

Научная статья

Статья в открытом доступе

УДК 331.101.1: 004:896

doi: 10.30987/2658-4026-2024-3-374-385

Ожидаемые проблемы человеко-машинного взаимодействия в системах Искусственного Интеллекта

Сохин Игорь Георгиевич^{1✉}

¹ АО «Научно-исследовательский институт авиационного оборудования»;
Московская область, Жуковский, Россия

¹ isokhin@yandex.ru

Аннотация.

На примере опыта управления интеллектуальными космическим роботами анализируются проблемы человеко-машинного взаимодействия, которые ожидаются в перспективных эргатических системах искусственного интеллекта. В качестве объекта исследования выбран интеллектуальный космический робот. Анализируется понятие «интеллект», его информационная природа, сходство и отличия искусственного и естественного интеллектов, механизм формирования информационных моделей искусственного интеллекта. Рассматриваются два альтернативных подхода к созданию и обучению искусственного интеллекта – символьный и нейросетевой, их свойства, сильные и слабые стороны. Представлена структурно-функциональная схема космического робота с человеком-оператором в контуре управления и тремя уровнями управления. Показано, что оптимальное взаимодействие человека-оператора с роботом заключается в передаче роботу уровня управления, наилучшим образом соответствующего сложившейся ситуации. Анализ информационных связей в схеме управления интеллектуальным роботом показал тенденцию возрастания уровня опосредованного взаимодействия человека-оператора со средой, вместо реальных объектов управления используются их виртуальные образы. Все это должно привести к трансформации деятельности человека. В заключении сформулированы ожидаемые проблемы взаимодействия человека с системами искусственного интеллекта: сохранить лидерские позиции человека в отношении искусственного интеллекта; научиться жить и работать в виртуальном пространстве, не теряя при этом связи с физической реальностью; разработать методы критического оценивания адекватности поведения систем искусственного интеллекта, опирающиеся на человеческие ценностные ориентиры и здравый смысл; трансформировать систему подготовки человека-оператора к взаимодействию с системой искусственного интеллекта.

Ключевые слова: Искусственный интеллект, ИИ-технологии, СИИ, машинное обучение, кодирование, информационная модель, человеко-машинное взаимодействие, адаптивность, космический робот.

Для цитирования: Сохин И.Г. Ожидаемые проблемы человеко-машинного взаимодействия в системах Искусственного Интеллекта // Эргодизайн. №3 (25). 2024. С. 374-385. <http://dx.doi.org/10.30987/2658-4026-2024-3-374-385>.

Original article

Open access article

Expected Problems of Human-Machine Interaction in Artificial Intelligence Systems

Igor G. Sokhin^{1✉}

¹ JSC Research Institute of Aviation Equipment; the Moscow Region, Zhukovsky, Russia

¹ isokhin@yandex.ru

Abstract.

In the context of the digital transformation of the economy, the share of intellectual labour is significantly increasing, especially in science and education. In these conditions, it is important to take into account indicators of professional reliability, which are directly related to the employee's performance quality. The essence of the concepts of "intellectual work" and "professional reliability" is revealed. The functional system of intellectual work is substantiated based on

considering the professional reliability indicators. The specificity of the information impact as a production factor and stressor in professional activity is shown. Ideas about information load and information hygiene are given. The paper analyses and discusses the results of scientific research on identifying factors of intellectual work in the context of education digitalization, which negatively affect health and cause teachers' professional diseases. Recommendations are theoretically substantiated and given for implementing areas of information hygiene and mental hygiene of intellectual work as a condition for the teacher's professional reliability in the context of education digitalization.

Key words: digital transformation of education, digitalization of education, intellectual work, professional reliability, functional system of intellectual work, information load, information hygiene

Для цитирования: Sokhin I.G. Expected Problems of Human-Machine Interaction in Artificial Intelligence Systems // Ergodesign. 2024;3(25): 374-385. <http://dx.doi.org/10.30987/2658-4026-2024-3-374-385>.

Введение

Ускоренное развитие и внедрение наукоемких сквозных технологий машинного обучения и искусственного интеллекта во отраслях экономики и социальной сферы является одним из наиболее значимых больших вызовов для научно-технологического развития РФ. Успехи в развитии ИИ-технологий, главным образом, в области компьютерного зрения и обработки естественного языка, вселяют надежды на существенное повышение уровня автоматизации множества технологических процессов в производстве, транспорте, быту и других сферах. Вместе с тем, внедрение этого приоритетного направления научно-технологического развития в эргатические системы сопряжено с принципиально новыми рисками. Возникает ряд актуальных методологических проблем взаимодействия человека-оператора с системами искусственного интеллекта (СИИ): оценивание рисков принятия ошибочных решений ИИ, обеспечение доверия человека к решениям ИИ, необходимость трансформации системы подготовки и деятельности человека-оператора в новых условиях и т.п.

В силу новизны данных проблем, с которыми ранее не приходилось сталкиваться, для их решения целесообразно на первых этапах опираться на имеющийся научно-методический задел и экспертный анализ опыта решения подобных проблем в смежных отраслях. Это позволит более четко сформулировать проблемы и наметить направления их решения.

