

На пути к исследованию КОГНИТИВНОЙ ЭВОЛЮЦИИ

НИИ системных исследований РАН

Редько Владимир Георгиевич
vgredko@gmail.com

Основная тема доклада

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. Задача моделирования когнитивной эволюции**
- 2. Заделы исследований когнитивной эволюции: модели автономных агентов**
- 3. Биологические эксперименты по элементарному мышлению животных**
- 4. Пример модели автономных агентов: бионическая модель поискового поведения**
- 5. Начальные шаги моделирования когнитивной эволюции**
- 6. Попутные работы по автономным агентам**
- 7. Что моделировать дальше?**
- 8. Перспективы моделирования происхождения логики, мышления, интеллекта**

Задача моделирования КОГНИТИВНОЙ ЭВОЛЮЦИИ

Гносеологическая проблема: почему логический вывод, сделанный человеком, можно использовать в познании природы?

Математики доказывают теоремы. Делается это формальным логическим путем, казалось бы совсем не связанным с реальной физической природой.

Математика эффективно применяется в физике. Почему результаты, полученные формальным логическим путем, применимы к физическим объектам в природе?

Почему математика применима к физике?

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?

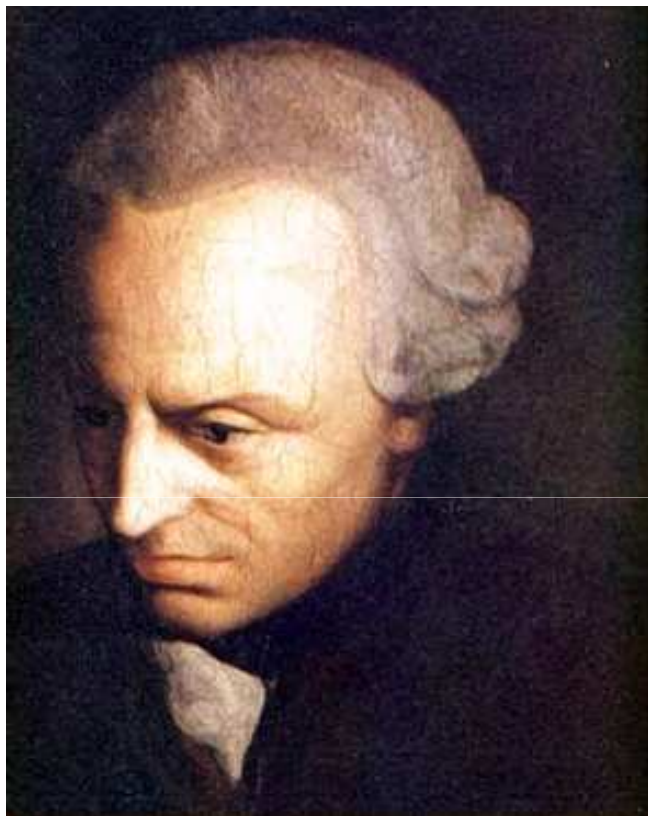
«...хотя вначале это звучит странно, но тем не менее верно, если я скажу: *рассудок не черпает свои законы (a priori) из природы, а предписывает их ей*»

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

«Не были бы законы разума, необходимые для априорного мышления, иными, если бы они сформировались иным историческим способом и если бы мы, следовательно, были оснащены иным типом нервной системы?

И вообще, возможно ли, чтобы законы нашего когнитивного аппарата не были связаны с законами реального внешнего мира?»

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



Иммануил Кант
1724-1804



Конрад Лоренц
1903-1989

Конрад Лоренц и его гусята

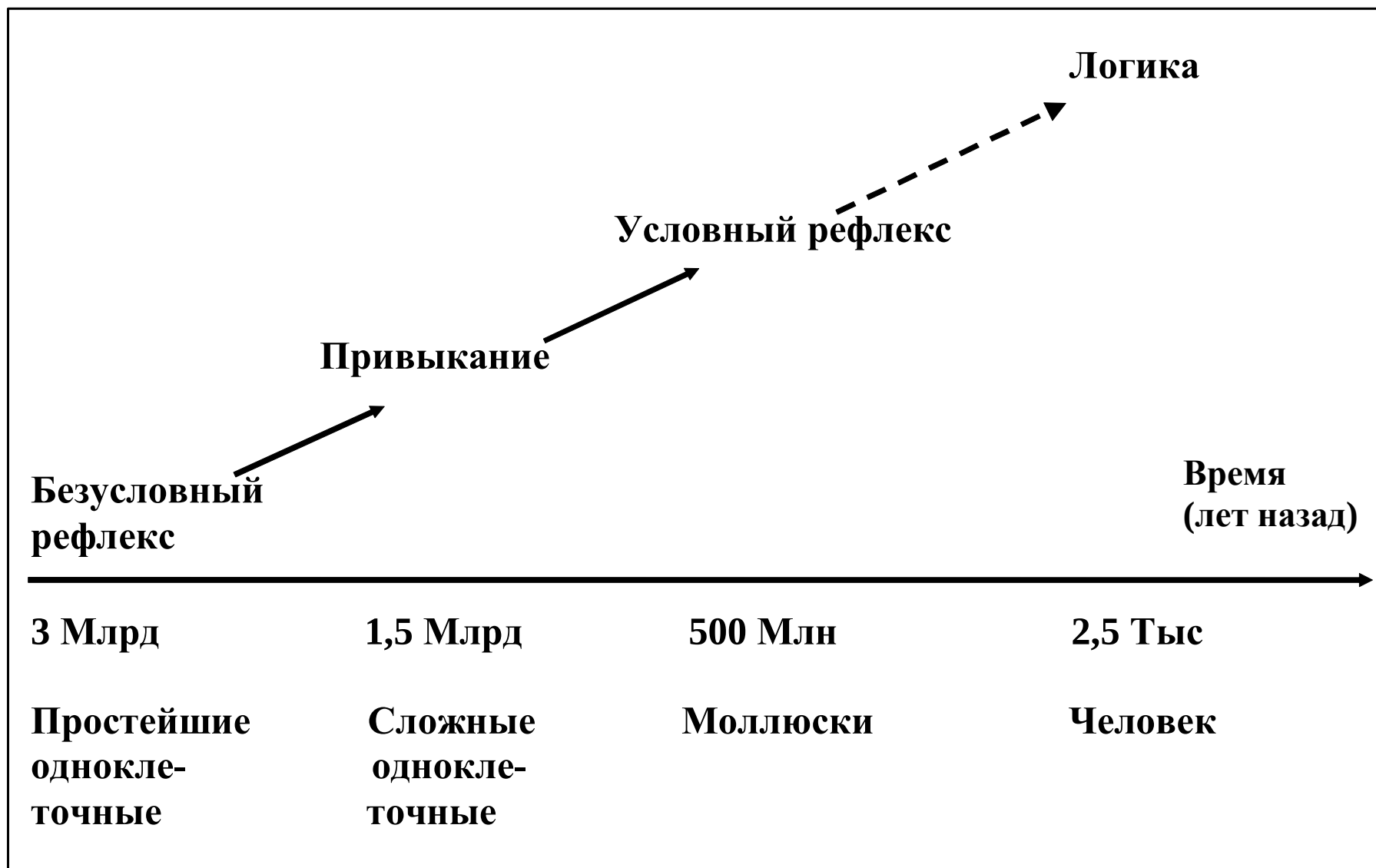


Подход к исследованию гносеологической проблемы

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

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

Собака (после выработки условного рефлекса):

$$\frac{УС, УС \rightarrow БС}{БС}$$

УС – условный стимул, БС – безусловный стимул

Правила логического вывода достаточно хорошо формализованы

Математическая теория логического вывода (под ред. А.В. Идельсона и Г.Е. Минца). М.: Наука, 1967.

