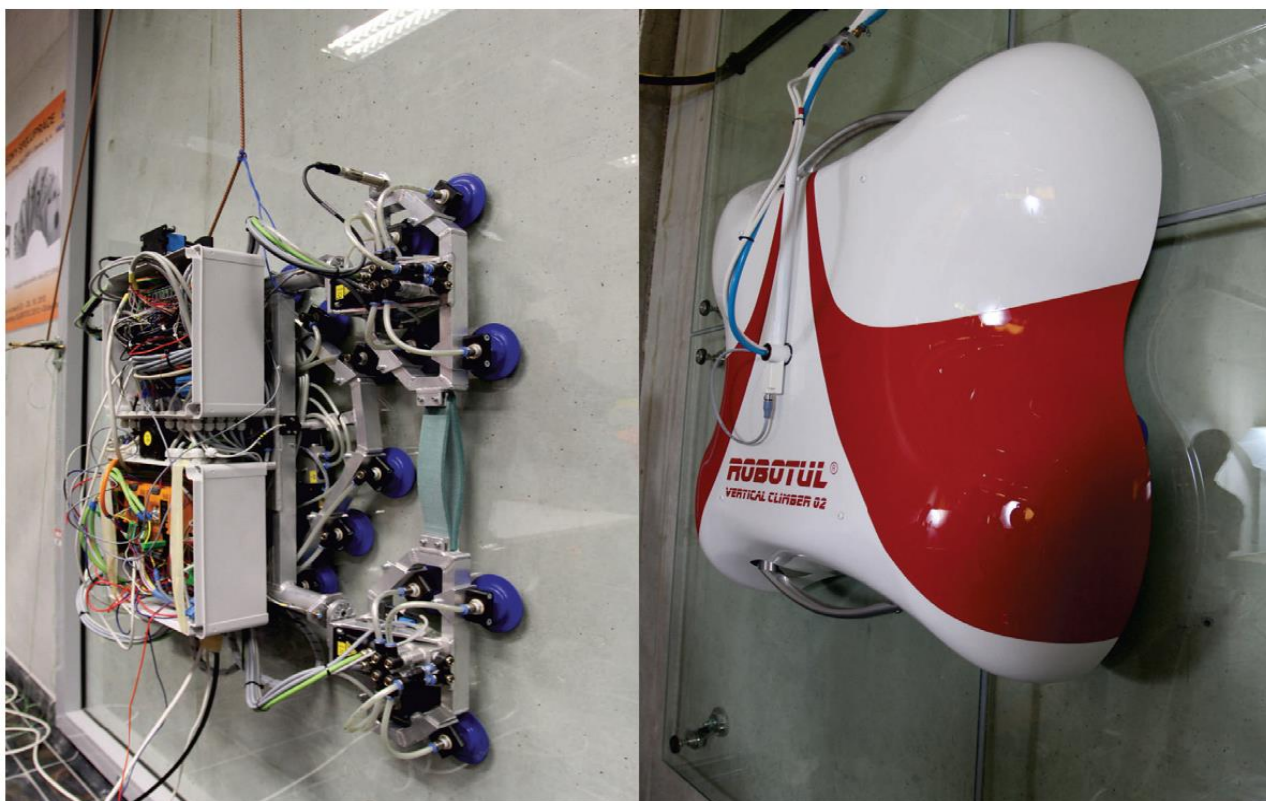


Рис. 24. Модели роботов вертикального перемещения [22]

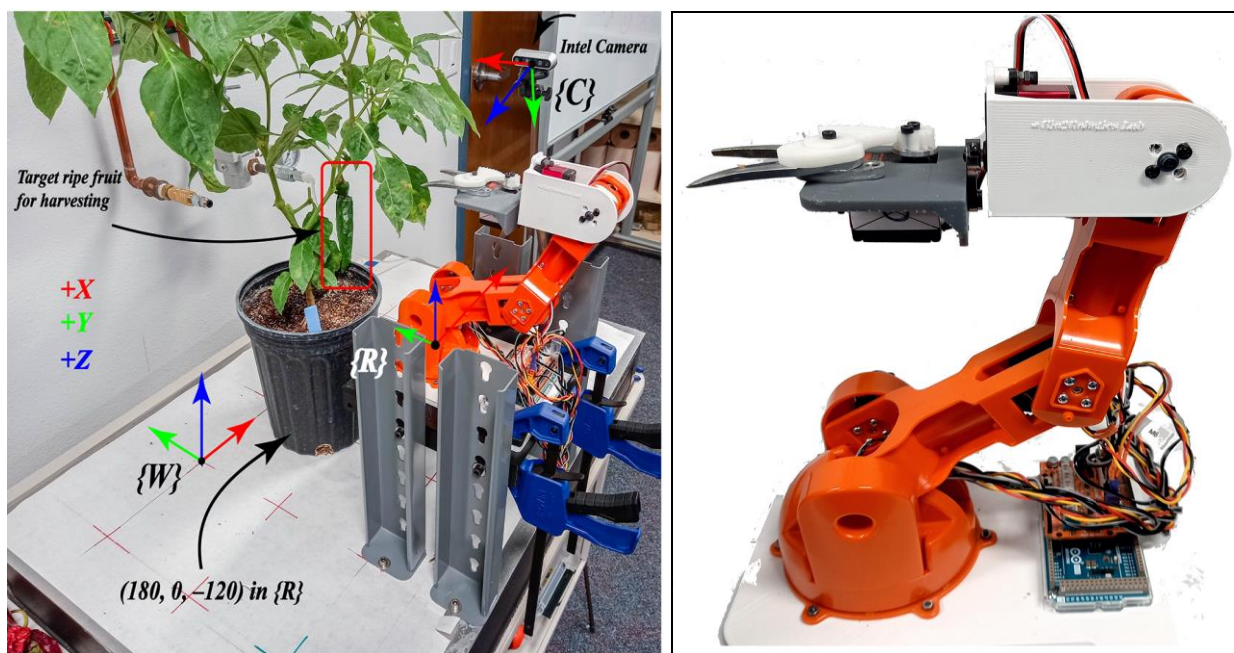


Рис. 25. Робот для сбора урожая перца [23]



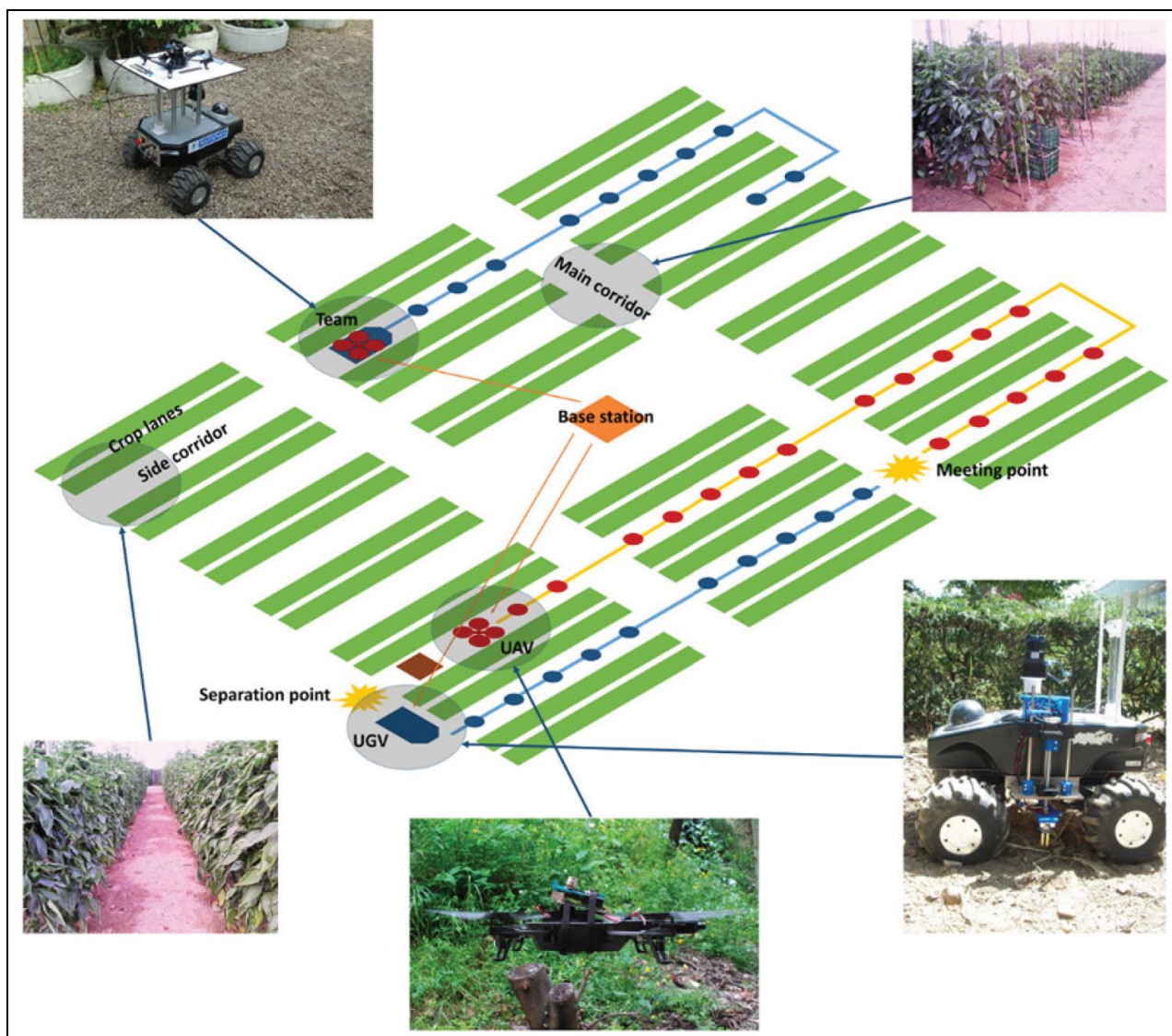


Рис. 26. Модели роботов для сельскохозяйственных работ [24]



Многороторные дроны широко используются во многих областях, таких как инспекция, доставка, наблюдение, сельское хозяйство и развлечения. Среди различных типов многороторных беспилотных летательных аппаратов квадрокоптеры наиболее популярны благодаря своей простой конструкции и относительно высокой аэродинамической эффективности. Однако из-за меньшего резервирования ротора квадрокоптеры также более уязвимы для отказов двигателя.

Отказоустойчивое управление имеет решающее значение для критически важных систем безопасности, таких как квадрокоптеры. Современные контроллеры полета могут стабилизировать квадрокоптер и управлять им, даже если он полностью потерял свою часть. Однако эти методы полагаются на внешние датчики, такие как *GPS* или системы захвата движения, для оценки состояния. Этого еще не удалось достичь с помощью только бортовых отказоустойчивых датчиков. Был предложен алгоритм на основе бортового технического зрения для достижения управления положением квадрокоптера, подвергнувшегося полному отказу одного из роторов. Экспериментальные проверки показывают [25], что подход позволяет точно контролировать положение квадрокоптера во время отказа двигателя без помощи внешних датчиков. Основная проблема при оценке состояния на основе зрения проистекает из неизбежного высокоскоростного рысканья (более 20 рад / с) поврежденного квадрокоптера, вызывающего размытость считывающего изображения в камерах, что пагубно сказывается на визуальной инерциальной одометрии (*VIO* – *visual inertial odometry*). Сравнивались два типа визуальных входов с алгоритмом оценки состояния на основе зрения: стандартные кадры и текущие визуальные события. Экспериментальные результаты показывают преимущество использования камеры для событий, особенно в условиях низкой освещенности, благодаря присущему ей высокому динамическому диапазону и высокому временному разрешению.

Данный подход сделает автономные квадрокоптеры более безопасными как в условиях отсутствия *GPS*, так и в условиях ухудшения качества. Полный отказ ротора приводит к быстрому вращению ( $\sim 20$  рад / с) квадрокоптера. Это высокоскоростное движение представляет собой серьезную проблему для бортовых методов оценки состояния. Во-первых, в оценке состояния на основе зрения это вызывает размытие при движении (нижний левый график на рис. 26, а).

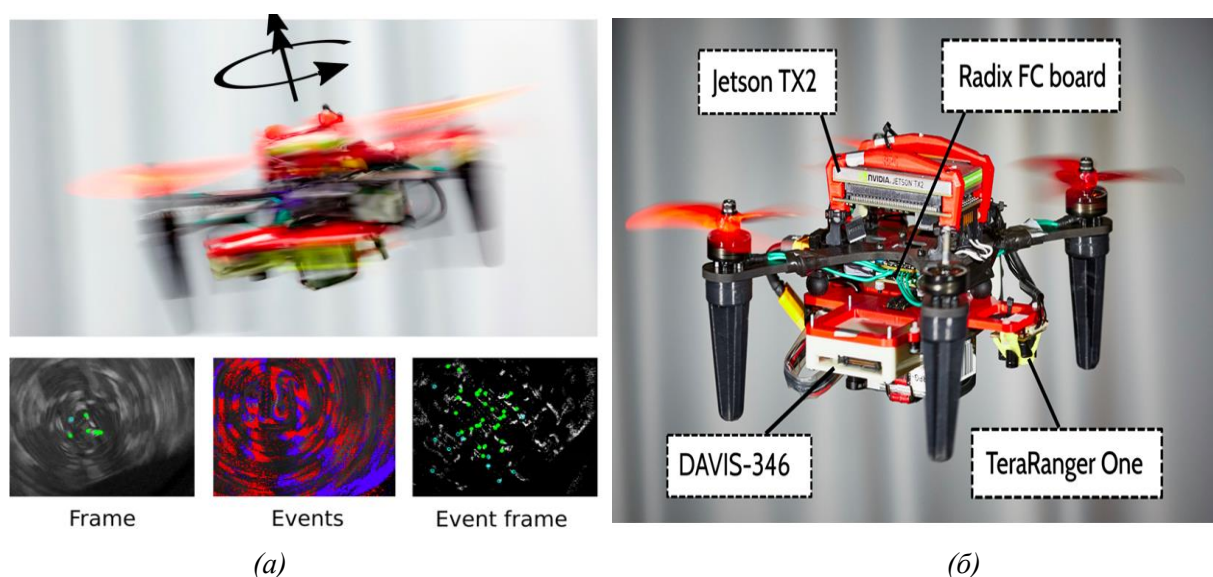


Рис. 27: (а) Управление квадрокоптером при полном выходе из строя ротора вызывает быстрое рысканье, более 20 рад / с (верхний рисунок); (б) Фотография летящего квадрокоптера с тремя роторами, на котором для оценки состояния используется камера съема текущих событий [25]

**Примечание.** На нижних рисунках (см. рис. 26, а) показан стандартный кадр и текущие события, снятые бортовой камерой съема событий. Внизу слева: стандартный кадр с размытием движения. Внизу по центру: только события (синий: положительные события, красный: отрицательные). Внизу справа: кадр обобщенного события, созданный из текущих событий. Синие круги обозначают обнаруженные объекты, зеленые точки обозначают отслеживаемые объекты.

Как отмечалось, такое размытие движения ухудшает, во-первых, скорость обнаружения и отслеживания и, следовательно, ухудшает работу визуальной инерциальной одометрии (*VIO*), особенно в

условиях низкой освещенности. Во-вторых, большое центробежное ускорение, считываемое инерциальным измерительным блоком (*IMU – inertial measurement unit*), часто вызывает большие ошибки в обычно используемых оценках ориентации. Эти проблемы необходимо решить для обеспечения автономного полета квадрокоптера, несмотря на отказ ротора, с использованием только бортовых датчиков.

Как показано на рис. 26, б тестируемый квадрокоптер состоит из рамы (из углеродного волокна) и деталей, напечатанных на 3D-принтере. Питается от четырех двигателей 2400KB с 5-дюймовыми гребными винтами. Оценка состояния и алгоритм управления выполняются системой Nvidia Jetson TX2, который содержит четырех-ядерный ЦП с частотой 2,0 Гц и двух-ядерный ЦП с частотой 2,5 Гц, работающий под управлением системы *Ubuntu 18.04* и *ROS*. Команды тяги двигателя из алгоритма управления отправляются на двигатели через плату управления полетом *Radix FC*, на которой работает встроенное программное обеспечение, которое также отправляет измерения *IMU* на *TX2* с частотой 200 Гц. Для измерения расстояния до поверхности земли использовали светодиодный датчик времени полета *TeraRanger One* как линза поля зрения видеокamеры. В качестве камеры событий использовалась система *Inivation DAVIS-346* с разрешением 346 x 240 пикселей.

В качестве стандартной камеры применялась *mvBlueFox-220w* с разрешением 376 x 240 пикселей, намеренно выбранным так, чтобы быть близким к разрешению камеры событий, и обеспечить объективное сравнение. И стандартная, и событийная камеры обращены вниз. Они обе используют время экспозиции, которое имеет важное значение для уменьшения размытости изображения при движении в стандартных кадрах. Предложенный алгоритм проверен в натурном эксперименте на открытом воздухе с естественным освещением и текстурами.