Предлагается рассмотреть эти проблемы на основе научно-методического задела, полученного при управлении космическими роботами, в которых ИИ используется в качестве обязательного элемента. Здесь под космическими роботами понимаются не только манипуляторы и антропоморфные космические роботы, но и все пилотируемые и автоматические космические аппараты [1]. Все существующие космические роботы проектируются как интеллектуальные

автоматы, способные адаптироваться к неопределенной среде функционирования. Традиционно их искусственный интеллект создавался посредством алгоритмического программирования, т.е. «прямого переноса» знаний естественного интеллекта разработчиков (кодирования) в информационную модель робота. Человеко-машинное взаимодействие в эргатической системе «космический робот-экипаж-Центр управления полетами» проявляется как на уровне непосредственного управления космическим роботом со стороны экипажа, так и удаленного из Центра управления полетами. Таким образом, управление космическим роботом может служить показательным примером человеко-машинного взаимодействия с интеллектуальным техническим агентом.

Ожидается, что анализ имеющегося опыта управления космическими роботами в совокупности с учетом возможностей и ограничений современных ИИ-технологий позволит выявить основные тенденции и методологические проблемы человеко-машинного взаимодействия в перспективных эргатических системах искусственного интеллекта.

Естественный и искусственный интеллект в технических системах

Национальная стратегия развития искусственного интеллекта [2] определяет искусственный интеллект как комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Данное определение рассматривает развитие до 2030 года только слабого (узкого) ИИ, т.е. ИИ первого уровня. Поэтому в данной статье мы также ограничимся рассмотрением слабого ИИ.

Чтобы разобраться в проблемах взаимодействия естественного интеллекта человека с искусственным интеллектом

робота, следует, прежде всего, определиться с понятием «интеллект». Что мы подразумеваем под интеллектуальной деятельностью, когда говорим об управлении техническим объектом? Какими отличительными свойствами характеризуется интеллектуальное управление? Как формируется искусственный интеллект? Какие существуют ограничения систем искусственного интеллекта, определяемые современным уровнем развития ИИ-технологий?

Понятие интеллект является зонтичным термином, охватывающим широкий круг явлений. Определить его в широком смысле достаточно сложно. По этой причине в научно-технической литературе по ИИ некоторые термины и понятия, связанные с ИИ, не согласованы между собой. Это отрицательно сказывается на понимании сути ИИ.

В данной статье интеллект рассматривается в узком смысле, как свойство управления техническими объектами, в частности, космическими роботами. Автор надеется, что такой подход позволит лучше понять внутренние механизмы ИИ. В теории автоматического управления интеллект ассоциируется со свойством адаптивности поведения как способности живого организма или технической системы приспосабливаться к изменяющимся условиям окружающей среды для достижения оптимального функционирования. Таким образом, в основе как естественного интеллекта человека-оператора, так и искусственного интеллекта технической системы лежит восприятие и обработка информации об окружающей среде для формирования адаптивных управленческих воздействий. Понятия «интеллектуальность» и «адаптивность» часто рассматриваются как синонимы. В частности, в робототехнике «адаптивный (интеллектуальный) робот определяется как способный выполнять задания, воспринимая внешнюю среду и/или взаимодействуя с внешними источниками и адаптируя свое поведение» [3]. Поэтому, анализируя проблемы человеко-машинного взаимодействия с искусственным интеллектом, необходимо обратиться к свойствам информации и информационных моделей (ИМ), используемых в СИИ для представления информации.

Информация и информационные модели систем искусственного интеллекта

В современном мире понятие «информация» употребляется повсеместно. Однако, несмотря на широкое использование понятия информации, ученые до конца не смогли договориться о едином понимании ее сущности. Существуют самые противоположные мнения: от понимания информации как всеобщего свойства материи до заверений, что это всего лишь омонимия – использование одного и того же термина в различных отраслях знаний со своим специфическим значением. Н. Винер, один из основоположников кибернетики, впервые выразил мнение, что информация – это не материя и не энергия, что она является самостоятельной сущностью, которая отражает порядок материального мира. Согласно Н. Винеру, действующий объект поглощает информацию из внешней среды и использует ее для выбора правильного поведения [4]. Попытки связать информацию с энтропией также оказались безуспешными, хотя они продолжаются до сих пор. Поэтому вопрос об определении понятия и сущности информации остается открытым [5].

Современные представления указывают на отсутствие материального, т.е. вещественного или полевого происхождения информации. Сегодня мы знаем, что информация, закодированная в молекуле ДНК, не только определяет форму всех живых организмов, но и передает наследуемые свойства родителей. То же наблюдается и в неживой природе. Информация, как некая нематериальная сила, заставляет материальные объекты принимать строго определенную форму. Это свойство мира было замечено ещё античными философами. Например, Аристотель рассматривает существование некой активной нематериальной силы, которую он называет формой и которая определяет свойства всех объектов материального мира. Материя по Аристотелю – это лишь возможность, без которой не может существовать никакая вещь, она сама по себе никакими качествами и лишена всякой определенности. Для того чтобы возникла конкретная вещь, необходимо соединение материи с неким другим активным началом – формой, которая придает материи определенность – качество этой вещи, превращает возможность в действительность. Далее, процесс превращения возможности в действительность, соединения материи с формой есть движение, которое носит «осмысленный», а не случайный характер. Эта парадигма дуализма материального и нематериального в природе была одной из

первых попыток понять сущность информации. Однако, в последствии эта идея не нашла своего развития.

