О моделировании когнитивной эволюции

НИИ системных исследований РАН

Редько Владимир Георгиевич
vgredko@gmail.com

Основная тема доклада

В докладе будут рассмотрены перспективы моделирования когнитивной эволюции, т.е. эволюции познавательных способностей биологических организмов, в результате которой произошло логическое мышление человека, используемое в научном познании

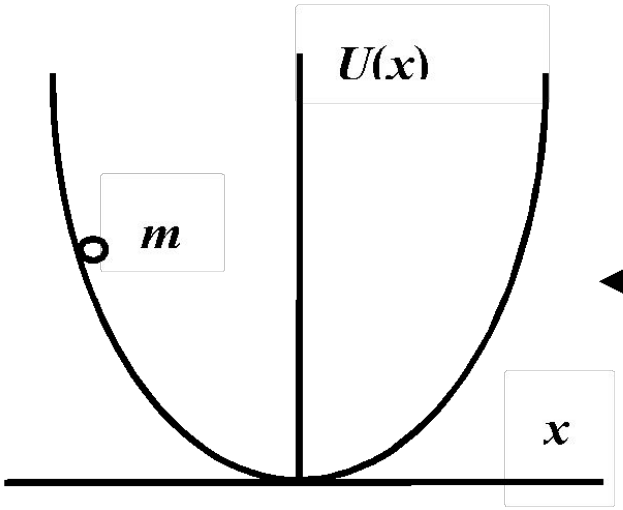
План доклада

1. **Ключевая проблема – почему формальный логический вывод, сделанный человеком, можно использовать в познании природы?**
2. **Заделы исследований когнитивной эволюции: модели автономных агентов**
3. **Биологические эксперименты по элементарному мышлению животных**
4. **Начальные шаги и перспективы моделирования когнитивной эволюции**

Ключевая проблема – почему формальный логический вывод, сделанный человеком, можно использовать в познании природы?

Гносеологическая проблема

Математики доказывают теоремы. Почему результаты, полученные формальным логическим путем, применимы к физическим объектам в природе?



$$m\, d^2x/dt^2 = -\partial U(x)/\partial x$$

?

$$m(dx/dt)^2/2 + U(x) = const$$

Почему логический вывод, сделанный человеком, применим к реальному объекту в природе?

Кто думал над такими вопросами

Иммануил Кант – провел исследование познавательных процессов в приближении фиксированного мышления взрослого человека («Критика чистого разума», 1781 г.)

Конрад Лоренц – от кантовской доктрины априорного к эволюционной теории познания (1941 г.)