На рис. 28 показан снимок летящего квадрокоптера, а вид сверху на него.

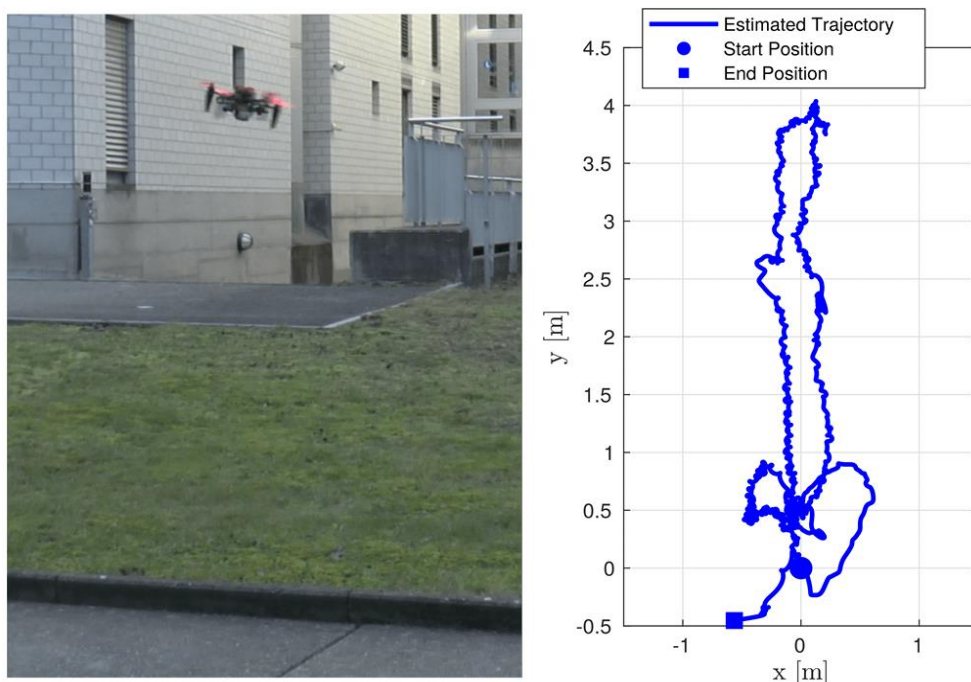


Рис. 28. Открытый полет для проверки предложенного отказоустойчивого решения полета.

Слева: снимок квадроцикла, летящего только с тремя гребными винтами.

Справа: вид сверху траектории полета вперед-назад, заданной устройством оценки состояния

Квадрокоптеру было дано задание пролететь вперед на 4 метра, а затем вернуться к исходной точке. В этом полете освещенность окружающей среды составляла 2000 люкс. В таких ярких условиях надежны как стандартные кадры, так и события. Следовательно, в этом полете использовалась стандартная камера.

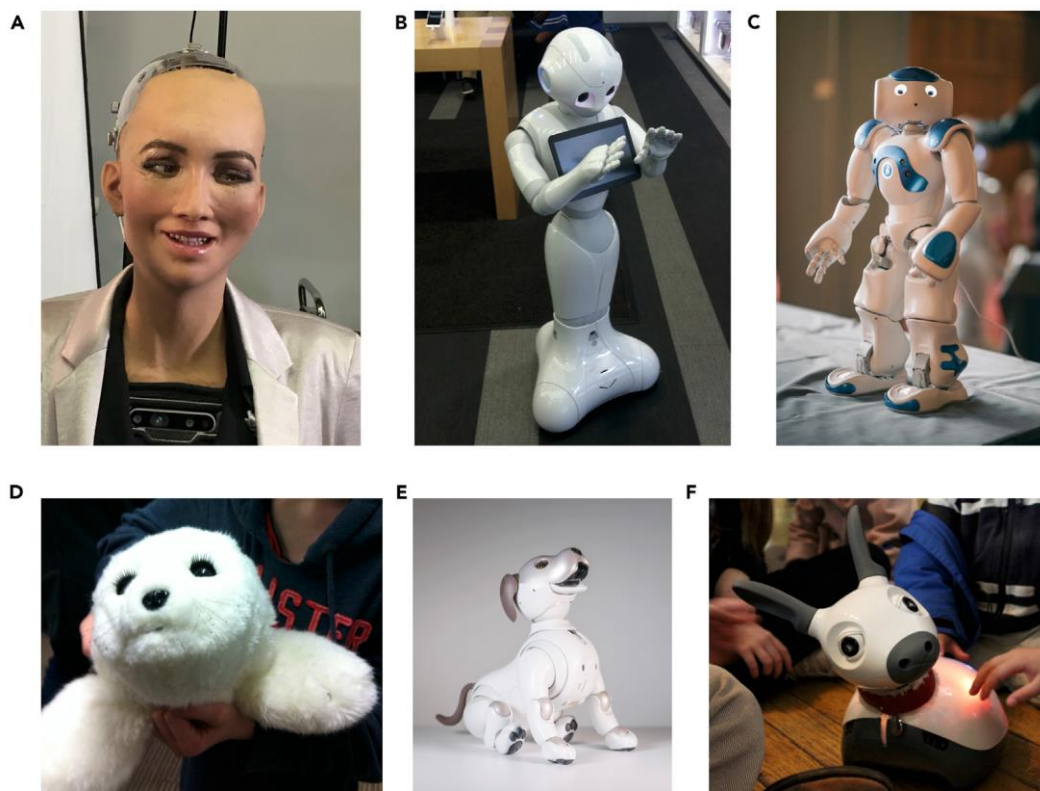
В заключении данной Части 1 статьи рассмотрим важный вопрос выбора социальных роботов для социотехнической системы с позиции оценки взаимоотношений человека-пользователя и робо-

том, функционирующем в самой социальной среде. Применение решений выбора типа и вида социальных роботов рассмотрено в Части 2 данной серии статей.

### ***Преимущества и риски взаимоотношений человека и робота в социальных системах***

В упрощенном понимании, роботы – это физические машины, в которых встроены элементы вычислительного интеллекта, которые позволяют им вести себя автономно, часто со способностью работать в течение длительных периодов времени без прямого контроля или наблюдения со стороны человека - оператора. Подмножество роботов является «социальным» или «социально вспомогательным», поскольку они обычно объединяют некоторую способность к слуховой (например, разговорной) и не слуховой коммуникации, часто имеющий внешний вид человека или животного (гуманоид и животное, соответственно), который представляет людям знакомые социальные сигналы, такие как лицо с четко видимыми «глазами», а также дисплеи или исполнительные механизмы, которые поддерживают выражение и жесты.

Хорошо известные примеры социальных роботов, показанные на рис. 29.



*Рис. 29. Примеры социальных роботов-гуманоидов и животных [26]*

*(A) София (Hanson Robotics), (B) Хао (Softbank), (C) Пеннер (Softbank), (D) Паро (Paro Robotics), (E) Aibo четвертого поколения (Sony Corporation), (F) MiRo-е (Последовательная робототехника).*

*Авторы: (A, B, D и F) Тони Прескотт; (C) Университет Шеффилда; (E) Пол Киллин*

Чтобы безопасно взаимодействовать и эффективно общаться с людьми, социальные роботы должны иметь встраиваемую бортовую систему управления или «когнитивную архитектуру», которая включает возможности для слухового и не слухового общения, анализа сцены, человека, обнаружение и распознавание объектов, знание мира, память, действия и планирование взаимодействия. Социальные роботы существенно различаются по конструкции и конфигурации этой архитектуры, многие из которых были непосредственно вдохновлены психологией человека или животных и нейробиологией.



Когнитивные архитектуры могут включать внутренние модели физической морфологии робота и близлежащей окружающей среды, отслеживающие его текущее положение и позу, чтобы обеспечить безопасное планирование движения и физического взаимодействия человека и робота. Также были разработаны когнитивные архитектуры, которые хранят воспоминания о прошлых событиях и создают модели действий, убеждений, желаний и намерений других, тем самым обеспечивая некоторые когнитивные способности такие нормы, как автобиографическая память и «теория разума», лежащие в основе человеческого социального познания.

Социальная робототехника может использовать более широкие достижения в области ИИ и робототехники, включая улучшения в области машинного обучения, компьютерного восприятия, обработки естественного языка и управления роботами. Хотя современные платформы социальных роботов имеют ограниченную встроенную обработку, они также все чаще могут использовать преимущества облачных вычислений для поддержки более интенсивных форм вычислений; это поможет предоставить более широкие возможности социального взаимодействия с учетом контекста. Социальные роботы также могут получать доступ к таким ресурсам, как внешние датчики и базы данных, которые позволят им предоставлять формы социальной поддержки, которые менее просты и естественны для людей. Таким образом, вместо того, чтобы копировать или подражать формам взаимодействия человека с человеком, можем ожидать, что отношения с роботами будут иметь несколько иные, а зачастую и взаимодополняющие качества.

Социальные роботы, которые могут взаимодействовать и общаться с людьми, становятся все популярнее и используются дома, в сфере обслуживания клиентов, в образовании и здравоохранении. Хотя все больше данных свидетельствует о том, что совместные и эмоционально ориентированные социальные роботы могут приносить пользу пользователям на протяжении всей жизни, продолжают споры об этических последствиях этих устройств и их потенциальном вреде. В этой перспективе исследуется баланс между выгодой и риском через призму взаимоотношений человека и робота. Рассматриваются определения и цели социальных роботов, исследуется их философский и психологический статус и соотносится с исследованиями отношений «человек – человек» и «человек – животное» по отношению «человек – человек – робот».

Отстаивая реляционный, а не эссенциалистский взгляд, рассматривается баланс пользы и вреда, который может возникнуть в результате различных типов отношений с социальными роботами, и рассматривается роль исследователей в понимании этического и социального воздействия социальной робототехники.

Социальные роботы, определяемые как роботы, которые взаимодействуют и общаются с людьми или другими агентами, демонстрируя социальное поведение и следуя нормам, в последние годы приобрели взрывную популярность благодаря быстрому развитию исследовательских прототипов и коммерциализации устройств.

В частности (см. рис. 29), включают Софию *Hanson Robotics*, похожую на человека робота, которая привлекла внимание средств массовой информации в 2017 году, когда ему был присвоен статус «почетного гражданина» в Саудовской Аравии, а также *Pepper* (2018) и *Nao* (2009) из *Softbank*, два широко доступных робота-человека, которые используются в исследованиях в различных дисциплинах и для коммерческих приложений при оказании помощи клиентам и обучении. К негуманоидным роботам относятся решения, похожие на животных, такие как игровая роботизированная собака *Sony Aibo* (2000) (переработанная и перезапущенная в 2018 году), похожий на тюленя *Paro*, разработанный для успокаивающего воздействия на жителей, нуждающихся в долгосрочном уходе. объектов (2011) и *Miro-e*, биоинспирационного робота, разработанного для применения в образовании и терапии.

Интерактивные и динамические особенности социальных роботов, а также их способность понимать человеческие эмоции и реагировать на них делают их кандидатами на решения для областей приложений, требующих социальной активности и комфорта. В сфере здравоохранения социальные роботы исследуются и внедряются в качестве инструментов для оказания помощи пациентам посредством эмоциональной поддержки на протяжении всей жизни из педиатрических популяций пожилым. В образовании социальные роботы используются в качестве наставников или соучеников для достижения когнитивных и эмоциональных результатов, включая развитие социальных навыков. Особое внимание уделяется группам населения, которые испытывают проблемы с общительностью и вниманием, например детям с расстройствами аутистического спектра (РАС).

Все чаще социальные роботы используются в качестве домашних компаньонов в качестве альтернативы домашним животным, и было обнаружено, что они оказывают положительное влияние на переживание одиночества. Некоторые из этих областей применения вызывают озабоченность, например, в ситуациях, когда маленькие дети или пожилые люди, живущие с когнитивными нарушениями, не понимают, является ли социальный робот живым существом, или в ситуациях, когда оказалось, что люди чрезмерно доверяют роботу-гиду. Растущее использование антропоморфных роботов в качестве сексуальных компаньонов также вызвало серьезные споры. Хотя использование роботов в большинстве этих условий только начинается, дебаты об этических рисках, которые они создают, уже идут полным ходом и часто освещаются в основных средствах массовой информации.

*Примечание.* Рассмотрим пространство на рис. 30 воображаемых результатов, в которое входят одни из них, вредные для человеческого благосостояния (левый кружок), а другие – полезные (правый кружок). Некоторые исходы сильно вероятны (центральный круг), другие – маловероятны (за пределами центрального круга).

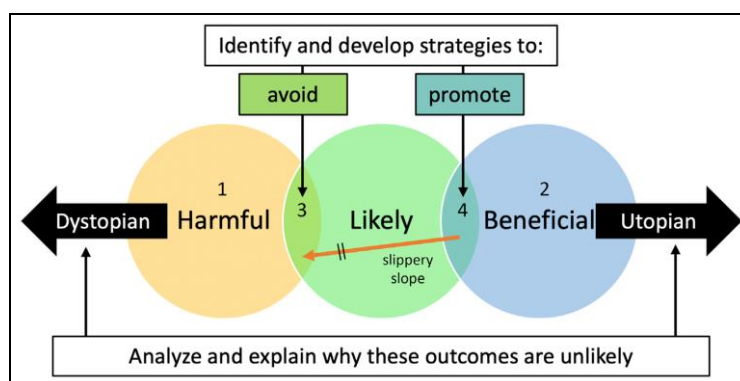


Рис. 30. Стратегия исследования этических и социальных последствий социальных роботов

Эта простая схема приводит к четырем категориям результатов: (1) антиутопические видения будущих миров, которые очень маловероятны, даже если их можно вообразить. Роль исследователей состоит в том, чтобы проанализировать эти результаты, оценить, какие шаги потребуются для их возникновения, и объяснить, почему эти сценарии вряд ли будут реализованы. (2) На противоположной крайности находятся неправдоподобные, хотя и вообразимые утопические видения. Опять же, роль исследователей состоит в том, чтобы анализировать и предостеречь от нереалистичных утверждений. Многие сценарии научной фантастики попадают в эти две категории. (3) Пересечение слов терминов «вредный» и «вероятный» означает негативные последствия, которые могут произойти. Установив эти вероятные сценарии, необходимо активно разрабатывать стратегии, которые могут смягчить их последствия. (4) И наоборот, пересечение терминов «выгодного» и «вероятного» представляет собой положительные результаты, которые могут произойти, и мы должны разработать стратегии, способствующие этим случаям. Рисунок также иллюстрирует «скользкую дорожку», когда то, что можно было бы рассматривать как положительный результат, оказывается вершиной сползания в категорию вредных. Скользкие склоны заслуживают анализа, и должны быть разработаны меры предосторожности (обозначены на рис. 29 как « $\backslash$ ») против тех, у кого есть убедительная причинно-следственная цепочка (ордер). Однако, необходимо также скептически относиться к скользким склонам, мотивированным глубоко антиутопическими видениями, поскольку количество необходимых шагов и потенциальных победителей может сделать такие сценарии крайне маловероятными.

Стоит отметить особенности использования аргументов «скользкой дорожки» в риторике, окружающей некоторые из этих социальных проблем. Например, опасения по поводу использования социальных роботов, ограничивающих доступ к контактам с людьми, и связанный с этим психологический ущерб часто прогнозируются на основании предположения о ненадлежащем и чрезмерном использовании роботов в, например, в учреждениях по уходу за детьми или престарелыми, где роботов можно представить как полностью или полностью заменяющие межличностные контакты. Чтобы оценить такие риски, необходимо определить причинные цепочки, посредством которых внедрение социальных роботов приведет к этим наихудшим исходам. Что касается системных «роботов – нянь», например, рассматривая ряд рисков и средств защиты (таких как юридическая ответственность), обнаружено, что использование социальных роботов в уходе за детьми «не более опасно, чем другие артефакты и уход за детьми. практики, уже присутствующие в нашем обществе. Это делается не для

того, чтобы игнорировать риск, а для того, чтобы признать, что такие проблемы, как зависимость, чрезмерная зависимость и их влияние на человеческие отношения представляют собой угрозы, которые социальные роботы разделяют с другими аспектами нашей жизни, все более активно связанной с цифровыми технологиями, от потоковых сервисов до социальных сетей и смартфонов. Поэтому влияние будущих отношений с роботами необходимо рассматривать наряду с изучением более широкой картины изменений в социальной взаимосвязанности людей, вызванных новыми технологиями.

Сбор эмпирических и теоретических данных о роли и влиянии социальных роботов может быть довольно разрозненным делом. Инженеры разрабатывают и совершенствуют компоненты аппаратного и программного обеспечения, продвигают интеграцию ИИ в социальных роботов и измеряют эффективность взаимодействия человека с роботом среди других целей. Философы и другие ученые-гуманитарии исследуют природу и мораль взаимоотношений человека и робота в связи с более широкими вопросами о состоянии человека. Социальные психологи, социальные экологи и специалисты по взаимоотношениям изучают динамику отношений и сетей социальных связей и исследуют их влияние на качество жизни, включая переживание одиночества. Специалисты по технологической этике отвечают на вопросы, связанные с приемлемостью для конечных пользователей, и с этическими проблемами, такими как последствия для достоинства, конфиденциальности и автономии.

