

Аннотации докладов:

К ВОПРОСУ О ФОРМАЛИЗАЦИИ ЭТИЧЕСКИХ УЧЕНИЙ ДЛЯ КОЛЛАБОРАТИВНОГО «ПОНИМАЮЩЕГО» РОБОТА

Леушина Влада Вячеславовна

лаборант-исследователь лаборатории робототехники, Национальный исследовательский центр «Курчатовский институт»

Королева Мария Николаевна

к.т.н., доцент кафедры «Компьютерные системы автоматизации производства» МГТУ им. Н.Э. Баумана

Создание понимающих агентов является одним из перспективных направлений в искусственном интеллекте. В частности, такие агенты востребованы в коллаборативной робототехнике, ведь человек-оператор ожидает, что в процессе взаимодействия с роботом-партнером он сможет понимать причины поступков своего «напарника». В определенном смысле человек ожидает от робота проявление свойств морального агента (например, принимать решения самостоятельно, основываясь на принципах морали). Понимание в данном случае заключается в частичном совпадении картин мира человека-оператора и коллаборативного «понимающего» робота, в том числе в части затрагивания понятий «моральность поведения», «целеполагание» и др. Архитектура такого робота имеет много общих черт с «архитектурой» живых существ. Она состоит из системы нижнего и верхнего уровней управления. Нижний (животный) уровень реализует рефлекторную деятельность и определяется эмоционально-потребностной схемой поведения. Высокий, или как его еще по-другому называют, когнитивный уровень управления является некоторой надстройкой над нижним уровнем, а его функционирование заключается в искажении и коррекции сенсорного восприятия, изменении значимости нужд. Именно за счет этого уровня в картине мира робота появляется один из важных компонентов – знак «Я» (познающий субъект), который позволяет реализовывать такие феномены, как: подражание, социальное обучение, эмпатия. Для определения цели и характера поведения здесь же вводится понятие «моральное поведение», основанное на концепции философского учения. Именно формирование картины мира робота на основе формального представления этического учения в определенной степени должна повлиять на целенаправленность поведения робота и сформировать схему объяснения его действий. В докладе основное внимание уделено организации морального поведения на когнитивном уровне управления, представлению схемы выбора поведенческих процедур, обоснованию выбора гедонизма как основного философского учения и его формализации для коллаборативного робота, описанию получившейся картины мира.

Этическая валидация автономных и интеллектуальных систем – нормативное регулирование и современное состояние исследований

Готовцев Павел Михайлович, *к.т.н., доцент, заместитель начальника отдела биотехнологий и биоэнергетики, НИЦ «Курчатовский институт»*

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.