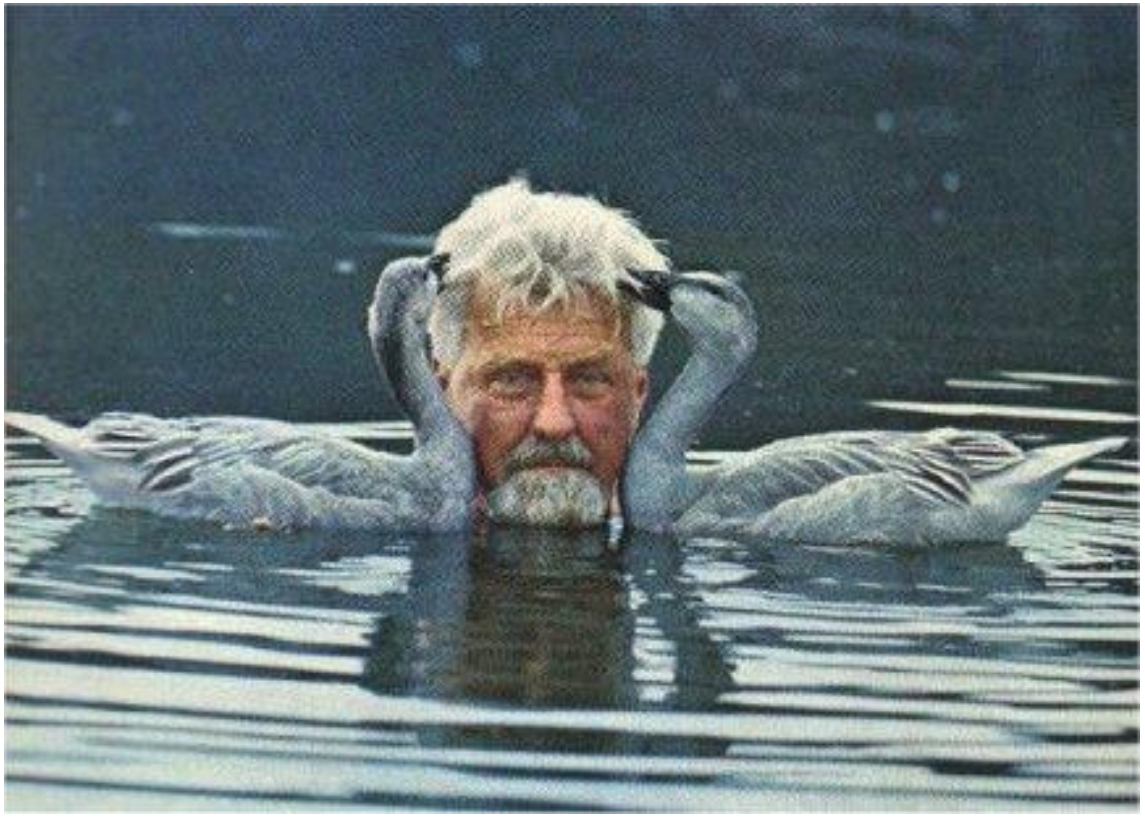
Может ли человек познавать законы природы?

«…хотя вначале это звучит странно, но тем не менее верно, если я скажу: *рассудок не черпает свои законы* (a priori) *из природы, а предписывает их ей*»

И. Кант. Пролегомены ко всякой будущей метафизике, могущей появиться как наука, 1783 г.

«…наши категории и формы восприятия, зафиксированные *до* индивидуального опыта, адаптированы к внешнему миру в точности по тем же причинам, по которым копыто лошади адаптировано к степному грунту еще до того, как лошадь рождается, а плавник рыбы – к воде до ее появления из икринки».

К. Лоренц. Кантовская доктрина априорного в свете современной биологии, 1941 г.



