

Материалы семинара

20-21 сентября 2022 г., г. Переславль-Залесский, Институт программных систем им. А. К. Айламазяна Российской академии наук

Научная программа семинара

20 сентября

Доклады

1. Разин Александр Владимирович «Естественная эволюция и компьютерная революция» (д.филос.н., профессор, заведующий Кафедрой этики философского факультета МГУ им. М.В. Ломоносова)

Аннотация. В докладе показывается принципиальная разница процессов творчества на естественной основе, которые осуществляются без заранее определенной цели и имеют характер небольших приращений, случайных приобретений, и творчества, которое осуществляется как целевое, происходит на основе концепций и искусственно изобретенных средств. Деннет отражает эти процессы в терминах творчество по восходящей и по нисходящей. С его точки зрения творчество по нисходящей содержит такие инструменты, которые упрощают условия творческого процесса, и, в то же время, делают его гораздо более эффективным по сравнению с эволюцией в природе. Деннет называет несколько видов инверсий, в которых были осуществлены принципиальные изменения представлений об условиях творчества и отношении человека к творческому процессу: инверсия Дарвина (можно создавать новое не имея цели), инверсия Тьюринга (реальность может быть редуцирована до таких простых элементов, которые могут быть заменены машиной, в результате возникает возможность умения без разумения), инверсия Юма (причинность как иллюзия нашего разума). Компьютеру не нужно доходить до понимания конечных причин, к чему стремится человек. Он работает по принципу сравнения информации в смысле теории смены состояний (неклассические виды детерминации). Это определяет эффективность и быстроту его работы, но это же означает отсутствие понимания тех операций, которые он совершает. Далее, важно подчеркнуть, что детали компьютера (транзисторы) не имеют собственных целей, в отличие от нейронов головного мозга, которые представляют собой клетки, деятельность которых имеет целевую направленность, и которые способны объединяться друг с другом по принципу сетевых связей. Возможно именно с этим связаны те способности мозга, которые мы пока не можем эмитировать в искусственном интеллекте, скажем - свобода воли. Деннет показывает, что создание чего-то подобного нейрону на кремниевом носителе в принципе возможно, но крайне сложно и дорого. Здесь возникает интересный вопрос: а не можем ли мы создать нечто лучшее, чем нейрон, ведь культурно опосредованное творчество уже создало то, что природа в принципе создать никогда бы не смогла, например - колесо, реактивный двигатель, в котором используется топливо и окислитель, и т.д. Для подобных созданий нужны такие качественные скачки, на которые природа неспособна. А каковы в этом плане будущие возможности искусственного интеллекта?

2. Готовцев Павел Михайлович «Этическая валидация автономных и интеллектуальных систем – нормативное регулирование и современное состояние исследований» (к.т.н., доцент, заместитель начальника отдела биотехнологий и биоэнергетики, НИЦ «Курчатовский институт»)

Аннотация. В докладе рассмотрены основные подходы к валидации автономных и интеллектуальных систем, которые отражены сегодня в нормативно-технической и научной литературе. Выделены основные сложности подобной валидации и показаны перспективные направления исследований в данной области.

3. Карпов Валерий Эдуардович 1, Королева Мария Николаевна 2 «Этика поведения

коллоративного робота» (1 к.т.н., доцент, заведующий Лабораторией робототехники НИЦ “Курчатовский институт”, вице-президент РАИИ; 2 к.т.н., доцент кафедры “Компьютерные системы автоматизации производства” МГТУ им. Н.Э. Баумана, ученый секретарь РАИИ)

Аннотация. Поведение робота должно отвечать этическим представлениям человека-оператора. Это связано с необходимостью решения роботом проблемы выбора, а также субъективизмом оценок поведения робота-партнера человеком. Рассматриваются некоторые проблемы при формализации этических аспектов поведения. Делается вывод, что в условиях принципиально неустранимых противоречий и неполноты при формализации моральных учений, основная задача этических схем – быть основой для объяснения постфактум, а вовсе не являться мотивирующим фактором.

4. Леушина Влада Вячеславовна «Онтология гедонизма» (лаборант-исследователь лаборатории робототехники НИЦ “Курчатовский институт”)

5. Асеева Ирина Александровна «Социальные роли искусственного интеллекта: этика человеко-машинного взаимодействия» (д-р.филос.наук, проф., в.н.с. ИНИОН РАН)

6. Буданов Владимир Григорьевич «Интегральная этика и перспективы формирование антропоморфного ИИ» (д-р.филос.наук, к.ф.-м.н., г.н.с., Заведующий сектором междисциплинарных проблем научно-технического развития. Институт философии РАН)

7. Шелекета Владислав Олегович «Проблема определения природы нарративов сознания и принципы объективной нравственности в методологическом ракурсе теории сознания Д. Деннета» (д-р.филос.наук, профессор кафедры теории и методологии науки БГТУ им. В. Г. Шухова, заместитель председателя Белгородского регионального отделения Научный совет РАН по методологии искусственного интеллекта и когнитивных исследований, член-корреспондент РАЕН)

8. Гаврилина Елена Александровна «Делегирование действия интеллектуальным автономным системам и редукция человеческой агентности» (к.ф.н., доцент, ИНИОН РАН)