В 1996 году А.И. Демин на основании сделанного им открытия сформулировал общую теорию информации [6], в которой попытался дать ключ к пониманию сущности информации и информационных взаимодействий. Приведем основные тезисы этой теории. Информация представляет собой всеобщее свойство материального мира, которое выступает как организованное по определенным правилам пространственное размещение материи – выделение порядка из хаоса. В природе объективно существует первичная информация в виде форм структуры вещественных тел и вторичная информация в виде отражения этой первичной информации в фундаментальных физических полях (электромагнитном, гравитационном), окружающих все вещественные тела. Вторичная информация в природе всегда является отражением первичной информации, характеризуя ее свойства, структуру и форму. Эта информация существует объективно, вне зависимости от воли и сознания людей. Однако, в процессе эволюции человек приобрел уникальную способность перекодировать объективную вторичную информацию. Воспринимая природную вторичную информацию с помощью органов чувств, человек формирует первичную (образную) субъективную информацию, которая непосредственно отражает существующее свойства объектов материального мира (образ мира). Далее эта первичная субъективная информация в процессе обучения и мышления перекодируется во вторичную субъективную информацию (абстрактную информацию о свойствах объектов материального мира, часто называемую знаниями). Вторичная субъективная (интеллектуальная) информация может многократно перекодироваться (представляя одни и те же данные в форме изображений, звуков, текстов, различных языков, теорий и других информационных моделей) и переноситься на другие носители информации без потери своего содержания. Так формируется естественный интеллект человека. Благодаря интеллекту человек обладает возможностью опосредованного информационного взаимодействия с объектами материального мира. Однако, кодирование вторичной информации, наряду с многими достоинствами, обладает и серьезным

недостатком – возможности генерации ложных выводов, не имеющих отношения к действительности. Это замечание относится как к естественному интеллекту, так и искусственному.

Возможность кодирования вторичной информации и возможность её передачи на искусственные носители привели к появлению технологий искусственного интеллекта. Действительно, в результате перекодирования порции вторичной субъективной информации можно создать некоторую информационную модель фрагмента реального мира. Эта информационная модель может быть размещена в памяти компьютера. Если еще подключить к компьютеру искусственные сенсоры, способные принимать информацию из внешнего мира, мы получим СИИ. Подобный робот, реагируя на внешние сигналы, будет действовать в соответствии с заложенной в него информационной моделью. Если эта модель окажется удачной, робот продемонстрирует поведение, которое покажется нам разумным, и мы называем его интеллектуальным. Хотя, в действительности, это просто результат имитации естественного человеческого интеллекта. В общих чертах, таким образом представляется механизм формирования искусственного интеллекта.

Качество системы искусственного интеллекта определяется её информационной моделью. И вот здесь возникает проблема, которую в ИИ называют «проблемой знаний». Не вдаваясь в подробности проблемы знаний, заметим, что знания являются упорядоченной информацией. В ИИ знания рассматриваются как данные, структурированные для систематического применения. Формула «данные – это новая нефть» образно характеризует связь между данными и знаниями. По аналогии с нефтью данные – это сырье, а знания – продукты переработки данных, то есть полезная информация. Проблема знаний заключается в том, чтобы: 1) получить знания из данных (обучение ИМ); 2) представить знания в структуре информационной модели; 3) использовать знания для решения интеллектуальных задач.

Обучение информационных моделей систем искусственного интеллекта

Цель обучения – приобретение знаний. Это в равной степени относится как к естественному интеллекту, так и искусственному. Ограничения статьи не позволяют подробно описать технологии машинного обучения ИМ СИИ. Поэтому для

наглядности автор воспользовался аналогиями с методами обучения естественного человеческого интеллекта.

Рассмотрим для примера обучение космонавтов управлению космическим аппаратом. Целью обучения является формирование управленческого интеллекта космонавтов, т.е. способности управлять кораблем в ситуациях неопределенности, которые могут возникнуть в полете. С этой целью обучение космонавтов осуществляется на двух уровнях: теоретическом (приобретение знаний) и практическом (тренировки на тренажерах). При теоретической подготовке космонавтам передается структурированная информация о системах космического аппарата в виде готовых знаний. При этом космонавты учатся применять полученные знания для рационального принятия решений в условиях неопределенности, используя логическое мышление. В ходе тренировок на тренажерах у космонавтов формируются автоматизмы реагирования на стереотипные ситуации. При этом на тренировках инструктор неожиданно для космонавтов вводит нештатные ситуации, заставляя их переключаться от автоматизмов к логическому принятию решений. На обоих уровнях обучения решающую роль играет обратная связь, получаемая космонавтами от инструкторов. Она позволяет разделить действия космонавтов на правильные и ошибочные. Следует заметить, что описанная выше система обучения космонавтов удивительным образом соответствует современной теории дуального процесса мышления человека, включающего ассоциативное мышление (система 1) и рассуждения (система 2) [7], хотя она сформировалась на много ранее появления теории.

Подобно двухуровневой системе обучения космонавтов существуют два альтернативных подхода к созданию и обучению ИМ СИИ. Предложенные Аланом Тьюрингом еще в прошлом веке эти подходы получили название «сверху-вниз» и «снизу-вверх».

Подход «сверху-вниз» (символьный) заключается в прямом переносе человеческих знаний в машину. Этот подход, называемый также *MDAI* (*Model-driven AI*, *MDAI*), ориентирован на информационную модель. Центральная идея символьного подхода – представить информационные сущности символами (перекодировать) и в дальнейшем выполнять с ними формализованные логические операции (индукция, дедукция,

абдукция). Это, по мнению разработчиков, позволяет создавать машины, способные делать логические выводы подобно человеку. Для организации процесса передачи знаний в информационную модель использовались самые разные приемы, в теории называемые традиционными методами машинного обучения. В традиционном подходе к машинному обучению основные усилия затрачиваются на проектирование признаков. Все они страдают общими непреодолимыми недостатками – они чрезвычайно трудоемки в процессе создания и поддержки, узко специализированы для конкретных задач, зависят от человеческих ошибок при выборе признаков. Представителями СИИ символьного подхода являются, в частности, экспертные системы, реализующие продукционные формы реляционной модели знаний. Для экспертных систем источники знаний – это эксперты. Для представления знаний используется продукционная форма в виде имплицативных отношений «Если-То» и связок И, ИЛИ между фактами. Вывод в такой системе знаний осуществляется на основе математической логики и процедур эвристического поиска. Это аналог системы 2 человеческого мышления, основанной на размышлениях. СИИ, созданные на основе символьного подхода, хорошо справляются с решением логических задач, но мало эффективны в задачах распознавания образов в перцептивных приложениях с зашумленными данными (изображения, текст, речь).