Иммануил Кант
1724-1804

Конрад Лоренц
1903-1989

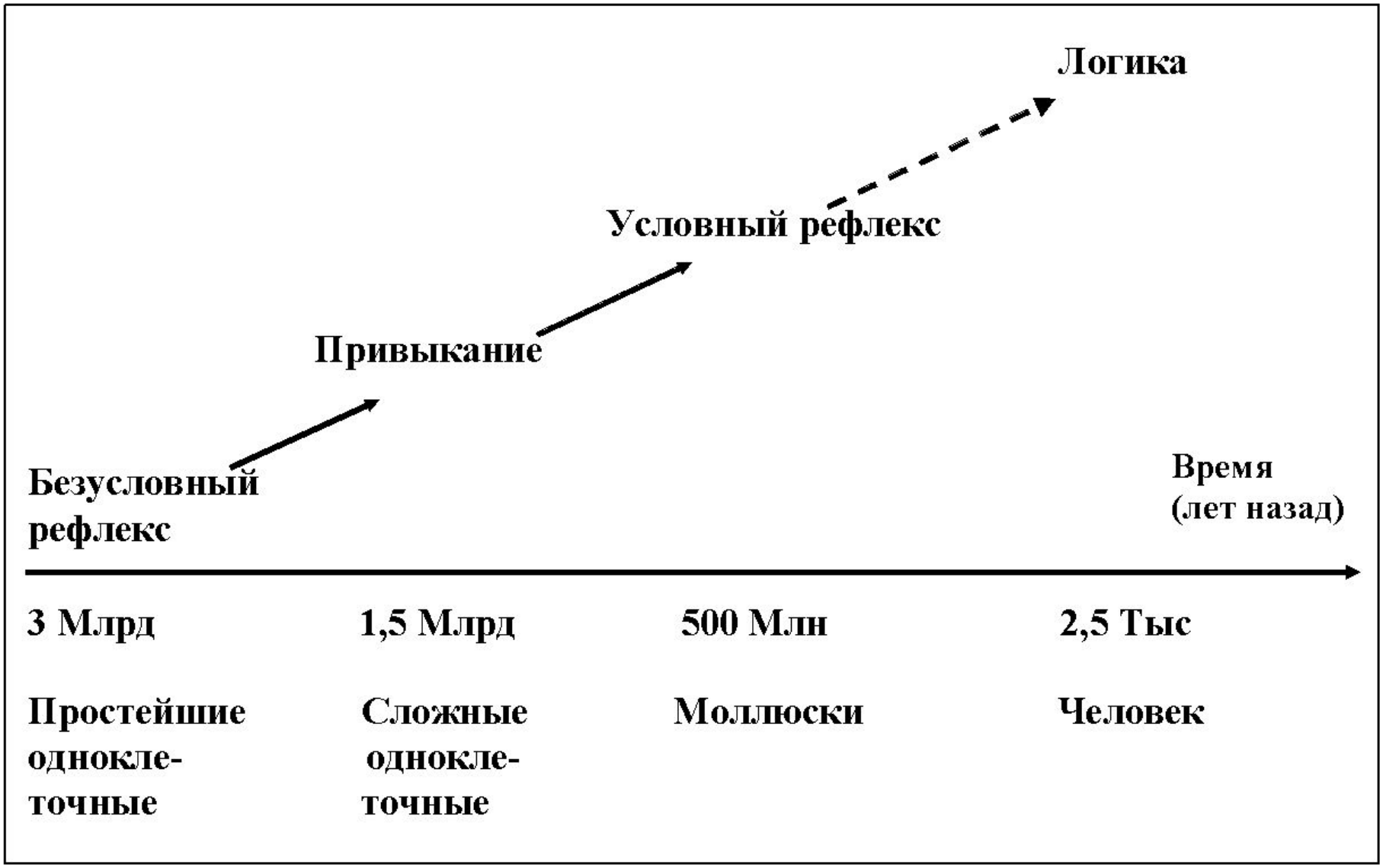
Конрад Лоренц и его гусята

Подход к исследованию гносеологической проблемы

Исследовать происхождение логического мышления, интеллекта человека путем построения математических и компьютерных моделей когнитивной эволюции.

Проследить весь путь биологической эволюции от простейших до человека, анализируя с помощью моделей, как на этом пути возникали способности познания закономерностей природы.

Проблема происхождения мышления



Аналогия

Математик (при доказательстве теорем): правило modus ponens: «если имеет место *A*, и из *A* следует *B*, то имеет место *B*», или

$$\frac{A, A \to B}{B}$$

Собака (после выработки условного рефлекса):

$$\frac{\text{УС}, \text{УС} \to \text{БС}}{\text{БС}}$$

УС – условный стимул, БС – безусловный стимул

Заделы исследований когнитивной эволюции

Адаптивное поведение

- Первая конференция: Париж, 1990 г. (Ж.-А. Мейер, С. Вильсон)
- Основной подход – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов» (аниматов, агентов), способных приспосабливаться к внешней среде
- ANIMAL + ROBOT = ANIMAT
- Программа-минимум – исследовать архитектуры и принципы функционирования, которые позволяют животным или роботам жить и действовать в переменной внешней среде
- Программа-максимум – попытаться проанализировать эволюцию когнитивных (познавательных) способностей животных и эволюционное происхождение человеческого интеллекта
- Предшественники: М.Л. Цетлин, М.М. Бонгард.
  Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. М.: Наука, 1987. М.: УРСС, 2004, 2011.