**В основе этих правил – элементарные правила, такие как
*modus ponens***

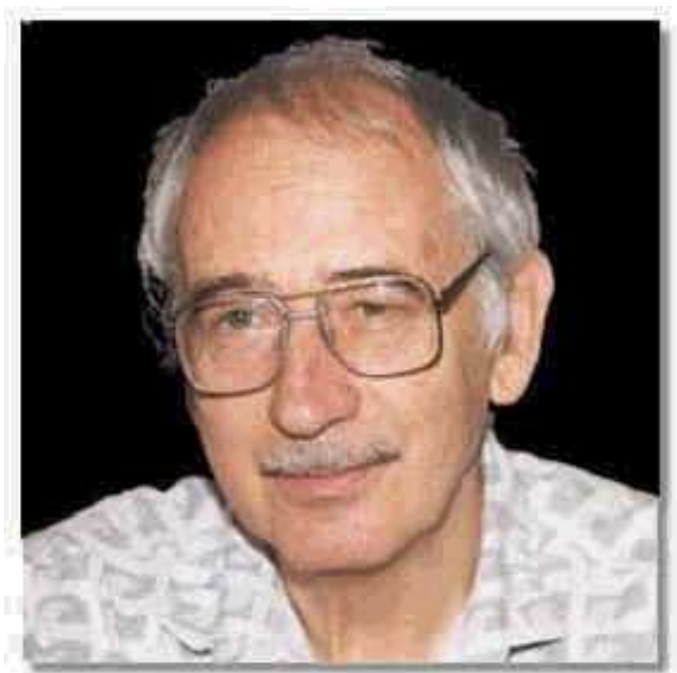
Пример исследования:

**Генцен Г. Исследования логических выводов.
Непротиворечивость чистой теории чисел (1930-е годы) //**
**Математическая теория логического вывода. М.: Наука, 1967. С. 9-76,
77-153**

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: модели
автономных агентов**

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.

Адаптивное поведение

Методы:

Нейронные сети

Эволюционное моделирование

Обучение с подкреплением (Reinforcement Learning)

Достоинство направления «Адаптивное поведение»:
исследование конкретных животных и конкретных роботов

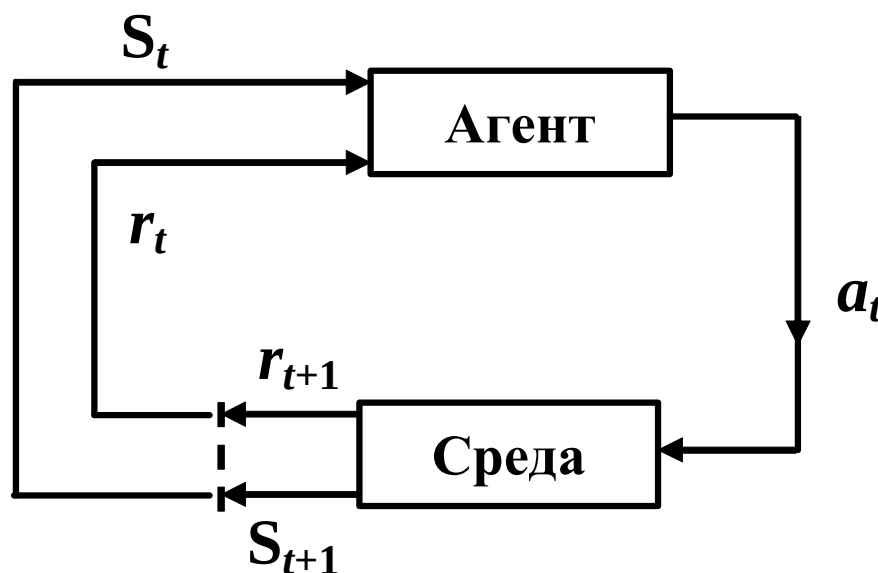
**Моделирование адаптивного поведения – бионический
подход к искусственному интеллекту**

Обучение с подкреплением

(Reinforcement Learning,

Richard Sutton, Andrew Barto, Массачусетский университет)

Схема обучения с подкреплением



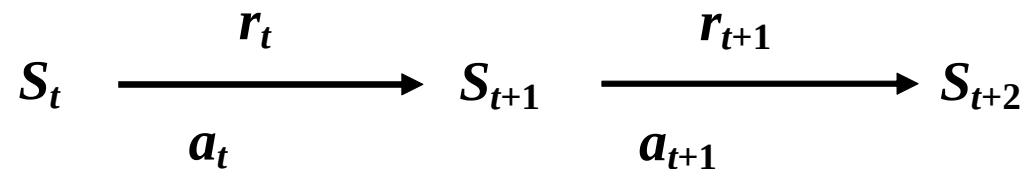
$t = 1, 2, \dots$ – время, S_t – ситуация, a_t – действие, r_t – подкрепление (поощрение, $r_t > 0$ или наказание, $r_t < 0$)

Цель агента – максимизировать суммарное подкрепление, которое можно получить в будущем, $R_t = \sum_k \gamma^k r_{t+k}$, $0 < \gamma < 1$

Метод SARSA

Итеративно формируются оценки $Q(S_t, a_t)$ суммарной величины награды, которую получит агент, если в ситуации S_t он выполнит действие a_t .

Число ситуаций и действий конечно.



Исходные $Q(S_t, a_t)$ произвольны.

Выбор действия:

=> в момент t с вероятностью $1 - \varepsilon$ выбирается действие с максимальным значением $Q(S_t, a_t)$: $a^* = \arg \max_a \{Q(S_t, a_t)\}$,

=> с вероятностью ε выбирается произвольное действие, $0 < \varepsilon \ll 1$.

Обучение, переоценка величины $Q(S_t, a_t)$:

$$\Delta Q(S_t, a_t) = \alpha [r_t + \gamma Q(S_{t+1}, a_{t+1}) - Q(S_t, a_t)] ,$$

α – параметр скорости обучения, $\alpha > 0$, $0 < \gamma < 1$, $1 - \gamma \ll 1$.

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 архитектуры и принципы функционирования познающих систем, которые можно использовать в искусственном интеллекте

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

Newel A. Unified Theory of Cognition, Cambridge, 1990
SOAR (State, Operator And Result)
<http://sitemaker.umich.edu/soar/home>

Самсонович А.В. Biologically Inspired Cognitive Architectures (BICA).

Журнал “Biologically Inspired Cognitive Architectures”

Международные конференции с 2010 года: BICA-2013 (Киев).

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. P. 151-180.

Заделы исследований когнитивной эволюции

Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова // Нейроинформатика (электронный 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 журнал). 2008. Т. 3. № 1. С. 79-108.

<http://www.niisi.ru/iont/ni/Journal/V3/N1/DeminVityaev.pdf>

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**)

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

Заделы исследований когнитивной эволюции

Witkowski M. An action-selection calculus // Adaptive Behavior, 2007. V. 15. No. 1. P. 73-97.

Butz M.V., Sigaud O., Pezzulo G., Baldassarre G. (Eds.).
Anticipatory Behavior in Adaptive Learning Systems: From
Brains to Individual and Social Behavior. LNAI 4520, Berlin,
Heidelberg: Springer Verlag, 2007.