Подход «снизу-вверх», называемый коннекционизмом или нейросетевым подходом, ориентирован на данные (*Data-driven AI*, *DDAI*). Информационными моделями в данном подходе являются искусственные нейронные сети (ИНС), а извлечение скрытых знаний из предъявляемых данных машина (компьютер) выполняет самостоятельно, используя методы глубокого обучения нейросетей. Ключевым моментом, отличающим этот подход от символьного является опора на параллельные вычисления и на статистику, а не на последовательные системы и логические правила. По сути, глубокое обучение является разновидностью обучения по прецедентам, когда по конечной совокупности прецедентов — пар «объект, ответ», называемой обучающей выборкой, требуется построить алгоритм, способный для любого возможного входного объекта выдать достаточно точный классифицирующий ответ. Упрощенная схема

глубокого машинного обучения представлена на рис.1. Работа современной обучаемой искусственной нейронной сети (ИНС) разделена на два этапа: обучение (*training*) и логический вывод (*inference*). На этапе обучения подбирают «правильные» веса сети таким образом, чтобы на обучающей выборке получать правильные ответы сети. Математически — это оптимизационная задача аппроксимации некоторой неизвестной функции $Y=F(X)$. ИНС решает эту задачу путем подбора таких весовых коэффициентов связей между нейронами сети, чтобы минимизировать некоторую статистическую функцию от ошибок/потерь (например, среднеквадратическую ошибку всех итераций обучения). Обучение ИНС требует больших вычислительных ресурсов суперкомпьютеров и длительного времени: часы, дни, недели в

зависимости от сложности архитектуры, задачи, объема обучающих данных. На этапе логического вывода обученная нейронная сеть переносится на другую платформу (например, контроллер робота) для решения поставленной задачи. Для этого параметры обученной сети - архитектура, веса связей обученной сети, конфигурацию обучения (функция ошибки, тип оптимизатора), состояние оптимизатора конвертируются в файлы специального формата (например, *HDF5 - Hierarchical Data Format v5*) и сохраняют. После этого эти конвертированные файлы загружаются специальные фреймворки компьютера, и обученная сеть используется для логического вывода (например, распознавания изображений).

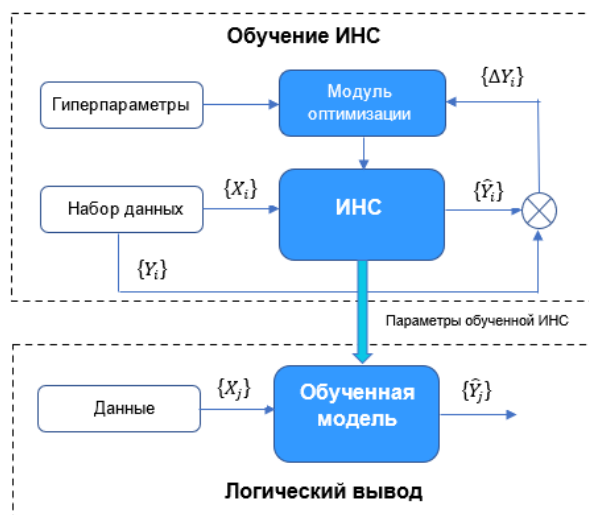


Рис.1. Схема глубокого машинного обучения
Fig.1. Deep machine learning scheme

СИИ, созданные на основе нейросетевого подхода, являются аналогом системы 1 ассоциативного мышления человека. Они хорошо справляются с задачами распознавания образов. На сегодняшний день нейросетевой подход является преобладающим в ИИ-технологиях компьютерного зрения и обработки естественного языка, в которых преимуществом обладает ассоциативное мышление. Слабыми сторонами подхода, основанного на глубоком машинном обучении, являются необходимость наличия репрезентативного, релевантного и корректно размеченного набора обучающих и тестовых данных, непрозрачность выводов, зависимость выводов от обучающей выборки и, как следствие, риски дезинформации.

Результаты сравнительного анализа свойств СИИ, созданных на основе символического и нейросетевого подходов, представлены в таблице 1 [1]. Как видно из таблицы, каждый из подходов к созданию ИИ имеет сильные и слабые стороны, которые должны учитываться при создании СИИ. Также следует учитывать особенности естественного человеческого и искусственного интеллектов. Преимущество человеческого интеллекта – в способности делать логические выводы и решать стоящие перед ним проблемы в условиях неопределенности, ограниченного и порой неполного объема данных. Ему помогает предшествующий опыт, интуиция, наблюдательность. Но человеческий интеллект слаб при работе с большими объемами данных, он попросту не

подготовлен к этому виду деятельности в процессе эволюции. И напротив, узкий ИИ, используя мощные вычислительные ресурсы, успешно справляется с извлечением полезной информации из колоссальных объемов данных, но к творчеству не приспособлен. Поэтому при проектировании человеко-

машинного взаимодействия с СИИ выбор приоритетов между человеческим интеллектом, символьным и нейросетевым ИИ должен определяться с учетом характера решаемой интеллектуальной задачи, сильными и слабыми сторонами каждого из интеллектов.

