



Задачный подход в когнитивных науках. Искусственный Интеллект 5.0 и Сознание.

д.ф.-м.н. Витяев Евгений Евгеньевич

Институт математики им. С.Л.Соболева

Продолжение

Е.Е.Витяев, С.С.Гончаров, Д.И.Свириденко

О ЗАДАЧНОМ ПОДХОДЕ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ И КОГНИТИВНЫХ НАУКАХ

Искусственный Интеллект 5.0 и Сознание

На пути к фундаментальной теории работы мозга. Сознание.

Лекция Президента МАКИ К. В. Анохин



Прямое соотнесение – теоретический тупик

Смотреть позже Поделиться

ПРОМЕЖУТОЧНОЕ ЗАКЛЮЧЕНИЕ:

Проблема заключается вовсе не в том, что существующие сегодня нейрофизиологические теории несовершенны и требуют дальнейшего развития. Используемые в них коррелятивные подходы просто не могут дать удовлетворительного ответа на вопросы о природе разума и субъективного опыта.

Если разум и сознание нельзя объяснить редукционно, то для решения этих вопросов требуется *нередукционная* теория. Такая теория должна описывать субъективный опыт как фундаментальный феномен, появляющийся лишь у определенных уровней организации материи.

Фундаментальные феномены не могут быть объяснены в терминах чего-то более простого. Там, где есть фундаментальные феномены всегда требуется фундаментальная теория.

Сегодня нам недостает именно такой *фундаментальной теории*.

Проблема не в том, что существующие нейрофизиологические теории несовершенны ... Используемые в них коррелятивные подходы просто не могут дать ответа на вопросы о природе разума и субъективного опыта ... Для этого нужна **нередукционная фундаментальная теория**.



Принципы построения фундаментальной теории работы мозга

«Сложилась одна универсальная закономерность в приспособлении организмов к внешним условиям, которая в дальнейшем бурно развивалась на протяжении всей эволюции живого мира: в высшей степени **быстрое отражение медленно разворачивающихся событий внешнего мира**» П.К. Анохин.

I. Это **информационная теория** предвосхищения событий внешнего мира. Нейрофизиологические, физиологические и психологические теории должны описывать, как они обеспечивают те или иные информационные процессы отражения и предвосхищения реальности.

Если бы внешний мир был случаен, то никакое его предвосхищение было бы невозможно. Но наш мир хорошо структурирован в нем есть как минимум причинность.

Для эффективного предвосхищения реальности мозг должен был эволюционно сформировать такие нейрофизиологические механизмы, которые бы автоматически улавливали и использовали эту структуру внешнего мира для прогноза событий.

II. Информационная теория отражения должна основываться на **законах структуры внешнего мира** и описывать одновременно как структуру реальности, так и нейрофизиологические и другие механизмы, которые обеспечивают отражение этой структуры.

Причинность, прогноз и их нейрофизиологические механизмы

В философии науки причинность сводится к предсказанию и объяснению:

«причинное отношение означает предсказуемость ... в том смысле, что, если полная предыдущая ситуация известна, событие может быть предсказано ..., если будут даны все относящиеся к событию факты и законы природы».

Таким образом, причинное отношение сводится к выводу предсказаний из имеющихся фактов и законов.

Существуют две модели предсказания: *Дедуктивно-номологическая и индуктивно-статистическая*, соответственно для детерминированных и статистических законов.

Причинное объяснение основано на индуктивных законах, полученных в результате обучения и опыта. Но для таких законов возникает проблема статистической двусмысленности.

L_1, \dots, L_m	[r]
C_1, \dots, C_n	
G	

Пример: «Если философ – то не миллионер», «если держатель приисков – миллионер».

Для известного философа, который имеет прииски получим противоречие.

Гемпель предложил *требование максимальной специфичности* для статистических законов: правила должны содержать максимум наличной информации.

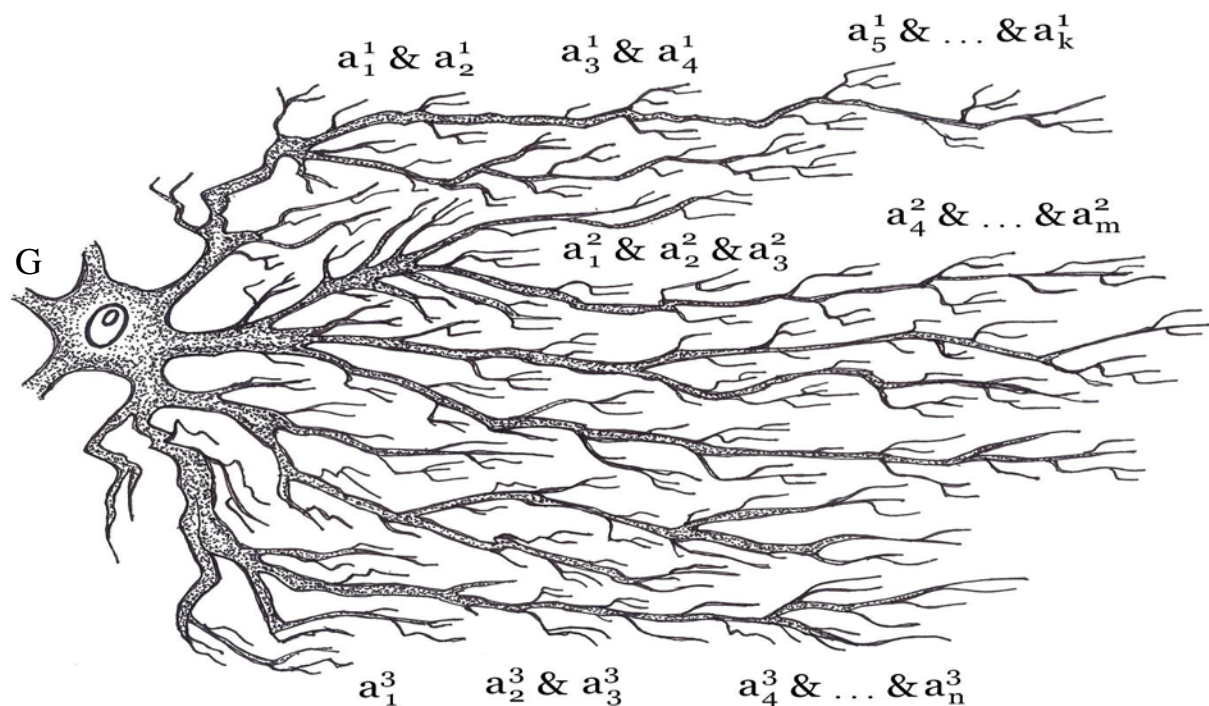
«Если философ не держатель приисков, то ещё более вероятно, что он не миллионер»

«Если держатель приисков не философ, то ещё более вероятно, что он миллионер»

Решение проблемы статистической двусмысленности.

Vityaev, E., Odintsov, S. How to predict consistently? // Trends in Mathematics and Computational Intelligence In: Studies in Computational Intelligence, 796, María Eugenia Cornejo (ed), 2019, 35-41. (обнаружение *максимально специфических причинных связей*)

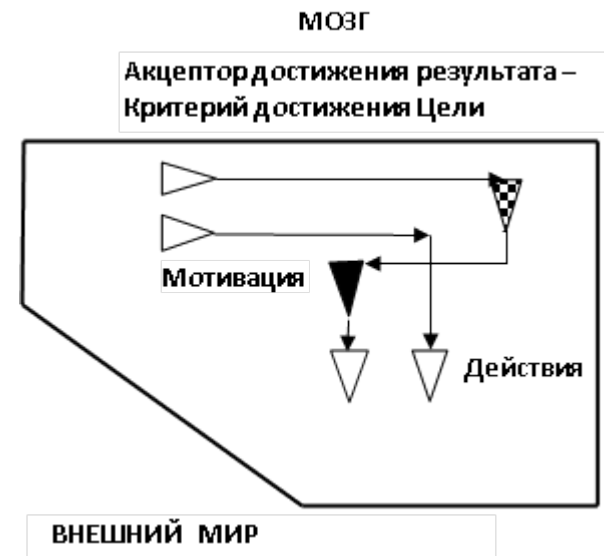
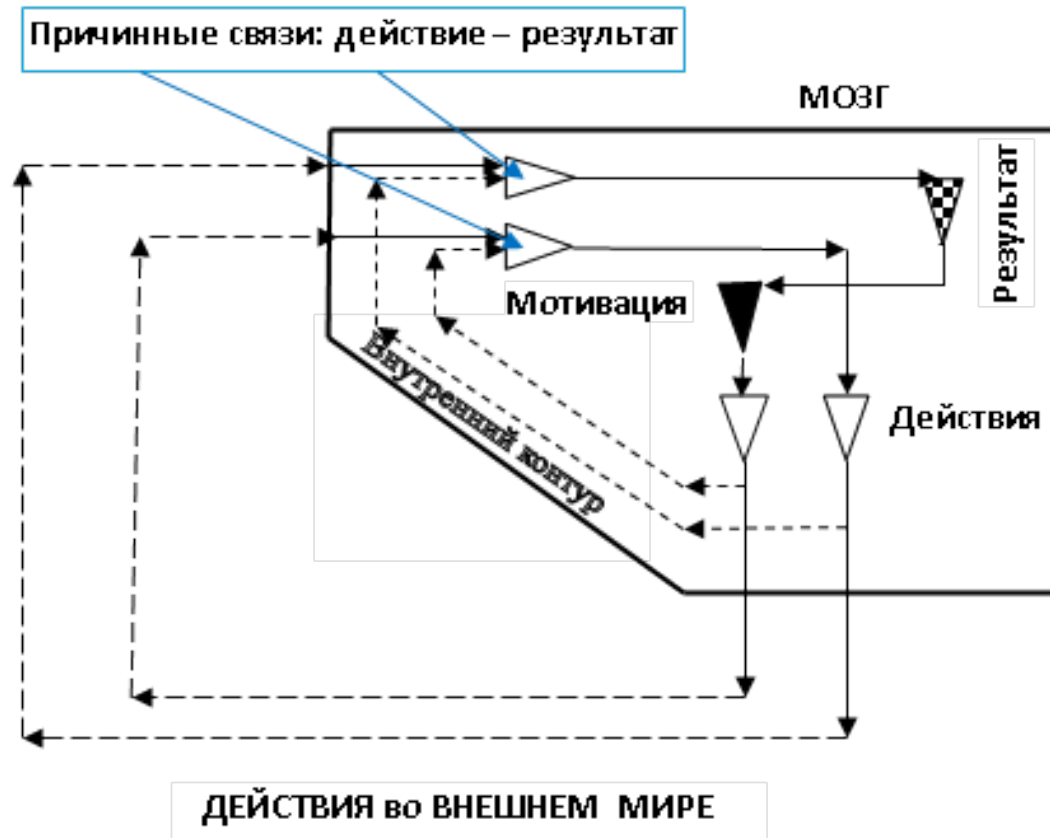
Формальная модель нейрона, обнаруживающая максимально специфические причинные связи и удовлетворяющая правилу Хебба.