Чернавский Д.С., Чернавская О.Д., Карп В.П., Никитин А.П.
Подходы к моделированию процесса мышления с позиций
динамической теории информации

Агент-ориентированные экономические модели (ЦЭМИ РАН)

**А.Р. Бахтизин. Агент-ориентированные модели экономики.
М.: Экономика, 2008**



**В.Л. Макаров. Социальный кластеризм.
Российский вызов. М.: Бизнес Атлас, 2010**

**В.Л. Макаров., А.Р. Бахтизин. Социальное
моделирование – новый компьютерный
прорыв (агент-ориентированные модели).
Москва: Экономика, 2013**

Лаборатория искусственных обществ:

<http://www.artsoc.ru/>

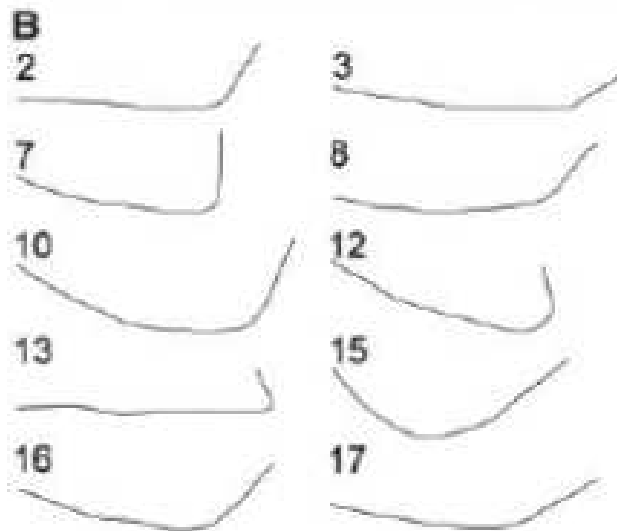
Интернет-журнал «Искусственные общества»

**Биологические эксперименты
по элементарному мышлению
ЖИВОТНЫХ**

Новокаледонские вороны могут **изобретать** **способ** изготовления орудий труда



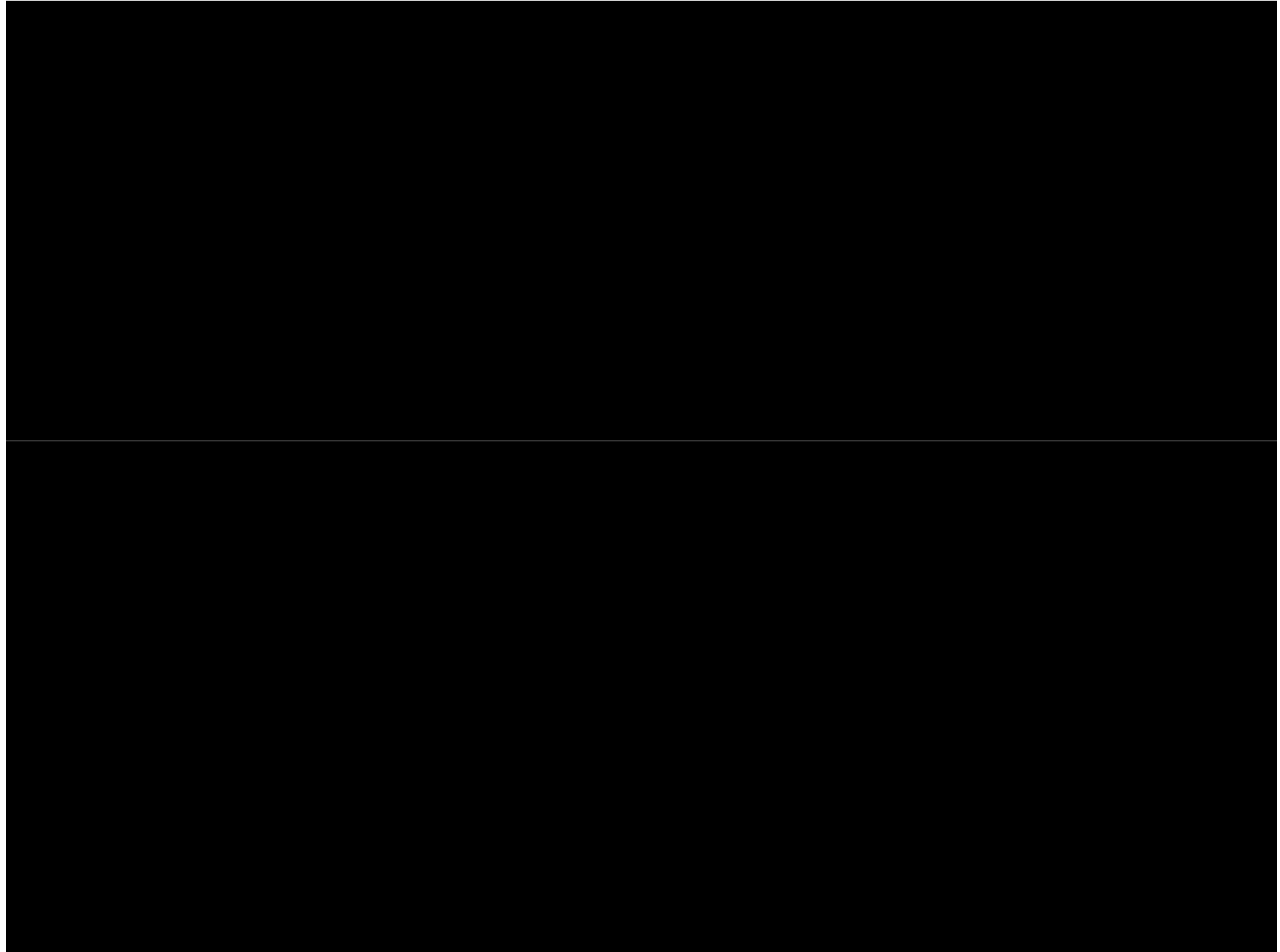
Двум воронам (молодой самке и самцу постарше) предлагали добывать ведро с пищей прямой проволокой и проволокой, согнутой крючком. Вороны сразу поняли, что ведро можно вытащить с помощью крючка.



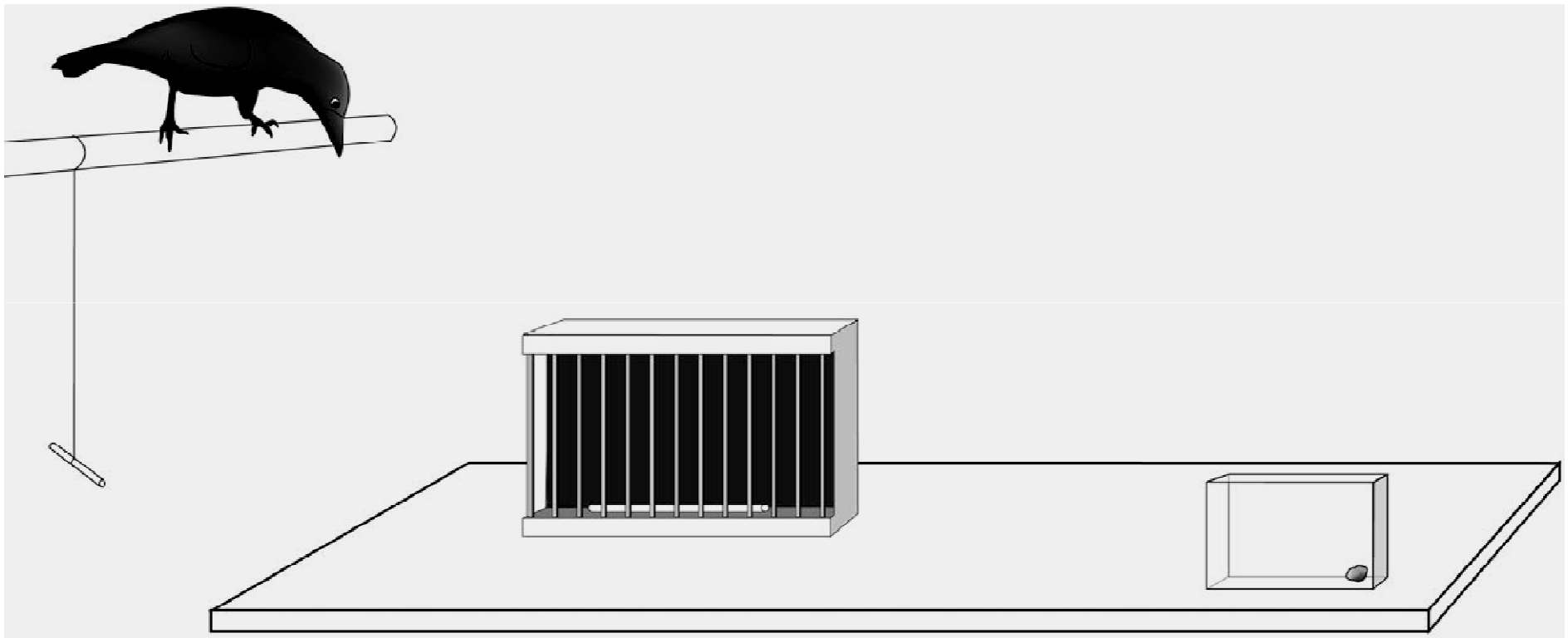
Однажды самец утащил крючок. Тогда самка научилась делать из прямой проволоки крючок, зажимая один конец проволоки в щели и загибая проволоку.

