

Натуральные вычисления и искусственный интеллект

И. Ф. Михайлов

Предмет исследований

Междисциплинарные проблемы когнитивной психологии, когнитивной и социальной нейронауки и искусственного интеллекта

Горизонт исследований

Согласно изначальному замыслу, когнитивные науки включают философию сознания, теорию познания, когнитивную психологию, нейронауку, когнитивную лингвистику и теорию искусственного интеллекта. Часто этот термин понимается предметно: к когнитивным наукам относят все дисциплины, изучающие сознание и познание, принятие решений, внимание и т. п., вне зависимости от их методологических ориентаций. Однако «когнитивная революция», состоявшаяся во второй половине XX в., имела вполне определённые методологические очертания: она опиралась на так называемую «компьютерную метафору» — понимание материальных структур, обеспечивающих когнитивные функции, как информационных процессоров, совершающих вычислительные операции над символическими репрезентациями — формами представления данных. Этот подход позволил, уйдя от различных вариантов бихевиоризма, вернуться к исследованию «внутренней» ментальной жизни, причём на вполне научных, принципиально фальсифицируемых основах.

Вместе с тем, с течением времени внутри классической когнитивной парадигмы стали накапливаться проблемы, появилась внешняя конкуренция со стороны альтернативных когнитивных подходов — коннекционизма, энактивизма (и других концепций воплощённого познания), «теории динамических систем» и др. Некоторые исследователи заговорили о наступлении «пост-когнитивной» эры, широкое распространение получили антикомпьютационалистские (не опирающиеся на вычислительный дискурс) и антирепрезентационалистские взгляды. Одновременно мощный импульс к развитию получили нейронауки, которые, по-своему адаптируя вычислительные подходы, сегодня составляет сильную конкуренцию привычной «высокоуровневой» когнитивистике.

Более того, определённый ренессанс вычислительных подходов к когнитивной сфере произошёл благодаря широкому распространению концепции предиктивного процессинга (известной также как «*предсказывающий разум*»), основанной на принципе минимизации свободной энергии и связывающей воедино термодинамический, теоретико-информационный и математико-статистический взгляды на широкий класс сложных автономных систем, включающий биологические, нейрокогнитивные и социальные. Сложившаяся к настоящему моменту ситуация ставит перед исследователями теоретические и методологические проблемы, которые вряд ли могут быть решены в рамках отдельных дисциплин, без междисциплинарного взаимодействия. С этой целью Институт философии РАН инициирует создание Центра философии сознания и когнитивных исследований для исследования и обсуждения концептуальных и методологических проблем, стоящих перед исследователями в очерченных предметных областях.

Основные направления исследований и темы дискуссий

Философия и когнитивные науки. Такие традиционные разделы философии, как философия сознания и эпистемология, безусловно, могут и должны быть частью общего когнитивного предприятия. Однако в реальности философы и учёные часто говорят на разных языках, не достигая необходимого уровня взаимопонимания. Для преодоления этого барьера необходимы усилия со стороны ещё одного традиционного раздела философского знания — философии науки, которая, имея опыт анализа конкретных научных теорий, должна помочь решить концептуальные и методологические проблемы когнитивного комплекса или, по крайней мере, наладить коммуникацию между различными исследовательскими областями. В последнее десятилетие *философия когнитивных наук* порождает вал англоязычных публикаций. Что такое когнитивная теория, в чём её специфика по сравнению с привычным естествознанием и психологией? Какова роль вычислительных и динамических моделей, меняют ли они традиционное соотношение теории и эмпирии? Помогут ли когнитивные методологические подходы преодолеть традиционные философские затруднения, вроде «парадокса гомункула»? Эти и подобные вопросы составляют философский уровень обсуждений.

Проблема вычислений в когнитивных науках. Когнитивная революция случилась как закономерное следствие компьютерной (вычислительной) революции середины XX в. Прорывные работы Гилберта, Тьюринга и Чёрча совпали с ускорившимся технологическим прогрессом, и вскоре вычислительные машины стали непременным атрибутом промышленных предприятий, офисов и научных лабораторий. Компьютерное моделирование во многих случаях заменило традиционное эмпирическое исследование. Следующий шаг — обнаружение вычислительных процессов в самой природе: репликация РНК и синтез протеинов, обработка визуальной информации, квазары и другие природные явления дали начало переводу ряда естественных наук на вычислительные рельсы. Конечно, когнитивные процессы — первый кандидат на рассмотрение с точки зрения вычислительного подхода. Но насколько сам этот подход релевантен? Что такое вычисления с точки зрения сегодняшних взглядов на природу сложных структурных процессов? Есть ли у них особый онтологический статус? Продвигает ли нас компьютеризация в каком-либо варианте к искомой цели — окончательному разрешению философских загадок?