Искусственная жизнь

Направление исследований «Искусственная жизнь» сформировалось в конце 1980-х годов. Основной мотивацией исследований искусственной жизни служит желание понять и промоделировать формальные принципы организации биологической жизни.

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”).

Когнитивные архитектуры

Под когнитивными архитектурами понимаются структуры и принципы функционирования познающих систем, которые можно использовать в искусственном интеллекте.

Langley P., Laird J.E., Rogers S. Cognitive architectures: Research issues and challenges // Cognitive Systems Research. 2009. V.10. No. 2. PP. 141-160.

Laird L.E. The Soar Cognitive Architecture. Cambridge et al.: The MIT Press, 2012.

Самсонович А.В. Biologically inspired cognitive architectures – проведение международных конференций 2010, 2011 гг.

Конференции BICA: Палермо (2012), Киев (2013).

Автономные агенты

Автономные агенты имеют свою собственную систему управления.

Автономные агенты вполне могут рассматриваться как объединяющее понятие для отмеченных направлений.

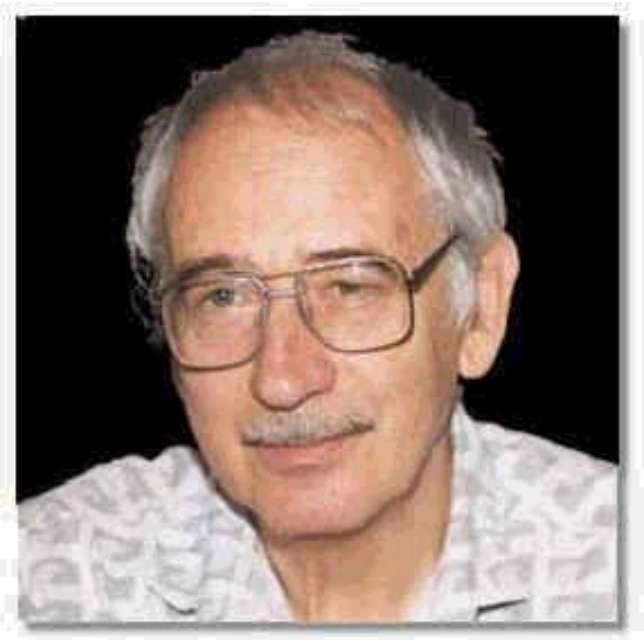
Работы по автономным агентам ведутся как со стороны биологических наук, так со стороны вычислительных наук.

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, special issue on Autonomous Mental Development, 2007. V. 11. No. 2. PP. 151-180.

Попытка пересмотра оснований математики

Анализ возможности построения теории предиктивных логических правил в контексте теории множеств и кибернетического подхода к обоснованию математики:

Turchin V.F. A constructive interpretation of the full set theory // Journal of Symbolic Logic, 1987. V. 52. No. 1. PP. 172 -201



В.Ф. Турчин. Феномен науки:

Кибернетический подход к эволюции – М.: Наука, 1993 (1-е изд.). М.: ЭТС, 2000 (2-е изд.):

http://www.refal.net/turchin/phenomenon/

Заделы исследований когнитивной эволюции

Работы по моделированию целенаправленного поведения (Г. С. Осипов), нестандартным логикам (В.К. Финн, А.П. Еремеев, В.Н. Вагин), многоагентным системам (В.Б. Тарасов), мобильным роботам (В.Э. Карпов, В.Е. Павловский, А.С. Ющенко)

Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова // Нейроинформатика (электронный рецензируемый журнал). 2008. Т. 3. № 1. С. 25-78.
http://www.niisi.ru/iont/ni/Journal/V3/N1/Vityaev.pdf

Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // Нейроинформатика (электронный рецензируемый журнал). 2008. Т. 3. № 1. С. 79-108.
http://www.niisi.ru/iont/ni/Journal/V3/N1/DeminVityaev.pdf

Заделы исследований когнитивной эволюции

Станкевич Л.А. Искусственные когнитивные системы // Научная сессия НИЯУ МИФИ - 2010. XII Всероссийская научно-техническая конференция "Нейроинформатика-2010". Лекции. М.: НИЯУ МИФИ, 2010. С. 106-160.