Таблица 1.

Сравнение свойств символьного и нейросетевого подходов

Table 1.

Comparison of properties of symbolic and neural network approaches

Свойство	Символьный подход	Нейросетевой подход
Представление информации	Символьное представление знаний	Наборы размеченных данных (изображения, текст, речь)
Способы обработки информации	Последовательные рассуждения/логические операции над символами	Параллельная распределенная обработка данных
Архитектура	База знаний + машина логического вывода (для экспертных систем)	Искусственная нейронная сеть (многообразие структур)
Сильные стороны	Логические рассуждения Объяснение логических выводов	Глубокое машинное обучение по прецедентам Распознавание образов (изображения, текст, речь)
Слабые стороны	Сложности приобретения знаний, поддержания больших баз знаний Узкая специализация баз знаний предметной области	Непрозрачность выводов, отсутствие объяснений Риски дезинформации Трудоемкость разметки обучающих данных

Использование систем искусственного интеллекта в космической робототехнике

Развитие космической робототехники началось вместе с зарождением космонавтики. Разработчики космической техники, опираясь на естественный интеллект, стремились в максимальной степени автоматизировать операции, выполняемые в космосе. При этом создаваемые СИИ ракетно-космической техники были основаны на символьном подходе. Это относится как к автоматическим космическим аппаратам, так и пилотируемым. Для автоматических космических аппаратов главной целью автоматизации было и, по-прежнему остается, обеспечение автономности их существования и адаптивности поведения в условиях чрезвычайно высокой неопределенности среды функционирования и удаленности от наземных центров управления. Для пилотируемых космических аппаратов автоматизация была направлена на

повышение надежности и безопасности космических полетов, снижение рисков для космонавтов, освобождение космонавтов от рутинных операций с тем, чтобы увеличить долю полетного времени на выполнение исследовательских задач.

Современные тенденции развития космонавтики, включающие переход к практическому использованию ближнего космоса, освоению космического пространства за пределами околоземных орбит, пилотируемым полетам к Луне и в дальний космос, развитию посещаемых космических станций, предполагают широкое использование средств космической робототехники [8]. Основное назначение космических роботов (КР) – выполнение роботизированных полетных операций космической деятельности автономно или во взаимодействии с космонавтами. КР должны будут автономно или под удаленным управлением человека-оператора выполнять

множество операций в космосе, которые ранее выполнялись космонавтами вручную, а также осуществлять ассистентскую поддержку деятельности космонавтов. Для реализации данной роли СИИ должны обеспечивать необходимый уровень адаптивности и автономности КР, сравнимую с интеллектом человека. Следует сделать важное для темы человеко-машинного взаимодействия замечание относительно автономности КР. Согласно ГОСТу [3] под автономностью робота понимается его способность функционировать без вмешательства человека. Применительно к КР автономность достаточно условна, поскольку даже в случае самостоятельного выполнения полетной операции КР будет находиться под контролем человека, будь то космонавт или оператор наземного Центра управления полетом. Это

обусловлено, с одной стороны, высокими рисками безвозвратных потерь в случае критических ошибок КР, с другой стороны, ограничениями адаптивности СИИ. В случае невозможности обеспечения необходимого уровня адаптивности системой ИИ робота, в контур управления должен включиться интеллект человека-оператора. Таким образом, в результате оптимального человеко-машинного взаимодействия должен быть обеспечен компромисс между автономностью КР и необходимой адаптивностью полетной операции.

Чтобы представить роль СИИ в КР рассмотрим общую структурную схему интеллектуального робота с использованием бионической терминологии (см. рис.2).



Рис. 2. Структурная схема интеллектуального робота
Fig. 2. Block diagram of an intelligent robot

Робот в общем случае состоит из 4 основных систем: информационно-измерительной («Сенсорика»), управляющей («Мозг»), исполнительской («Моторика») и системы связи с другими роботами, человеком или внутренними системами робота («Язык»).

«Сенсорика» – это искусственные органы чувств робота, которые предназначены для восприятия и преобразования информации о состоянии внешней среды и состоянии самого робота в соответствии с потребностями управляющей системы робота.

«Мозг» предназначен, во-первых, для выработки закона управления приводами (двигателями) механизмов исполнительской системы, используя сигналы обратной связи от сенсорики; во-вторых, для общения робота с человеком на каком-либо языке.

«Моторика» служит для выполнения управляющих сигналов (программы), формируемых мозгом, а также воздействия на окружающую среду.

«Язык» робота предназначен для организации обмена информацией между системами робота, между роботом и человеком или другими роботами на понятном им языке.

Каждая из рассмотренных систем робота в перспективе должна включать СИИ.

Ожидается, что в рамках узкого ИИ в космической робототехнике найдут применение две сквозные нейросетевые ИИ-технологии: компьютерное зрение и обработка текста на естественном языке. Технологии компьютерного зрения будут использоваться для создания сенсоров, в частности системы технического зрения робота, самой универсальной интеллектуальной сенсорной системы. Системы технического зрения КР найдут применение в таких задачах как манипулирование объектами, визуальная навигация мобильных роботов, распознавание звуков и голосов. Технологии обработки

текста на естественном языке будут использоваться для создания интеллектуальных многомодальных интерфейсов взаимодействия человека-оператора с роботом на естественном языке (текст, речь, жесты, мимика), справочно-информационной, ассистентской и психологической поддержки космонавтов в

полете, анализа данных, поддержки принятия управленческих решений.

