

НА ПУТИ К МОДЕЛИРОВАНИЮ МЫШЛЕНИЯ*

В.Г. РЕДЬКО, З.Б. СОХОВА

Научно-исследовательский институт системных исследований РАН
117218, Москва, Нахимовский просп., 36, корп. 1
E-mail: niisi@niisi.msk.ru

В работе характеризуются подходы к моделированию мышления. Основное внимание уделяется возможности моделирования когнитивной эволюции, той эволюции, в результате которой сформировалось мышление человека, которое используется в научном познании. Предложена предварительная программа будущих исследований когнитивной эволюции и характеризуются начальные модели, соответствующие этой программе. Исследование когнитивной эволюции предполагает использование моделей автономных агентов, характеризующих познавательные свойства биологических организмов.

Ключевые слова: на пути к теории происхождения мышления, когнитивная эволюция, автономные агенты, познавательные способности биологических организмов.

ВВЕДЕНИЕ

В последние годы начаты работы по исследованию подходов к моделированию мышления [1]. Один из важных подходов – эволюционный, направленный на исследование когнитивной эволюции [2]. В данном подходе изучается эволюция познавательных способностей биологических организмов. Результатом такой эволюции является мышление человека, которое используется в научном познании. Моделирование когнитивной эволюции тесно связано с проблемой происхождения мышления человека и с проблемами научного познания природы. Важно, что в настоящее время имеются серьезные научные заделы моделирования когнитивной эволюции: 1) исследования компьютерных моделей автономных агентов, отражающих свойства как живых, так и модельных организмов, и 2) биологические исследования познавательных способностей животных. В настоящей работе кратко характеризуются эти заделы, предлагается предварительная программа будущих исследований когнитивной эволюции и характеризуются начальные модели, соответствующие этой программе. Также предлагается схема модели агента-физика, способного к самостоятельному открытию законов природы. Особое внимание уделяется связи с познавательными процессами, используемыми человеком в научном познании природы.

ЗАДЕЛЫ МОДЕЛИРОВАНИЯ КОГНИТИВНОЙ ЭВОЛЮЦИИ

В вычислительных науках сложилось несколько направлений исследований, которые связаны с моделированием происхождения мышления и которые могут быть использованы при таком моделировании. Охарактеризуем кратко эти направления.

Адаптивное поведение. Направление развивается с начала 1990-х годов [3]. Исследователи в данном направлении создают и изучают искусственные «организмы» (в виде компьютерной программы или робота), которые способны адаптироваться к окружающей среде.

Искусственная жизнь. «Искусственная жизнь» является достаточно молодым междисциплинарным направлением исследований. Впервые термин *Artificial Life* был использован К. Лангтоном на конференции в Санта Фе в 1987 г. Сторонники данного

* Работа выполнена в рамках бюджетного проекта № 35.16. «Моделирование процессов формирования когнитивных способностей автономных агентов» (№ 0065-2014-0016).

направления исследуют эволюционные сценарии не только «жизни, какая она есть», но и более общие формы «жизни, какой она могла бы быть».

Когнитивные архитектуры. Когнитивная архитектура представляет собой структуру и принципы функционирования систем познания, которые можно использовать в искусственном интеллекте. Исследования в этой области важны, так как они поддерживают основную цель искусственного интеллекта и когнитивной науки: создание и понимание искусственных агентов, которые обладают теми же возможностями, что и человек [4]. Еще одно направление, которое развивается в последние годы, – биологически инспирированные когнитивные архитектуры (Biologically Inspired Cognitive Architectures, BICA) [5].

Близкие работы по научным основам искусственного интеллекта. Здесь стоит отметить работы, которые включают изучение нестандартных логик, исследования когнитивных систем роботов и познавательных методов, используемых человеком.

Интеллектуальные автономные агенты. Исследователи данного направления большое внимание уделяют биологически обоснованным автономным агентам и компьютерным моделям агентов, обладающих когнитивными и некоторыми интеллектуальными свойствами.