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

Новокаледонские вороны могут изобретать способ изготовления орудий труда



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

	trial									
crow	1	2	3	4	5	6	7	8	9	10
Sam (I)	green	green	green	orange	green	green	green	green	green	green
Caspar (I)	lightgreen	green	green	green	green	lightgreen	green	green	green	lightgreen
Maya (I)	blue	blue	orange	green	green	green	green	green	green	green
Djinn (I)	lightblue	blue	blue	green	green	green	green	green	green	orange
Lazlo (C)	green	green	green	green	green	green	green	green	green	green
Chocho (C)	green	green	lightgreen	lightgreen	green	green	green	green	green	green
Korben (C)	green	lightgreen	green	green	green	lightgreen	green	green	green	green

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

Пример модели автономных агентов

Бионическая модель поискового поведения

Непомнящих В.А., Попов Е.Е., Редько В.Г. Бионическая модель адаптивного поискового поведения // Известия РАН. Теория и системы управления. 2008. № 1. С.85-93.

Ручейник и его чехол-домик



Личинки ручейников ведут поиск крупных частиц для строительства чехла-домика

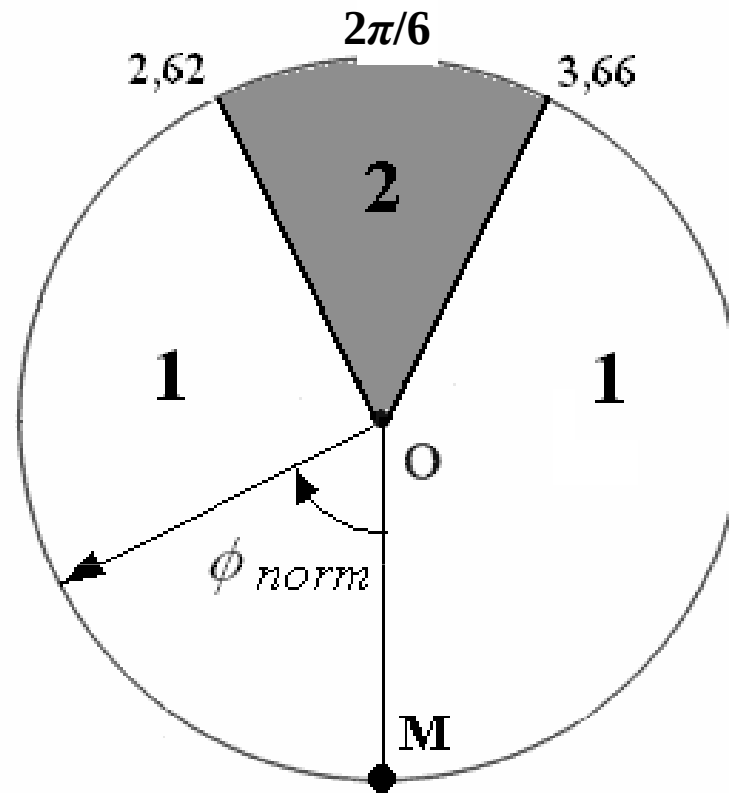
Биологический эксперимент

(В.А. Непомнящих, 2002)



Личинок ручейников помещали в кольцевой коридор с водой, дно которого было покрыто сплошным слоем песка. На небольшом участке коридора ($1/6$ площади коридора) были еще крупные плоские частицы (скорлупки). Эксперимент показал, что личинки ручейников находят участок с крупными частицами, где они строят домик в основном из скорлупок.

Геометрия кольцевого коридора



Личинка помещалась в точку М.

Крупные частицы (скорлупки) расположены на участке 2.

Вероятность встречи крупной частицы на участке 2 равна 0.2.

На участке 1 расположены только мелкие частицы (песчинки)

Результаты биологического эксперимента

- Эксперимент проводился с 40 личинками, которые наблюдались в течение 1 часа с момента первого прикрепления частицы
- Все личинки двигались по коридору и посещали участок 2 со скорлупками
- Личинки преимущественно (36 из 40) вели строительство на участке 2, где они собирали домик из крупных частиц – скорлупок
- Среднее число прикрепленных частиц для 36 личинок, покидавших и возвращавшихся на участок 2, составило 5 скорлупок и 2 песчинки. 4 личинки, оставшиеся на участке 1, в среднем прикрепили по 7 песчинок

Компьютерная модель

- Поведение регулируется мотивацией к прикреплению $M(t)$
- Возможны три действия личинки:
 - 1) **прикрепление** протестированной частицы к домику,
 - 2) **тестирование** частицы,
 - 3) **блуждание**, поиск нового места
- **Прикрепление** происходит при превышении мотивацией $M(t)$ порога, пропорционального площади последней прикрепленной частицы, при $M(t) > Th = k_0 S_{attach}$, $k_0 > 0$
- **Тестирование** происходит при $Th > M(t) > 0$
- **Блуждание** – при $0 > M(t)$

Динамика мотивации к прикреплению $M(t)$

$$M(t) = k_1 M(t-1) + \xi(t) + I(t) ,$$

время t дискретно, шаг по времени $\Delta t = 1$ с, k_1 – параметр, характеризующий инерционность ($0 < k_1 < 1$, $1-k_1 \ll 1$)

$\xi(t)$ – нормально распределенная случайная величина со средним 0 и средним квадратическим σ

$I(t)$ – интенсивность раздражителя

При *тестировании*:

$$I(t) = k_2 (S_{curr} - S_{last}) / S_{last} ,$$

$k_2 > 0$, S_{curr} , S_{last} – площади тестируемой и последней протестированной частицы