Функциональная схема управления движениями исполнительного механизма робота, представленная на рис.3, демонстрирует уровни интеллектуального управления. Эта схема также заимствована из бионических представлений о движениях человека [9], [10].

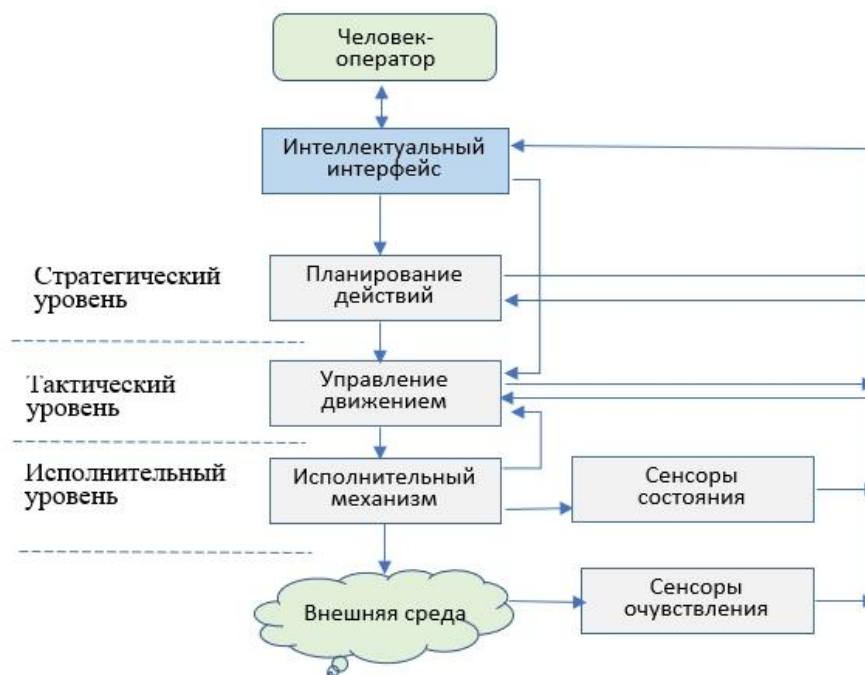


Рис. 3. Функциональная схема управления интеллектуального робота
Fig. 3. Functional control diagram of an intelligent robot

На нижнем, *исполнительном уровне* управления осуществляется программное управление приводами манипулятора для выполнения типовых движений. Информация от датчиков положения, скоростей, ускорений, моментов используется для управления динамическими характеристиками приводов. Это параметрическая адаптация – самонастройка варьируемых параметров программного управления.

На *тактическом уровне* осуществляется уже адаптивное управление приводами. На этом уровне используются система технического зрения, силомоментные, тактильные, локационные сенсорные системы для обеспечения необходимой силы захвата, самонаведения приводов в заранее неизвестном пространстве, обхода непредвиденных препятствий и т.п. Это программная адаптация - выбор адаптивной программы из готового набора типовых программ.

На *стратегическом уровне* осуществляется планирование действий робота с учетом сенсорной информации, получаемой в реальном масштабе времени. На этом уровне осуществляется анализ сенсорной информации с помощью информационной модели среды, используя все сенсоры для реализации программ выполнения типовых операций. Это структурная адаптация (самоорганизация) – изменение алгоритма управления, на основе оценки текущей обстановки. На этом уровне через интеллектуальный интерфейс может осуществляться выдача заданий человеком-оператором:

- целеуказаний на выполнение отдельных операций;
- оказание помощи в выполнении отдельных функций КР в диалоговом режиме;
- контроль за функционированием КР и оперативное взятие управления на себя, когда робот автономно не может выполнить какую-то сложную операцию либо в нештатных ситуациях.

Оптимальное взаимодействие человека-оператора с КР заключается в передаче роботу уровня управления, наилучшим образом соответствующего сложившейся ситуации. В простых ситуациях роботу доверяется стратегический и все нижестоящие уровни управления. В наиболее сложных и ответственных ситуациях робот реализует только исполнительный уровень, все остальные функции вышестоящих уровней управления берет на себя человек-оператор.

Ожидаемая трансформация деятельности человека-оператора в системах искусственного интеллекта

Рассматривая человека-оператора в контуре управления СИИ (в частности, интеллектуального робота, см. рис.3), легко заметить, что человек взаимодействует с физической средой опосредованно. Между ним и средой, на которую он, собственно, и воздействует, располагаются человеко-машинный интерфейс и три уровня управления. Первичная информация от среды многократно кодируется и к человеку приходит более упорядоченная, в целом облегчая принятие управленческих решений. Вместе с тем, ограничения, связанные с глубоким машинным обучением СИИ, а также многократное кодирование информации увеличивают риски искажения первичной информации и, как следствие, может приводить к ошибкам управления. Опосредованное управление присуще практически любому человеко-машинному взаимодействию. Однако, в случае взаимодействия с СИИ уровень опосредованности возрастает многократно – теперь человек вынужден взаимодействовать с виртуальным образом объекта управления и рискует «оторваться» от реальности. Попробуем оценить, как эта реальная угроза отразится на деятельности человека-оператора в СИИ.