Vityaev E.E. A formal model of neuron that provides consistent predictions // *Biologically Inspired Cognitive Architectures 2012. Proceedings of the Third Annual Meeting of the BICA Society. In Advances in Intelligent Systems and Computing*, v.196, Springer. 2013, pp. 339-344.

1. В процессе эволюции нейроны могли прийти именно к этой модели (естественный ИИ).
2. Если в процессе эволюции нейроны не пришли к этой модели, то возможен сверх сильный ИИ.
3. Принципиально точнее нейронных сетей, но и вычислительно сложнее.
4. Разработана система Discovery, обнаруживающая такие правила, успешно применявшаяся для решения большого числа задач.

Причинность и Теория Функциональных Систем работы мозга



Цель нельзя достичь не имея критерия ее достижения, иначе всегда можно считать, что цель уже достигнута. Действие также всегда целенаправленно, иначе непонятно, чем закончить действие. Если нет цели действия, то всегда можно считать, что цель уже достигнута и делать ничего не надо.



С точки зрения принципов построения фундаментальной теории :

III. Активные действия в среде и целенаправленная деятельность всегда должны сопровождаться критерием достижения цели и без этих критериев она смысла не имеет.

В ИИ и в науке понятия цели формулируется как задача.

Задача осмыслена тогда и только тогда, когда у нас есть **критерий решенности задачи**.

ЗАДАЧА: Семантическая/информационная МОДЕЛЬ предметной области.

Исходные данные + ЗАПРОС (ЦЕЛЬ).

КРИТЕРИЙ решения задачи (достижения цели).

Понятие задачи в основаниях математики

В математических теориях критерием решения задачи считается наличие доказательства. Но этот критерий применим только если в рамках самой формальной системы мы имеем, как доказательство решения задачи, так и возможность убедиться средствами самой же формальной системы, что данное доказательство действительно является решением задачи. Было доказано, что только в «слабых» формальных системах (для которых не проходит теорема Гёделя) мы можем средствами самой формальной системы определить, является ли некоторый текст доказательством решения задачи или нет.

Акад. РАН Ю.Л. Ершов, д.фил.-н. К.Ф. Самохвалов. Современная философия математики, Новосибирск, 2007.

В результате программа Гильберта обоснования математики формулируется иначе: не нужно для всей математики доказывать её непротиворечивость – это ненужно. Нужно формулировать и решать задачи в рамках слабых формальных систем.

Понятие задачи в Искусственном Интеллекте

Стюарт Рассел, Питер Норвиг. Искусственный интеллект: современный подход, Киев, 2006, Издательский дом “Вильямс”, 1409 стр.

Агентный подход к ИИ должен быть переписан с точки зрения задачного подхода, как классификация задач, решаемых агентами.

Понятие задачи в когнитивных науках. Теория Функциональных Систем

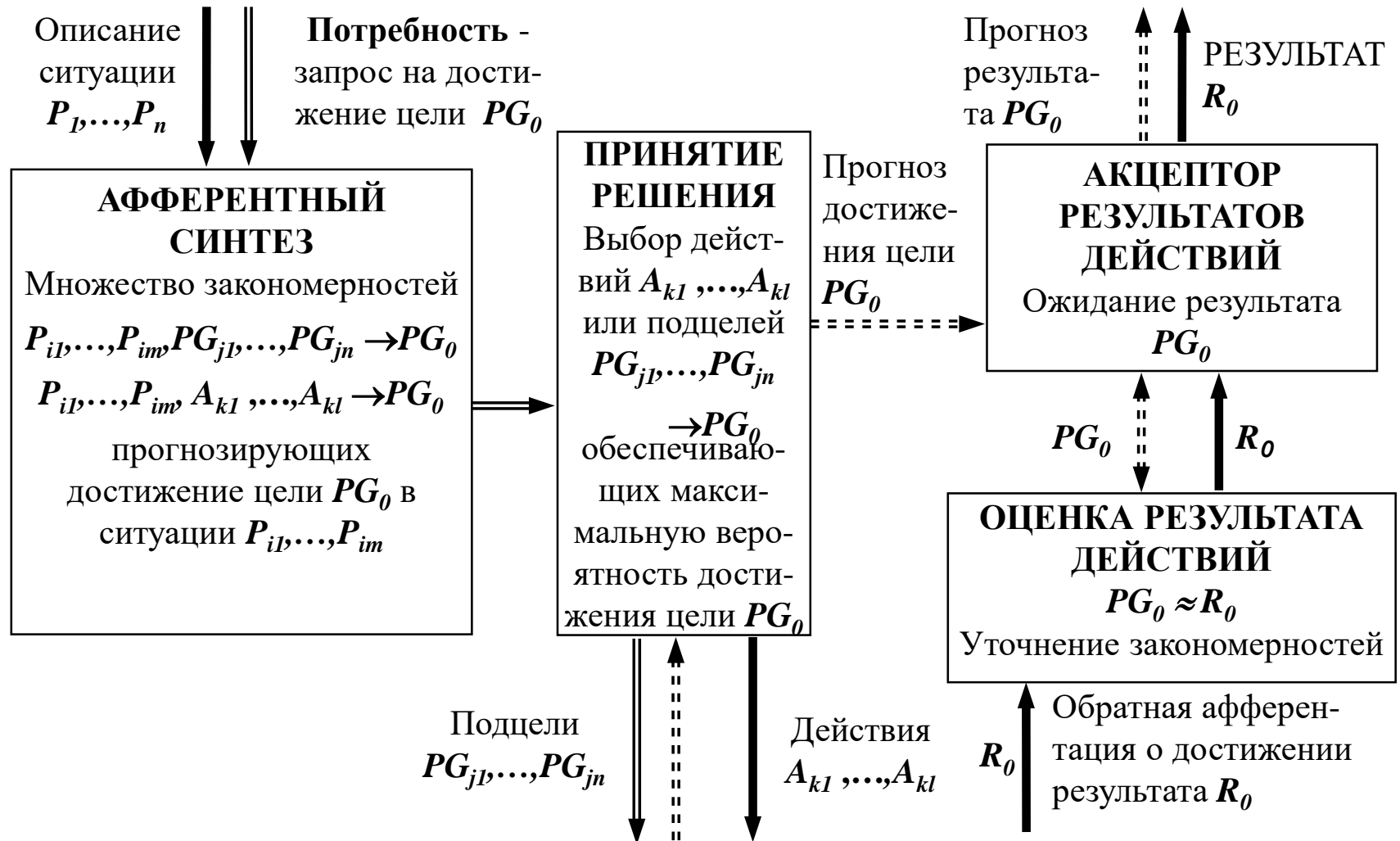
«Пожалуй, одним из самых драматических моментов в истории изучения мозга как интегративного образования является фиксация внимания на самом действии, а не на его **результатах** ... мы можем считать, что результатом «хватательного рефлекса» будет не само хватание как действие, а та **совокупность афферентных раздражений**, которая соответствует признакам «схваченного» предмета» П.К. Анохин.

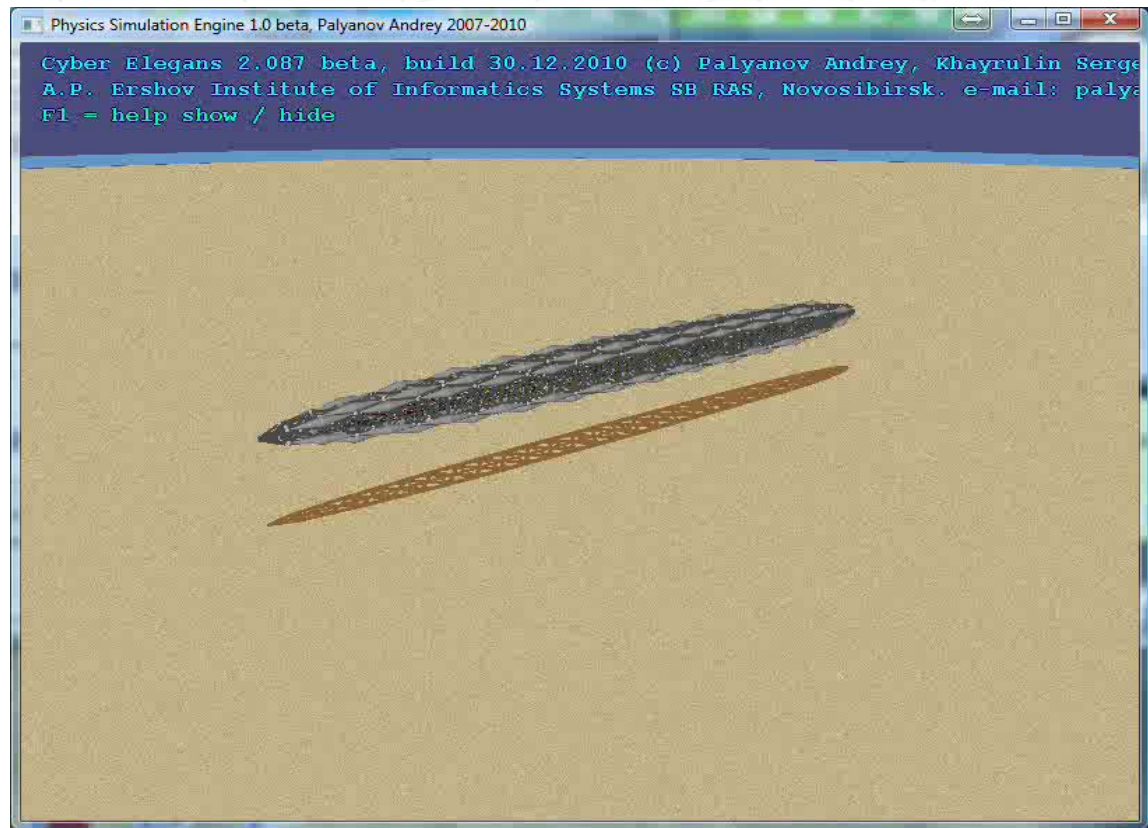
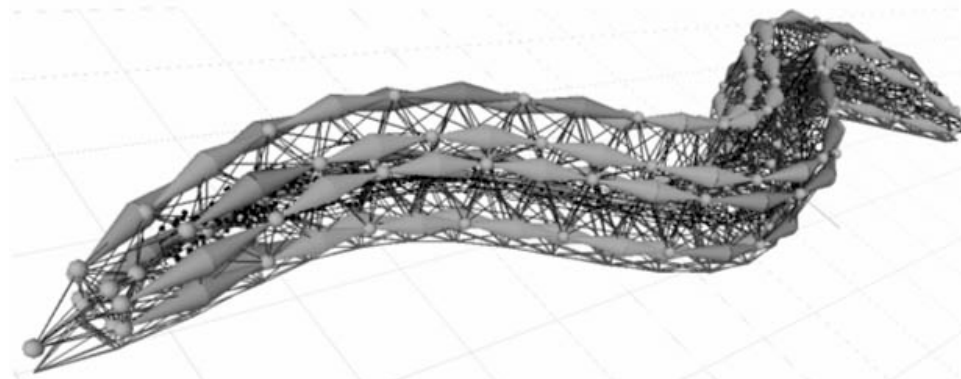
«совокупность афферентных раздражений» и есть критерий достижения цели в ТФС.