При *перемещении и прикреплении* $I(t) = 0$

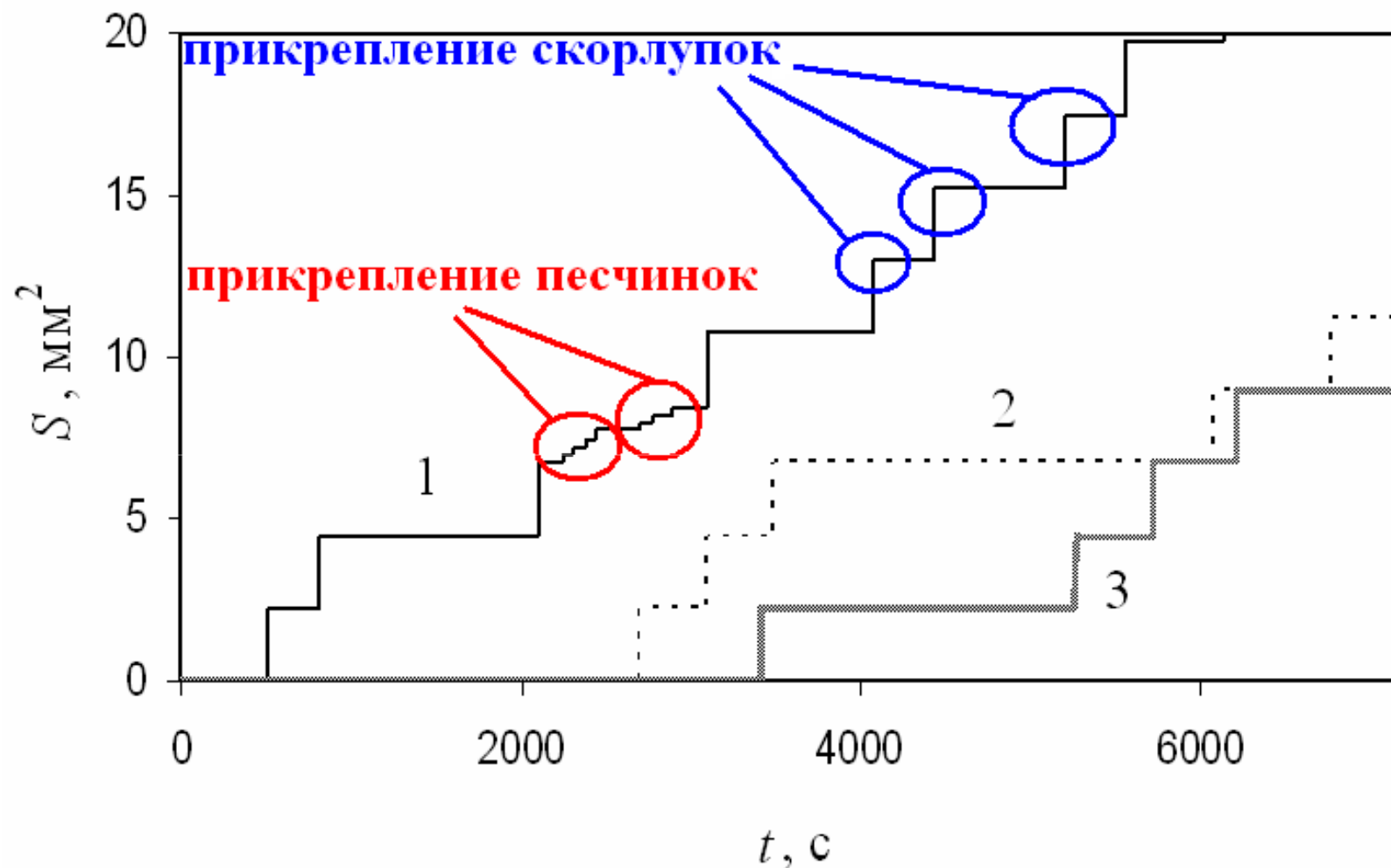
Параметры расчета

- Диаметр коридора $d = 90$ мм
- Размер песчинки = 0.5 мм ($S = 0.25$ мм²)
- Размер скорлупки = 1.5 мм ($S = 2.25$ мм²)
- Величина перемещения за один такт времени $L = 2$ мм
- $k_0 = 1$, $k_1 = 0.99$, $k_2 = 0.007$, $\sigma = 0.05$
- Время тестирования / прикрепления = 5/60 с (для песчинки), 10/120 с (для скорлупки)
- Расчет проводился в течение 7200 с (2 часа) для 40 личинок, аналогично биологическому эксперименту
- Исходная мотивация к прикреплению $M(0) = 0$
- Сначала личинка помещалась в центр участка 1

Результаты моделирования

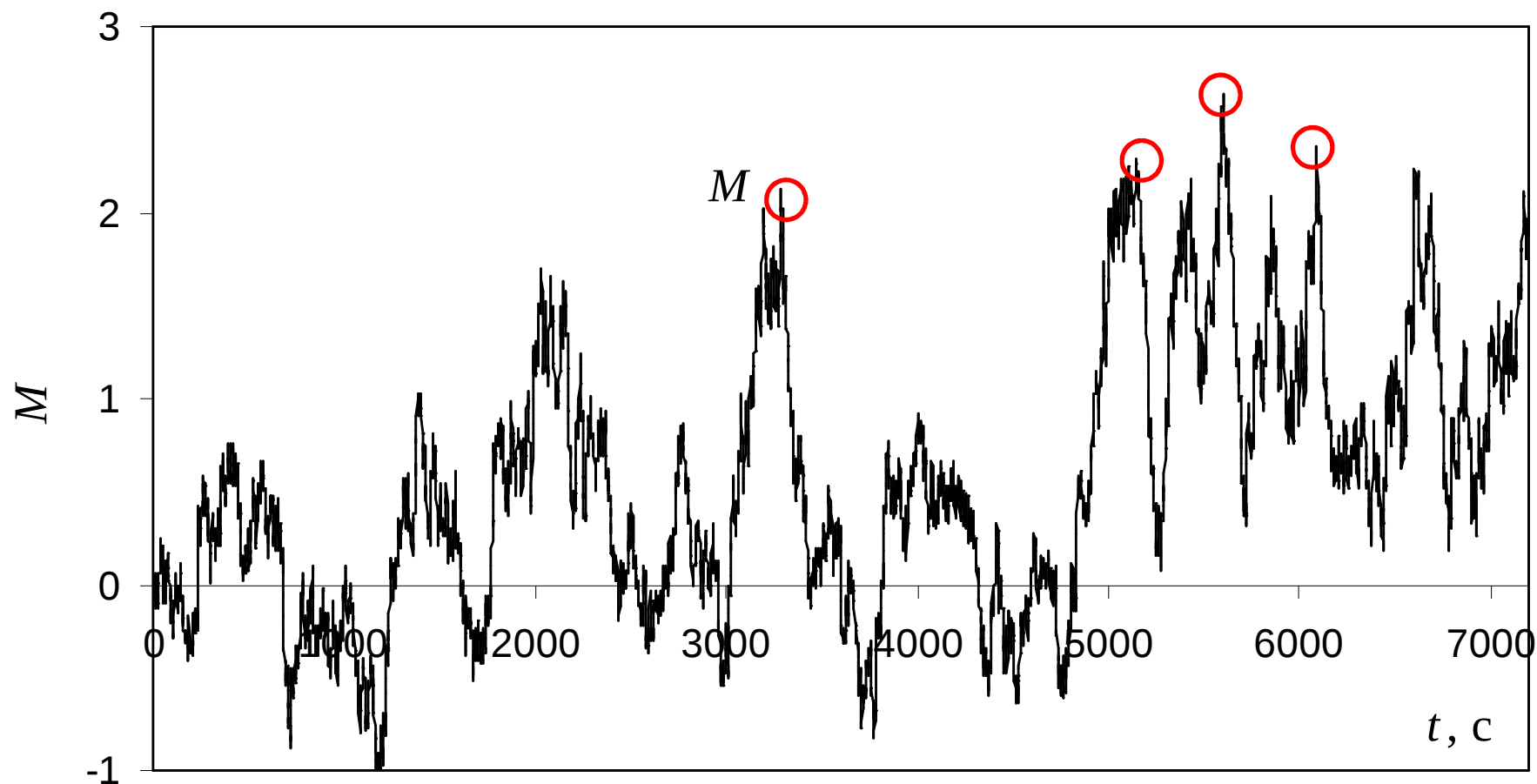
- Почти во всех случаях (в 39 из 40) первой прикреплялась крупная частица (скорлупка)
- Количество частиц в течение часа после прикрепления первой частицы в среднем по 40 расчетам составило: 4.2 крупные частицы (среднее квадратическое отклонение 1.68) и 0.6 мелких частиц (среднее квадратическое отклонение 1.53)
- Прикрепление мелких частиц наблюдалось только в 8 расчетах из 40
- Среднее время начала прикрепления первой частицы 1815 с (среднее квадратическое отклонение 872 с)

Динамика площади домика $S(t)$



Преимущественно прикрепляются крупные частицы. Есть сильный разброс числа прикрепляемых частиц и момента начала прикрепления. Число прикреплений невелико.

Динамика мотивации к прикреплению $M(t)$



Прикрепление начинается в моменты $t = 3288, 5140, 5595, 6090$ с.