Репрезентации и сенсомоторные паттерны в когнитивных исследованиях. В первые годы когнитивного перелома казалось, что репрезентации — физические структуры, символически представляющие некие данные для обработки — являются необходимым элементом когнитивных процессов. Однако появление альтернативных парадигм — коннекционизма, энактивизма и др. — позволило говорить о «субсимвольных», «распределённых» репрезентациях или о возможности вовсе обойтись без этого понятия. Представители различных когнитивных дисциплин, по нашему замыслу, могут встретиться в Центре для того, чтобы обменяться мнениями о релевантности репрезентаций как концепции, объясняющей исследуемые ими явления. Оперирует ли мозг репрезентациями? Необходимы ли особые подсистемы представления данных для конструирования искусственных интеллектуальных систем? Составляет ли «динамический» подход реальную альтернативу репрезентационализму? Связан ли когнитивный репрезентационализм с одноимённым направлением в философии сознания? И, с другой стороны, можно ли усмотреть концептуальные связи между философским инференциализмом (Роберт Брэндом) и представлением об *'active inference'* (Карл Фристон), составляющим важную часть теории предиктивного процессинга?

Когнитивные исследования и искусственный интеллект. Специалисты в области технических наук часто определяют когнитивные устройства как устройства, способные рационально адаптировать своё поведение к меняющимся условиям среды на основе обратной связи. В таких устройствах, как правило, используются те или иные технологии ИИ, будь то различные механизмы машинного обучения или мультиагентные алгоритмы.

Естественным образом возникает вопрос: в какой мере эти и подобные решения могут быть моделями «естественной» психической жизни? Объективно возникает необходимость междисциплинарного взаимодействия специалистов в области ИИ с психологами и нейрофизиологами. Тем более, что результаты последних оказали в последние десятилетия существенное влияние на развитие технологий: достаточно назвать нейросети, мультиагентные системы, генетические и эволюционные алгоритмы и т. п. Таким образом, одна из целей семинара — служить площадкой для продуктивного взаимодействия специалистов в области искусственного интеллекта и исследователей интеллекта естественного.

Характер и принципы объяснения в когнитивных науках. С учётом очень быстрых и разноплановых изменений, происходящих в области когнитивных исследований, все более острым оказывается вопрос о приемлемых (или желательных) способах объяснения получаемых эмпирических результатов. Отчасти это связано с чрезвычайным предметным многообразием получаемых данных, затрагивающих часто самые неожиданные стороны когнитивного функционирования. Однако не меньшую сложность составляет и разноплановый характер исследований (скажем, серьезную проблему представляет сопоставление результатов классических поведенческих экспериментов, компьютерного моделирования, визуализации работы мозга в ходе решения каких-то задач и т.д.). Обсуждение возможных стратегий объяснения в сложившихся условиях — также одна из целей планируемых междисциплинарных исследований.

Предполагаемая тематика мероприятий на ближайшие три года

Философские проблемы когнитивных наук:

1. Психофизическая проблема в философии и когнитивных науках
2. Ментальная каузальность в физической картине мира и её преломление в когнитивных науках
3. Когнитивная наука как натуралистический проект
4. Нейроалармизм *vs.* нейроредукционизм: обсуждение перспектив и необходимых условий междисциплинарного сотрудничества когнитивной психологии и нейронаук

Проблема вычислений в когнитивных науках:

5. Байесовские модели как методология описания психики и возможная основа построения когнитивных теорий (Bayesian Mind)
6. Предсказательное кодирование (Predictive Mind) как новая парадигма в когнитивных исследованиях

Репрезентации и сенсомоторные паттерны в когнитивных исследованиях:

7. Включение феноменального опыта в модели функционирования когнитивных процессов (проблемы современных исследований сознания и эмоций)
8. Телесная укоренённость и моторная активность в человеческом познании
9. Новая материальность и роль артефактов в человеческом познании
10. Когнитивные исследования повседневности: в направлении экологической валидности (от фундаментальных исследований запоминания дверей до прикладных исследований в юзабилити и гейм-дизайне)

Когнитивные исследования и искусственный интеллект:

11. Машинное обучение и нейросети как инструменты решения технических проблем современных когнитивных исследований
12. Мультиагентные системы и когнитивные теории социальности

Характер и принципы объяснения в когнитивных науках:

13. Идеи (пост)неклассической науки в психологии в сравнении с когнитивными подходами
14. Проблема реплицируемости (локализация и последующее нивелирование источников изменчивости) в современных экспериментальных исследованиях
15. Иерархические vs. "плоские" модели когнитивных процессов: проблемы и возможности
16. Параметрические модели в когнитивной психологии (бедность и нищета вербального теоретизирования)
17. Специфика совместного внимания и совместного действия с точки зрения когнитивной архитектуры