Единственной физиологической теорией, в которой целенаправленная деятельность фактически рассматривается решение мозгом ЗАДАЧИ по удовлетворению некоторой потребности, а также рассматриваются физиологические механизмы решения этой задачи.

Evgenii E. Vityaev Purposefulness as a Principle of Brain Activity // Anticipation: Learning from the Past, (ed.) M. Nadin. Cognitive Systems Monographs, V.25, Chapter No.: 13. Springer, 2015, pp. 231-254.

СХЕМА РАБОТЫ ФУНКЦИОНАЛЬНЫХ СИСТЕМ В ТЕРМИНАХ МС-ПРАВИЛ





ПОДХОДЫ К МОДЕЛИРОВАНИЮ МЫШЛЕНИЯ



НАУКИ об
ИСКУССТВЕННОМ



Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления. *Нейроинформатика*, 2008, том 3, № 1, стр. 79-107

Витяев Е.Е. Логика работы мозга. Подходы к моделированию мышления. (сборник под ред. д.ф.-м.н. В.Г. Редько). УРСС Эдиториал, Москва, 2014г., стр. 120-153.

Evgenii E. Vityaev Purposefulness as a Principle of Brain Activity // *Anticipation: Learning from the Past*, (ed.) M. Nadin. *Cognitive Systems Monographs*, V.25, Chapter No.: 13. Springer, 2015, pp. 231-254.

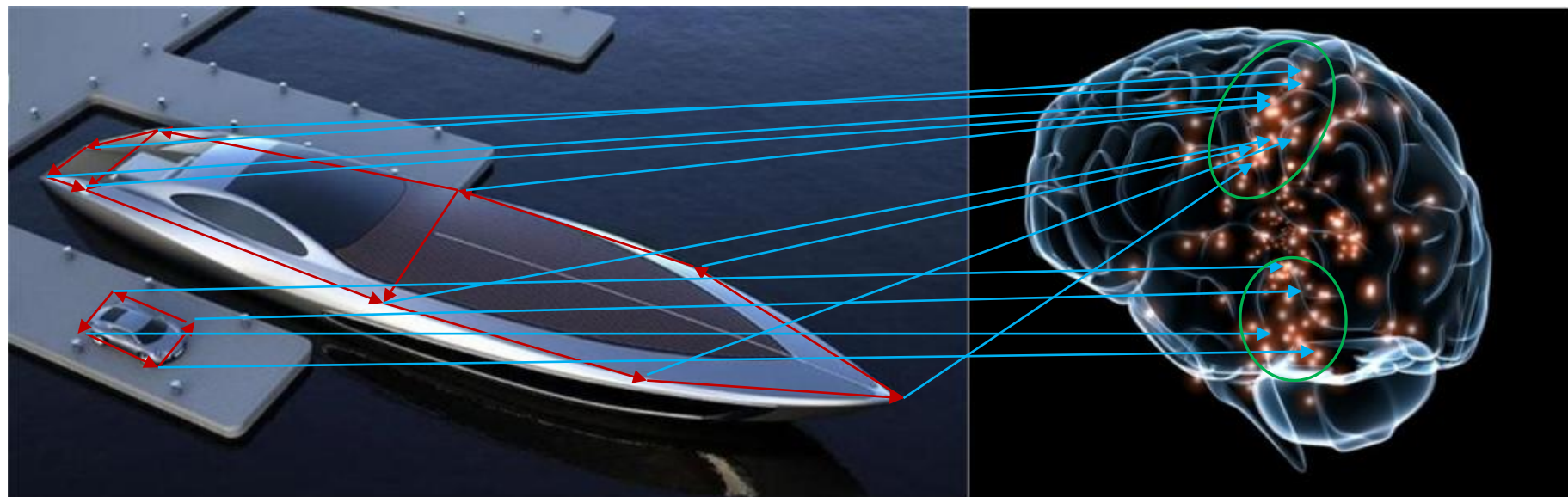
Витяев Е.Е., Неупокоев Н.В. Формальная модель восприятия и образа как неподвижной точки предвосхищений. Подходы к моделированию мышления. (сборник под ред. д.ф.-м.н. В.Г. Редько). УРСС Эдиториал, Москва, 2014г., стр. 155-172.

Витяев Е.Е. Формализация когнитива // *Нейроинформатика-2016*, МИФИ, Москва 2016.

IV. Принцип: отражение высоко коррелированной структуры объектов внешнего мира в «естественной» классификации, «естественных» понятиях и сознании, которая улавливается МС-причинными связями.

**«ЕСТЕСТВЕННАЯ»
КЛАССИФИКАЦИЯ
ОБЪЕКТОВ
ВНЕШНЕГО МИРА**

**«ЕСТЕСТВЕННЫЕ»
ПОНЯТИЯ И СОЗНАНИЕ КАК
ИНТЕГРИРОВАННАЯ
ИНФОРМАЦИЯ**





«ЕСТЕСТВЕННАЯ» КЛАССИФИКАЦИЯ

По Дж. Ст. Миллю «искусственные» классификации отличаются от «естественных» тем, что они могут быть основаны на любом одном или несколькими признаках, по которым и различаются классы.

«Естественные» классы животных или растений отличаются столь большим (потенциально бесконечным) количеством свойств, что их нельзя перечислить.

Смирнов Е.С.: «Таксономическая проблема заключается в “индикации”: от бесконечно большого числа признаков нам нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки».

В бывшем СССР было классификационное движение, которое анализировало, что такое «естественная» классификация и каковы ее свойства. Именно это направление пыталось понять, что такое «естественная» классификация **как закон действительности – закон высоко коррелированной структуры объектов внешнего мира**. Было отмечено, что «естественные» классификации являются генетическими – естественные объекты всегда имеют некоторый генезис своего происхождения. Рассмотрим генезис с точки зрения принципа физического детерминизма и причинности: «знание состояния некоторой физической системы позволяет, в принципе, предсказать все последующие её состояния», поэтому комплекс параметров начального состояния объекта в процессе его генезиса определяет весь комплекс параметров результирующего «естественного» объекта.



Для определения 128 классов бинарными признаками требуется всего 7 независимых бинарных признака, т.к. $2^7 = 128$.

Свойство переопределенности Кожара В.Л. называл “таксономической насыщенностью”.

На данных с 124 признаками таксономическая насыщенность была 23, т.е. классификация по любым случайно выбранным 23 признакам была одинаковой. Это значит, что есть $(124 - 23) \cdot C_{124}^{23}$ закономерностей, предсказывающих значения $(124 - 23)$ признаков по выбранным 23-м.

Для пространств признаков большой размерности для методов Machine Learning возникает проблема «проклятия размерности», когда при увеличении размерности признакового точность методов падает.

Можно показать, что в разработанном нами методе «естественной» классификации, учитывающем переопределённость информации, точность отнесения к классу растет с увеличением числа признаков.

Витяев Е.Е., Мартынович В.В. Формализация "естественной" классификации и систематики через неподвижные точки предсказаний // Сибирские электронные математические известия, Том 12, Институтом математики им. С. Л. Соболева СО РАН, 2015, стр. 1006-1031.



«Естественные» понятия в когнитивных науках

В работах Eleanor Rosch были сформулированы принципы категоризации «естественных» категорий, подтверждающие высказывания Дж. Ст. Милля и естествоиспытателей:

Структура воспринимаемого мира: “воспринимаемый мир – не является неструктурированным множеством равновероятно встречающихся свойств, наоборот, объекты воспринимаемого мира имеют ... высоко коррелированную структуру. ... комбинации того, что мы воспринимаем как атрибуты реальных объектов не встречаются равномерно. Некоторые пары, тройки и т.д. достаточно вероятны ... другие редки; другие логически или эмпирически не встречаются”.

Непосредственно воспринимаемые объекты (basic objects) – информационно богатые связки наблюдаемых и функциональных свойств, которые образуют естественную разрывность, создающую категоризацию.

«Категории могут быть рассмотрены в терминах их чистых случаев, если воспринимающий обращает внимание на **корреляционную структуру воспринимаемых атрибутов** ... Под прототипами категорий мы в общем случае имеем ввиду чистые случаи принадлежности категории».

В дальнейшем теория «естественных» понятий Eleanor Rosch получила название прототипической теории понятий (prototype theory).



В дальнейших исследованиях было обнаружено, что моделей, основанных на признаках, сходстве и прототипах, недостаточно для описания классов.

Необходимо учитывать причинные и онтологические знания, относящиеся к объектам классов. Например, люди не только знают, что птицы имеют крылья, могут летать и вить гнезда на деревьях, но также и то, что птицы выют гнезда на деревьях, потому что могут летать, и летать, потому что они имеют крылья.

Учитывая эти исследования, Bob Rehder выдвинул теорию причинных моделей (causal-model theory), в которой отношение объекта к категории основывается уже не на множестве признаков и близости по признакам, а на основании *сходства порождающего причинного механизма*.

Для описания причинных моделей Bob Rehder использовал «развертку» причинных моделей, использующую Байесовские сети. Однако Байесовские сети не поддерживают циклов и поэтому не могут моделировать циклические причинные связи.

Нами предполагается *новый математический аппарат* для определения «естественной» классификации, «естественных» понятий, основанный на формализации циклических причинных связей.

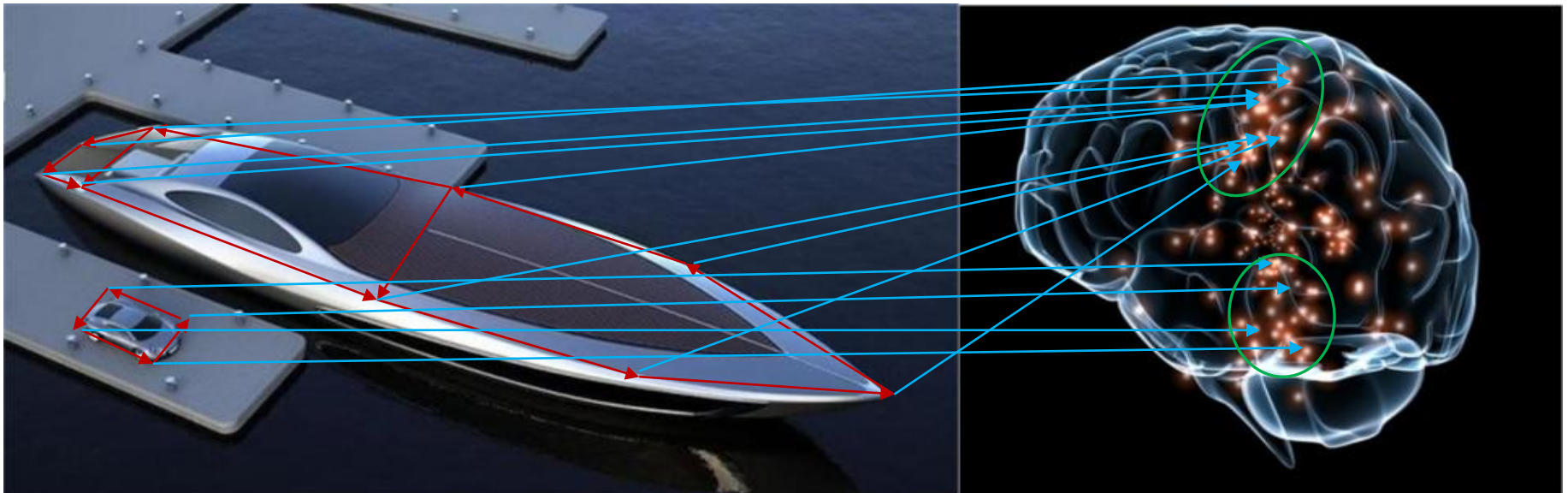
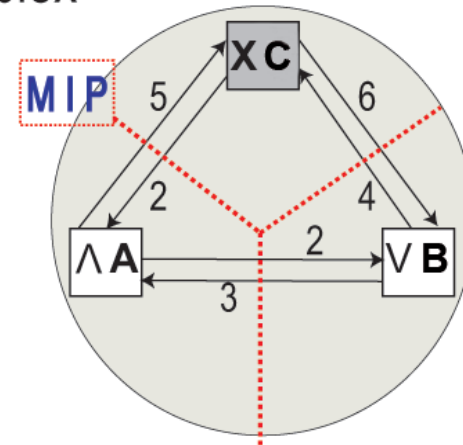
Эти причинные связи при восприятии «естественных» объектов замыкаются на себя образуя определенный «резонанс» циклических причинных связей, что дает причинные модели.