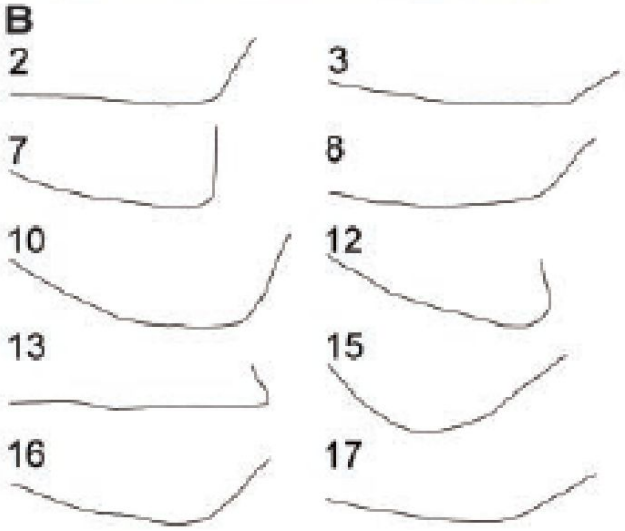
Автономные агенты, модельный футбол, антропоморфные роботы

Вайнцвайг М.Н., Полякова М.П. О моделировании мышления // От моделей поведения к искусственному интеллекту М.: УРСС, 2006. С. 280-286.

Жданов А.А. Автономный искусственный интеллект. М.: Бином. Лаборатория знаний, 2009