Важные проблемы, возникающие при взвешивании выгод и рисков взаимоотношений человека и социального робота, приводят к отсутствию эффективного регулирования и управления в этой сфере. Исследования могут сыграть ключевую роль в формировании как политики, так и практики. Как показано на рис. 29, задачи включают (1) рассмотрение положительного и отрицательного потенциального воздействия социальных роботов, (2) определение возможных результатов, которые являются правдоподобными, и (3) разработку стратегий для поощрения положительного воздействия и предотвращения отрицательного. Обращение к этим направлениям исследований с использованием трансдисциплинарного подхода, целенаправленного взаимодействия с более широким обществом, будет иметь решающее значение для продвижения этой области в будущем.

Сравнить и изучить отношения между сознанием выбора социальных роботов типа сервисного обслуживания с ИИ и *IRS* среди японских и тайваньских студентов робототехнических колледжей с помощью анкетного опроса было целью [27].

Отметим, что синдром немедленного ответа (*IRS – Immediate Response Syndrome*) вызван навязчивой идеей, что необходимо отвечать немедленно, как только получают электронное письмо, и это состояние, когда не могут отпустить свой мобильный телефон даже на мгновение. В результате системы обработки анкетных данных *ANOVA* было обнаружено, что тайваньские студенты часто пользуются Интернетом для общения с японскими студентами, и их склонность к *IRS* высока. Что касается роботов с ИИ, японские студенты колледжей, как правило, были по активности выше, чем на Тайване, в отношении терминов [*Ожидания*] [*Тревога*] [*Интерес*] [*Забота*] [*Волнение*] [*Полезность*]. Что касается «ожидаемых» в перспективе роботов, активность японских студентов колледжей выше соответствующих студентов из Тайваня о роботах таких как, действующих во время стихийных бедствий, о роботах-спасателях, таких как спасатели, роботы, выполняющие опасные работы, роботы для приготовления пищи, роботы безопасности и роботы транспортной отрасли. Тайваньские студенты ждут роботов-уборщиков, роботов-автоответчиков, роботов, противостоящих детям, сельскохозяйственных рабочих роботов, роботов для обслуживания клиентов и семейных роботов.

Что касается разработки роботов с ИИ, высокие группы *IRS* показали более высокие значения «ожиданий» и «интересных» вещей, а группы низкого уровня *IRS* показали более высокие значения «беспокойства» / «обиды». Что касается «ожидаемого» робота, группа *IRS* (уровень высокий – *high*) показала более высокие значения для предметов, таких как роботы для приготовления пищи, роботы-уборщики, роботы-автоответчики, роботы, заботящиеся о детях, роботы для обслуживания клиентов и семейные роботы. Что касается роботов-спасателей, таких как роботы, активные во время стихийных бедствий, спасателей, таких как спасатели, и роботов, выполняющих опасные работы, низкая группа *IRS* показала более высокие значения.

Этот результат заключается в том, что чем выше тенденция *IRS*, тем выше ожидание и полезность ИИ-роботов, показанных в [27]. Как в Японии, так и на Тайване, высокая тенденция *IRS*, похоже, зависит от Интернета и не является «здоровой», и исходит из высоких ожиданий и высоких интересов к ИИ-роботам. Кроме того, группа низкого уровня *IRS* ожидает, что робот ИИ как объект будет

выполнять такую работу, как опасная работа, которую он / она не хочет выполнять. Высокие группы IRS знакомы с направлениями как кулинария, стирка, автоответчик, няни, семья. Из этого можно сделать вывод, что высокие группы IRS заинтересованы в новых вещах и ощущают близость и проявляют интерес к новым роботам.

Свойство «привязанность / любопытство» к роботам ИИ воспринимается как свойство «пристрастие / любопытство» к новым вещам. В дополнение к группе высокой и низкой IRS на рис. 30 показана взаимосвязь между IRS и сродством к новым вещам. Раньше учителя младших и старших классов средней школы выясняли, есть ли способ запретить соцсети в полночь. Проводился анализ ответов типа: «Я опаздывал спать, потому что меня удивили соцсети посреди ночи, я спал на уроке, я не мог сосредоточиться на уроке, и я страдал от учеников, у которых были проблемы с дневной жизнью». Такой ученик будет расположен в зоне «B» на рис. 31.

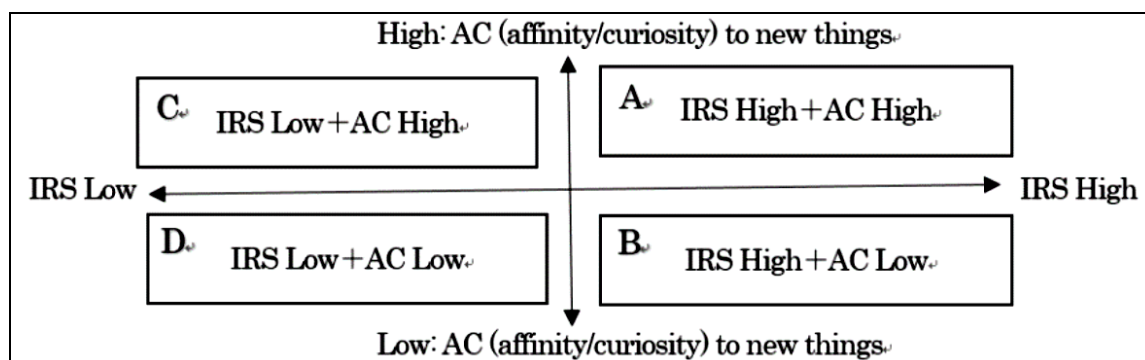


Рис. 31. Близость / любопытство – AC (affinity/curiosity) – к новым вещам и IRS

**Вывод:** «если любопытство к новому явлению велико, вы не будете спать на уроках, на которых можно быстро получить информацию о новых интересных разработках. Кроме того, молодые люди с низкой склонностью к IRS, расположенные в секторе «D», низкая склонность к новому, низкое любопытство, могут затормозить у них новую эру разработок в будущем».

**Решение:** Робот сервисного обслуживания детей школьного возраста включает в свою базу знаний такого рода выводы и может пробудить у школьника дополнительный стимул интереса к рассматриваемой проблемно-ориентированной предметной области знаний.

## Роботы сервисного обслуживания в социотехнических системах

Люди с ограниченными возможностями часто работают из-за своих возможностей, вызывая неудовлетворенность своей профессией, поскольку они чувствуют недостаток самореализации. Проект «Следующее поколение – разработка инклюзивной работы с гибкими роботизированными решениями» направлен на поддержку и расширение индивидуальных рабочих возможностей с помощью инновационных технологий. Вследствие этого люди с ограниченными возможностями будут рассматриваться как часть первичного рынка труда и будут признаны ценными членами рабочего сообщества и общества. Для достижения этой цели рабочие места в промышленности должны быть адаптированы индивидуально в соответствии с физическими и когнитивными способностями людей с ограниченными возможностями. Идея состоит в том, чтобы внедрить взаимодействие человека и робота (HRC) на рабочих местах. Основная цель коллаборативного робота – поддерживать человека, позволяя ему выполнять рабочие задачи независимо от его физических недостатков. HRC представлен как новая вспомогательная технология для людей с ограниченными возможностями для оказания физической помощи. Вспомогательные технологии полезны только в том случае, если они адаптированы к индивидуальным потребностям.

Таким образом, распределение задач между людьми и роботами является центральным аспектом процесса проектирования и программирования рабочего места HRC. Для людей с ограниченными возможностями распределение должно основываться на их индивидуальных возможностях. Следовательно, требуется метод для определения рабочей нагрузки задач, в которых конкретный человек может быть перегружен и нуждается в помощи и поддержке совместного робота.

В дополнение к перечисленным требованиям рабочее место должно быть спроектировано таким образом, чтобы можно было представить множество различных процессов сборки и тестирования,



которые встречаются в промышленных приложениях. Кроме того, в дизайне должны быть реализованы принципы эргономики. По этой причине настоящая концепция приводит к созданию рабочего места, основанного на типовой сборочной станции, состоящей из регулируемого по высоте стола, монитора и освещения, а также фиксированных мест для инструментов, материалов и готовой продукции. Участник и робот, поддерживающий *HRC*, должны иметь доступ к инструментам, материалам и хранилищу готовой продукции. Кроме того, мобильная платформа может помочь, доставляя материалы на рабочее место. Концепция показана на рис. 32.

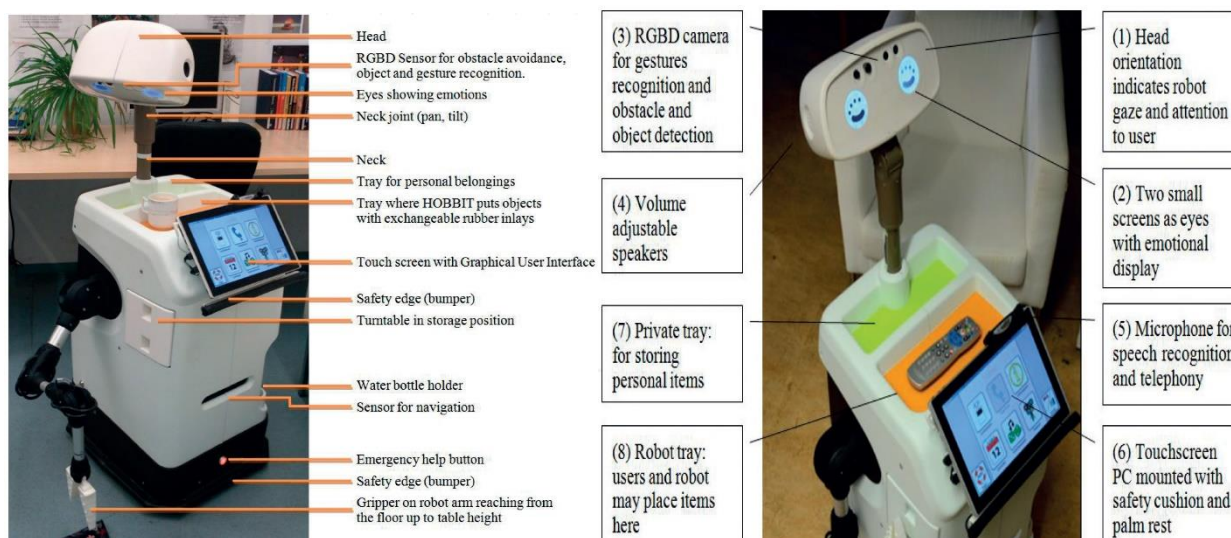
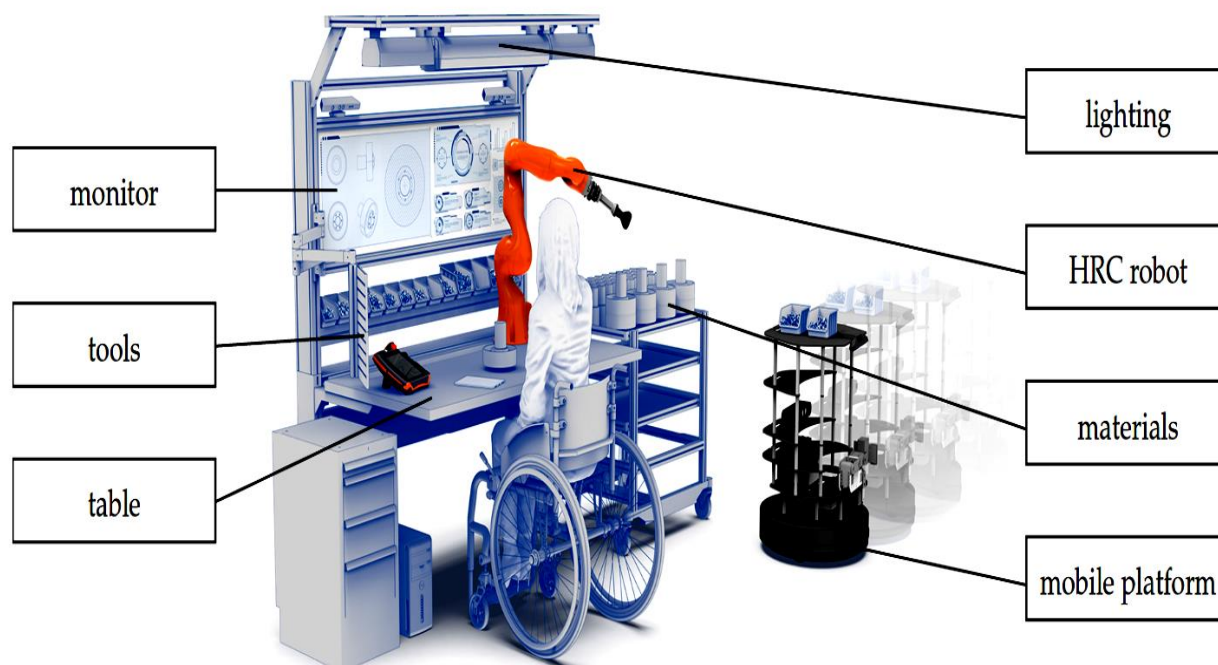


Рис. 32. Концепция рабочего места человек – робот для людей с ограниченными возможностями [20]

Физическая реализация влияет на взаимодействие человека-робота, поскольку люди взаимодействуют с гуманоидами иначе, чем с негуманоидами. Помимо ожиданий, морфология робота играет жизненно важную роль в его удобстве использования, приемлемости и выразительности. Поэтому требуется, чтобы морфология робота соответствовала его предполагаемой функциональности. Например, роботы, предназначенные для выполнения задачи, подобные человеку, должны быть оснащены соответствующим образом; визуальное сходство с человеком может не иметь большого значения в таких случаях, как в случае с *ATLAS* и подобными гуманоидами (рис. 33, а). С другой стороны, те, которые предназначены для целей взаимодействия, должны быть более похожи на людей, с отчет-

ливыми выражениями лица (например, София) (рис. 33. б) или с эмоциональными речевыми способностями (например, Перец) (рис. 33, с).

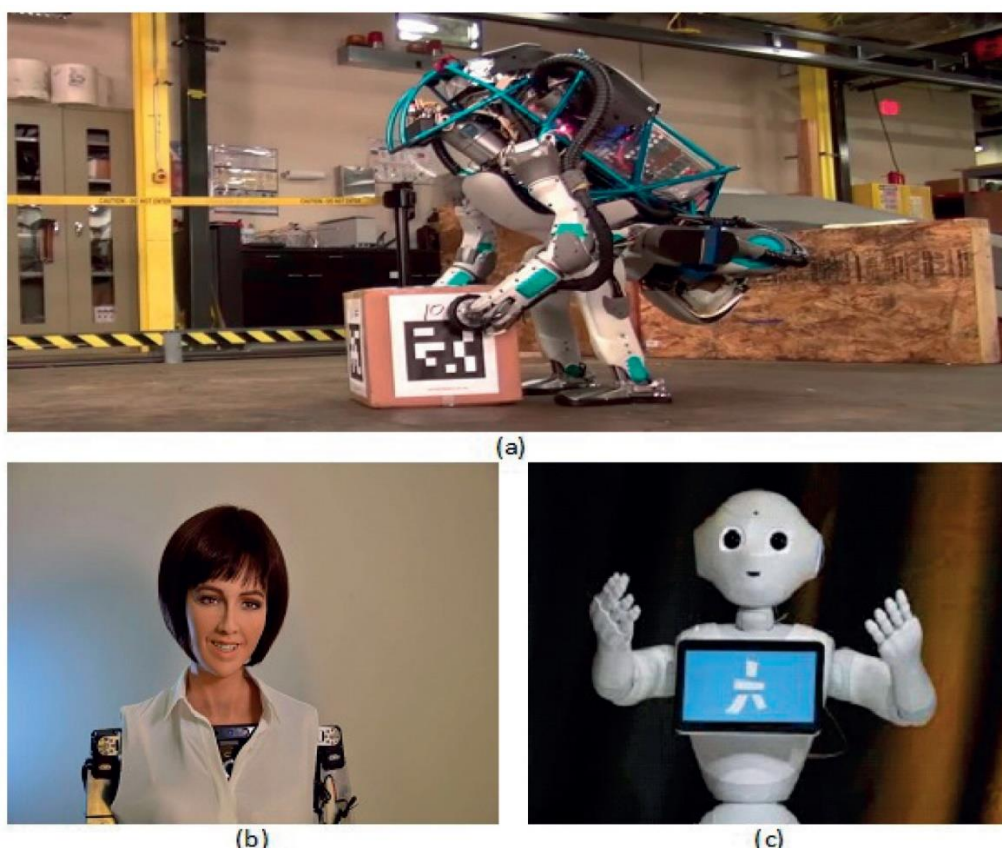


Рис. 33. Продвинутые гуманоиды: (а) АТЛАС, (б) София, (с) Пеппер [28]

Стремясь добиться натуралистического воплощения, дизайнеры черпают вдохновение в самой природе. Морфологический дизайн естественных социальных роботов можно отнести к антропоморфизму. В зависимости от области применения, морфологическое вдохновение для мировоззрения робота также может быть взято из зооморфизма (например, домашние животные или существа), карикатуры (например, анимация или вымышленные персонажи) и функциональных ожиданий (например, вспомогательных или служебных роботов и т. д.). Тем не менее, большинство социальных роботов предназначены для работы с людьми; таким образом, общая идея состоит в том, чтобы придать им внешний вид, похожий на человека. Поэтому больше внимания уделяется антропоморфизму.