G. Tononi: Consciousness as integrated information

Finding the *integrated information* generated by a system requires analyzing it into complexes. Any system of elements can be decomposed into its constituent complexes.

У G.Tononi нет внешнего мира и интегрированная информация – внутренняя характеристика системы.

A. Complex





Анализ формальных понятий (formal concept analysis)

Формальный контекст $K = \langle G, M, I \rangle$, $I \subseteq G \times M$

G – объекты, M – признаки, I – отношение объект-признак

Операторы вывода $A \subseteq G$, $A' = \{m \in M \mid \forall g \in A (gIm)\}$

$B \subseteq M$, $B' = \{g \in G \mid \forall m \in B (gIm)\}$

Формальное понятие в контексте $K = \langle G, M, I \rangle$ это пара

(A, B) , $A \subseteq G$, $B \subseteq M$, $A' = B$, $B' = A$

A – множество всех объектов из G , имеющих все признаки из B , и B – множество всех признаков из M , которыми обладают все объекты из A .

ANIMALS	preying	flying	bird	mammal
LION	×			×
FINCH		×	×	
EAGLE	×	×	×	
HARE				×
OSTRICH			×	



Вероятностное обобщение формальных понятий

Дж. Ст. Милль писал, что: «Естественные группы ... определяются признаками ... Однако при этом принимаются во внимание не только признаки, безусловно общие для всех включаемых в группу предметов, но вся совокупность тех признаков, из которых все встречаются в большинстве этих предметов, а большинство – во всех».

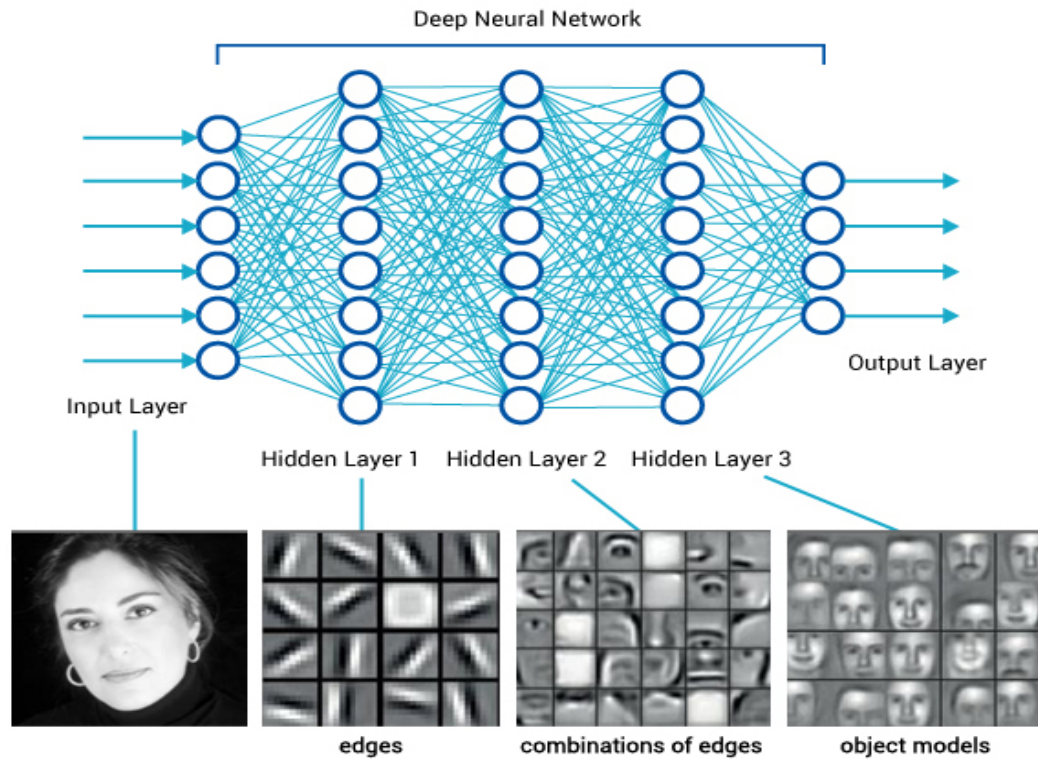
Этапы вероятностного обобщения формальных понятий :

1. Определить формальные понятия как неподвижные точки импликаций.
2. Импликации заменили максимально специфическими причинными связями.
3. Определить вероятностные формальные понятия как неподвижные точки максимально специфических причинных связей.
4. Доказать, что эти неподвижные точки логически непротиворечивы.
5. Определить “естественную” классификацию и “естественные” понятия как вероятностные формальные понятия.

Alexander Demin, Denis Ponomaryov, Evgeny Vityaev. Probabilistic Concepts in Formal Contexts // Lecture Notes in Computer Science, Vol. 7162, Springer Verlag, 2012, p. 394-410

E. E. Vityaev, V. V. Martinovich. Probabilistic Formal Concepts with Negation // A. Voronkov, I. Virbitskaite (Eds.): PCI 2014, LNCS 8974, 2015, pp.385-399.

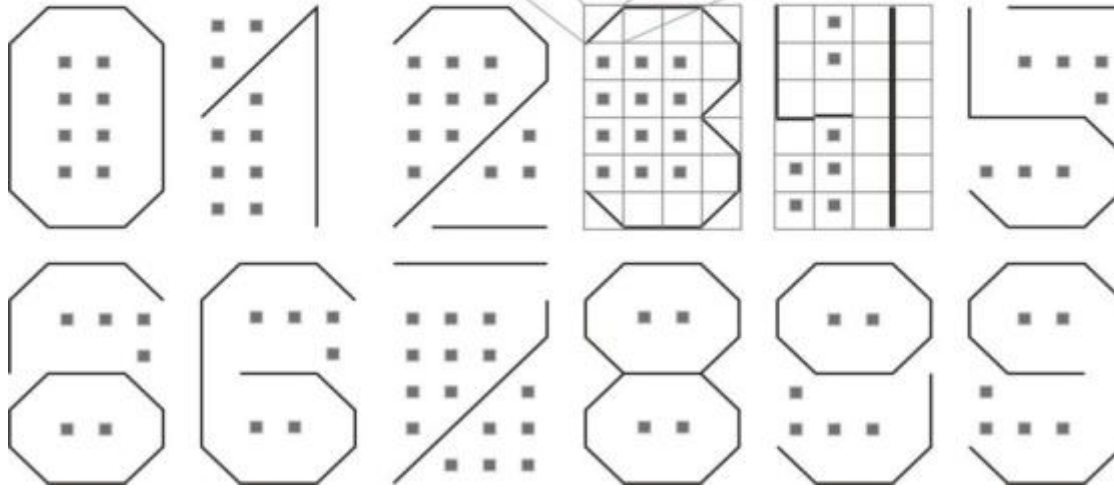
Probabilistic explainable deep learning.



- автоматически обнаруживает иерархию вторичных признаков;
- максимально точно и непротиворечиво формирует классы;
- образ мира – наибольшая неподвижная точка;
- непрерывно во времени и в пространстве делает предсказания и отслеживает их подтверждение.

“Естественная” классификация и “естественные” понятия как вероятностные формальные понятия

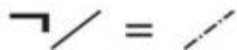
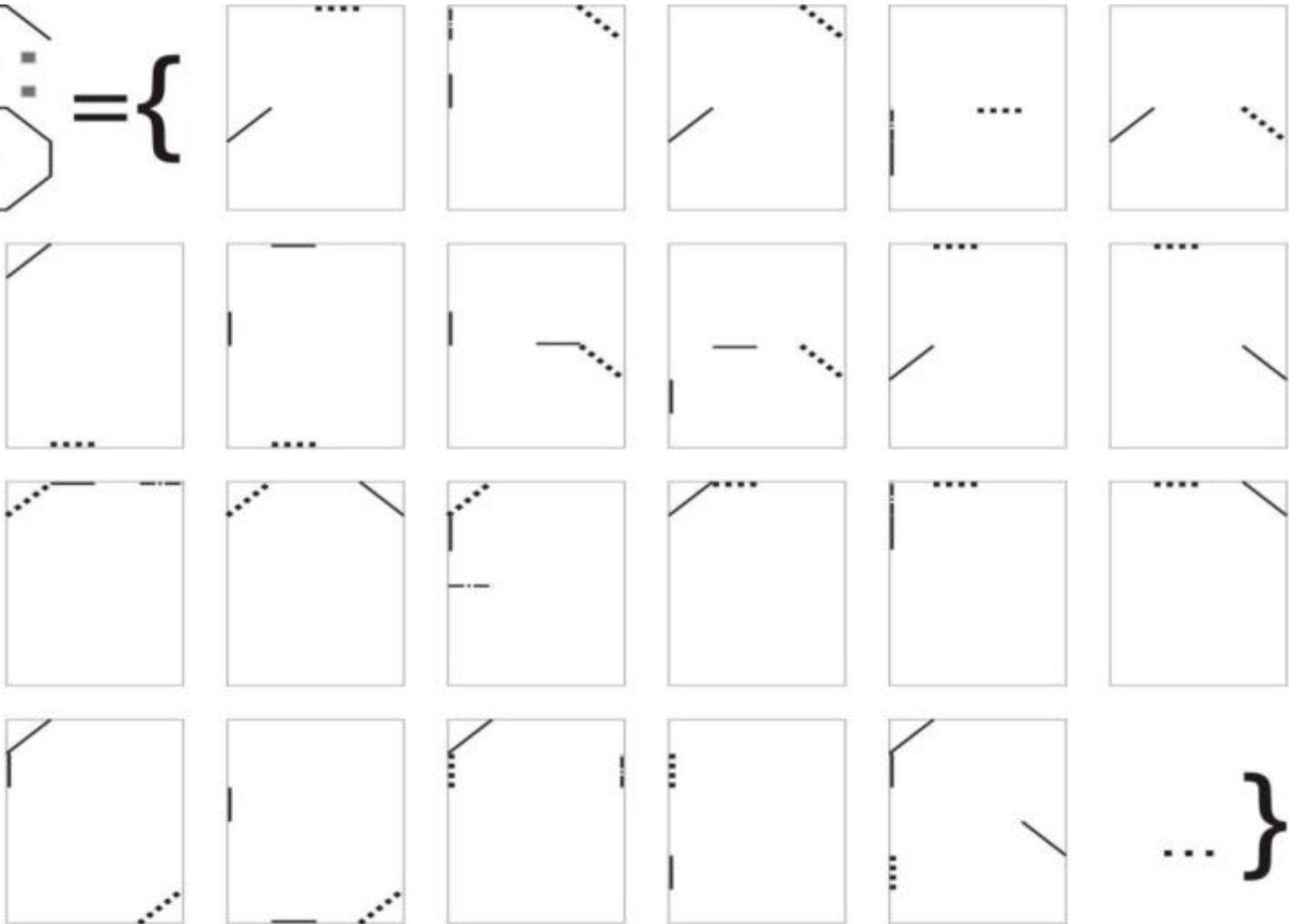
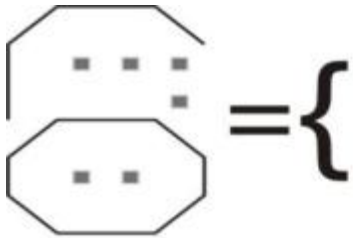
1	2	3	4	5	6	7