Следует отметить, что автономные агенты вполне могут рассматриваться как объединяющее понятие для указанных выше направлений.

Перейдем от заделов исследования когнитивной эволюции в вычислительных науках к краткому изложению заделов со стороны биологии.

Биологические эксперименты по «элементарному мышлению животных». Рассмотрим биологические исследования, которые показывают, что элементарными формами мышления обладают не только высшие животные, но и достаточно простые. Подчеркивая, что изучаются биологические предшественники мышления человека, авторы этих исследований характеризуют такие познавательные способности, как «элементарное мышление животных» [6]. Остановимся на экспериментах по когнитивным способностям новокаледонских воронов.

Новокаледонские вороны могут изобретать способ изготовления орудий труда. Новокаледонские вороны, которые живут на природе, умеют обкусывать веточки таким образом, что получаются простые орудия (заостренные палочки или крючки). На такую палочку ворона может насадить личинки насекомых, а с помощью крючка может вытаскивать личинки из-под коры.

Ученые из университета Оксфорда проводили исследования с новокаледонскими воронами, которые долгое время находились в неволе [7]. Эксперимент заключался в следующем. Двум воронам (самке и самцу) предлагали добывать ведро с пищей со дна прозрачного вертикального цилиндра. Около цилиндра были прямая проволока и проволока, согнутая крючком. Для того чтобы вытащить ведро, нужно было использовать крючок. До этого с проволокой вороны дела не имели, несмотря на это, они сразу поняли, что ведро можно достать с помощью крючка. Интересен тот факт, что, когда самец утащил крючок, самка сначала пыталась вытащить ведро с помощью прямой проволоки, и когда у нее не получилось, она догадалась зажать один конец прямой проволоки в одной из щелей экспериментальной установки и загнуть проволоку. А затем с помощью изготовленного крючка уже доставала ведро с пищей. После такого эксперимента, если крючок убирала, а оставляли прямую проволоку, ворона уже сразу делала крючок и доставала пищу. Заметим, что самец только наблюдал за ней и иногда отнимал у нее пищу, но опыт не перенял. Таким образом, ворона самостоятельно изобрела способ изготовления орудия труда.

Новокаледонские вороны могут мысленно составлять планы цепочек целенаправленных действий. Еще один интересный эксперимент с новокаледонскими воронами провели ученые из Новой Зеландии [8]. В течение эксперимента воронам надо

было добыть пищу, которая лежала в одном из контейнеров. Но до пищи можно было дотянуться только с помощью длинной палочки, которая лежала в другом зарешеченном контейнере. Длинную же палочку можно было достать только с помощью короткой палочки, которая висела на шнуре. Таким образом задача, которую предлагали решить воронам, состояла из трех частей: 1) подтянуть к себе шнур и освободить от шнура маленькую палочку, 2) с помощью маленькой палочки достать длинную палочку, 3) с помощью длинной палочки достать пищу. Сначала вороны тренировались в более простых условиях, т. е. они могли использовать инструменты по отдельности или же имели возможность поработать только с частью инструментов. После этого, когда воронам предлагали выполнить полное задание, состоящее из трех действий, те вороны, которые имели опыт использования всех трех инструментов по отдельности, выполняли задание с первого раза. Те же вороны, которые были знакомы только с частью инструментов, не всегда справлялись с первого раза. Но тем не менее они тоже достаточно быстро научились решать поставленную задачу. Таким образом, вороны научились продумывать план решения новой задачи, мысленно связывая в плане ранее освоенный опыт.

ПРЕДВАРИТЕЛЬНАЯ ПРОГРАММА БУДУЩИХ ИССЛЕДОВАНИЙ КОГНИТИВНОЙ ЭВОЛЮЦИИ

Отталкиваясь от очерченных работ по заделам, рассмотрим контуры программы будущих исследований, ориентированные на моделирование когнитивной эволюции. При этом выделим наиболее важные этапы, ведущие к логическому мышлению. Ниже представлены пункты программы.