Биологические эксперименты по элементарному мышлению животных

Новокаледонские вороны могут изобретать и изготовлять орудия труда

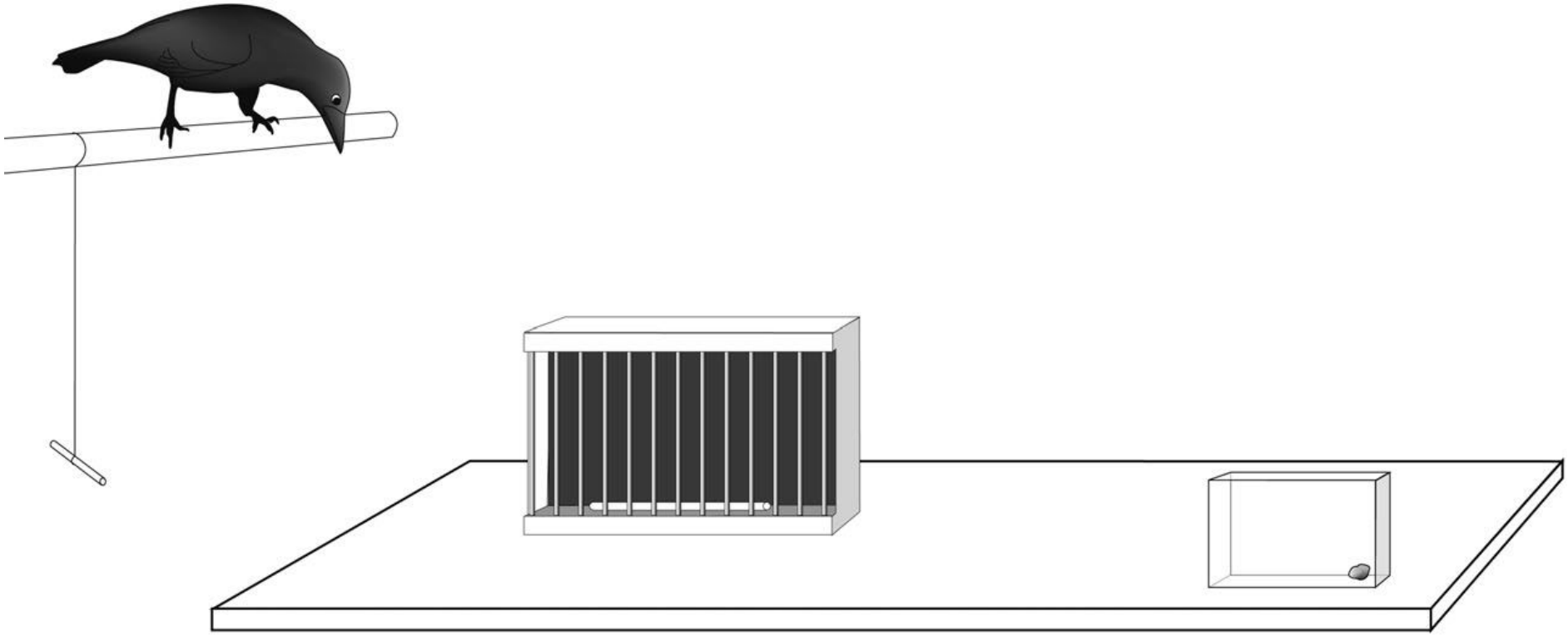


Двум воронам (молодой самке и самцу постарше) предлагали добывать ведерко с пищей прямой проволочкой и проволочкой, согнутой крючком. Вороны сразу поняли, что ведерко можно вытащить с помощью крючка.

Однажды самец утащил крючок. Тогда самка быстро научилась делать из прямой проволоки крючок, зажимая один конец проволоки в щели и загибая проволоку.