1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24

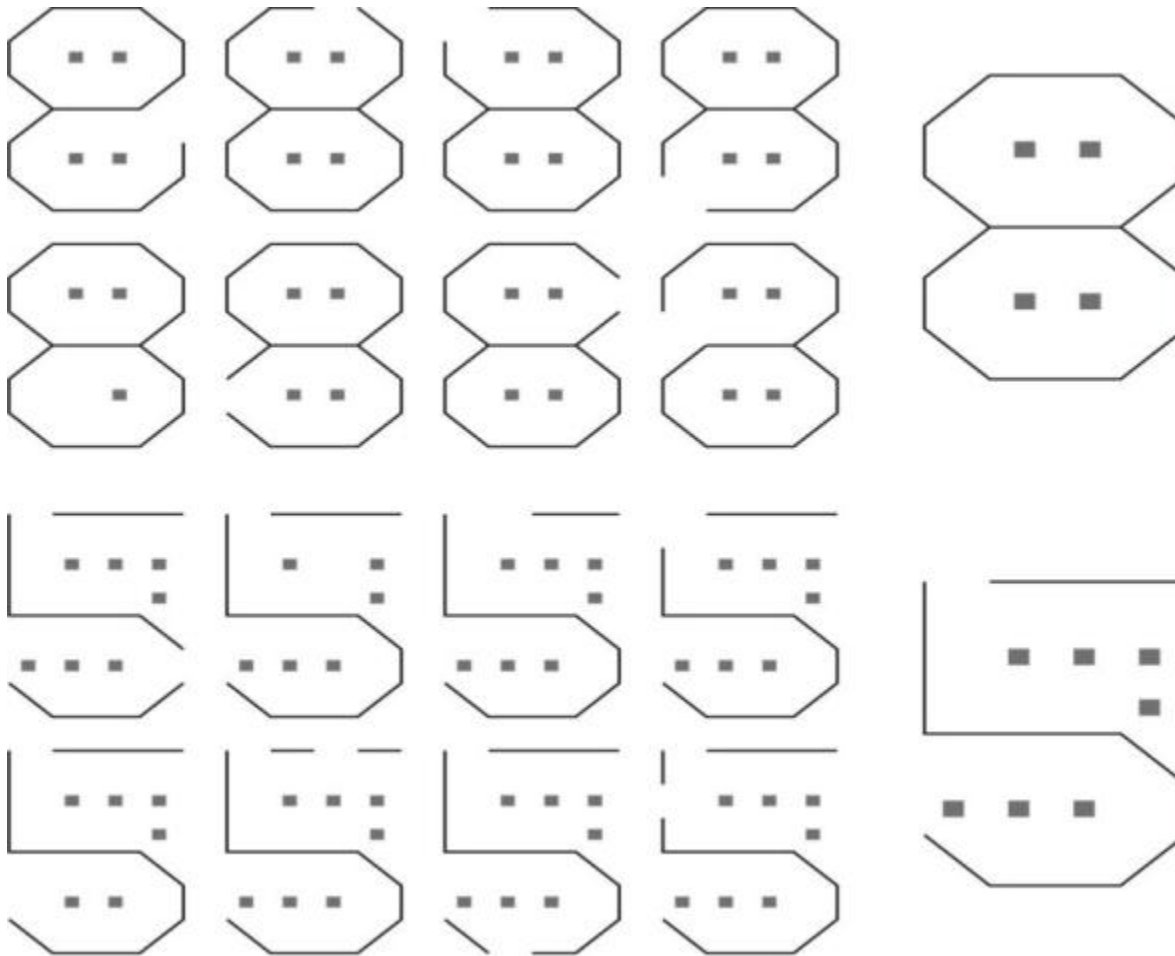


Fix-point



Regularities use an attribute negation

Fix-point



Digits 5 and 8 with altered attributes.

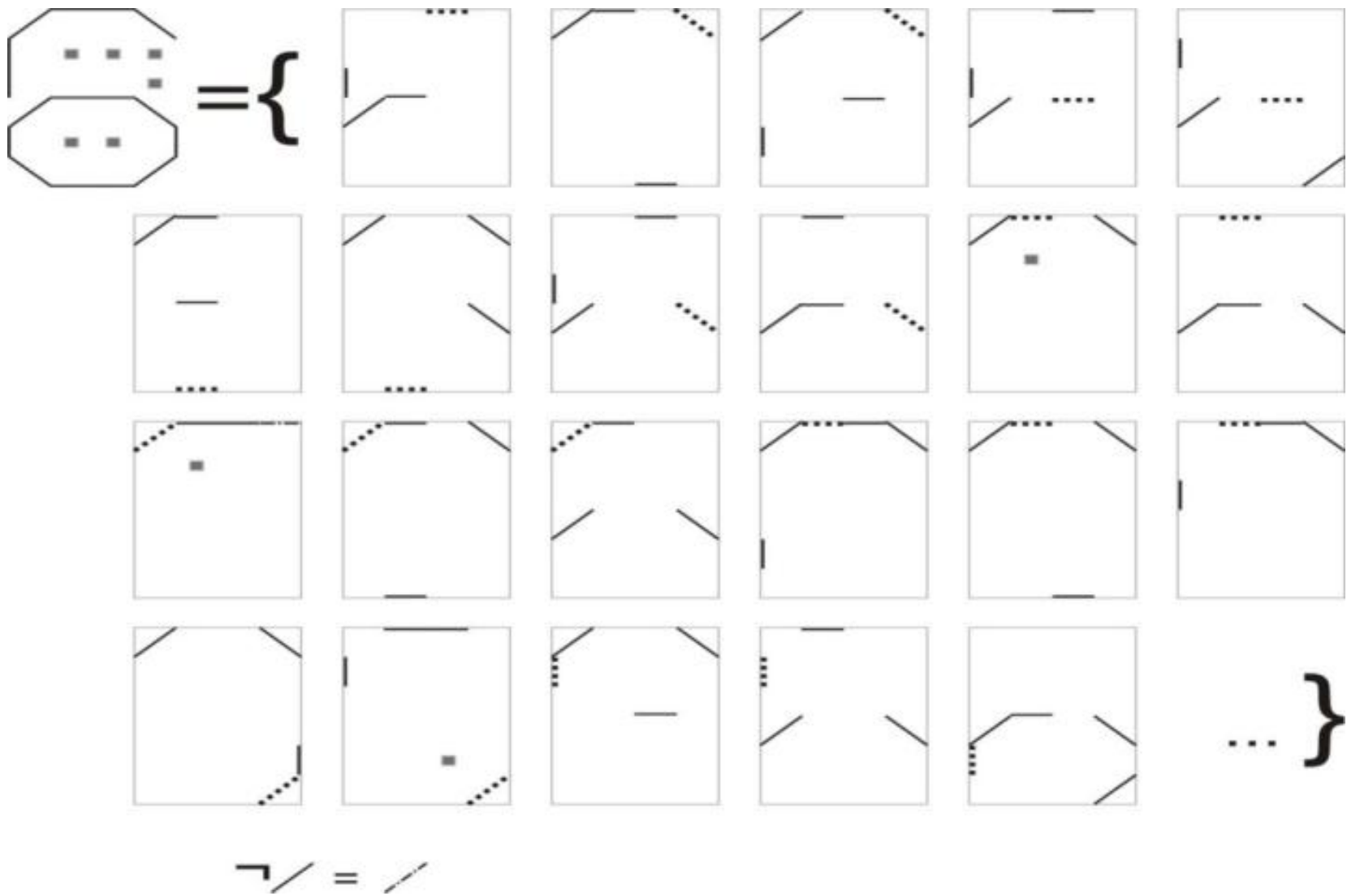
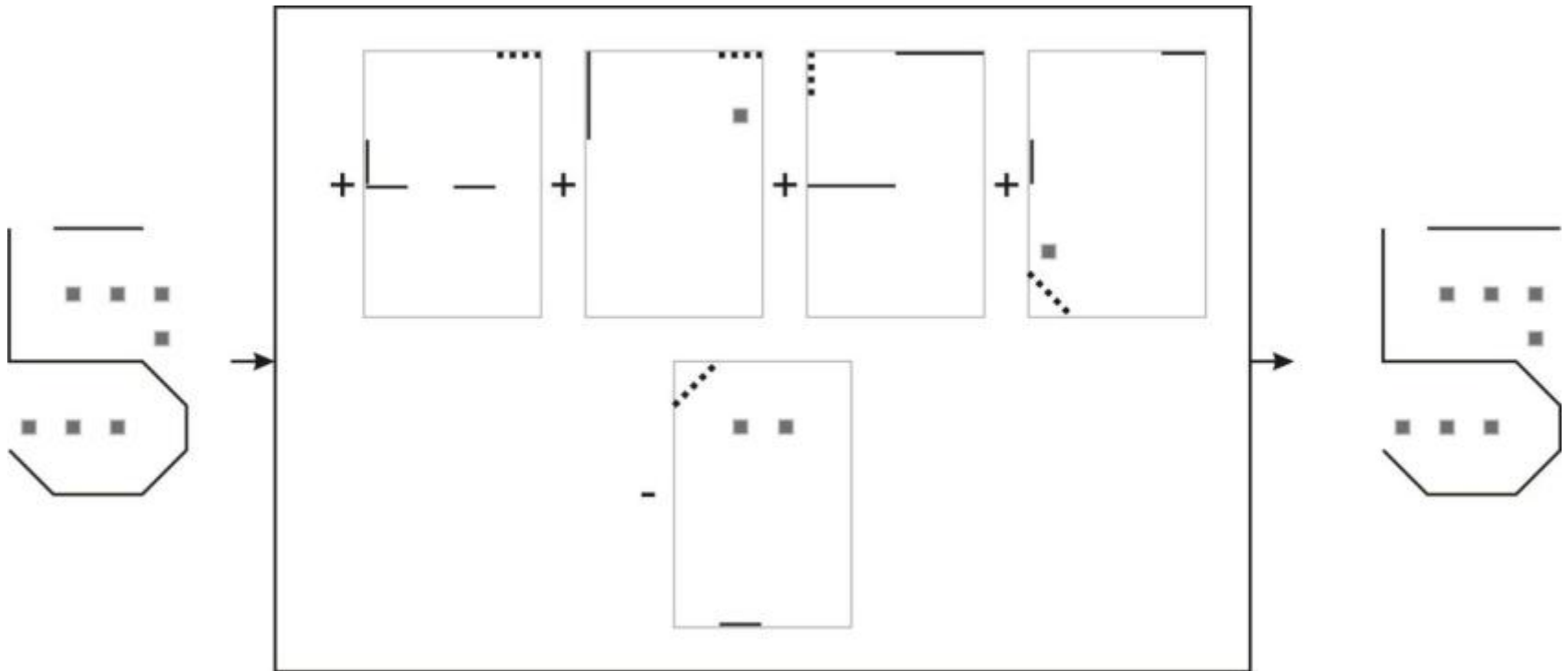