А. Разработка моделей адаптивного поведения автономных агентов с несколькими естественными потребностями: питания, размножения, безопасности. Здесь могут представлять интерес работы, моделирующие естественное поведение простых модельных организмов. В данном направлении уже ведутся работы (см. ниже).

Б. Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов. Здесь можно рассматривать вопросы, связанные с появлением в «сознании» животного свойства «понятие».

В. Исследование процессов формирования причинных связей в памяти животных. Должно быть, одним из ключевых свойств активного познания закономерностей окружающей среды является запоминание причинно-следственных связей между событиями во внешней среде и разумное использование этих связей в поведении.

Г. Исследование процессов формирования логических выводов в «сознании» животных. Можно утверждать, что уже на базе классического условного рефлекса животные способны делать «логический вывод» вида: «Если имеет место условный стимул, и за условным стимулом следует безусловный, то нужно ожидать появления безусловного стимула». В сущности, такие выводы имеют сходство с выводами математика, который доказывает теоремы. Поэтому интересен вопрос: насколько адаптивна логика поведения животных и насколько она подобна нашей, человеческой логике.

Д. Исследование коммуникаций, возникновения языка. Несомненно, что язык и мышление тесно взаимосвязаны. Поэтому целесообразно исследовать, каким образом развитие коммуникаций и языка способствовало развитию логики, мышления, интеллекта человека; как возникновение языка повлияло на познавательные возможности живых существ.

Перечисленные пункты очерчивают круг исследований от моделирования простейших форм поведения к логическим правилам дедуктивного вывода, используемым в математике.

Пункт *А* соответствует общим свойствам достаточно простого организма и вместе с тем характеризует довольно полноценное поведение организма.

Пункты *Б*, *В*, *Г* характеризуют следующие друг за другом и причинно связанные между собой этапы эволюции познавательных способностей биологических организмов. Действительно, сначала должны сформироваться обобщающие сенсорную информацию

понятия, затем с помощью понятий можно формализовать причинные связи между событиями во внешнем мире и причинные связи между действиями организма и вызываемыми этими действиями результатами. А затем, базируясь на причинных связях, можно делать логические выводы о будущих событиях, строить планы достижения нужных результатов.

Пункт *Д* характеризует дополнительные к предыдущим пунктам коллективные свойства организмов, которые естественно должны были возникать в процессе эволюции.

Можно ожидать, что предлагаемые по данной программе исследования помогут понять причины возникновения существенных познавательных свойств организмов в процессе эволюции.

Конечно же, перечисленные пункты характеризуют не все нетривиальные когнитивные способности. Например, важные познавательные свойства – поиск нового, открытие нового путем случайного поиска, путем сопоставления уже известных знаний или путем озарения, неожиданного нахождения нужного решения. Можно, скажем, вспомнить впечатления А. Пуанкаре, который описывал, как он довольно долго безуспешно искал нужное математическое решение и «держал в голове» сведения о решаемой проблеме и необходимости нахождения решения, и только внезапно (как бы случайно) нашел требуемое решение [9].

Поэтому перечисленные пункты претендуют только на начальное приближение, полноценное исследование когнитивной эволюции целесообразно вести шаг за шагом методом последовательных приближений. Тем не менее эти пункты характеризуют путь от простейших форм адаптивного поведения к логическим правилам вывода, используемым в математических доказательствах [10], т.е. эти пункты намечают первое приближение всего пути исследований от начала до конца.

НАЧАЛЬНЫЕ ШАГИ МОДЕЛИРОВАНИЯ КОГНИТИВНОЙ ЭВОЛЮЦИИ

Модель автономных агентов с несколькими естественными потребностями. В работе [11] исследована компьютерная модель адаптивного поведения автономных агентов, обладающих потребностями питания, размножения, безопасности, аналогичными основным потребностям биологических организмов. Модель соответствует пункту *А* изложенной выше программы. В модели продемонстрировано формирование циклов поведения, в которых последовательно удовлетворяются потребности питания, безопасности и размножения.

