

Моделирование когнитивной эволюции: заделы, первые результаты, перспективы¹

Редько В.Г.

Научно-исследовательский институт системных исследований РАН, Москва

В настоящей работе обсуждается новое интересное направление исследований – моделирование когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, приведшей к происхождению мышления человека. Актуальность исследований когнитивной эволюции связана с глубокой гносеологической проблемой: почему формальное логическое человеческое мышление, казалось бы, совсем не связанное с реальным физическим миром, применимо к познанию природы? Подробнее об этой проблеме см. [1]. В данной работе характеризуются заделы, первые результаты и перспективы исследований когнитивной эволюции. Очертим направления, в которых развиваются заделы моделирования когнитивной эволюции.

1. Заделы моделирования когнитивной эволюции

Адаптивное поведение. Направление «Адаптивное поведение» развивается с начала 1990-х годов [2]. Основной подход направления – конструирование и исследование искусственных «организмов» (в виде компьютерной программы или робота), способных приспосабливаться к внешней среде. Эти организмы часто называются «аниматами» (от англ. animal и robot: animal + robot = animat), агентами, автономными агентами. Исследователи адаптивного поведения разрабатывают такие модели, которые применимы к описанию поведения как реального животного, так и искусственного анимата. Дальняя цель этих работ – анализ эволюции когнитивных способностей животных и происхождения интеллекта человека. Эта цель близка к задаче моделирования когнитивной эволюции.

Искусственная жизнь. Близким направлением к «Адаптивному поведению» является интересное направление исследований «Искусственная жизнь». Основной мотивацией исследований искусственной жизни служит желание понять и промоделировать формальные принципы организации биологической жизни. Как сказал руководитель первой международной конференции по искусственной жизни К. Лангтон, «основное предположение искусственной жизни состоит в том, что "логическая форма" организма может быть отделена от материальной основы его конструкции».

Сторонники направления «Искусственная жизнь» часто считают, что они исследуют более общие формы жизни, чем те, которые существуют на Земле. Т.е. изучается жизнь, какой она могла бы в принципе быть (“life-as-it-could-be”), а не обязательно та жизнь, какой мы ее знаем (“life-as-we-know-it”). Обзор ранних работ по искусственной жизни содержится в [3].

Когнитивные архитектуры. Под когнитивными архитектурами понимаются структуры и принципы функционирования познающих систем, которые можно использовать в искусственном интеллекте [4]. Пример когнитивной архитектуры – система Soar (от англ. State, Operator And Result) [5]. Основная цель работ по Soar – создание системы функционирования интеллектуальных агентов, работающих в широкой области: от простейших форм до оперирования в сложных, заранее не предсказуемых условиях. Систему Soar предложили специалисты в области искусственного интеллекта еще в 1980-х годах, тогда ее инициировали, как попытку построить унифицированную теорию познания. В последние годы большое внимание уделяется биологически инспирированным когнитивным архитектурам [6].

¹ Работа выполнена при финансовой поддержке РФФИ, проект № 13-01-00399.

Интеллектуальные автономные агенты. Это близкое к когнитивным архитектурам направление исследований, в котором большое внимание уделяется биологически обоснованным автономным агентам и компьютерным моделям агентов, обладающих когнитивными и некоторыми интеллектуальными свойствами. Обзор исследований по автономным агентам содержится в работе [7]. Необходимо подчеркнуть, что автономные агенты вполне могут рассматриваться как объединяющее понятие для указанных направлений исследований: агенты вполне естественно могут моделировать биологические организмы и искусственные аналоги организмов. Вполне естественно изучать познавательные свойства и знания автономных агентов.

Близкие работы по научным основам искусственного интеллекта. Такие работы включают исследования когнитивных систем роботов, изучение нестандартных логик, исследование познавательных методов, используемых человеком.

Биологические эксперименты по «элементарному мышлению животных». Помимо исследований в области вычислительных наук интересные работы по когнитивным способностям живых организмов ведутся биологами. В последние годы проведены интересные биологические исследования, показывающие, что элементарные формы мышления присущи не только высшим животным, но и достаточно простым, например, насекомым (пчелам, муравьям), а также врановым птицам.