Fig. Digit 6 closure

Активность модели – процесс идеализации



$$\text{Krit}(X) = \sum_{R \in S(X)} \mu(R) - \sum_{R \in F(X)} \mu(R)$$

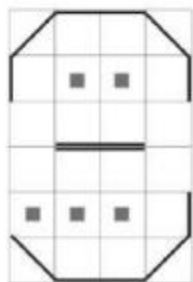
Сравнение с сетями Хопфилда

\mathcal{A} — 360 объектов (30 копий 12-ти цифр)

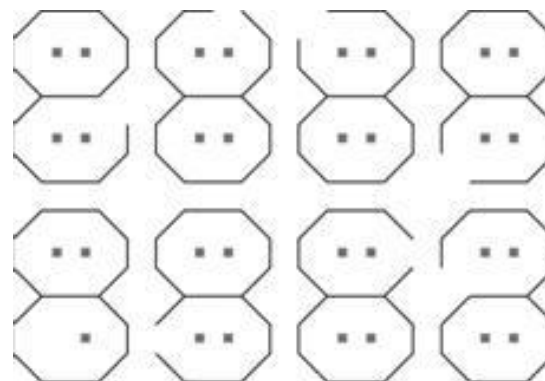
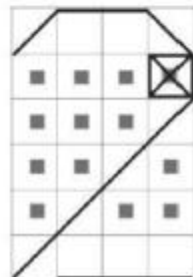
\mathcal{B} — 360 объектов: 30 копий 12-ти цифр, в каждой из которых отсутствует один произвольный признак

Тип	Обучение	Результат
Неподвижные точки	\mathcal{A}	12 классов 55089 закономерностей
	\mathcal{B}	14 классов 73458 закономерностей
Сеть Хопфилда	\mathcal{A}	1 класс
	\mathcal{B}	1 класс
newhor из MATLAB	\mathcal{A}	12 классов
	\mathcal{B}	для каждого объекта порождается свой класс

а)



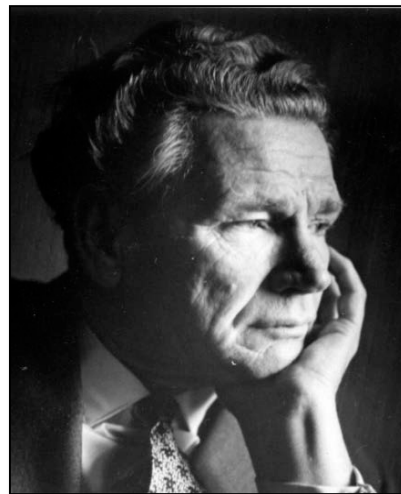
б)



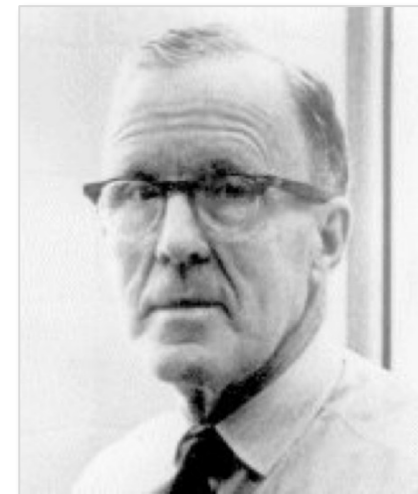
Супер ИИ, Моделирование когнитома, Модель сознания.

КОГНИТОМ - когнитивная реальность, проявляющаяся в феноменологии и поведении. **Когнитом** состоит из **когов** – **когнитивных групп нейронов** (**COgnitive Group** – **COG**)

1. **когов** обобщающих представления теории **функциональных систем** П.К.Анохина и
2. **когов** феноменального опыта, обобщающих **клеточные ансамбли** Д. Хебба.



П.К.Анохин
(1898-1974)



D.O.Hebb
(1904-1985)

Витяев Е.Е. Формализация когнитома // Нейроинформатика-2016, МИФИ, Москва 2016.

Витяев Е.Е. Сознание как логически непротиворечивая прогностическая модель реальности // МАТЕРИАЛЫ Всероссийской конференции "КОГНИТИВНЫЕ ИССЛЕДОВАНИЯ НА СОВРЕМЕННОМ ЭТАПЕ" (КИСЭ-2017), Казань, 2017, 20 стр.

Evgenii Vityaev. Consciousness as a logically consistent and prognostic model of reality // Cognitive Systems Research, 2019 Elsevier, 59 (2020), 231–246.



Спасибо за внимание

e-mail: vityaev@math.nsc.ru

Scientific Discovery website

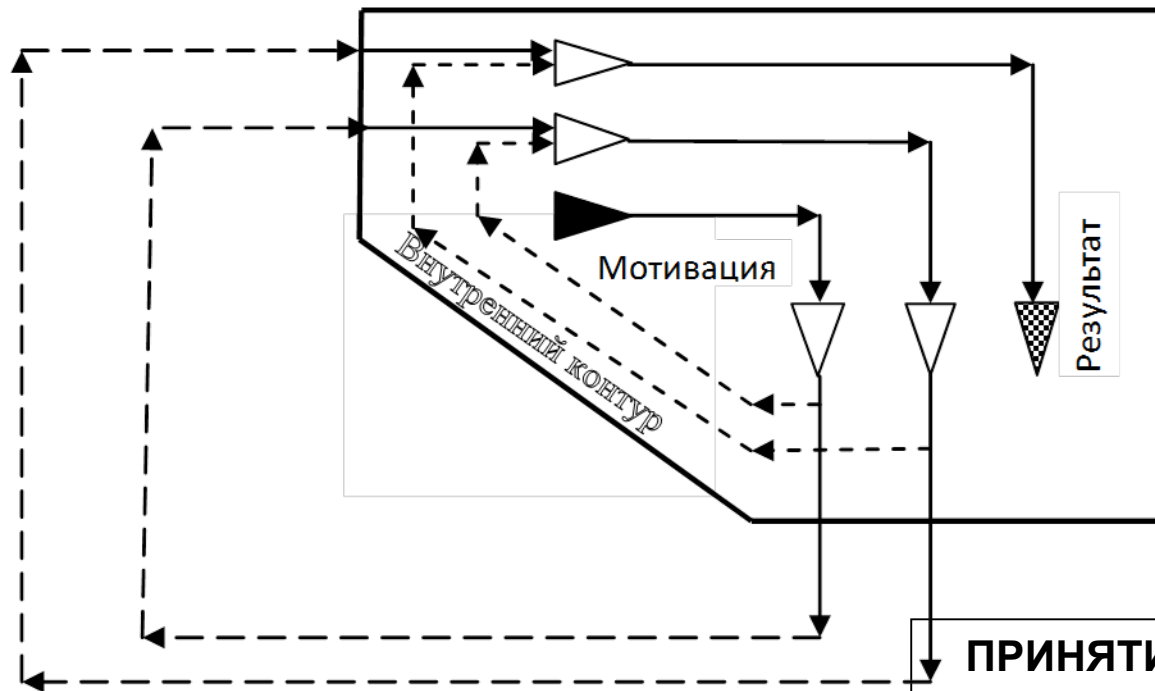
<http://www.math.nsc.ru/AP/ScientificDiscovery>

Открытый код для разработки аниматов:

<http://math.nsc.ru/AP/ScientificDiscovery/soft/FS.html>



МОЗГ



Целенаправленная деятельность в Теории Функциональных систем

ДЕЙСТВИЯ во ВНЕШНЕМ МИРЕ

АФФЕРЕНТНЫЙ СИНТЕЗ
Синтез
мотивационного
возбуждения, памяти
и информации об
окружающей среде