Недавно были разработаны различные гуманоидные роботы для реагирования на стихийные бедствия с новыми подходами в теории управления и другими механизмами для преодоления неровностей местности. Традиционно роботы-гуманоиды преодолевали эти препятствия, шагая, или перелезая, но этим стратегиям не хватает эффективности, особенно для опасных сред, таких как непреодолимые препятствия и геологические разломы. Для выполнения неотложных задач в сложных реальных сценариях ожидается, что роботы-гуманоиды будут обладать аэродинамическими навыками, такими как прыжки в высоту или в длину, полеты на короткие дистанции и парение, которые в несколько раз превышают длину тела.

На рис. 34 поток планирования включает в себя многослойную локальную карту для вычисления проходимости, вспомогательную задачу и всенаправленную функцию Ляпунова. Вместо обычной стратегии следования по путевым точкам или отслеживания пути реактивный поток справляется с отклонением робота, устраняя негладкие движения с помощью векторного поля (определяемого методом обратной связи), которое предоставляет команды управления в реальном времени роботу. Выходной сигнал контроллера походки определяется как функция мгновенной позы робота.



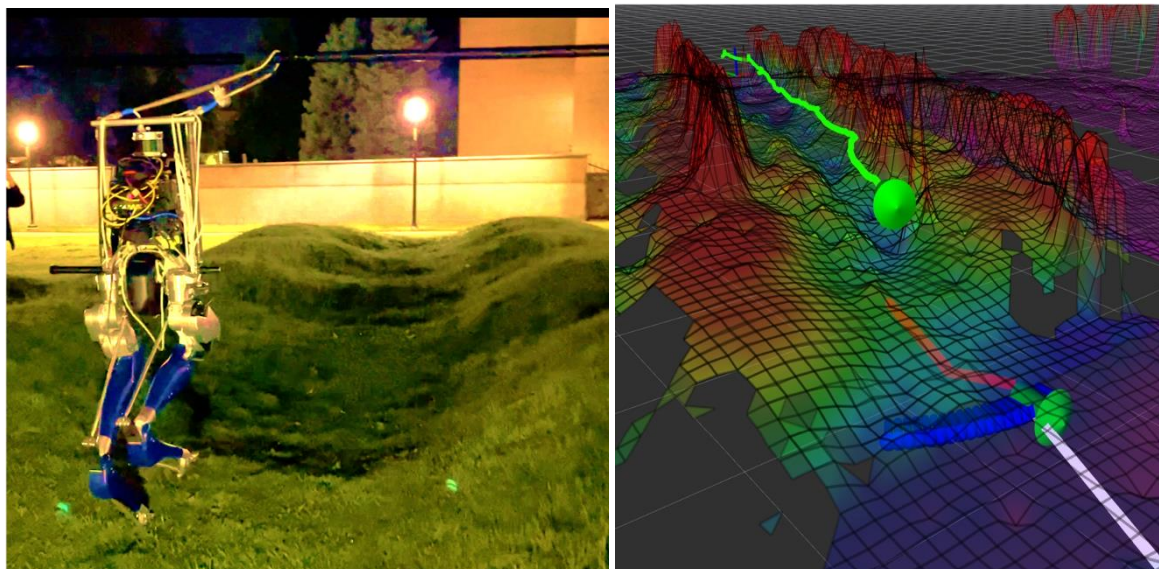
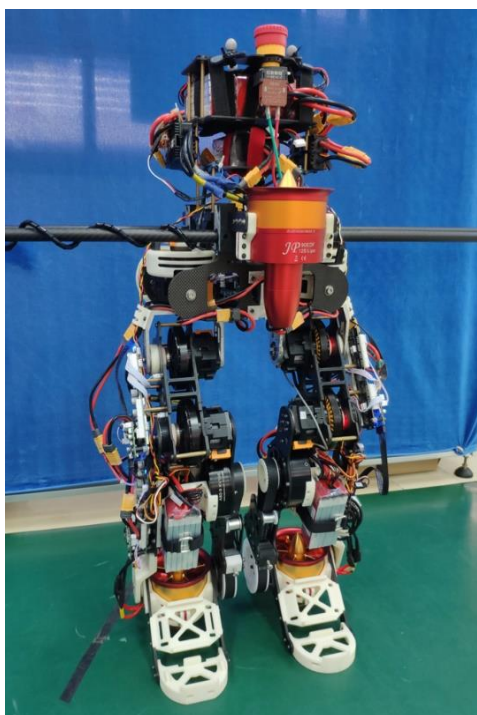
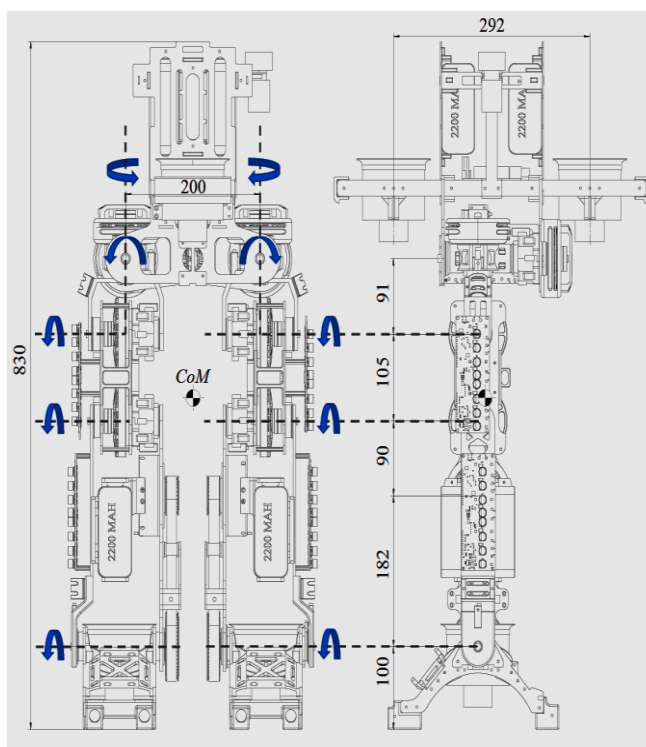


Рис. 34. Робот «Кэсси Блу» автономно пересекает Волновое поле с помощью предлагаемой системы реактивного планирования, которая состоит из потока планирования и потока реагирования

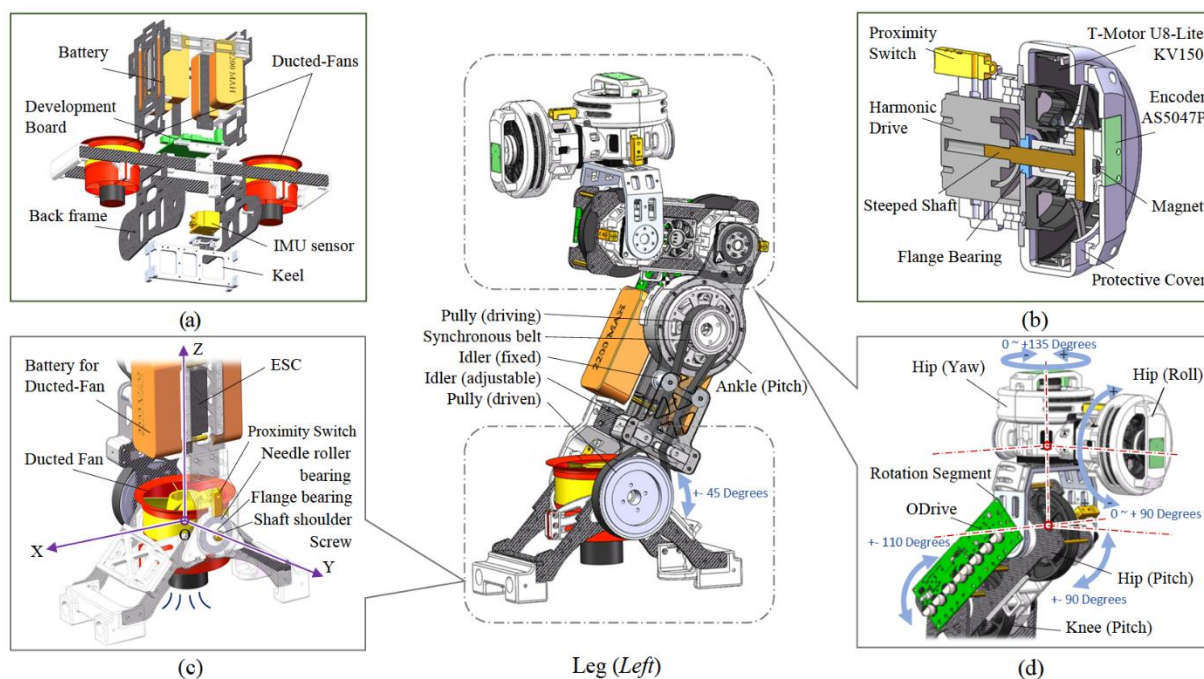
Достижение полета на короткие расстояния помогает повысить эффективность роботов-гуманоидов, перемещающихся в сложных условиях (например, преодолевая большие препятствия или достигая высоты) для быстрых ликвидаций последствий аварийных ситуаций. Была предложена конструкция летающего робота-гуманоида на основе вытяжных вентиляторов под названием *Jet-HR2* (рис. 35).



Вид летающего робота-гуманоида *Jet-HR2*



Конфигурация канального вентилятора *Jet-HR2*



Сборка механических компонентов ноги робота: (а) Талия, (б) Модульный сустав, используемый в бедре и колене, (с) Ступни, (д) Бедро

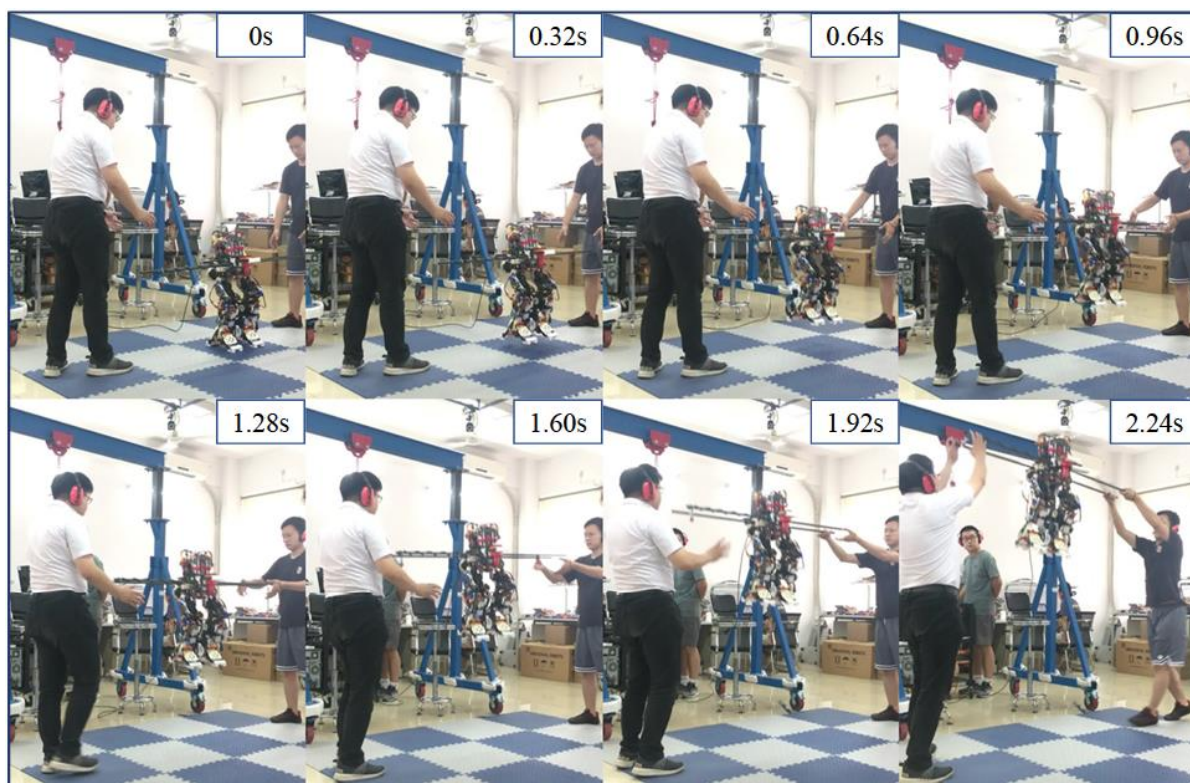


Рис. 35. Jet-HR2: летающий робот-гуманоид [29]. Результат эксперимента с различными стратегиями управления, в которых: а) взлет с контроллерами тангажа и рыскания

Робот имеет 10 шарнирно соединенных связей, приводимых в движение бесщеточными двигателями и двигателями для синхронно плавного движения соединений. Для достижения устойчивого взлета в условиях малой тяги к весу, конструкция робота была разработана на основе концепции вектора тяги. Силовая установка состоит из четырехканальных вентиляторов, то есть два закреплены на поясе робота, а два других установлены на опорах для управления вектором тяги. Вектор тяги контролируется путем регулировки положения стопы во время полета. Была предложена упрощенная



модель и стратегия управления для решения проблемы неустойчивости ориентации, вызванной массовыми ошибками и ошибками положения суставов во время взлета. В частности, модель создана для анализа сил во время взлета в фиксированном положении и определения условия эффективного регулирования положения робота в воздухе в условиях низкого отношения тяги к весу. Для упрощения модели робота и целенаправленного решения проблем пикирования и рыскания во время взлета делаются следующие допущения: 1) робот сохраняет фиксированное положение во время взлета и считается твердым телом и 2) ноги робота остаются параллельными друг с другом. Другими словами, компоненты ног робота по оси  $z$  и оси  $x$  остаются неизменными в раме тела. 3) Угол наклона канальных вентиляторов можно независимо регулировать в раме корпуса.

Результаты экспериментов показали (см. рис. 35), что вращение робота и его поведение в пикировании во время взлета эффективно подавлялись за счет управления вектором тяги вытяжного вентилятора на ноге. Робот успешно выполнил взлет с соотношением тяги к весу 1,17 (17 кг / 20 кг) и сохранил стабильное положение, достигнув высоты взлета более 1000 мм.

Пример развития когнитивных способностей автономного робота с учетом приведенного аффорденс исчисления приведен на рис. 36.

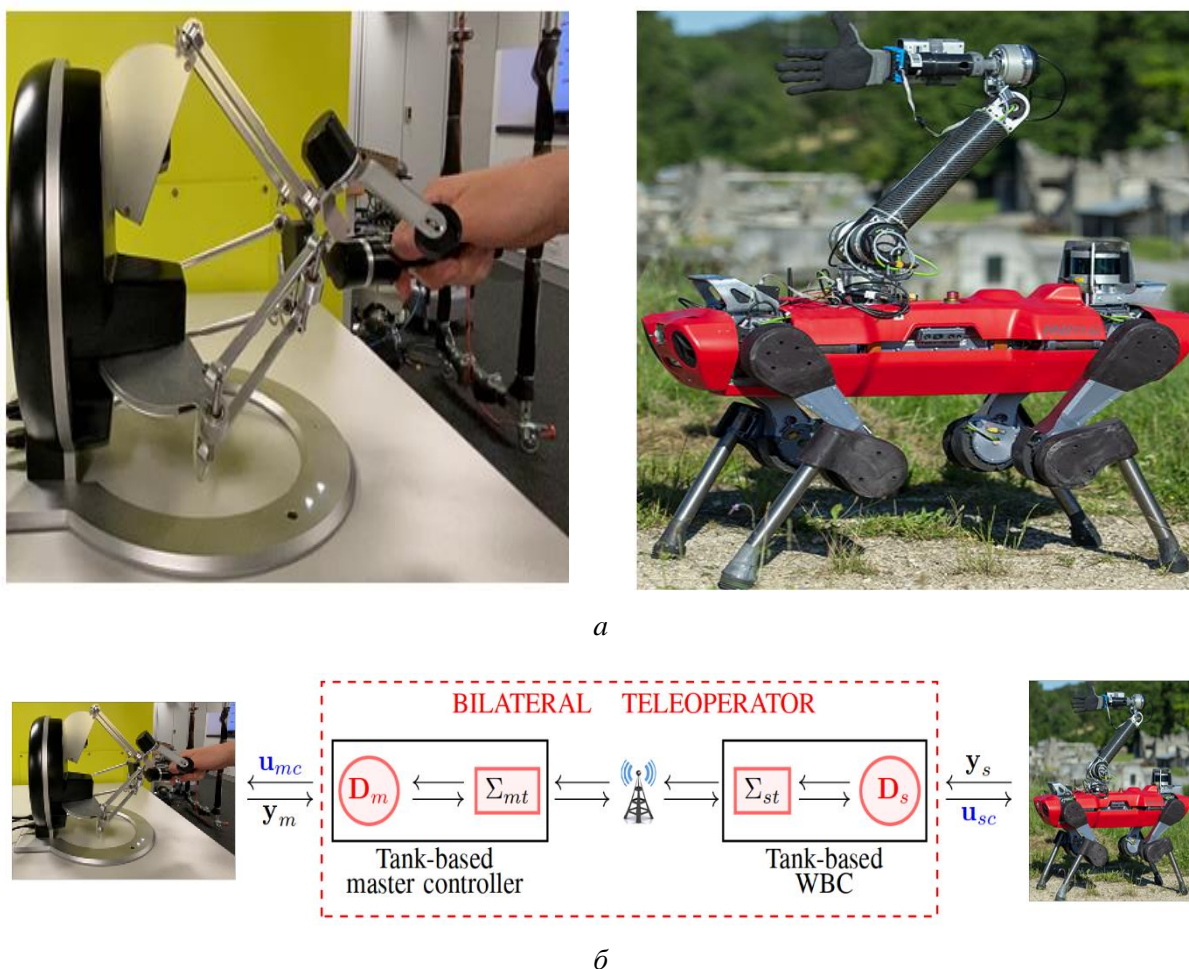


Рис. 36. (а) Экспериментальная установка с Omegaab в качестве тактильного устройства на ведущей стороне (слева) и платформа ANYmal C, оснащенная роботизированной рукой на ведомой стороне (справа); (б) Схема взаимодействия основных элементов слева направо: человек-оператор и тактильное устройство, двусторонний телеоператор и окружающая среда [30]

В случае структурной схемы управления одноколесным роботом со встроенным манипулятором (слева на рис. 37) в контур управления введены традиционный ПД-регулятор (позиционное управление одноколесным роботом) и когнитивный регулятор, вырабатывающий силу управления манипулятором совместно с информацией о позиции робота перед дверью.