Weir A.A.S., Chappell J., Kacelnik A. Science. 2002. V. 297. P. 981-983

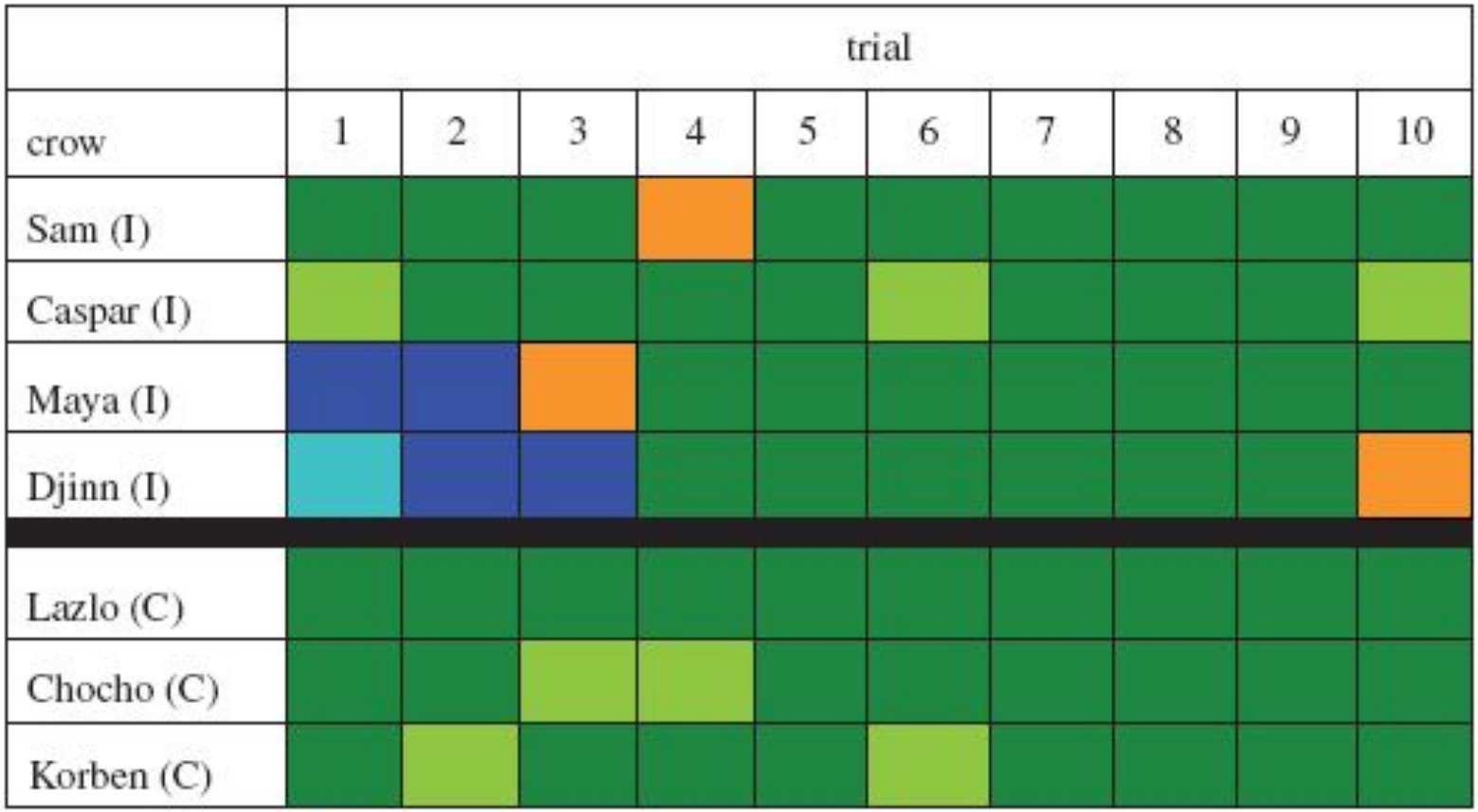
Вороны мысленно составляют планы цепочек целенаправленных действий



Taylor A.H., Elliffe D., Hunt G.R., Gray R.D. Proc. R. Soc. B. 2010. V. 277. P. 2637-2643.

Действия: 1) Подтянуть шнуром маленькую палочку. 2) Маленькой палочкой достать длинную. 3) Длинной палочкой достать пищу.

Результаты составления плана воронами



3 вороны, имевшие опыт со всеми действиями отдельно, сразу решают задачу. 4 вороны, имевшие опыт с частью действий, решают задачу, но не всегда сразу. Зеленый цвет – решение находится, синий – не находится, оранжевый – находится при малых поисковых вариациях

Контуры программы будущих исследований когнитивной эволюции

- **Исследование моделей адаптивного поведения агентов с несколькими естественными потребностями: питания, размножения, безопасности**
- **Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов, уровню понятий (аналогов слов)**
- **Исследование процессов формирования причинной связи в памяти животных. Например, связи между условным стимулом (УС) и следующим за ним безусловным стимулом (БС). Анализ роли прогнозов в адаптивном поведении**
- **Исследование процессов формирования логических выводов в «сознании» животных**

  {УС, УС --> БС} => БС – аналог modus ponens

Начальные шаги моделирования когнитивной эволюции

Модель автономных агентов с естественными потребностями: питание, размножение, безопасность

Коваль А.Г., Редько В.Г. Поведение модельных организмов, обладающих естественными потребностями и мотивациями // Математическая биология и биоинформатика, 2012, Т. 7. № 1. С. 266-273.
URL: http://www.matbio.org/2012/Koval2012(7_266).pdf

Модель агентов с потребностями питания, размножения, безопасности

Система управления агента – набор правил вида: $S \rightarrow A$ (в ситуации S нужно выполнить действие A).