ПРИНЯТИЕ РЕШЕНИЯ
Выбор одного
конкретного
способа
действий

АКЦЕПТОР РЕЗУЛЬТАТОВ ДЕЙСТВИЙ
Модель
параметров
ожидаемого
результата

Действия

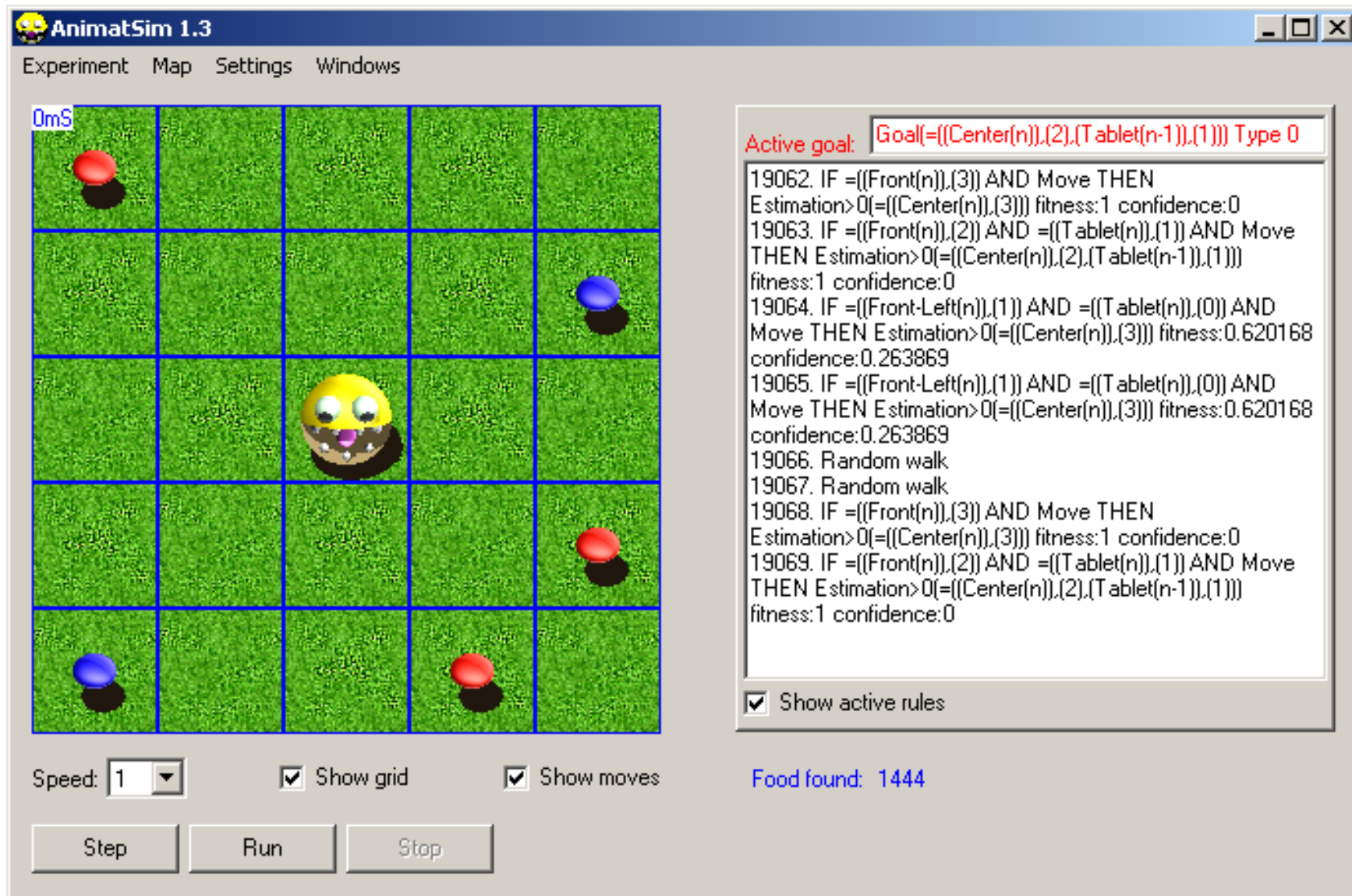
ВЫПОЛНЕНИЕ ДЕЙСТВИЙ

Обратная афферентация

Принципиальные отличия от G. Tononi

- 1) Причинные связи, которые обнаруживает нейрон, должны быть максимально специфическими, решающими проблему статистической двусмысленности и максимально учитывающие имеющуюся информацию.
- 2) Комплексы надо искать, как неподвижные точки – тогда они сами себя обнаруживают.
- 3) Сознание/quale – отражает структуру предсказаний внешнего мира и формируется автоматически и объясняет причину интеграции информации;
- 4) Сознание/quale – не только устанавливает наличие взаимосвязи между признаками комплекса, но и проверяет *соответствие модели – неподвижной точки предсказаний – на соответствие внешнему миру*;
- 5) Основа соответствия внешнему миру – *правильность всех предсказаний свойств внешнего мира в неподвижной точке* предсказаний, а не information matching. (A measure that should be sensitive to how well an intrinsic information structure ‘resonates’ with the extrinsic information structure of the environment is information matching. This is given by the change in the average integrated information generated by a complex when it interacts with its environment, compared to when it is exposed to uncorrelated noise);
- 6) Модель активна, а не пассивна – вытормаживает альтернативы.

ПРОГРАММА, МОДЕЛИРУЮЩАЯ АНИМАТ И СРЕДУ ЕГО ОБИТАНИЯ



ОПИСАНИЕ АНИМАТА

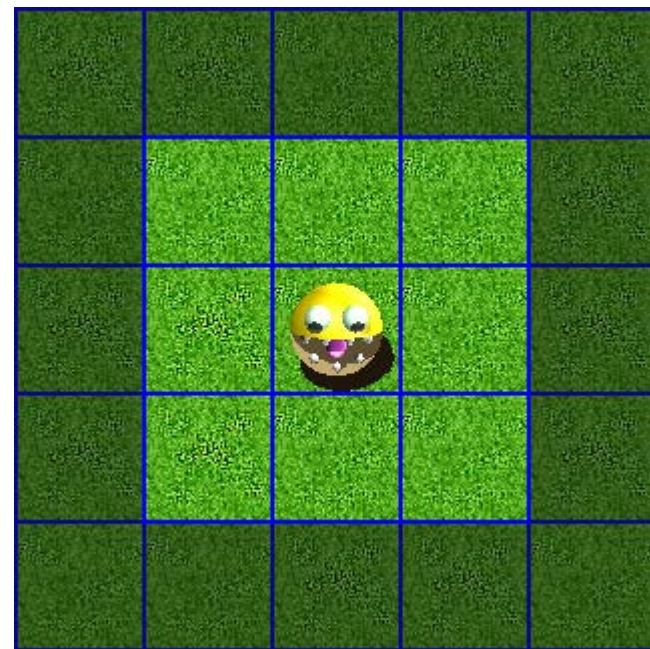
Действия

Шаг вперед, Поворот на 90° на лево, Поворот на 90° направо.

Сенсоры

Девять сенсоров, информирующих анимат о состоянии окружающих клеток. Сенсоры принимают значения «пусто», «препятствие», «еда» или «таблетка».

Один сенсор «есть таблетка», информирующий анимат о наличии таблетки и принимающий значения «да» или «нет».



Расположение сенсоров

Предикаты

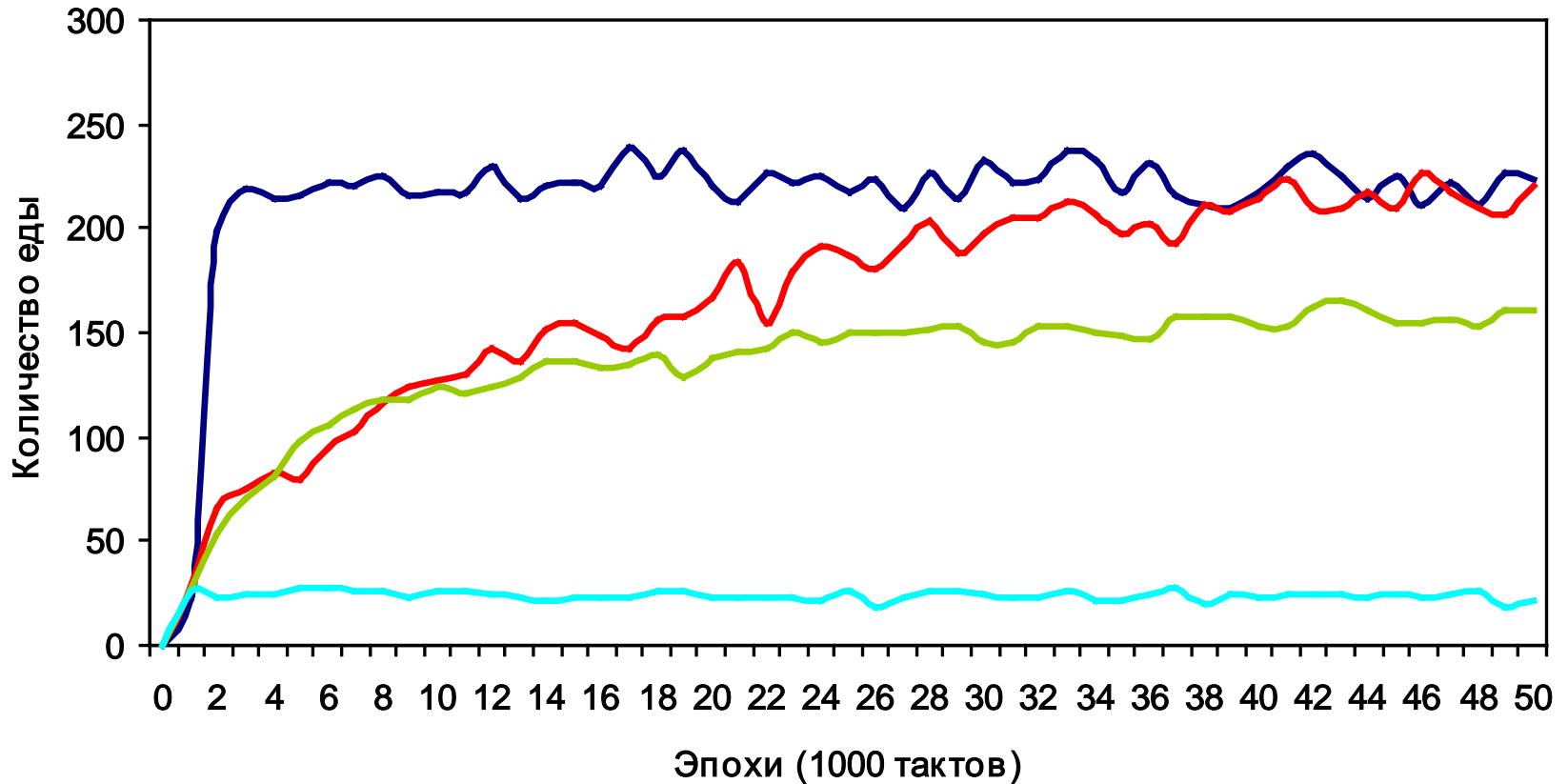
$\langle s = \text{«пусто»} \rangle$, $\langle s = \text{«препятствие»} \rangle$, $\langle s = \text{«еда»} \rangle$, $\langle s = \text{«таблетка»} \rangle$, где s – окружающая клетка, и $\langle \text{«есть таблетка»} = \text{«да»} \rangle$.

Основная цель

Достижение ситуации одновременного наличия таблетки и нахождения еды в центральной клетке.

Предикат-цель: $\langle \text{«центральная клетка»} = \text{«еда»} \text{ И } \text{«есть таблетка»} = \text{«да»} \rangle$.