Модель формирования обобщенных понятий автономными агентами. В работе [12] изучена компьютерная модель, в которой автономный агент самостоятельно проводит обобщения, в результате которых формируются понятия. Эта модель соответствует пункту *Б* программы моделирования когнитивной эволюции. Рассматривалось поведение автономного агента в двумерной клеточной среде, в клетках которой случайным образом размещены порции пищи.

Показано, что в процессе обучения агент способен самостоятельно формировать эвристики (правила повеления), которые приводят к повышению ресурса агента. Также агент способен формировать следующие полезные понятия: «имеется пища в моей клетке», «имеется пища в клетке впереди меня», «имеется пища в клетке справа/слева от меня». То есть агент способен самостоятельно формировать понятия, обобщающие сенсорную информацию.

Модели познавательных способностей рыб, изучающих лабиринты [13]. Модели основаны на биологическом эксперименте, в котором изучалось поведение рыб в простых лабиринтах. Была построена компьютерная модель накопления рыбами знаний о коридорах. Считалось, что агент (модельная рыба) имеет определенное знание K_i о каждом из коридоров лабиринта. Предполагалось, что в момент посещения агентом i -го коридора величина K_i становится равной 1, а для остальных коридоров величины K_i постепенно уменьшаются. Каждый коридор соответствовал определенной ситуации, в которую попадал агент. Агент мог сам ставить себе цель: двигаться в тот коридор, знание о котором минимально, т.е. в тот коридор, в котором он давно не был. Была также построена модель,

в которой рассчитывалась уверенность агентов в предсказании будущих ситуаций при выполнении тех или иных действий.

Кроме того, была построена модель, предполагавшая, что модельная рыба уже хорошо освоила лабиринт, т.е. агент уже имеет сформированные знания для всех возможных ситуаций и предсказания будущих ситуаций при выполнении определенных действий. Используя эти знания и предсказания, агент может создать план движения к целевому месту.

Агент формирует план, сначала мысленно анализируя пути движения от целевой ситуации к исходной. В результате у агента создается простая *база знаний*, которая характеризует ситуации, действия, прогнозы результатов действий, оценку расстояний между текущими ситуациями и целевой ситуацией (в наших моделях это расстояние равно числу действий, необходимых для перехода из текущей ситуации в целевую). Используя эту базу знаний, агент формирует план движения от исходной ситуации к целевой.

Модель планирования новокаледонскими воронами. Модель основана на биологическом эксперименте с новокаледонскими воронами [8]. В этом эксперименте вороны предварительно обучались выполнению отдельных элементов довольно сложного поведения. После обучения воронам предлагалось решить следующую трехзвенную проблему: (1) подтянуть привязанную на шнуре короткую палочку и освободить ее от шнура, (2) с помощью короткой палочки достать длинную палочку из зарешеченного контейнера, (3) длинной палочкой извлечь пищу из второго глубокого прозрачного контейнера.

Итак, вороны обучались отдельным элементам поведения, направленного на решение трехзвенной проблемы. А затем вороны самостоятельно мысленно планировали полностью все это поведение, мысленно связывая отдельные элементы в единую цепочку.

Была построена модель мысленного формирования плана поведения, нацеленного на решение трехзвенной проблемы [14]. Модель предполагала, что механизм формирования плана решения трехзвенной проблемы включает в себя следующие этапы:

- прямое мысленное рассмотрение от исходной ситуации к целевой (до тех пор, пока не будет найдена правильная цепочка действий, ведущих к цели);
- проверка найденного решения путем тестирующего обратного и прямого мысленного рассмотрения;
- генерация стереотипа целенаправленного поведения после тестирования.

После первого нахождения результата происходит его проверка, при этом при тестирующем обратном мысленном рассмотрении оцениваются расстояния между текущими ситуациями и целевой, при успешной проверке формируется база знаний, характеризующая ситуации, действия, прогнозы результатов действий, оценки расстояний между текущими ситуациями и целевой ситуацией. На основе базы знаний строится план решения трехзвенной проблемы.