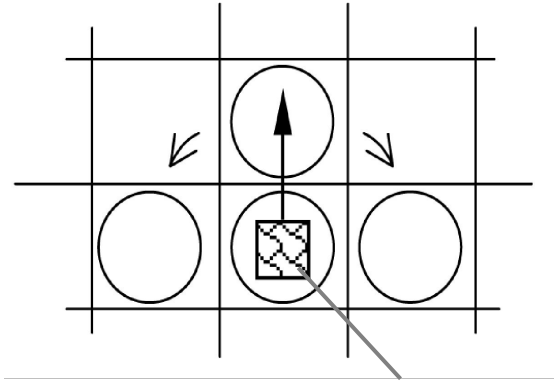
Ситуация S : 1) активность хищника рядом с агентом, 2) предыдущее действие агента, 3) приоритетная потребность.

Действия A: 1) поиск пищи, 2) питание, 3) подготовка к размножению, 4) размножение, 5) оборона от хищника, 6) покой.

В результате обучения агентов происходит формирование циклов, в которых последовательно удовлетворяются потребности питания, безопасности и размножения.

Формирование внутренних понятий

Описание модели



Агент

Стрелка показывает направление вперед, кружки – поле зрения агента

Автономный агент ведет поиск пищи в двумерной клеточной среде.

Система управления агента – набор правил вида: $S \rightarrow A$.

Ситуация S определяется наличием или отсутствием пищи в поле зрения агента.

Действия агента A: двигаться вперед, поворачиваться направо или налево, питаться, отдыхать.

Веса правил настраиваются методом обучения с подкреплением. Действия выбирались в соответствии с весами.

Формирование внутренних понятий

Результаты моделирования

Поведение обученного агента таково:
Действие *питание* выполняется, если имеется пища в той клетке, в которой находится агент. Действия *перемещение вперед* либо *поворот направо/налево* выполняются, если нет пищи в той клетке, в которой находится агент, но имеется пища в клетке впереди агента либо в клетке справа/слева от агента.

Вычислялось среднее число применений определенного действия для той или иной ситуации.
Показано, что автономный агент формировал внутренние понятия «*имеется пища в моей клетке*», «*имеется пища в клетке впереди меня*», «*имеется пища в клетке справа/слева от меня*».

Перспективы моделирования когнитивной эволюции

Контуры программы будущих исследований когнитивной эволюции

- **Исследование моделей адаптивного поведения агентов с несколькими естественными потребностями: питания, размножения, безопасности**
- **Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов, уровню понятий (аналогов слов)**
- **Исследование процессов формирования причинной связи в памяти животных. Например, связи между условным стимулом (УС) и следующим за ним безусловным стимулом (БС). Анализ роли прогнозов в адаптивном поведении**
- **Исследование процессов формирования логических выводов в «сознании» животных**

  {УС, УС --> БС} => БС – аналог modus ponens

Картина когнитивной эволюции пока только чуть-чуть видна

Пункты программы очерчивают круг исследований от моделирования простейших форм поведения к логическим правилам, используемым в математике.

Анализ известных моделей показывает, что уже имеются отдельные элементы, соответствующие каждому из пунктов.

Образно говоря, у нас уже есть некоторые небольшие фрагменты картины, но мы еще не видим всей картины.

Четкой последовательности серьезных, канонических моделей, которые показывали бы общую картину происхождения логического мышления, пока еще нет.

Моделирование когнитивной эволюции – перспективное междисциплинарное направление исследований

Это направление связано с широким кругом дисциплин:

- **с основаниями математики,**
- **с теорией познания,**
- **с анализом познавательных способностей биологических организмов,**
- **с когнитивными исследованиями,**
- **с научными основами искусственного интеллекта.**

Моделирование когнитивной эволюции – перспективное фундаментальное направление исследований

Наиболее серьезные когнитивные процессы – процессы научного познания

Исследования когнитивной эволюции нацелены на понимание причин применимости логического мышления в научном познании, на укрепление фундамента науки

Спасибо за внимание!