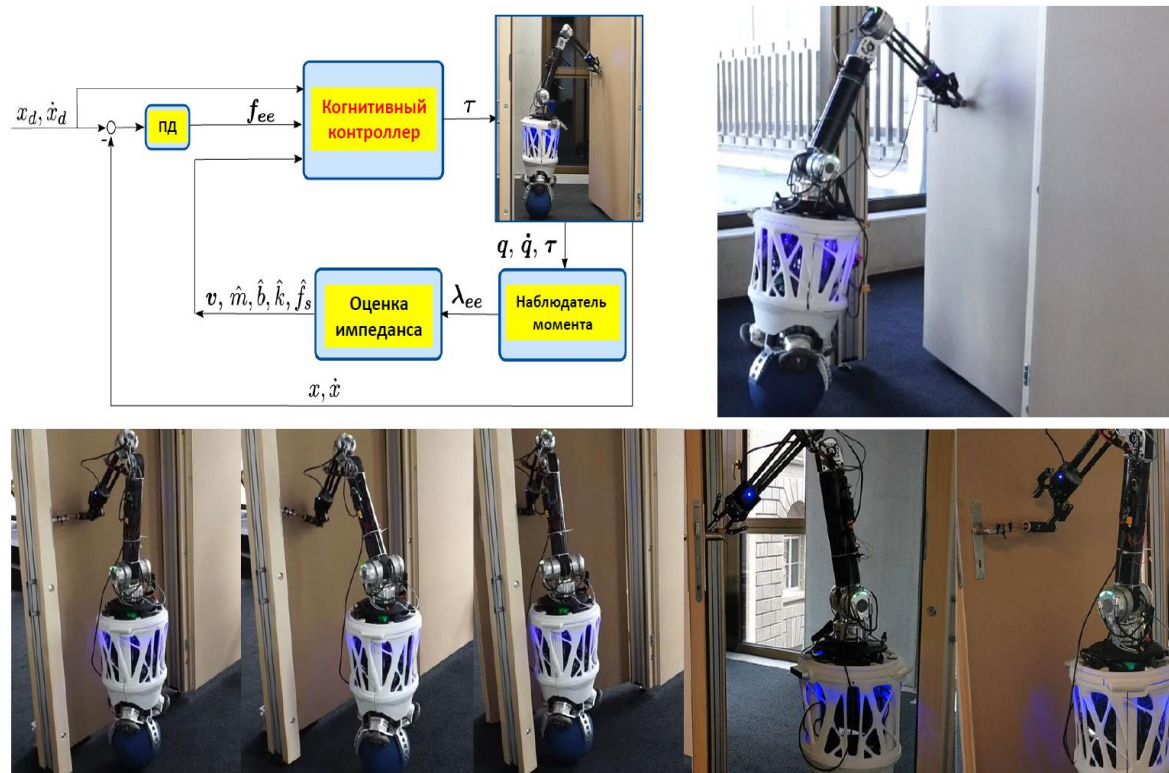


Рис. 37. Применение когнитивного силового управления одноколесным роботом с манипулятором для открытия двери с оцувствленным захватным элементом

При этом силовое управление осуществляется по информации от внешнего источника (датчика контакта захватывающего устройства) с объектом захвата (ручкой двери), роль которого выполняет блок «Оценка импеданса». При этом автономный одноколесный робот представляет собой глобально динамически неустойчивый нелинейный объект с несущим на борту манипулятором [31].

### Роль информационно – термодинамического описания моделей социотехнических систем: взаимосвязь с принципом свободной энергии

Следует подчеркнуть, что структурно социотехническая система как физическая система представляет собой самоорганизующуюся (термодинамически неравновесную и информационно открытую) систему. В этом случае существенное значение имеет физическая платформа построения и анализа потоков информации и материальных ресурсов в структурах таких самоорганизующихся систем, оптимизации перераспределения потоков и связей, взаимодействующих внутри подсистем, и взаимодействия с внешней средой.

*Примечание.* Различие между равновесным стационарным и неравновесным стационарным (*NESS – non-equilibrium steady state*) распределением является важным. С математической точки зрения стабильное состояние равновесия находится там, где сохраняется свойство детализированного баланса. Это означает, что любой переход между состояниями равновесия с такой же вероятностью будет идти как «вперед», так и «назад». По сути, динамика полностью симметрична времени, и поэтому в таких системах нет понятия стрелы времени. И наоборот, неравновесное стационарное состояние, при котором детализированный баланс не является неизменным, существует направленность динамики и, следовательно, стрела времени, даже несмотря на то, что фактическое распределение по состояниям остается постоянным. С термодинамической точки зрения достижение установившегося состояния равновесия является неизбежным с точки зрения второго закона термодинамики, поскольку это состояние с максимальной энтропией. И наоборот, *NESS* не является решением с максимальной энтропией, поскольку направленность динамики движения означает, что в системе обеспечивается предсказуемость, которая может в соответствии с теорией использоваться для производства

полезной работы. Неравновесное установившееся состояние является характеристикой термодинамической системы, требует внешнего источника движущей энергии в качестве постоянного входа в систему, которая затем рассеивается во внешнее окружение и дает *NESS* положительную скорость производства энтропии.

*Пример.* В качестве интуитивного примера представим себе термодинамическое равновесие чашки кофе со сливками. Устойчивое состояние равновесия (*ESS – equilibrium steady state*) – это когда кофе и сливки полностью растворились друг в друге, так что пропорция сливок остается постоянной по всей кофейной чашке. Это будет неизбежным результатом (согласно второму закону термодинамики) добавления в кофе ложки сливок с изначально низкой энтропией и высокой концентрацией. С другой стороны, мы можем думать о неравновесном устойчивом состоянии (*NESS*) как о том, когда сливки и кофе равномерно распределены повсюду, но кто-то постоянно помешивает кофе в определенном направлении. Здесь мы находимся в устойчивом состоянии, потому что концентрации сливок и кофе не меняются со временем, но, тем не менее, есть направленность динамики в направлении перемешивания. Эта направленность сохраняется только благодаря постоянному подаче энергии в систему (перемешивание). Поток, вызываемый перемешиванием, называется «соленоидальным потоком» и математически обязательно ортогонален градиенту распределения в установившемся состоянии. Это необходимо для того, чтобы соленоидальный поток не поднимался или не спускался по градиенту плотности и, таким образом, не изменял распределение установившегося состояния, которое, как установившееся состояние, по определению не может измениться. Биологические самоорганизующиеся системы часто считаются «диссипативными структурами», или неравновесными стационарными состояниями с точки зрения термодинамики, поскольку они поддерживают относительно устойчивое состояние с течением времени, для поддержания которого требуется постоянный приток энергии.

Принцип свободной энергии (*FEP – free energy principle*) постулирует глубокую связь между стохастической термодинамикой самоорганизации и обучением через вариационный логический вывод. В частности, утверждается [32], что любая самоорганизующаяся система, которая может быть статистически отделена от окружающей среды и которая поддерживает себя в неравновесном устойчивом состоянии, может быть сконструирована как система, минимизирующая теоретико-информационную функциональную (вариационная свободная) энергию и, таким образом, выполняющая вариационный байесовский вывод для определения скрытого состояния окружающей среды.

Этот принцип также широко применяется в нейробиологии и начинает проникать в глубокое квантовое машинное обучение, стимулируя создание новых и мощных квантовых алгоритмов, с помощью которых действия, восприятие и обучение могут быть объединены для достижения единой цели. Технически принцип свободной энергии утверждает, что любая «подсистема», которая достигает одного равновесного устойчивого состояния, может быть истолкована как выполнение элементарного вида байесовского вывода.

*Примечание.* Неформально это можно охарактеризовать как логически самоочевидное утверждение; а именно, «все, что существует, активно ищет доказательства своего существования». Тогда форма существования жизни — это доказательство ее собственного существования. Это утверждение подтверждается тем фактом, что градиентные потоки, которые обеспечивают неравновесное установившееся состояние, могут быть выражены как градиентный спуск и отражают факты наличия «неожиданности» (*surprisal*) или обладание собственной информацией (*self-information*). Данное явление математически эквивалентно градиентному возрастанию данных логарифмической модели или предельной вероятности. Принимается утверждение, что существование данного эффекта подразумевает постоянную плотность, интерпретируемую как предельная вероятность неявной модели, а динамические потоки, которые поддерживают это установившееся состояние, направляются к областям с высокой плотностью вероятности. В байесовской статистике предельная вероятность известна как «модельное освидетельствование или подтверждение». Подразумевается, что диссипативные потоки, которые максимизируют «свидетельства» для модели (т. е. определяют стационарную плотность) предопределяют факт существования и могут быть интерпретированы как «поиск свидетельств» существования. Это обеспечивает формальную основу для телеологического представления о «самоорганизации» в диссипативных структурах как о «самоочевидности». Модели самоорганизующихся социотехнических систем представляют собой, как отмечалось, неравновесные системы с потоками физических и логических данных, которые могут быть описаны представлениями принципа свободной энергии [33, 34].



По построению, это (обобщенный) градиентный поток «само-информации» или «неожиданности» (см. рис. 38).

В терминах марковского бланкета [32] внутренние состояния системы независимы от его внешних состояний, когда они обусловлены его состояниями бланкета («лоскутного одеяла»). Состояния бланкета можно далее разделить на активные и сенсорные состояния, на которые не влияют внутренние и внешние состояния соответственно. Это разделение не является частью определения марковского бланкета, но определяет способ описания состояний бланкета. Для любого данного «одеяла», состоящего из соединенных «лоскутов» (подсистем), набор активных или сенсорных состояний может быть пустым. Динамическая связь между этими состояниями бланкета и внутренним и внешним состояниями асимметрична. Однако асимметрия возникает конкретно в предположении разреженной связи. Исходя из ланжевеновской формулировки описания случайной динамической системы, соответствующая динамика плотности может быть выражена как решения уравнения Фоккера-Планка. Важно отметить, что, поскольку (условно) предполагается неравновесное установившееся состояние, стационарное решение системы Фоккера-Планка позволяет выразить динамику или поток состояний в терминах разложения Гельмгольца. Это разложение делит поток на диссипативные градиентные потоки (по определению самоинформации) любого состояния (то есть его отрицательную логарифмическую вероятность в установившемся состоянии) и бездивергентный (или соленоидальный) поток. Когда это решение может быть разложено в соответствии с марковским бланкетным разбиением, неявные условные зависимости требуют исчезновения определенных соленоидальных условий связи. Это означает, что можно выразить динамику автономных состояний чего-либо (т. е., внутренних и активных состояний) как функцию (если и только если) от «лоскутного» (покровного) и внутреннего состояний.

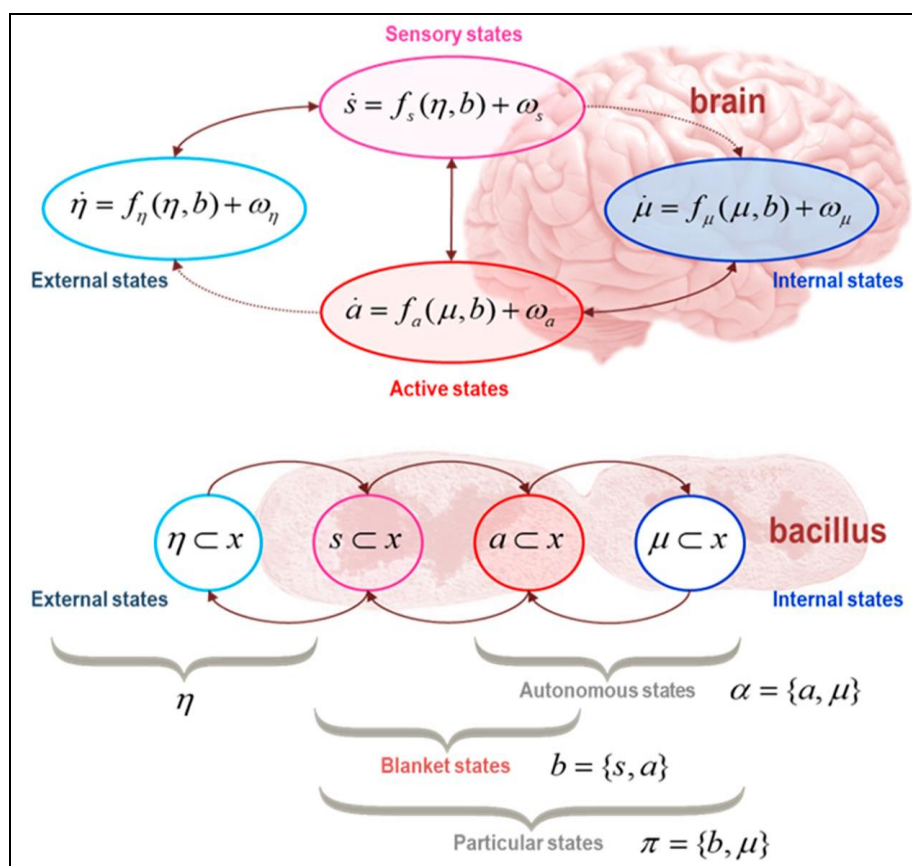


Рис. 38. Интуитивное представление Марковского бланкета [32, 33]

*Примечание.* Мозг (или «бацилла») состоит из внутренних состояний  $\mu$ , которые отделены от внешнего мира (внешние состояния  $\eta$ ) общими состояниями  $b$ , которые сами могут быть разделены на сенсорные состояния  $s$ , представляющие сенсорный эпителий, и на которые непосредственно влияют внешние состояния и активных состояний  $a$ , представляющих эффекторы организмов, которые находятся под прямым влиянием внутренних состояний и действуют на внешние состояния. Восприятие касается минимизации свободной энергии внутренних состояний, в то время как дей-



ствие касается минимизации ожидаемой свободной энергии активных состояний. Напомним, что  $x$  является векторной величиной, поэтому  $f$  является векторной функцией. Индексы в нижнем индексе идентифицируют элементы этого вектора. На рис. 38 показаны наборы состояний, обозначенные каждым нижним индексом. Этот (разреженный) поток исключает соленоидальную связь между автономными и неавтономными состояниями и делает внутренние и внешние состояния условно независимыми. Однако это не исключает соленоидальной связи между внутренним и активным состояниями или между внешним и сенсорным состояниями. Примеры этих форм соленоидальной муфты могут быть обычными. Например, соленоидная связь между внешними и сенсорными состояниями может проявляться колебательными взаимодействиями между внешней средой и органами чувств, которые реагируют на вибрации (например, барабанной перепонкой уха). Точно так же соленоидная связь между внутренним и активным состояниями может быть повсеместной у неодушевленных существ с мозгом - в форме создателей пространства и генераторов центральных паттернов, которые производят стереотипное поведение, такое как фони́рование или дыхание.