Отметим, что в модели была введена проверка полученного результата путем обратного и прямого тестирующего мысленного рассмотрения. Это тестирование подобно быстрому обратному (а, возможно, и прямому) возбуждению клеток места в гиппокампе животных после достижения результата путем перемещения в пространстве [15, 16]. Возможно, что такая проверка играет важную роль в надежном запоминании полученного решения.

После проверки формируется стереотип целенаправленного поведения. При выполнении действий в соответствии со стереотипом уже не нужны мысленные усилия для формирования плана поведения, следовательно, стереотипное поведение должно происходить быстро.

Было проведено компьютерное моделирование по приведенной схеме формирования планов воронами. Результаты моделирования качественно согласовывались с биологическим экспериментом [8].

Таким образом, построены и исследованы биологически инспирированные модели когнитивных автономных агентов, познающих закономерности взаимодействия с внешним

миром. Обобщая свойства модельных рыб и модельных ворон, можно выделить следующие характерные черты рассмотренных когнитивных агентов:

- 1). Оценка *знаний* о каждой из возможных ситуаций.
- 2). Движение к ситуации, знания о которой минимальны. Тем самым *агент может самостоятельно формировать новые текущие цели своего поведения*. Например, рыба в лабиринте может решить двигаться в тот коридор, для которого знания минимальны, т.е. в тот коридор, который она давно не посещала.
- 3). Формирование *предсказаний* о результатах действий. Оценка *уверенности в предсказаниях*, стремление к надежным предсказаниям, уверенность в которых высока.
- 4). Формирование *простой базы знаний*, характеризующей ситуации, действия, прогнозы результатов действий, оценки расстояний между текущими ситуациями и целевой ситуацией.
- 5). Использование базы знаний при планировании целенаправленного поведения.
- 6). *Прямые и обратные мыслительные процессы* (от исходной ситуации к целевой и обратно).
- 7). Формирование *стереотипа* целенаправленного поведения.

СХЕМА МОДЕЛИ АГЕНТА-ФИЗИКА

Интересный подход к воссозданию интеллекта человеческого уровня был предложен в работе [17]. Было предложено создать такого интеллектуального робота, поведение которого было бы настолько «разумно», что люди могли бы выбрать этого робота президентом страны.

Хотя с точки зрения анализа когнитивных способностей человека, используемых в научном познании, можно было бы подойти к воссозданию интеллекта человеческого уровня другим образом. В работе М.Н. Вайнцвайга [18] анализировались модели автономных агентов, которые познают закономерности механики, наблюдая за столкновениями тел, например, шаров. Исследование таких автономных агентов в конечном итоге может быть нацелено на то, чтобы агенты пришли к самостоятельному открытию трех законов Ньютона. Ну, чем это не интеллект человеческого уровня (конечно, пока только в замыслах) – открытие законов природы?!

И используя наши знания о научной работе Исаака Ньютона, можно представить, каков мог бы быть «интеллект» автономного агента, чтобы агент соответствовал уровню ученого. У агента должна быть база знаний, должно быть стремление к получению новых знаний и к обобщению знаний. Должна быть любознательность, направляющая агента к постановке вопросов о внешнем мире и решению этих вопросов путем постановки экспериментов. Агент должен учитывать многочисленные связи между уже имеющимися знаниями. Должен быть коллектив агентов, исследующий внешний мир, и должны быть коммуникации между агентами. Должно быть самосознание агентов, эмоциональная самооценка результатов своей деятельности и стремление агента достигнуть наиболее высоких результатов в коллективе агентов. Агенты должны иметь стремление к получению наиболее ясных, четких и компактных знаний, таких как законы Ньютона или аксиомы Евклида. И, конечно, агенты должны освоить возможности логических выводов, позволяющих получить многочисленные следствия законов и аксиом.

Понятно, что путь к интеллекту такого уровня весьма непрост но, все же он просматривается.

ЗАКЛЮЧЕНИЕ

Итак, предложен подход к исследованию происхождения мышления человека на основе моделирования когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, той эволюции, в результате которой сформировалось мышление человека, используемое в научном познании.