Есть эффект **частичного успеха**: есть рост мотивации $M(t)$, но мотивация не достигает порога и прикрепление не начинается

Модель и биологический эксперимент качественно согласуются (данные по 40 примерам)

- 1. Личинки (как живые, так и модельные) обследуют весь коридор: участок 1 и участок 2**
- 2. Преимущественно прикрепляются крупные частицы: 4.2 в модели, 5 в эксперименте. Малое число прикрепленных песчинок: 0.6 в модели, 2 в эксперименте**
- 3. Есть сильный разброс числа прикрепляемых частиц и момента начала прикрепления (как в модели, так и в эксперименте)**
- 4. Число прикрепленных частиц в обоих случаях невелико**

Существенные особенности модели

В модели есть **память** личинки о размерах последней протестированной и последней прикрепленной частиц

Динамика регулирующей поведение **мотивации $M(t)$** проста и эффективна. Она включает:

- 1) **инерцию** изменения $M(t)$
- 2) случайные вариации $M(t)$
- 3) направленное изменение $M(t)$

План доклада

- 1. Задача моделирования когнитивной эволюции**
- 2. Заделы исследований когнитивной эволюции: модели автономных агентов**
- 3. Биологические эксперименты по элементарному мышлению животных**
- 4. Пример модели автономных агентов: бионическая модель поискового поведения**
- 5. Начальные шаги моделирования когнитивной эволюции**
- 6. Попутные работы по автономным агентам**
- 7. Что моделировать дальше?**
- 8. Перспективы моделирования происхождения логики, мышления, интеллекта**

Контурь 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**
 $\{УС, УС \rightarrow БС\} \Rightarrow БС$ – 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
(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)**

Описание модели

Агенты имеют потребности питания, размножения, безопасности. Каждой потребности соответствовала своя мотивация. Имелись факторы потребностей F_F , F_S и F_R и пороги этих факторов T_F , T_S , T_R . При превышении фактором порога потребность удовлетворялась.

Иерархия мотиваций: 1) питание, 2) безопасность 3) размножение.

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

Ситуация S определялась наличием активного хищника, ведущей мотивацией, предыдущим действием агента. Ведущая мотивация выбиралась из неудовлетворенных в соответствии с иерархией.

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

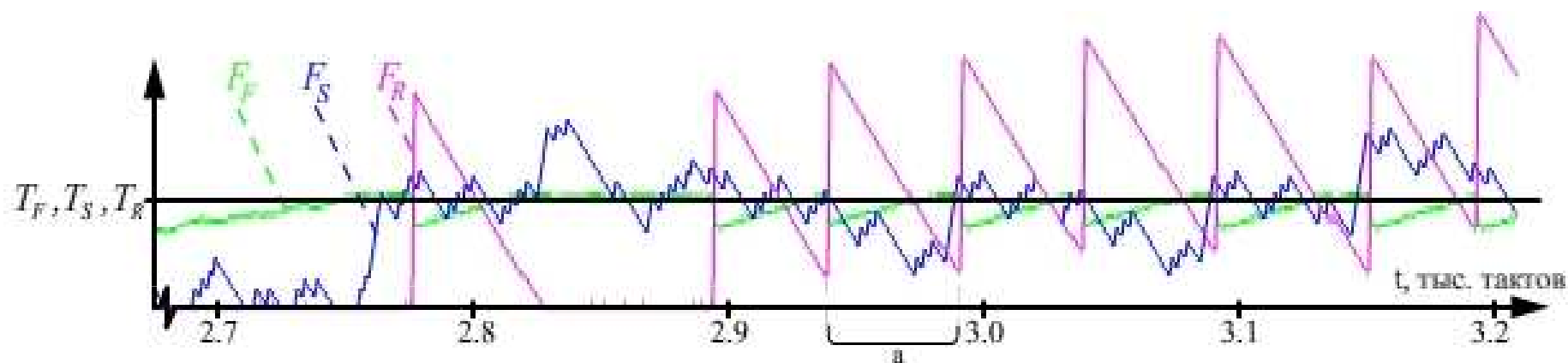
Факторы обычно медленно уменьшались, но возрастали при выполнении полезных действий, соответствующих ведущей мотивации.

Веса правил формировались методом обучения с подкреплением.

Вес правила увеличивался при увеличении фактора ведущей мотивации. Действия выбирались в соответствии с весами правил.

Результаты моделирования

Динамика факторов F_F , F_S и F_R , соответствующих потребностям агента; T_F , T_S , T_R – пороги факторов



Динамика ведущих мотиваций



Выводы по модели с несколькими потребностями и мотивациями

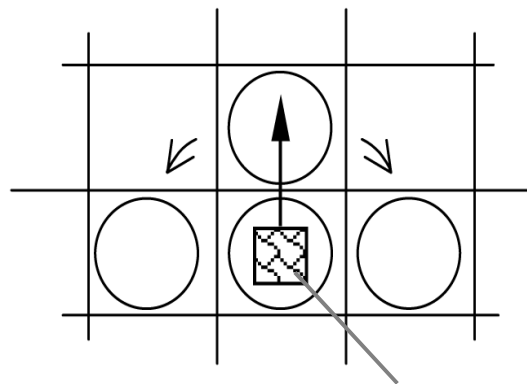
Продемонстрировано формирование циклов поведения, в которых последовательно удовлетворяются потребности питания, безопасности и размножения.

Необходим анализ вариантов модели и сопоставление с биологическими экспериментами.

Формирование внутренних понятий

Описание модели

Автономный агент ведет поиск пищи в двумерной клеточной среде.



Агент

Стрелка показывает направление вперед, кружки – поле зрения агента

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 самостоятельно обучающегося 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 (соответствующие ситуациям): *«имеется пища в моей клетке»*, *«имеется пища в клетке впереди меня»*, *«имеется пища в клетке справа/слева от меня»*.

Попутные работы по автономным агентам

Модель взаимодействия между обучением и эволюцией

Исследован эффект генетической ассимиляции: индивидуально приобретаемые навыки организмов могут стать наследуемыми.

Эффект работает в два этапа. На 1-м этапе эволюционирующие организмы (благодаря мутациям) приобретают свойство обучиться некоторому полезному навыку. Приспособленность таких организмов увеличивается, следовательно, они распространяются по популяции. Но обучение имеет свои недостатки, так как оно требует энергии и времени. Поэтому возможен 2-й этап: приобретенный полезный навык в течение многих поколений может быть «переизобретен» генетической эволюцией, в результате чего он записывается непосредственно в геном и становится наследуемым (организмы имеют его с рождения). **Полезный навык, который первоначально был приобретаемым, становится наследуемым, хотя эволюция имеет Дарвиновский характер.**

Редько В.Г. Модель взаимодействия между обучением и эволюционной оптимизацией // Математическая биология и биоинформатика (электронный журнал), 2012. Т.7. № 2. С. 676-691.

URL: http://www.matbio.org/2012/Redko_7_676.pdf

Модель честной рыночной экономики

Построена многоагентная модель прозрачной рыночной экономической системы, состоящей из **сообщества инвесторов и производителей, которые налаживают взаимодействие внутри сообщества с помощью легких агентов-посланников.**

Экономика честная: динамика капиталов открыта всему сообществу. Намерения инвесторов вложить тот или иной капитал в конкретных производителей также открыты.

Экономика рыночная – имеется конкуренция: слабые инвесторы и слабые производители могут вымирать.

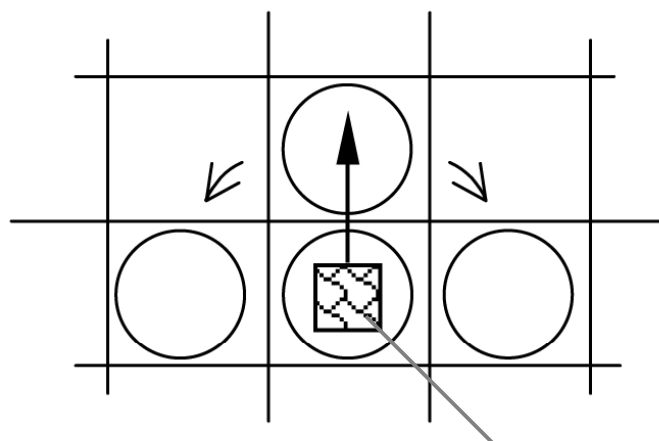
Нет откатов, нет коррупции

Редько В.Г., Сохова З.Б. Многоагентная модель прозрачной рыночной экономической системы // Труды НИИСИ РАН, 2013. Т. 3. № 2. С. 61-65.