*Пример: Ожидаемая свободная энергия и активный вывод.* В модели социотехнической системе рассмотрим только взаимосвязь между внутренним и внешним состояниями и отметим, что динамика внутреннего состояния может рассматриваться как выполнение вариационного градиентного спуска на параметрах вариационной плотности по внешним состояниям. Динамика внутреннего состояния точно следует вариационному градиентному спуску, если предположить, что внутренние состояния параметризуют «лапласовский» аппроксимативный апостериорный спуск (или приблизительно следуют вариационному градиентному спуску, если предполагается рассмотрение более широкого класса вариационных апостериорных функций). Исходя из этого предположения, можно интерпретировать динамику внутренних состояний как выполнение некоторого «перцептивного» вывода о причинах флуктуаций в общих состояниях, а именно, внешних состояниях. Рассмотрим активные состояния. Во-первых, из леммы приближенного байесовского вывода можно выразить динамику автономных состояний (активного и внутреннего) в терминах приближенного градиентного спуска на вариационной свободной энергии. Согласно лемме о маргинальном потоке, если игнорируется соленоидальная связь между внутренним и активным состояниями, то можно разделить этот спуск на отдельные (маргинальные) спуски на внутреннем и активном состояниях, что позволяет осуществить запись соответствующих уравнений, описывающих динамику активных состояний. Определенные условия соленоидальной связи исключаются, когда бланкет Маркова появляется из-за разреженной связи. Это наблюдение свидетельствует о построении потока системных состояний  $x = (\eta, s, a, \mu)$  в неравновесном стационарном состоянии. Составляющие  $x$  называют внешними, сенсорными, активными и внутренними состояниями соответственно. Этот поток можно выразить как решение уравнения Фоккера-Планка в терминах разложения Гельмгольца, также известного как фундаментальная теорема векторного исчисления:  $\dot{x} = f(x) + \omega$ ,  $f(x) = (Q - \Gamma)\nabla J(x)$ . Предполагается, что  $\omega$  – это белый (с нулевой автокорреляцией) гауссовский шум с нулевым средним таким, что  $\omega = N(x; 0, 2\Gamma)$ , где  $\Gamma$  – дисперсия шума. Это стохастическое дифференциальное уравнение также можно представить не в терминах динамически изменяющихся состояний, а в терминах динамически изменяющегося распределения вероятностей по состояниям. Это преобразование достигается с помощью уравнения Фоккера-Планка, с помощью которого изменение распределения по состояниям можно записать как,  $\frac{dp(x,t)}{dt} = \nabla_x f(x)p(x,t) + \Gamma \nabla_x^2 p(x,t)$ , где  $p(x,t)$  – мгновенное распределение по состояниям в данный момент времени  $t$ . Распределение  $p(x,t)$  начинается с распределения  $p(x_0, t) = N(f(x_0, 0), \Gamma)$  из-за начального шума  $\omega$  и начального состояния  $x_0$ . Здесь  $\nabla_x f(x, t)$  – функция градиента и просто обозначает вектор частных производных функции  $f$  по каждому элементу вектора  $x$ ;  $\nabla_x^2 f(x)$  представляет собой вектор вторых частных производных функции.

Представленное разложение лежит в основе принципа свободной энергии и большинства формулировок неравновесного стационарного состояния в нелинейных системах; от молекулярных взаимодействий до эволюции динамических систем. Здесь  $J(x) = -\ln p(x)$  является мерой эффекта «неожиданность» или самоинформацией, а антисимметричная матрица  $Q = -Q^T$  определяет соленоидальный поток. Плотность  $p(x)$  – это стационарная плотность. Положительно определенная матрица  $\Gamma \propto I$  (единичная матрица) играет роль тензора диффузии или ковариационной матрицы, описывающей амплитуду случайных колебаний  $\omega$  (предполагается, что это винеровский процесс). В этой фор-

ме поток можно разложить на диссипативные градиентные потоки  $-\Gamma \nabla J$  и бездивергентный или соленоидальный поток  $Q \nabla J$ . Недиagonalные члены  $\Gamma$  равны нулю, поскольку случайные колебания независимы. Однако эта независимость не исключает соленоидальной связи между состояниями. Для простоты предполагается, что амплитуда колебаний и соленоидальные матрицы течения меняются достаточно медленно с изменением  $x$ , поэтому их можно считать постоянными.

*Примечание.* Схематично рис. 38 иллюстрирует разделение состояний на внутренние состояния ( $\mu$ , синий) и скрытые или внешние состояния ( $\eta$ , голубой), которые разделены марковским «одеялом» ( $b$ ), содержащим сенсорное ( $s$ , в пурпурный) и активные состояния ( $a$ , красный). Верхняя панель показывает это разделение в том виде, в каком оно применяется к действию и восприятию в мозгу. Обратите внимание, что единственные недостающие влияния находятся между внутренним и внешним состояниями - и направленные влияния от внешних (соответственно внутренним) к активным (соответственно сенсорным) состояниям. Выжившие направленные воздействия выделены пунктирными соединителями. В этом случае самоорганизация внутренних состояний соответствует восприятию, в то время как активные состояния связывают внутренние состояния с внешними состояниями. На нижней панели показано то же разделение, но перестроенное таким образом, что внутренние состояния связаны с внутриклеточными состояниями *Bacillus*, где сенсорные состояния становятся поверхностными состояниями или клеточной мембраной, лежащей над активными состояниями (например, актиновыми филаментами цитоскелета). Здесь связь между сенсорным и внутренним, а также между активным и внешним состояниями была подавлена, чтобы выявить простую архитектуру связи, которая ведет к марковскому «одеялу». Автономные состояния ( $\alpha$ ) – это те состояния, на которые не влияют внешние состояния, в то время как отдельные состояния ( $\pi$ ) составляют частицу; а именно автономные и сенсорные состояния - или общие и внутренние состояния. Уравнения движения на верхней панели подтверждают условную независимость марковского одеяла, как описано в основном тексте. При этом не учитывается, что эти уравнения движения определяют марковское «одеяло» или даже что они необходимы для марковского «одеяла», которое может существовать в статических системах без динамики. Вместо этого, эти уравнения представляют потоки, которые при определенных допущениях приводят к возникновению марковского бланкета в установившемся состоянии.

Общий случай соответствует верхней панели на рис. 38, в которой активные состояния напрямую влияют на внешние состояния, а сенсорные состояния напрямую связаны с внутренними состояниями. В обоих случаях, автономные состояния являются функциями отдельных состояний и только от них, а внутренние состояния условно независимы от внешних состояний. Эти два условия служат основой леммы о свободной энергии. Таким образом, можно выразить поток автономных состояний как функцию отдельных состояний и только от них [32]:

$$\begin{bmatrix} f_{\eta}(\eta, b) \\ f_s(\eta, b) \\ f_{\alpha}(\eta, b) \end{bmatrix} = \begin{bmatrix} Q_{\eta\eta} - \Gamma_{\eta\eta} & Q_{\eta s} \\ -Q_{\eta s}^T & Q_{ss} - \Gamma_{ss} \\ & & Q_{\alpha\alpha} - \Gamma_{\alpha\alpha} \end{bmatrix} \begin{bmatrix} \nabla_{\eta} J \\ \nabla_s J \\ \nabla_{\alpha} J \end{bmatrix} \rightarrow (\mu \perp \eta) | b$$

Технически марковское «одеяло» – это набор переменных, которые, если они известны, отображают два других набора условно независимых. Это означает, что совместная плотность внутренних и внешних состояний, обусловленная состояниями бланкета, разлагается на множители следующим образом:  $(\mu \perp \eta) | b \Leftrightarrow p(\mu, \eta | b) = p(\mu | b) p(\eta | b)$

*Примечание.* Этот (разреженный) поток исключает соленоидальную связь между автономными и неавтономными состояниями и делает внутренние и внешние состояния условно независимыми. Однако, как отмечалось, это не исключает соленоидальной связи между внутренним и активным состояниями или между внешним и сенсорным состояниями. Примеры этих форм соленоидальной муфты могут быть обычными. Например, соленоидная связь между внешними и сенсорными состояниями может проявляться колебательными взаимодействиями между внешней средой и органами чувств, которые реагируют на вибрации (например, барабанной перепонкой уха). Точно так же соленоидная связь между внутренним и активным состояниями может быть повсеместной у неодушевленных существ с мозгом - в форме создателей пространства и генераторов центральных паттернов, которые производят стереотипное поведение, такое как фони́рование или дыхание.

Наиболее вероятный путь определения направления потоков соответствует результату применения вариационного принципа наименьшего действия, где *свободная энергия*  $F(\pi) \equiv F[q_\mu(\eta), \pi]$ , является функционалом от вариационной плотности  $q_\mu(\eta) \approx p(\eta|b)$ , параметризованная условным математическим ожиданием внутренних состояний. В общем, определяемая свободная энергия является верхней границей эффекта «неожиданности» (surprisal) отдельных состояний:

$$\begin{aligned}
 F(\pi) &= \underbrace{E_q[J(\eta, \pi)]}_{\text{энергия}} - \underbrace{H[q_\mu(\eta)]}_{\text{энтропия}} \\
 &= \underbrace{J(\pi)}_{\text{неожиданность}} + \underbrace{D[q_\mu(\eta) \| p(\eta|b)]}_{\text{связанность доказательств}} \\
 &= \underbrace{E_q[J(\eta|\pi)]}_{\text{неточность}} + \underbrace{D[q_\mu(\eta) \| p(\eta)]}_{\text{сложность}}
 \end{aligned}$$

Математические ожидания, в том числе неявные в Кульбак-Лейблер (*KL*)-дивергенциях, относятся только к внешним состояниям, так как свободная энергия является функцией отдельных состояний, а не их среднего значения. Этот функционал может быть выражен в нескольких формах; а именно, ожидаемая энергия за вычетом энтропии вариационной плотности, которая эквивалентна самоинформации, связанной с конкретными состояниями (то есть неожиданностью), плюс расхождение *KL* между вариационной и апостериорной плотностями (то есть, связанная с очевидностью). В свою очередь, это можно разложить на ожидаемую отрицательную логарифмическую вероятность конкретных состояний (то есть в точности) и расхождение *KL* между апостериорной и априорной плотностями (то есть сложность). Таким образом, свободная энергия представляет собой *функцию Ляпунова* для ожидаемого потока автономных состояний.

*Примечание.* Этот функционал называется свободной энергией, потому что он включает энтропию и ожидаемый потенциал. Тем не менее, это может быть переформулировано как неожиданность, связанная с конкретными состояниями (сама интерпретируемая как потенциал) и *KL*-дивергенция (т.е. относительная энтропия) между вариационной плотностью, параметризованной любой точкой на внутреннем многообразии, и апостериорной над внешние состояния при заданных состояниях бланкета. В этом случае ожидаемый потенциал является неожиданностью, а расхождение *KL* находится между вариационной плотностью - параметризованной любой точкой на внутреннем многообразии - и апостериорными по сравнению с внешними состояниями при заданных состояниях бланкета. В байесовской статистике и машинном обучении это расхождение известно как наличие доказательств, так как обеспечивает неотрицательную границу неожиданных или логарифмических данных. Чтобы разрешить интерпретацию неожиданности в терминах модельных свидетельств, необходимо выразить «неожиданность» в терминах генеративной модели; а именно, общая плотность общих причин и последствий. Здесь под генеративной моделью понимается установившаяся плотность распределения вероятностей по внешним (причинам) и бланкетным (последствиям) состояниям.

Существует два способа представить эту интерпретацию потоков на внутреннем (статистическом) многообразии. Первый и самый простой – это установить, что вариационная плотность, параметризованная внутренним состоянием или координатой на внутреннем многообразии, является апостериорной по сравнению с внешними состояниями. С этой точки зрения представленная выше граница выше стремится к нулю, и поток (условные ожидания, заданные бланкетные состояния) внутренних состояний можно рассматривать как выполнение точного байесовского вывода. Однако, эта интерпретация не может указать, как (условные ожидания) внутреннее состояние параметризует косвенную байесовскую уверенность в определении внешних состояниях. Необходимо определить функциональную форму для вариационной плотности и связать внутренние состояния с ее параметрами или достаточной статистикой (например, средним значением и точностью). Однако, как только определяют параметризованную форму, то теряется форма точного байесовского вывода и переходят в область приближенного байесовского вывода. Это связано с тем, что точная эквивалентность вариационной и апостериорной плотности по внешним состояниям больше не гарантируется. Это отменяет указанную выше оценку и приводит к вариационному Байесовскому выводу [33, 34].

*Примечание.* Этот вид вывода преобладает в статистической литературе и литературе по машинному обучению, потому что относительно просто вычислить вариационную свободную энергию,



заданную параметризованной формой для вариационной плотности. С другой стороны, оценить строгость доказательства байесовской модели практически невозможно. В некоторых случаях вариационного Байеса использование вариационной свободной энергии было введено в контексте формулировки интеграла по путям квантовой электродинамики. По сути, трудноразрешимую проблему интеграции переводят в решаемую проблему оптимизации. Другими словами, предоставляют вычисляемую целевую функцию, минимизация которой будет приближать минимизацию ошибки достоверности оценки, которые всегда ограничены сверху вариационной свободной энергией. Действительно, в машинном обучении вариационная свободная энергия часто упоминается как доказательная нижняя граница.

Это означает, что можно интерпретировать поток на внутреннем многообразии как точный байесовский вывод. В качестве альтернативы, если выбрать функциональную форму для вариационной плотности, это означало, что поток стремится к приближенному байесовскому выводу. Это различие было сформулировано в терминах разницы между частной и вариационной свободной энергией. Функциональная форма для вариационной плотности была гауссовой, что привело к повсеместной форме вариационного Байеса при предположении Лапласа. Ключевым моментом является то, что граница может быть равна нулю или нет. При этом размер (границы) значения не имеет. Это следует из того факта, что внутренние состояния выполняют не ожидаемый градиентный спуск. Если соответствующая вариационная свободная энергия имеет (приблизительно) одно и то же значение - с точностью до аддитивной константы  $c$  - динамика будет везде одинаковой (поскольку градиенты не зависят от константы). Другими словами, не имеет значения, является ли интерпретация приближения с малым или большим значением границы: приближенная интерпретация байесовского вывода требует только того, чтобы граница была (приблизительно) одинаковой везде на статистическом многообразии (см. рис. 39).

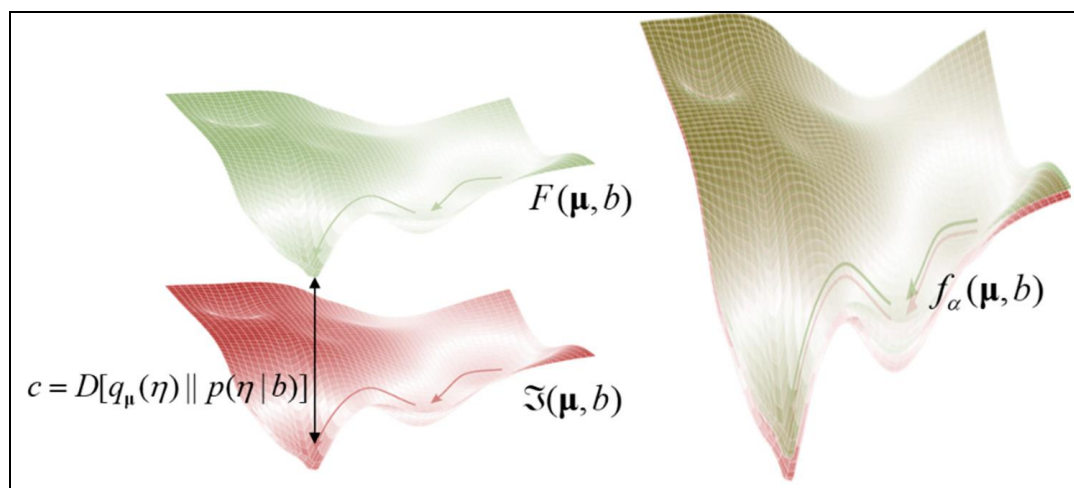


Рис. 39. Границы достоверности и градиентные потоки [32]

Схема на рис. 39 на интуитивном уровне показывает, что градиент течет в неожиданном направлении (обозначено розовым цветом) как функция некоторого статистического многообразия (здесь условные ожидания внутренних состояний при заданных состояниях бланкета) – то же самое, что градиентные потоки на свободном от вариаций энергии (зеленая), тогда и только тогда, когда  $KL$ -дивергенция или оценка достоверности сохраняются над многообразием. На левой панели показаны две функции сбоку, а на правой панели показывает вид сверху.

В более общем смысле, соленоидная связь может быть существенной для самоорганизации и активного вывода (особенно, если рассматривать ее в свете теории колебаний и кооперации).