ЛИТЕРАТУРА

1. От моделей поведения к искусственному интеллекту / Под ред. В.Г. Редько. М.: УРСС, 2006. 448 с.
2. Редько В.Г. Моделирование когнитивной эволюции: на пути к теории эволюционного происхождения мышления. М: ЛЕНАНД/URSS, 2015. 256 с.
3. From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior (Eds. Meyer J.-A., Wilson S.W.). Cambridge: MIT Press, 1991. 562 pp.
4. Langley P., Laird J.E., Rogers S. Cognitive architectures: Research issues and challenges // Cognitive Systems Research. 2009. Vol. 10. No. 2. PP. 141-160.
5. Samsonovich A.V. On a roadmap for the BICA Challenge // Biologically Inspired Cognitive Architectures. 2012. Vol. 1. PP. 100-107.
6. Зорина З.А., Поляева И.И. Зоопсихология. Элементарное мышление животных. М.: АспектПресс. 2001. 320 с.
7. Weir A.A.S., Chappell J., Kacelnik A. Shaping of hooks in New Caledonian crows // Science. 2002. Vol. 297. No. 5583. PP. 981-983.
8. Taylor A.H., Elliffe D., Hunt G.R., Gray R.D. Complex cognition and behavioural innovation in New Caledonian crows // Proceedings of the Royal Society B: Biological Sciences. 2010. Vol. 277. No. 1694. PP. 2637-2643.
9. Пуанкаре А. О науке. М.: Наука. 1990. 736 с. (См. в этом сборнике трудов: книга «Наука и метод», глава III «Математическое творчество»).
10. Математическая теория логического вывода / Под ред. А.В. Идельсона и Г.Е. Минца). М.: Наука. 1967. 351 с.
11. 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)
12. 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.
13. Red'ko V.G., Nepomnyashchikh V.A., Osipova E.A. Models of fish exploratory behavior in mazes // Biologically Inspired Cognitive Architectures. 2015. Vol. 13. PP. 9-16.
14. Red'ko V.G., Burtsev M.S. Modeling of mechanism of plan formation by New Caledonian crows // Procedia Computer Science. 2016. Vol. 88. P. 403-408. URL: <http://www.sciencedirect.com/science/article/pii/S1877050916317124>
15. Foster D.J., Matthew A., Wilson M.A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state // Nature. 2006. Vol. 440. No. 7084. PP. 680-683.
16. Diba K., Buzsaki G. Forward and reverse hippocampal place-cell sequences during ripples // Nature Neuroscience. 2007. Vol. 10. No. 10. PP. 1241-1242.
17. Chella A., Lebiere C., Noelle D.C., Samsonovich A.V. On a Roadmap to Biologically Inspired Cognitive Agents // Biologically Inspired Cognitive Architectures 2011. Proceedings of Second Annual Meeting of the BICA Society (Eds. Samsonovich A.V., Johansdottir K.R.). Amsterdam et al.: IOS Press, 2011. PP. 453-460.
18. Вайнцвайг М.Н. Мышление как механизм обучения организации поведения // В книге [1]. С. 203-218.

ON THE WAY TO THE MODELING OF THINKING

V.G. RED'KO, Z.B. SOKHOVA

Scientific Research Institute for System Analysis, Russian Academy of Sciences
Russia, 117218, Moscow, Nakhimovsky prospect, 36, building 1
E-mail: niisi@niisi.msk.ru

Approaches to the modeling of thinking are characterized. The main attention is paid to the possibility of modeling cognitive evolution, that is evolution of animal cognitive abilities. The result of this evolution is human thinking, which is used in scientific cognition. A preliminary program of future studies of

cognitive evolution is proposed and the initial models corresponding to this program are characterized. The study of cognitive evolution involves the use of models of autonomous agents that characterize the cognitive properties of biological organisms.

Keywords: origin of thinking, cognitive evolution, autonomous agents, cognitive abilities of biological organisms.

Работа поступила 12.11.2017 г.