Например, в университете Оксфорда проводили исследования с новокаледонскими воронами и показали, что ворона сама, без какого-либо обучения, без каких-либо инструкций, способна изобрести способ изготовления орудия труда. См. [8] и сайт исследователей новокаледонских ворон университета Оксфорда [<http://users.ox.ac.uk/~kgroup/index.html>, <http://users.ox.ac.uk/~kgroup/tools/introduction.shtml>].

2. Этапы будущих исследований когнитивной эволюции

Как же использовать очерченные заделы при моделировании происхождения мышления? Каковы эволюционные уровни, на которых стоит остановиться? Выделим наиболее существенные этапы, ведущие от простого поведения к логическому мышлению.

А. Моделирование адаптивного поведения автономных агентов с несколькими естественными потребностями: питания, размножения, безопасности. Это могло бы быть моделирование достаточно естественного и полноценного поведения простых модельных организмов. Моделирование в этом направлении уже начато, см. ниже.

Б. Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов. Такой переход можно рассматривать, как появление в «сознании» животного свойства «понятие». Обобщенные образы можно представить как мысленные аналоги наших слов, не произносимые животными, но реально используемые ими. Использование понятий приводит к существенному сокращению и требуемой памяти, и времени обработки информации, поэтому оно должно быть эволюционно выгодным.

В. Исследование процессов формирования причинных связей в памяти животных. По-видимому, запоминание причинно-следственных связей между событиями во внешней среде и адекватное использование этих связей в поведении – одно из ключевых свойств активного познания животным закономерностей внешнего мира. Естественный следующий шаг – переход от отдельных причинных связей к логическим выводам на основе уже сформировавшихся знаний.

Г. Исследование процессов формирования логических выводов в «сознании» животных. Фактически, уже на базе классического условного рефлекса животные способны делать «логический вывод» такого вида: «Если имеет место условный стимул, и за условным стимулом следует безусловный, то нужно ожидать появления безусловного стимула». В определенной степени такие выводы подобны выводам математика, доказывающего теоремы [1]. И целесообразно разобраться в системах подобных выводов, понять, насколько

адаптивна логика поведения животных и насколько она подобна нашей, человеческой логике.

Перечисленные пункты очерчивают круг исследований от моделирования простейших форм поведения к логическим правилам, используемым в математике. Опираясь на эти пункты, мы начали соответствующее моделирование.

3. Начальные шаги моделирования когнитивной эволюции

Компьютерная модель автономных агентов, имеющих естественные потребности. В [9] была построена и исследована компьютерная модель автономных адаптивных агентов, обладающих естественными для живых организмов потребностями: питания, безопасности и размножения. Предполагалось, что агент имеет определенный ресурс R , который уменьшался при выполнении действий и увеличивался при питании. Кроме того, ресурс агента существенно уменьшался, когда на агента нападал активный хищник. Потребностям агента соответствовали три фактора: фактор питания F_F , фактор безопасности F_S , фактор размножения F_R . Для каждого фактора имелся порог T_F , T_S и T_R , при превышении фактором которого удовлетворялась соответствующая потребность.

Между потребностями вводилась следующая иерархия: 1) пищевая потребность (наиболее приоритетна), 2) потребность безопасности, 3) потребность размножения (наименее приоритетна). Потребность считалась ведущей, если она была наиболее приоритетной из всех потребностей, для которых фактор меньше своего порога.

Система управления агента была основана на наборе правил вида: $S_k \rightarrow A_k$, где S_k – ситуация, A_k – действие. Ситуация S_k определялась активностью хищника рядом с агентом, предыдущим действием агента и ведущей потребностью. Агент мог выполнять одно из следующих действий A_k : 1) покой, 2) поиск пищи, 3) питание, 4) подготовка к размножению, 5) размножение, 6) оборона. Каждое правило имело свой вес W_k , веса правил настраивались методом обучения с подкреплением [10]. Подкрепления определялись изменением фактора ведущей потребности. Агент выполнял действия, соответствующие правилам с большими весами.

Моделирование показало, что в результате обучения происходит формирование циклов поведения, в которых последовательно удовлетворяются потребности питания, безопасности и размножения.