Человек будет, в основном, непосредственно взаимодействовать не с физическими объектами управления, а с их виртуальными двойниками. Посредником между ними будет интеллектуальный многомодальный интерфейс. Это обстоятельство выдвигает высокие требования к функционалу и качеству человеко-машинного интерфейса [11]. Одним из основных требований является обеспечение удобства и естественности работы человека-оператора в процессе управления. В связи с этим интерфейс «человек-робот» должен позволять человеку-оператору воспринимать

аудиовизуальную, речевую, силомоментную, тактильную, справочную и другие виды информации, которые в совокупности должны обеспечивать человеку-оператору эффект присутствия в операционной среде. Причем, эффект присутствия заключается не столько в создании психологически комфортных условий для человека, сколько в обеспечении ему доступа к альтернативным источникам информации. Помимо кодированной информации от СИИ человек должен иметь доступ к первичной информации об объектах управления, например, в виде визуального ряда реальной обстановки. В плане создания интеллектуального многомодального интерфейса перспективными являются ИИ-технологии обработки естественного языка и технологии виртуальной/смешанной реальности.

Человек, находясь в контуре управления робототехнического комплекса, должен сохранить за собой роль лидера. Это сделать не просто, т.к. человеку трудно проверить решения, которые ему предлагают СИИ. ИИ значительно превосходит человека в скорости обработки больших данных. При этом современные модели глубокого обучения не используют логику для определения причинно-следственных отношений между знаниями об окружающем мире, а используют корреляцию между наборами данных. Поэтому они не способны объяснить, как и почему были приняты те или иные решения. По-видимому, чтобы сформировать доверие человека-оператора к СИИ, потребуется существенная трансформация систем подготовки человека-оператора. Такая трансформация предполагает разработку методов оперативной верификации ИИ-решений в масштабе реального времени, обучение человека-оператора практическому использованию этих методов при взаимодействии с СИИ. Подобное происходило с подготовкой космонавтов при переходе от аналоговых систем управления космическими аппаратами к цифровым. Этот переход сопровождался существенным усложнением алгоритмов управления и, как следствие, возрастанием разнообразия откликов системы на различные ситуации. И хотя цифровые системы управления были существенно интеллектуальнее аналоговых и обеспечивали высокую степень автоматизации процессов, система подготовки космонавтов была переориентирована на обнаружении и парировании множества нештатных ситуаций,

связанных с неправильным функционированием автоматики.

Следует быть особенно осторожным при внедрении СИИ в критически опасные отрасли, где ущерб от неправильных решений и действий может быть очень велик. Для принятия решений об использовании СИИ рекомендуется [12] использовать риск-ориентированный подход, основанный на оценке размера потенциального ущерба с учетом вероятности его наступления по сравнению с потенциальным положительным эффектом от его внедрения.

Заключение

Современные достижения ИИ-технологий поражают воображение людей своей имитируемой «интеллектуальностью». В первую очередь это относится к последним чат-ботам типа ChatGPT, информационные модели которых (так называемые «большие лингвистические модели») построены на основе трансформерной архитектуры многослойных нейронных сетей и механизмов обучения без учителя. Благодаря своей архитектуре и обучению на больших объемах данных без явного учителя, ChatGPT способен генерировать качественные и связные тексты, отвечая на вопросы и поддерживая разговор. Вместе с тем, стоит осознать, что современные ИИ-технологии — это только инструмент, способный расширить когнитивные способности человека. Искусственные нейронные сети, ставшие основой современного ИИ, очень успешно реализовали процесс параллельных вычислений корреляционных связей между данными. Если раньше пользователи сети Интернет «вручную» с помощью поисковиков выполняли поиск необходимой информации,

то теперь обученные нейронные сети способны автоматизировать этот процесс. Учитывая большие вычислительные возможности современных компьютеров, это привело к ошеломляющим результатам в обработке больших данных.

Применение СИИ влечет за собой революционные перемены, связанные с новыми возможностями автоматизации процессов человеческой жизнедеятельности. Эти перемены затрагивают не только техносферу, но и наше мировоззрение, понимание роли человека в этой новой техносфере.

Ожидается, что при взаимодействии с ИИ человек столкнется с рядом проблем, которые ему предстоит решить:

- сохранить лидерские позиции человека в отношении ИИ, реализовать принципы антропоцентрического подхода при создании и применении СИИ;
- научиться жить и работать в виртуальном пространстве, не теряя при этом связи с физической реальностью;
- разработать методы критического оценивания адекватности поведения СИИ, опирающиеся на человеческие ценностные ориентиры и здравый смысл;
- трансформировать систему подготовки человека-оператора к взаимодействию с СИИ.

Перед человечеством стоит глобальная проблема «приручить» искусственный интеллект, сделав его слугой своего естественного интеллекта. При этом не забывать развивать свой природный интеллект, чтобы избежать опасности следования на поводу у своего слуги.

СПИСОК ИСТОЧНИКОВ

1. **Сохин И.Г., Богданов А.А., Дудоров Е.А. и др.** Искусственный интеллект в робототехнических системах космического назначения. Искусственный интеллект в космической технике. Состояние. Перспективы применения. Монография. Под ред. А.Н. Балухто. М.: Радиотехника, 2021. С. 237-316. DOI 10.18127/B9785931082042.
2. **Указ Президента Российской Федерации** от 10 октября 2019 г. № 490 «Национальная стратегия развития искусственного интеллекта на период до 2030 года».
3. **ГОСТ Р 60.0.0.4-2019/ИСО 8373:2012.** Роботы и робототехнические устройства. Термины и определения.
4. **Винер Н.** Кибернетика или управление и связь в животном и машине: [Пер. с англ.] 2-е изд. Москва: Сов. радио, 1968. 326 с.
5. **Чернавский Д.С.** Синергетика и информация (динамическая теория информации). Едиториал УРСС, 2004. 287 с. ISBN 5-354-00241-9.