Идея проекта на Нобелевскую премию мира

Модель взаимодействующих агентов

(М.С. Бурцев, 2002)



Агент

Стрелка показывает направление вперед, кружки – поле зрения анимата

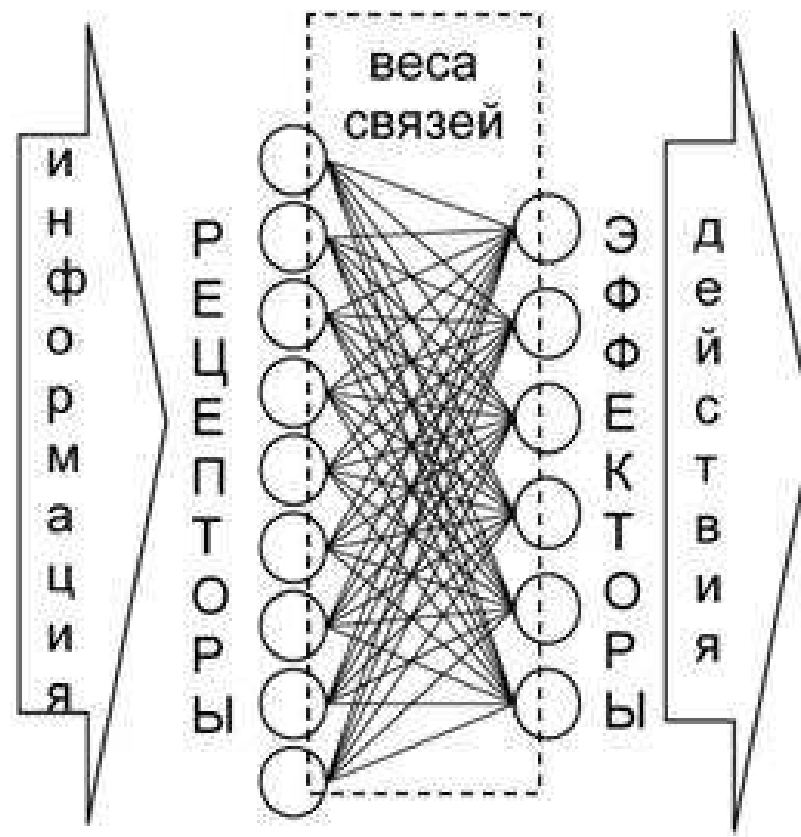
Действия агента: двигаться вперед, поворачиваться направо или налево, питаться, размножаться, **бороться с другими агентами**

Система управления агента – однослойная нейронная сеть

При питании агент увеличивает свой ресурс, при действиях (особенно при борьбе) – расходует. При борьбе агенты отнимают ресурс друг у друга.

Популяция агентов эволюционирует, есть отбор и мутации

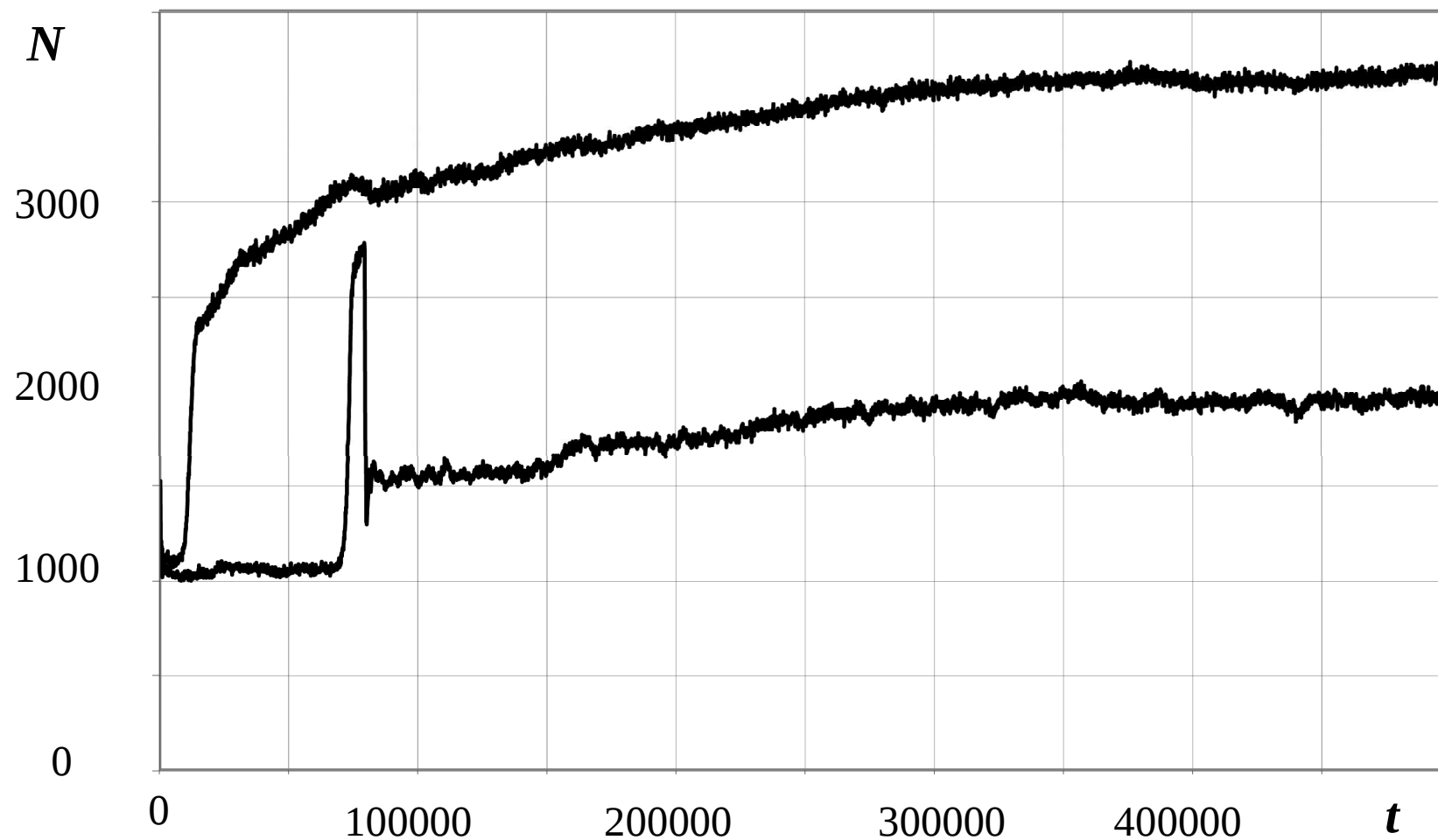
Структура нейронной сети агента



При мутациях меняются веса связей (веса синапсов нейронов) и **может происходить удаление или добавление сенсоров или эффекторов.**

При отборе выживают агенты, имеющие наибольший ресурс

Зависимость численности популяции N от времени t



Внизу – полная модель, вверху – эффекторы борьбы удалены

Идея проекта на Нобелевскую премию мира

Если рассмотреть эволюцию агентов как эволюцию государств, то можно предложить идею проекта на Нобелевскую премию мира «Разработка научных основ всемирного разоружения».

Подробнее:

Редько В.Г. Будущее России, будущее человечества // Сложность. Разум. Постнеклассика, 2013. № 3. С. 55-64.

См. также:

<http://spkurdyumov.ru/future/budushhee-rossii-budushhee-chelovechestva/>

Что моделировать дальше?

Биологические корни творческого поиска

Непомнящих В.А. Адаптация и автономия в поведении животных // XV Всероссийская научно-техническая конференция «Нейроинформатика-2013». Лекции по нейроинформатике. М.: НИЯУ МИФИ, 2013. С. 106-123.