### Ожидаемая свободная энергия и активный вывод

До сих пор рассматривалась только взаимосвязь между внутренним и внешним состояниями. Установлено, что динамика внутреннего состояния может рассматриваться как выполнение вариационного градиентного спуска по параметрам вариационной плотности по внешним состояниям. Динамика внутреннего состояния точно следует вариационному градиентному спуску, если предположить, что внутренние состояния параметризуют лапласовский аппроксимативный апостериорный

спуск, или они приблизительно следуют вариационному градиентному спуску, если предположить более широкий класс вариационных апостериорных функций. Исходя из этого, можно интерпретировать динамику внутренних состояний как выполнение некоторого «перцептивного» вывода о причинах флуктуаций в общих состояниях, а именно, внешних состояниях. Из приближенной леммы байесовского вывода следует, что можно выразить динамику автономных состояний (активного и внутреннего) в терминах приближенного градиентного спуска на вариационной свободной энергии (рис. 40).

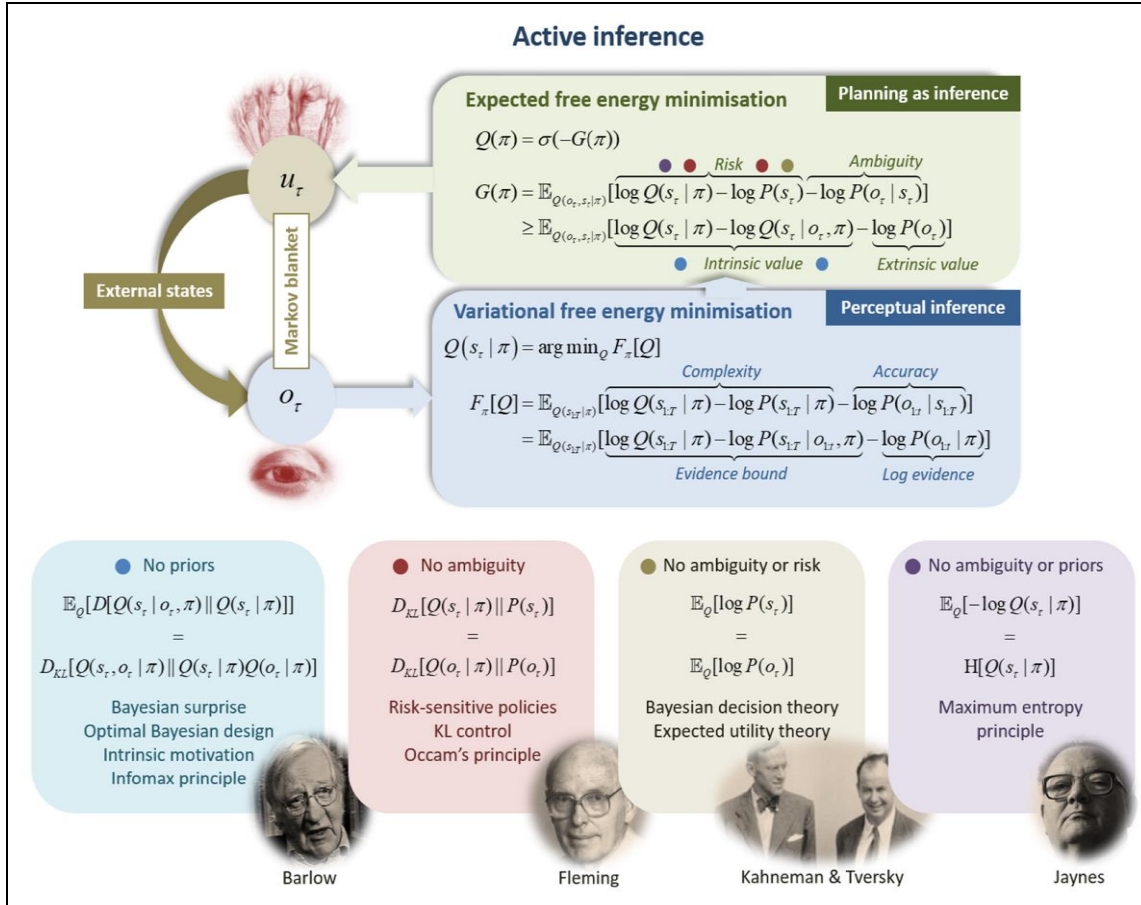


Рис. 40. Активный вывод и минимизация свободной энергии [35]

Вверху: активный вывод дискретного пространства состояний требует двух отдельных операций минимизации – одной вариационной свободной энергии для восприятия и другой ожидаемой свободной энергии для выбора действия. Внизу: функционал EFE можно разложить различными способами, чтобы получить множество важных целей.

Согласно лемме о маргинальном потоке, если игнорируем соленоидальную связь между внутренним и активным состояниями, то можем разделить этот спуск на отдельные (маргинальные) спуски на внутреннем и активном состояниях, что позволяет записать динамику активных состояний как  $\frac{da}{dt} \approx (\Gamma_{aa} - Q_{aa}) \nabla_a F(s, \alpha)$ , где  $\Gamma_{aa}, Q_{aa}$  являются блочными матрицами, соответствующими только взаимодействиям между активными состояниями в больших  $\Gamma$ - и  $Q$ -матрицах. Важно отметить, что из определения вариационной свободной энергии, имеем

$$F(i, s, \alpha) = D_{KL}[q(e; \mu) \| p^*(e, i, s, a)] = \underbrace{\mathbb{E}_{q(e| \mu)} [-\ln p^*(i, s, a)]}_{\text{неточность}} + \underbrace{D_{KL}[q(e; \mu) \| p^*(e)]}_{\text{сложность}}.$$

Важно отметить, что единственный член в этом разложении, который зависит от активных состояний  $a$ , – это первый член эффекта «неточность». Таким образом, можем записать динамику активных состояний как,  $\frac{da}{dt} \approx (\Gamma_{aa} - Q_{aa}) \nabla_a \mathbb{E}_{q(e| \mu)} [-\ln p^*(i, s, \alpha | e)]$ , откуда можем увидеть, что

динамика активных состояний эффективно минимизирует неточность (или максимизирует точность). Фактически можно интерпретировать динамику активных состояний в плотности *NESS*, чтобы гарантировать, что вариационные «убеждения», кодируемые «покрытиями» и внутренними состояниями системы, являются как можно более точными. Поскольку активные состояния могут влиять только на внешние состояния, но не на внутренние, то полученный результат достигается путем воздействия на внешние состояния, чтобы привести их в соответствие с «уверенностями», представленными внутренними состояниями - следовательно, реализуется *активный вывод*.

Приведенные условия обеспечивает хорошую характеристику динамики системы в состоянии равновесия, но часто также интересуются свойства динамических систем как самоорганизующихся к равновесию. В частности, стремятся охарактеризовать природу активных состояний во время этого процесса самоорганизации, чтобы можно было понять необходимые виды активного поведения, которые должна проявлять любая самоорганизующаяся система. Для понимания природы явления этой самоорганизации, определяют другую теоретико-информационную величину, ожидаемую свободную энергию (*EFE* – *expected free energy*), которая служит верхней границей эффекта «неожиданность» на протяжении всего процесса самоорганизации, с равенством только в состоянии своего равновесия. Поскольку есть эта верхняя граница, можем интерпретировать самоорганизующиеся системы вдали от равновесия, следя за их неожиданной динамикой как аппроксимирующей ожидаемой минимизацией свободной энергии, используя логику, прямо аналогичную приближенной лемме байесовского вывода. И наоборот, изменение этой логики позволяет строить самоорганизующиеся системы, определяя некоторую желаемую плотность *NESS*, а затем предписывая динамику, которая просто минимизирует величину *EFE*.

В результате, можем интерпретировать *EFE* как эффективную количественную оценку несоответствия между текущей прогнозируемой плотностью и равновесием. Из этого следует, что *EFE* также обязательно является функцией Ляпунова самоорганизующейся динамики, и имеет смысл интерпретировать самоорганизующуюся динамику под марковским «одеялом» как минимизирующую *EFE*. И наоборот, если определить набор динамики, которая самоорганизуется для некоторого заданного аттрактора  $p^*(e, i, s, a)$ , тогда нужно просто определить динамику, которая минимизирует *EFE* для достижения сходимости к равновесию (но может оказаться локальным минимумом).

В [36-38] показана взаимосвязь принципа свободной энергии с информационно-термодинамическим критерием взаимосвязи устойчивости (функции Ляпунова), управляемости и робастности динамической эволюции неравновесных процессов самоорганизации.

*Пример.* Рассмотренный стремительно развивающийся принцип свободной энергии (*Free-energy principle, Karl J. Friston*) проиллюстрирован на рис. 41.

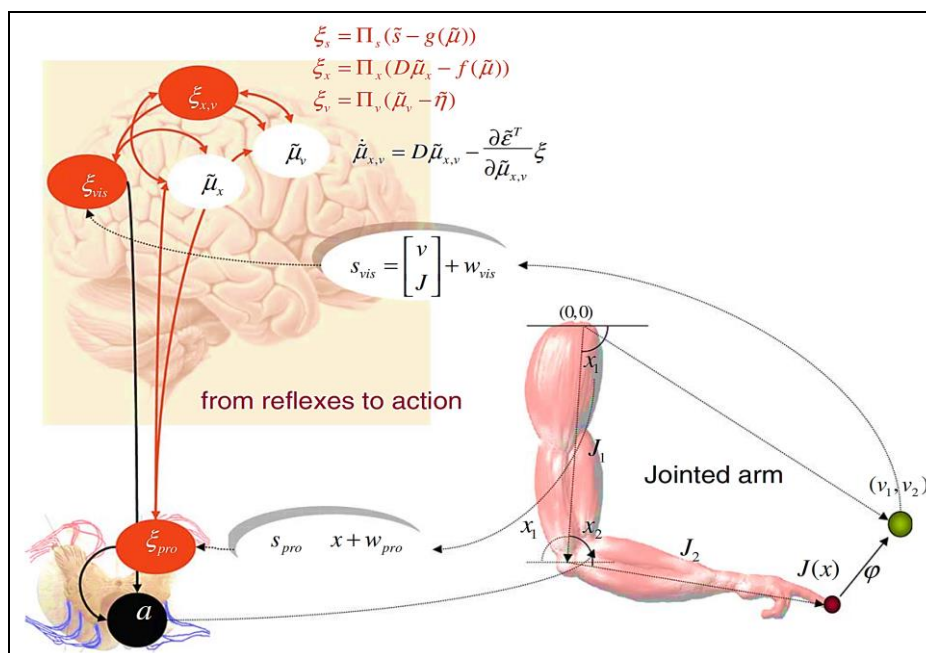


Рис. 41. Принцип свободной энергии как синергизм рефлексии и когнитивного процесса



Как отмечалось, выражение для свободной энергии  $F$  может быть определено тремя эквивалентными способами: 1) энергия минус энтропия; 2) Информационная дивергенция плюс сюрприз; 3) Сложность минус точность. Математически перечисленные три выражения имеют следующий вид:

$$F = -\langle \ln p(s, \psi | m) \rangle_q + \langle \ln q(\psi | \mu) \rangle_q \quad 1)$$

$$= D(q(\psi | \mu) \| p(\psi | s, m)) - \ln p(s | m) \quad 2)$$

$$= D(q(\psi | \mu) \| p(\psi | m)) - \ln p(s | \psi, m) \quad 3)$$

соответственно.

Здесь  $D(\square || \square)$  означает дивергенцию Кульбака-Лейблера соответствующих распределений вероятностей событий и Байесовский закон произведения вероятностей правдоподобия суждений  $p(s, \psi | m) = p(s | \psi, m) p(\psi | m)$ . Принцип минимизация свободной энергии при съеме информации датчиком (см. рис. 41) для логического вывода и принятия решения позволяет сформировать структуру когнитивного ПИД-регулятора с минимально необходимой управляющей силой и по своей структуре совпадает с результатом минимизации энтропии (минимизации потерь энергии) или максимизации извлекаемой полезной работы (максимально полезной работы при минимальной управляющей силе регулятора). Включение совместного применения интеллектуального и когнитивного регуляторов в контур управления, со спроектированными в соответствии с указанными принципами базами знаний позволяет гарантированно достигать роботу цели управления в условиях нештатных ситуаций.

Динамика взаимодействия человека с роботом и теоретические допущения [39] представлены на рис. 42.

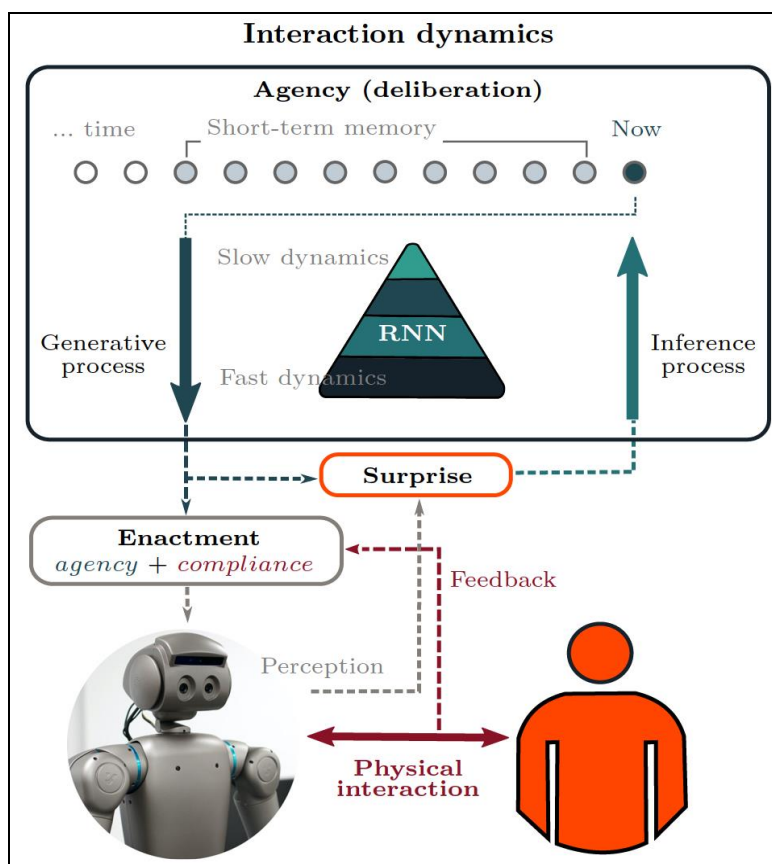


Рис. 42. Управление взаимодействием человека-оператора с роботом, основанном на способностях операторов действовать совместно и подчиняться принятым решениям

Способность действовать изучается через оптимизацию свободной энергии (неожиданности) в кратковременной памяти прямого восприятия. Следовательно, из уникального прошлого робота

предполагаемое будущее переосмысливается в иерархической рекуррентной нейронной сети (*RNN*) для управления осмыслением и проецируется в пространство восприятия. Способность к податливости модулируется как пропорциональная интегральная (*PI*) схема управления для автоматической адаптации тела к действиям, индуцированным партнером.

*Примечание.* Принять считать, что свободная энергия минимизируется двумя фундаментальными способами: посредством прогнозирующего кодирования для косвенного восприятия и посредством активного вывода для целенаправленного действия. В прогнозирующем кодировании эффект «неожиданность» сводится к минимуму в восходящем направлении, так как внутреннее состояние согласуется с генерируемыми более согласованными предсказаниями в отношении сенсорных восприятий данных. Этот принцип был исследован в теории психологических исследований социального познания. Иными словами, при активном выводе сенсорные предсказания пытаются выполнить по нисходящей траектории пути, предпринимая целенаправленные действия в окружающей среде.

Независимо от минимизации эффекта неожиданности, в теории *FEP* предполагается, что прогноз чувствительности сопровождается оценкой точности. Следовательно, неожиданность считается минимальной по отношению к точности, т.е., что свободная энергия минимизируется в большей степени, когда она связана с оценкой высокой точности (т.е., сильным убеждением). Практическое значение предыдущего заявления о том, что агент с сильным убеждением распространяется на активно-эгоцентрически по отношению к окружающей среде без изменения своего внутреннего состояния или намерения, тогда как агент со слабой мерой уверенности имеет тенденцию действовать адаптивно к окружающей среде, легко изменяя свое внутреннее состояние или намерение.

Робот на иерархической рекуррентной нейронной сети получает способность действовать (или осмысливать) и податливостью (реактивность, приспособляемость или адаптацией) по отношению к человеку. действия партнера. В соответствии с теорией принципом свободной энергии (*FEP – free energy principle*, *K. Friston* [32]), координационное управление роботом рассматривается как процесс вариационной оптимизации, который включает иерархическое динамическое представление, в котором поток информации сверху вниз (разработанный в процессе генерации) характеризует целенаправленное действие агента (или намерения вести себя) в контексте взаимодействия, и восходящий поток информации (процесс вывода) объясняет их последствия; поэтому конфликты, которые могут возникнуть между этими двумя процессами, пытаются уменьшить за счет минимизации свободной энергии как статистической величины.

Для того, чтобы наделить когнитивные вычислительные модели способностью функционировать в реалистичных средах в условиях неопределенности, необходимо рассмотреть вероятностные модели пространственного познания, реализуемые в головном мозге (ГМ). На рис. 43 показан обзор [40] всех используемых методов и реализации на основе них пространственной модели и механизмов.