REFERENCES

1. **Sokhin I.G., Bogdanov A.A., Dudorov E.A., et al.** Artificial Intelligence in Robotic Systems for Space Applications. Artificial Intelligence in Space Technology. Status. Application Prospects. Balukhto AN, editor. Moscow: Radiotekhnika; 2021. p. 237-316. DOI 10.18127/B9785931082042.
2. **Decree of the President of the Russian Federation** of October 10, 2019, no. 490 “National Strategy for the Development of Artificial Intelligence through 2030”.
3. **State Standard R 60.0.0.4-2019/ISO 8373:2012.** Robots and Robotic Devices. Terms and Definitions; 2019 Sep 01.
4. **Wiener N.** Cybernetics or Control and Communication in the Animal and the Machine. Translated from English. 2nd ed. Moscow: Sovetskoe Radio; 1968. 326 p.
5. **Chernavsky D.S.** Synergetics and Information (Dynamic Information Theory). Editorial URSS; 2004. 287 p.

6. **Демин А.И.** Парадигма дуализма: Пространство – время, информация – энергия. М.: Изд. ЛКИ, 2011. 320 с. ISBN 978-5-382-01294-0.

7. **Канеман Д.** «Думай медленно... Решай быстро». Изд. АСТ, 2011. 706 с. ISBN 978-5-17-080053-7.

8. **Сохин И.Г., Дудоров Е.А.** Предназначение и задачи робототехнических систем в российской лунной программе // Известия высших учебных заведений. Машиностроение. №12(729). 2020. С.3-15. DOI 10.18698/0536-1044-2020-12. EDN CGYYPB.

9. **Сырямкин В.И.** Информационные устройства и системы в робототехнике и мехатронике: учеб. пособие. (Серия: Интеллектуальные технические системы). Томск: Изд-во Том. ун-та, 2016. 524 с. ISBN 978-5-7511-2443-4.

10. **Юревич Е.Ю.** Управление роботами и робототехническими системами. Уч. пособие СПбГТУ, 2000 г., 171 с.

11. **Сохин И.Г., Дудоров Е.А., Богданов А.А. и др.** Эргономическое сопровождение разработки антропоморфных робототехнических систем космического назначения//Известия высших учебных заведений. Машиностроение. №1(730). 2021. С. 16-26. DOI 10.18698/0536-1044-2021-1-16-26. EDN VGPZVV.

12. **Распоряжение Правительства РФ** от 19 августа 2020 г. № 2129-р «Об утверждении Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники».

6. **Demin A.I.** Dualism Paradigm: Space – Time, Information – Energy. Moscow: LKI; 2011. 320 p.

7. **Kahneman D.** Thinking, Fast and Slow. AST; 2011. 706 p.

8. **Sokhin I.G., Dudorov E.A.** The Purpose and Tasks of Robotic Systems in the Russian Lunar Program. Proceedings of Higher Educational Institutions. Machine Building. 2020;12(729):3-15. DOI 10.18698 / 0536-1044-2020-12.

9. **Syryamkin V.I.** Information Devices and Systems in Robotics and Mechatronics. Series: Intelligent Technical Systems. Tomsk: Publishing House of Tomsk. University; 2016. 524 p.

10. **Yurevich E.Yu.** Control of Robots and Robotic Systems. Saint Petersburg: Peter the Great St. Petersburg Polytechnic University; 2000. 171 p.

11. **Sokhin I.G., Dudorov E.A., Bogdanov A.A., et al.** Ergonomic Support for the Development of Anthropomorphic Robotic Systems for Space Purposes. Proceedings of Higher Educational Institutions. Machine Building. 2021;1(730):16-26. DOI 10.18698/0536-1044-2021-1-16-26.

12. **Order of the Government of the Russian Federation** of August 19, 2020, no. 2129-r “On Approval of the Concept for the Development of Regulation of Relations in the Field of Artificial Intelligence and Robotics Technologies”.

Информация об авторах:

Сохин Игорь Георгиевич – доктор технических наук, доцент, действительный член академии космонавтики им. КЭ. Циолковского, главный научный сотрудник АО «Научно-исследовательский институт авиационного оборудования», международный идентификационный номер автора: Author-ID-РИНЦ 1663-7699.

Information about the authors:

Sokhin Igor Georgievich – Doctor of Technical Sciences, Associate Professor, Full Member of the K.E. Tsiolkovsky Academy of Cosmonautics, Chief Researcher of JSC Research Institute of Aviation Equipment, the author’s international identification number: Author-ID-RSCI: 1663-7699.

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации.

Contribution of the authors: the authors contributed equally to this article.

Авторы заявляют об отсутствии конфликта интересов.

The authors declare no conflicts of interests.

Статья поступила в редакцию 16.07.2024; одобрена после рецензирования 23.07.2024; принята к публикации 24.07.2024. Рецензент – Спасенников В.В., доктор психологических наук., профессор Брянского государственного технического университета, главный редактор журнала «Эргодизайн»

The paper was submitted for publication on the 16th of July, 2024; approved after the peer review on the 23rd of July, 2024; accepted for publication on the 24th of July, 2024. Reviewer – Spasennikov V.V. Doctor of Psychological Sciences, Professor of Bryansk State Technical University, Editor-in-Chief of the journal “Ergodesign”.