В поведении животных постоянно присутствует цель, независимая от частных задач: достижение определенного баланса между поиском новой, непредсказуемой стимуляции и стремлением к предсказуемости результатов своего поведения

Биологические корни творческого поиска

Пример поисковой активности

Скворцам предлагали миски с одинаковым количеством корма, смешанного с песком. Одна из мисок была прикрыта прозрачной крышкой, а другая – непрозрачной, скворцы могли легко пробить крышки клювом и достать корм.

Птицы предпочитали добывать корм в миске с непрозрачной крышкой, несмотря на то, что им приходилось искать там корм вслепую. Это предпочтение имело место, даже если корма в миске с непрозрачной крышкой было меньше.

Зачастую поисковая активность проявлялась в ущерб удовлетворению важной потребности

Предпосылки творческой активности

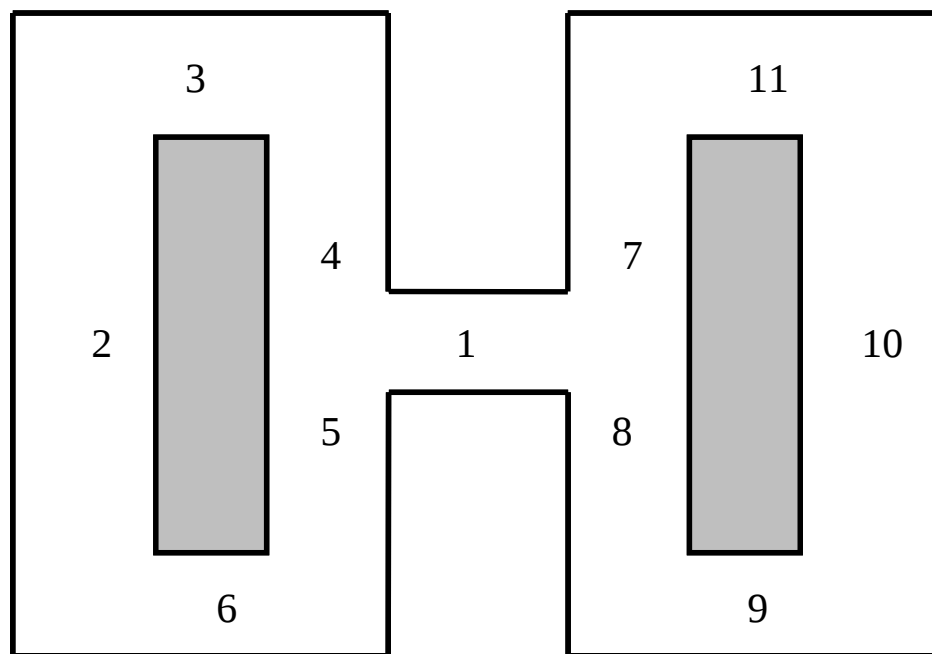
Мы с В.А. Непомнящих начали строить модель поисковой активности рыбок данио рерио, изучающих простые лабиринты и строящих когнитивные карты лабиринтов

Дальняя цель — связать предпосылки с творческой активностью ученого

Непомнящих В.А., Осипова Е.А., Редько В.Г., Шарипова Т.И., Бесхлебнова Г.А. Модель навигации животных в лабиринтах // XVI Всероссийская научно-техническая конференция «Нейроинформатика-2014». Ч. I. М.: НИЯУ МИФИ, 2014. С. 110-117.

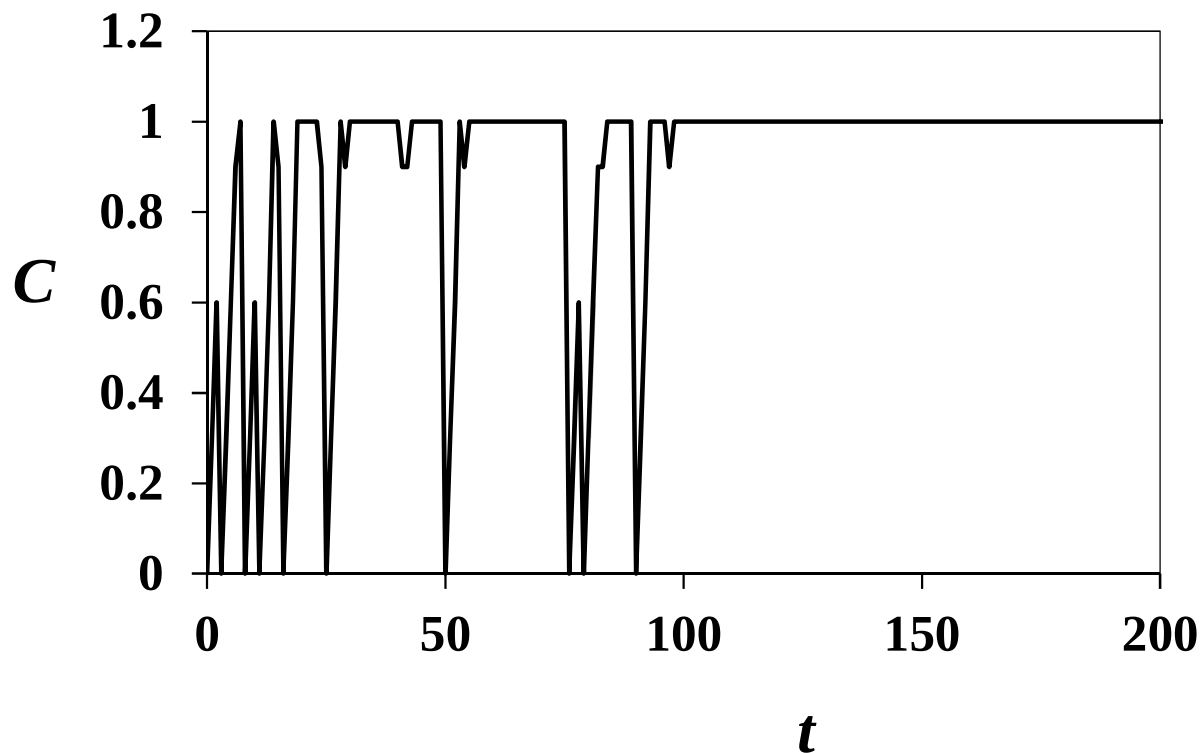
Предпосылки творческой активности

Экспериментальный лабиринт



Предпосылки творческой активности

Начальные результаты моделирования: динамика знания рыбы C о коридоре, в котором она находится



Проблема моделирования интеллекта человеческого уровня

На конференции BICA-2011 была предложена идея робота, подобного человеку (внешне и по интеллекту), которого население страны могло бы выбрать своим президентом

Chella A., Lebiere C., Noelle D.C., Samsonovich A.V. On a roadmap to biologically in-spired cognitive agents // Proceedings of BICA-2011. P. 453-460

Возможна ли модель ученого-физика?

М.Н. Вайнцвайг. Обучение управлению поведением в мире объектов пространства-времени // Лекция на конференции «Нейроинформатика-2011»

Предложены принципы компьютерного моделирования агента, познающего закономерности механики движущихся тел. И начато собственно моделирование (совместно с В.В. и П.В. Максимовыми, ИППИ РАН)

Возникает вопрос: может ли такой агент самостоятельно открыть три закона Ньютона?

Возможна ли модель ученого-физика?

Red'ko V.G. Principles of functioning of autonomous agent-physicist // Proceedings of BICA 2012. P. 265-266.

Компьютерный агент, нацеленный на открытие законов механики, должен обладать следующими свойствами (аналогично Исааку Ньютону):

- иметь стремление к новым знаниям и к формулировке этих знаний в четкой, компактной форме;**
- задавать вопросы о принципах взаимодействия тел и решать эти вопросы с помощью физических экспериментов;**
- обладать самосознанием и стремиться к достижению наивысшего уровня среди подобных агентов.**

**Перспективы моделирования
происхождения логики, мышления,
интеллекта**

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**
 $\{УС, УС \rightarrow БС\} \Rightarrow БС$ – 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 очень маленькие 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.**

**Видео доклада по близкой теме на
Поспеловских чтениях:**

<http://bicasociety.org/resources/russian/>

Спасибо за внимание!