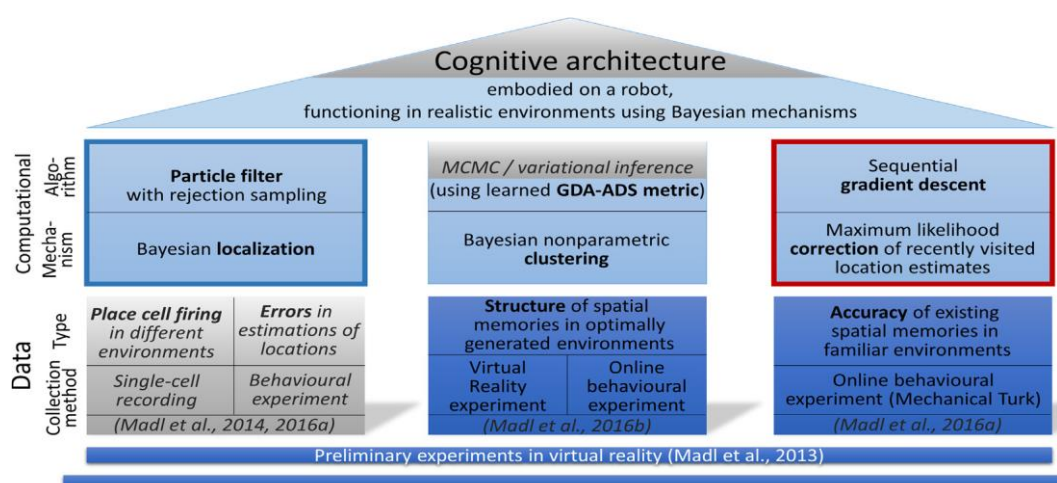


Рис. 43. Когнитивная архитектура робота, функционирующего в реалистичной среде с использованием байесовских механизмов

Для реализации возможности планирования новых путей достижения целей агент (будь то биологический или искусственный), должен обладать способностью определить себя, свою цель и воз-

можные препятствия даже в условиях зашумления. С вероятностной точки зрения эту проблему определения можно описать как байесовскую сеть. Во избежание необходимости вычислений для каждого посещенного места и каждого наблюдаемого ориентира, их можно разделить на подгруппы. В частности, в первом приближении решением этой проблемы может быть разделение на:

- применение байесовских сигналов для интеграции зашумленных наблюдений в оценку местоположения;
- байесовскую локализацию для поддержания этого местоположения во времени;
- коррекцию на основе максимального правдоподобия для запоминания последней оценки местоположения при его повторном посещении.

Эти вероятностные модели для определения собственного местоположения и местоположения объектов составляют основу когнитивного программного агента, способного функционировать в реалистичном симуляторе робота. На рис. 44 представлен обзор [40] того, как байесовские механизмы, описанные выше, могут быть реализованы в пространствах соответствующих областей ГМ.

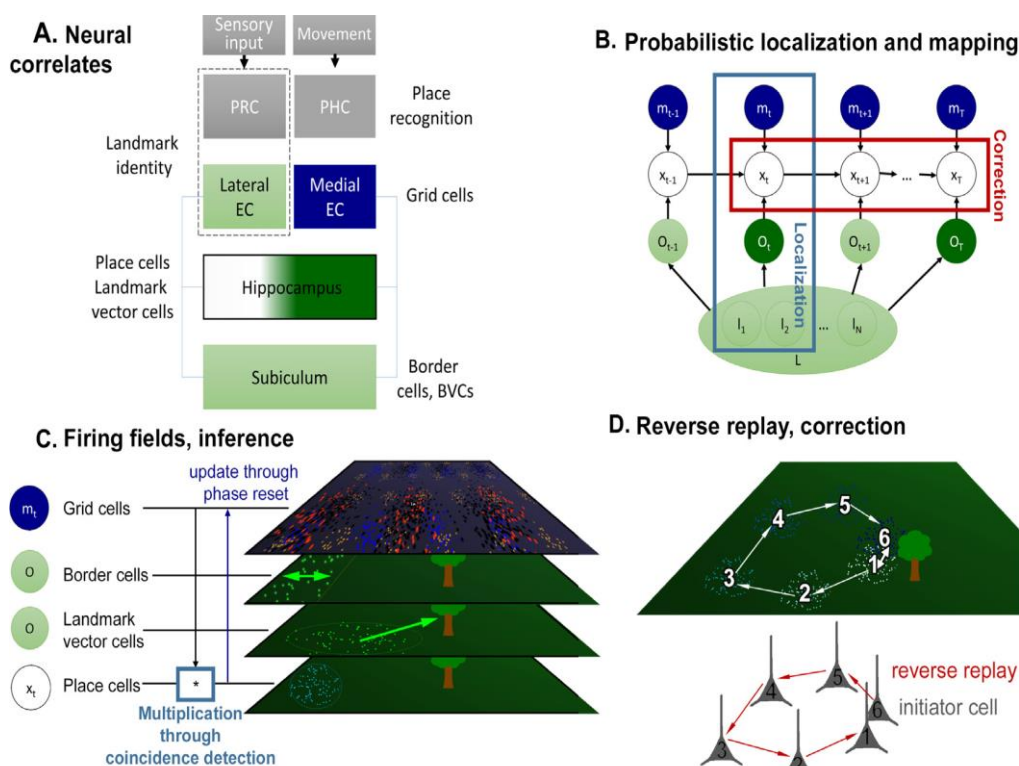


Рис. 44. Вероятностная пространственная локализация и отображение, реализуемое ГМ

На рис. 44, А представлены нейронные корреляты локализации, где *PRC* – периринальная кора (англ. *Perirhinal cortex*), *PHC* – парагиппокампальная кора (англ. *Parahippocampal cortex*), *EC* – энторинальная кора (англ. *Entorhinal cortex*). На рис. 44, В отображена вероятностная графическая модель проблемы одновременной локализации и отображения. Рисунок 44, С иллюстрирует поля возбуждения нейронов в момент локализации [41]. Цветные точки представляют собой всплески соответствующих нейронов в рецептивных точках. Нейронная решетка, а также информация о нейронах границ и нейронах векторов ориентиров интегрируются в нейроны места при обнаружении совпадений для получения оптимальной оценки местоположения. Эта оценка используется для обновления представлений нейронной решетки посредством обнуления фазы для борьбы с накоплением ошибок интеграции пути. Рисунок 44, D демонстрирует цикл (поля возбуждения нейронов 1 – 6), данные в котором возможно скорректировать путём распознавания тех же мест в позициях 1 – 6 с помощью обратного воспроизведения.

Рассмотренные особенности технологии когнитивных вычислений пояснены в Части 2 данной статьи на примере, имеющем самостоятельное теоретическое и практическое значение: интеллектуальный регулятор на технологии эмоционального вычисления типа *BELBIC*.



*Примечание.* Ряд других моделей [42] поведенческих реакций коллективов роботов составляют научный базис и программно-алгоритмическую платформу для разработки интеллектуальных когнитивных систем управления роботизированными социотехническими технологическими процессами и промышленными системами. В [37] приведены подробные технические детали решения рассматриваемых задач. Обсуждение особенностей технологии квантовых когнитивных вычислений приведены в Эпilogue к [37] и пояснены на примерах, имеющих также самостоятельное теоретическое и практическое значение: 1) оценка точности квантовых методов обработки ЭЭГ; 2) квантовые модели эмоционального поведения робота.

В последующих Частях 2, 3 и 4 данной работы рассмотрены также кратко физические и информационно-термодинамические ограничения на процессы самоорганизации баз знаний когнитивных и интеллектуальных регуляторов, встраиваемых в контур систем управления роботами за счет извлечения и передачи информации в виде управления исполнительными устройствами. Данные интеллектуальные устройства и средства когнитивного управления позволяют существенно расширить возможности интеллектуальных когнитивных роботов, погруженных в среду социотехнических систем.

## Заключение

Из приведенных результатов видно, что с развитием программного обеспечения и возможностей аппаратных средств робототехники, появления отраслей больших данных и ИИ, когнитивных вычислений, а также более общих квантовых когнитивных исчислений, в промышленных разработках стало уделяться значительное внимание развитию интеллектуальных когнитивных роботов самого широкого применения в социотехнических системах как в академических кругах, так и в промышленности.

Следует отметить особую роль применения информационно-термодинамических моделей квантовой самоорганизации неточных баз знаний в интеллектуальных когнитивных регуляторах, на основе которых проектируется интеллектуальное и умное производство, а также формирование физически обоснованной платформы проектирования робастных структур подобного типа производств на основе принципа свободной энергии.

## Список литературы

1. Agha A. et. al. NeBula: Quest for Robotic Autonomy in Challenging Environments; TEAM CoSTAR at the DARPA Subterranean Challenge (preprint version) // arXiv:2103.11470v1 [cs.RO] 21 Mar 2021.
  2. Hudson N., et al. Heterogeneous Ground and Air Platforms, Homogeneous Sensing: Team CSIRO Data61's Approach to the DARPA Subterranean Challenge // arXiv:2104.09053v1 [cs.RO] 19 Apr 2021.
  3. Jagatheesaperumal S.K. et. al. The Duo of Artificial Intelligence and Big Data for Industry 4.0: Review of Applications, Techniques, Challenges, and Future Research Directions // arXiv:2104.02425v1 [cs.AI] 6 Apr 2021.
  4. World Robotics 2016 Service Robots. International Federation of Robotics (IFR). 2016.
  5. CORBYS Cognitive Control Framework for Robotic Systems (FP7 — 270219). Deliverable D2.1 Requirements and Specification. 2011.
  6. Yang G.-Z. The grand challenges of *Science Robotics* // Sci. Robotics. — 2018. — Vol. 3. — P. 7650.
  7. Sniderman B. et al. Industry 4.0 and manufacturing ecosystems. — Deloitte's Univ. Press, 2020.
  8. White Paper: AI in Industrial Automation. — German Electrical and Electronic Manufacturers' Association, 2021.
- Dagnaw G. Artificial Intelligence Towards Future Industrial Opportunities and Artificial Intelligence Towards Future Industrial Opportunities and Challenges // The 6th Annual ACIST Proceedings (2020) African Conference on Information Systems and Technology 16. 2020. Available from: <https://digitalcommons.kennesaw.edu/acist/2020/allpapers/16>.

9. Yang F., Gu S. Industry4.0, a revolution that requires technology and national strategies // *Complex & Intelligent Systems*. — 2021. — Vol. 7. — Pp. 1311–1325. Available from: <https://doi.org/10.1007/s40747-020-00267-9>.
10. Lee Ch., Lim Ch. From technological development to social advance: A review of Industry 4.0 through machine learning // *Technological Forecasting & Social Change*. — 2021. — Vol. 167. — Pp. 120653.
11. Jamwal A. et al. Industry 4.0 Technologies for Manufacturing Sustainability: A Systematic Review and Future Research Directions // *Appl. Sci.* — 2021. — Vol. 11. — Pp. 5725. Available from: <https://doi.org/10.3390/app11125725>.
12. Wang B. et al. Smart Manufacturing and Intelligent Manufacturing: A Comparative Review // *Engineering*. — 2021. — Vol. 7. — Pp. 738–757.
13. Butepage J., Kragic D. Human-Robot Collaboration: From Psychology to Social Robotics // *arXiv:1705.10146v1 [cs.RO]* 29 May 2017.
14. Wu H. et al. Put the Bear on the Chair! Intelligent Robot Interaction with Previously Unseen Objects via Robot Imagination // *arXiv:2108.05539v1 [cs.RO]* 12 Aug 2021.
15. Wan J. et al. Artificial Intelligence-Driven Customized Manufacturing Factory: Key Technologies, Applications, and Challenges // *Proceedings of the IEEE* — October 2020. DOI: 10.1109/JPROC.2020.3034808.
16. Fujimaki R. The 6 Challenges of Implementing AI in Manufacturing // *Foundry Management & Technology*. — Mar 01, 2021 [[The%206%20Challenges%20AI.pdf](#)].
- Hentout A. et al. Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017 // *Advanced Robotics*. — 2019. Available from: <https://doi.org/10.1080/01691864.2019.1636714>.
17. Krakovski M. et al. “Gymmy”: Designing and Testing a Robot for Physical and Cognitive Training of Older Adults // *Appl. Sci.* — 2021. — Vol. 11. — Pp. 6431. <https://doi.org/10.3390/app11146431>.
18. Hüsing E. et al. Determining Robotic Assistance for Inclusive Workplaces for People with Disabilities // *Robotics*. — 2021. — Vol. 10. — Pp. 44. Available from: <https://doi.org/10.3390/robotics10010044>.
19. Wang Y. et al. The WL\_PCR: A Planning for Ground-to-Pole Transition of Wheeled-Legged Pole-Climbing Robots // *Robotics*. — 2021. — Vol. 10. — P. 96. Available from: <https://doi.org/10.3390/robotics10030096>.
20. Horák M. et al. Service Robots for Motion and Special Applications on the Vertical Oriented Walls // *In Service Robots (Chapter 5)*. — INTECH. — 2017. Available from: <http://dx.doi.org/10.5772/intechopen.70>.
21. Masood M.U., Haghshenas-Jaryani M. A Study on the Feasibility of Robotic Harvesting of Chile Pepper // *Robotics*. — 2021. — Vol. 10. — P. 94. Available from: <https://doi.org/10.3390/robotics10030094>.
22. Roldán J.J. et al. Robots in Agriculture: State of Art and Practical Experiences // *In Service Robots (Chapter 4)*. — INTECH. — 2017. Available from: <http://dx.doi.org/10.5772/intechopen.69874> [Downloaded from: <http://www.intechopen.com/books/service-robots>].
23. Sun S. et al. Autonomous Quadrotor Flight despite Rotor Failure with Onboard Vision Sensors: Frames vs. Events // *arXiv: 2102.13406v1 [cs.RO]* 26 февраля 2021.
- Prescott T.J., Robillard J.M. Are friends electric? The benefits and risks of human-robot relationships // *iScience*. — 2021. — Vol. 24. — P. 101993. Available from: <http://creativecommons.org/licenses/by/4.0/>.
24. Kanoh H. Immediate Response Syndrome and Acceptance of AI Robots --Comparison between Japan and Taiwan // *Procedia Computer Science*. — 2017. — Vol. 112. — Pp. 2486–2496.

25. Moetesum M. et al. Socially Believable Robots // IntechOpen. — 2018. Available from: <http://dx.doi.org/10.5772/intechopen.71375>.
26. Li Y. et al. Design of a Flying Humanoid Robot Based on Thrust Vector Control // arXiv:2108.11557v1 [cs.RO] 26 Aug 2021.
27. Risiglione M. et al. Passivity-based control for haptic teleoperation of a legged manipulator in presence of time-delays // arXiv:2108.07658v1 [cs.RO] 17 Aug 2021.
28. Minniti M.V. et al. Model Predictive Robot-Environment Interaction Control for Mobile Manipulation Tasks // arXiv:2106.04202v1 [cs.RO] 8 Jun 2021 2021.
29. Friston K.J. et al. Some Interesting Observations on the Free Energy Principle // Entropy. — 2021. — Vol. 23. — P. 1076. Available from: <https://doi.org/10.3390/e23081076>.
30. Millidge B. et al. A mathematical walkthrough and discussion of the free energy principle // arXiv:2108.13343v1 [cs.AI] 30 Aug 2021.
31. Parr, T., et al. Markov blankets, information geometry and stochastic thermodynamics // Philos. Trans. of the Royal Society A. — 2020. — Vol. 378(2164). — P. 20190159.
32. 35 Da Costa, L. et al. Active inference on discrete state-spaces: a synthesis // arXiv preprint arXiv:2001.07203. 2020.
33. Кореньков В.В., Ульянов С.В., Шевченко А.А., Шевченко А. В. Интеллектуальная когнитивная робототехника. Ч. 1: Технологии квантовых когнитивных вычислений. — М.: Курс, 2021.
34. Ulyanov S.V. Quantum Algorithm of Imperfect KB Self-organization Pt I: Smart Control – Information — Thermodynamic Bounds // Artificial Intelligence Advances. — 2021. — Vol. 3. — No 2. — Pp. 13–36.
35. Зрелов П.В., Кореньков В.В., Тятюшкина О.Ю. и др. Устойчивость, управляемость и робастность когнитивного управления: Термодинамика эволюционных информационных процессов // Системный анализ в науке и образовании. — 2021. — № 3.
36. Chame F.M. et al. Towards hybrid primary intersubjectivity: a neural robotics library for human science // arXiv:2006.15948v1 [cs.RO] 29 Jun 2020.
37. Madl T. et al. A computational cognitive framework of spatial memory in brains and robots // Cognitive Systems Research. — 2018. — Vol. 47. — Pp. 147–172. Available from: <http://dx.doi.org/10.1016/j.cogsys.2017.08.002>.
38. Baars B.J., Gage N.M. Cognition, brain, and consciousness: introduction to cognitive neuroscience. - Academic Press, Elsevier, 2010.
39. Hekmatmanesh A. et al. Review of the State-of-the-art on Bio-signal-based Brain-controlled Vehicles // arXiv:2006.02937v1 [eess.SP] 3 Jun 2020.