Модель формирования обобщенных эвристик и простых обобщенных образов. Формирование обобщенных эвристик и простых обобщенных образов в процессе обучения агентов при поиске агентами пищи в двумерной клеточной среде было продемонстрировано в модели [11]. Модель предполагала, что в части клеток были порции пищи. При съедании пищи ресурс агента пополнялся. Система управления агента была основана на правилах вида $S \rightarrow A$, веса правил оптимизировались методом обучения с подкреплением.

Компьютерное моделирование показало, что в результате обучения агент формировал следующие обобщающие эвристики: «Если имеется пища в той же клетке, в которой находится агент, то нужно съесть пищу». «Если в клетке агента нет пищи, но есть пища в одной из ближайших клеток, то нужно выполнить действие или цепочку действий, направленных на движение в клетку с пищей».

Кроме этого, в компьютерную модель вводилась процедура усреднения, в результате которой агент формировал внутренние понятия «имеется пища в моей клетке», «имеется пища в клетке впереди меня», «имеется пища в клетке справа/слева от меня». Именно эти понятия использовались в обобщающих эвристиках агента.

Таким образом, построены первые модели, соответствующие выделенным этапам A и B , хотя пока эти модели целесообразно рассматривать только как начало более полноценных исследований.

4. Междисциплинарные связи исследований когнитивной эволюции

Укажем потенциальные междисциплинарные связи будущих исследований когнитивной эволюции и тем самым подчеркнем перспективность этих исследований.

- Эти исследования связаны с основаниями науки, с основаниями математики, с серьезной проблемой: почему логические выводы, математические доказательства применимы к реальной природе.
- Данные исследования интересны с философской, эпистемологической точки зрения – они нацелены на прояснение причин применимости человеческого мышления в познании природы.
- В ряде направлений исследований вычислительного интеллекта и в биологических работах развивается задел по моделированию когнитивной эволюции.
- Эти исследования интересны с точки зрения развития когнитивных наук, так как они связаны с наиболее важными когнитивными процессами – процессами научного познания.

Литература

1. Редько В.Г. Моделирование когнитивной эволюции – перспективное направление исследований на стыке биологии и математики // Математическая биология и биоинформатика (электронный журнал). 2010. Т. 5. № 2. С. 215-229. URL: [http://www.matbio.org/downloads/Redko2010\(5_215\).pdf](http://www.matbio.org/downloads/Redko2010(5_215).pdf)
2. From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior. (Meyer J.-A., Wilson S.W., Eds.). Cambridge: MIT Press, 1991.
3. Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. М.: УРСС, 2005.
4. Langley P., Laird J.E., Rogers S. Cognitive architectures: Research issues and challenges // Cognitive Systems Research. 2009. V.10. No. 2. PP. 141-160.
5. Laird L.E. The Soar Cognitive Architecture. Cambridge et al.: The MIT Press, 2012.
6. Biologically Inspired Cognitive Architectures 2012. Proceedings of the Third Annual Meeting of the BICA Society (A. Chella, R.Pirrone, R. Sorbello, K.R. Johansdottir, Eds). Heidelberg et al.: Springer, 2012.
7. Vernon D., Metta G., Sandini G. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents // IEEE Transactions on Evolutionary Computation. 2007. V. 11. No. 2. PP. 151-180.
8. Weir A.A.S., Chappell J., Kacelnik A. Shaping of hooks in New Caledonian crows // Science. 2002. V. 297. No. 5583. PP. 981-983.
9. Коваль А.Г. Редько В.Г. Поведение модельных организмов, обладающих естественными потребностями и мотивациями // Математическая биология и биоинформатика (электронный журнал). 2012. Т. 7. № 1. С. 266-273. URL: [http://www.matbio.org/2012/Koval2012\(7_266\).pdf](http://www.matbio.org/2012/Koval2012(7_266).pdf)
10. Саттон Р.С., Барто Э.Г. Обучение с подкреплением. М.: Бином, 2011.
11. Бесхлебнова Г.А., Редько В.Г. Модель формирования обобщенных понятий автономными агентами // Четвертая международная конференция по когнитивной науке: Тезисы докладов: В 2-х томах. Томск: ТГУ, 2010. Т. 1. С. 174-175.