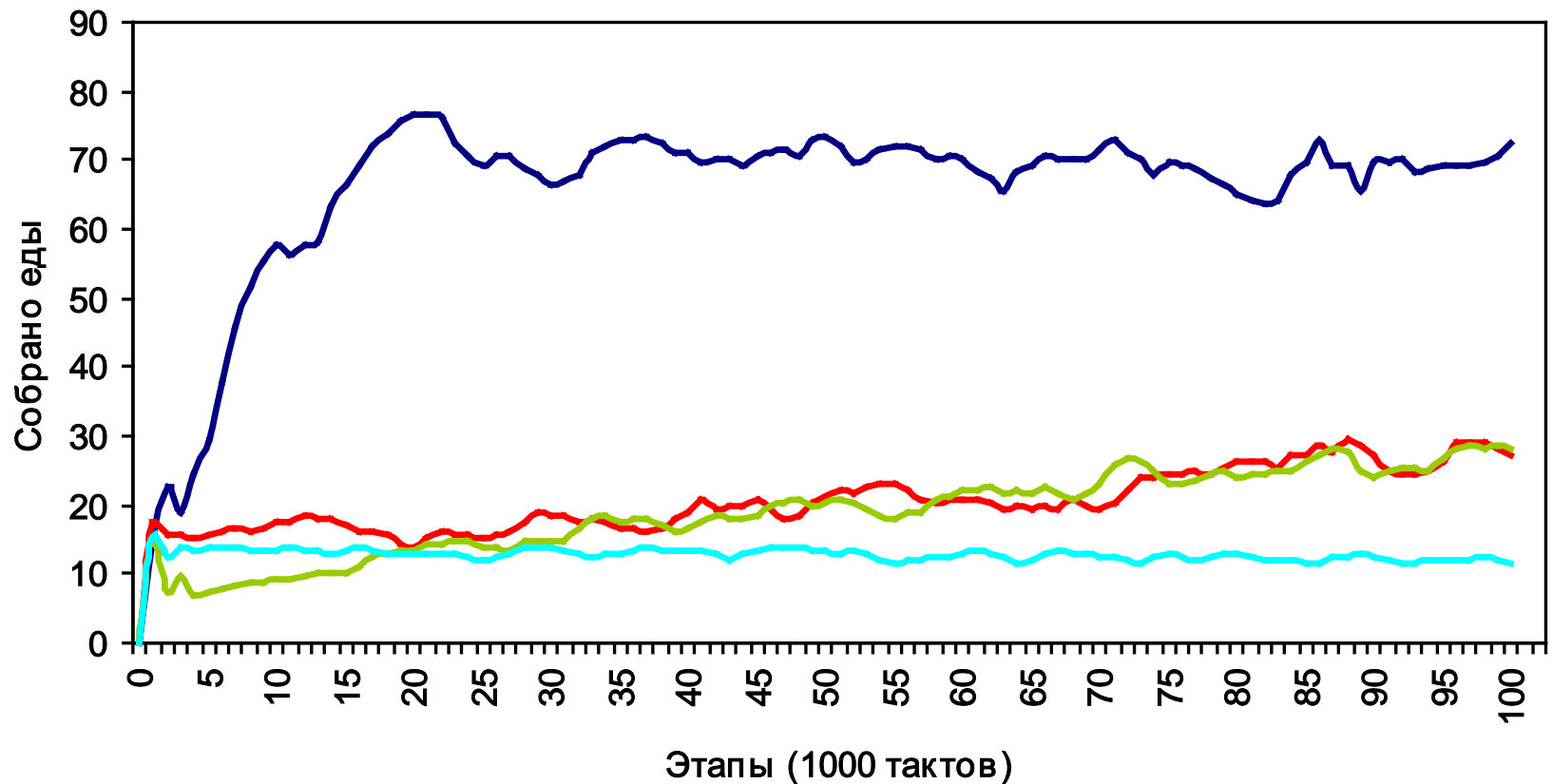
РЕЗУЛЬТАТЫ СРАВНЕНИЯ НА ПРИМЕРЕ РЕШЕНИЯ КЛАССИЧЕСКОЙ ЗАДАЧИ ФУРАЖИРОВАНИЯ



— Семантический вывод — Q-Neural Net — Q-Lookup Table — Random Walk

**Количество «еды», собранное аниматором в разных
системах управления.**

РЕЗУЛЬТАТЫ СРАВНЕНИЯ НА ПРИМЕРЕ РЕШЕНИЯ ЗАДАЧИ ФУРАЖИРОВАНИЯ С ФОРМИРОВАНИЕМ ПОДЦЕЛЕЙ



— Семантический вывод — Q-Neural Net — Q-Lookup Table — Random Walk

Количество «еды», собранное аниматором с разными системами управления.

Переключающая функция эмоций

Семантический вероятностный вывод формализует также понятия *вероятностного прогнозирования* и *вероятности*, введённые в работах И.М. Фейгенберга и информационной теории эмоций П. В. Симонова.

«Суммируя результаты собственных опытов и данные литературы, мы пришли ... к выводу о том, что эмоция есть отражение мозгом человека и животных какой-либо актуальной **потребности** (её качества и величины) и **вероятности** (возможности) её удовлетворения... » П. В. Симонов.

Эмоции – **необходимый** критерий выбора между различными способами достижения цели, учитывающий вероятность достижения цели, сложность и трудоёмкость, а также санкционирующую афферентацию от удовлетворения потребности.

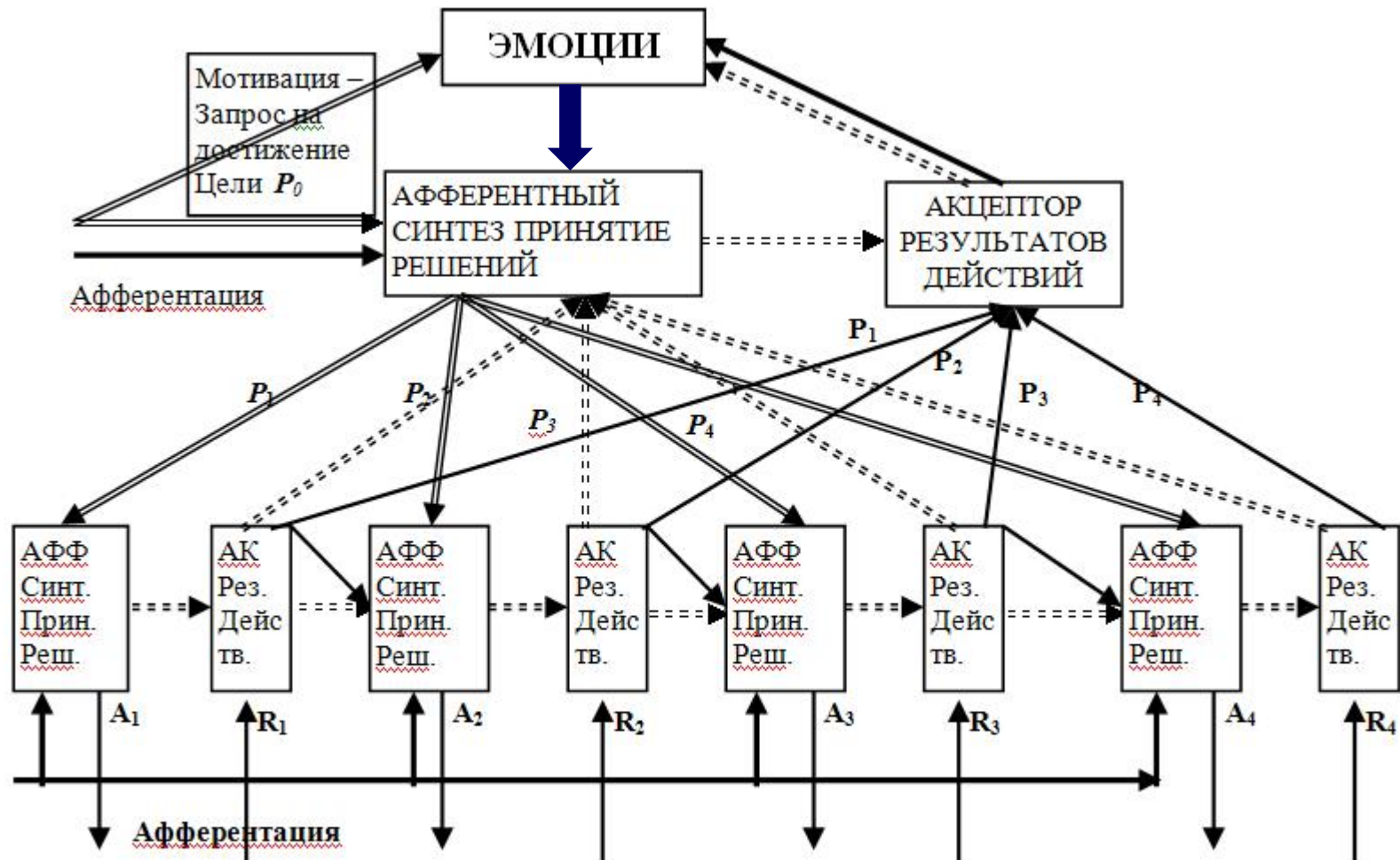
В Искусственном Интеллекте для интеллектуальных агентов такой критерий называется **полезностью**.

В математике есть направление исследований, занимающееся полезностью:

«Теория полезности для принятия решений».

В соответствии с информационной теорией эмоций на стадии **принятия решений**, выбор конкретного способа достижения цели осуществляется **переключающей функцией эмоций**.

Принятие решений. Переключающая функция эмоций.





Иерархичность структуры предвосхищений в сложной вероятностной среде.

В работах Н.А. Бернштейна исследована многоуровневая организация движений: «Движения уровня С пространственного поля имеют прежде всего ясно выраженный целевой характер они ведут "откуда-то" "куда-то" и "зачем-то», «Ведущим мотивом на уровне действий D является не предмет сам по себе, как геометрическая форма, а смысловая сторона действий с предметом ...».

Дж. Гибсон «Экологическая теория восприятия». Мир устроен иерархически: «мелкие элементы содержатся в более крупных ... поэтому я ввожу для него специальный термин: **встроенность**». «Следует отметить ... что на любом участке земной поверхности встречаются, в общем-то, одни и те же элементы. Размер песчинок, где бы они ни встретились, всегда приблизительно один и тот же. Стебли травы также везде более или менее одинаковы. То же самое можно сказать о камнях, пучках травы, кустах и т.п. ... И хотя их повторяемость лишена метрической регулярности, они все же обладают **регулярностью** стохастической, то есть они регулярны в вероятностном смысле».

Отсюда следует **локальность** и **иерархичность** структуры предсказаний внешнего мира:

- 1) **локальность** – хорошо предсказываются в процессе восприятия или целенаправленного действия только целостные части объектов, событий, когнитивных карт ...
- 2) **результатами** локальных предсказаний являются **значения** этих целостностей, фиксирующих факт опознания и классификации данных целостностей как элементов **стохастической регулярности (неподвижной точкой)**;
- 3) Предсказание не может распространяться за пределы целостности (части) объекта поскольку целостные части объектов могут комбинироваться по-разному в разных объектах и предсказать по одной части объекта, соседние невозможно.

3) предсказание возможно в рамках некоторой большей целостности, объединяющей эти частные целостности – большем объекте (стуле), масштабном событии и т.д.

По ножке стула нельзя предсказать какая часть объекта находится рядом не зная, что это ножка именно стула, а не стола, дивана, кресла и т.д.;

Из этого **с необходимостью** следует, что предсказания автоматически формируют:

- 1) **значения (результаты распознавания)** разных уровней общности;
- 2) **значения** организованы иерархически как иерархия объектов и событий.

Уточним понятие результата **значения** так, чтобы результаты могли формироваться автоматически в процессе обучения в соответствии с иерархией объектов и вложенности событий:

- 1) результат увеличивает вероятность достижения конечного результата;
- 2) результат обладает свойством ветвления: если в процессе действий получен некоторый результат, то дальнейшие действия могут развиваться по-разному в соответствии с иерархией объектов или событий.
- 3) результат должен содержать набор признаков – значений, фиксирующих законченность цепочки действий и возможность перехода к следующей цепочке действий.

Формальное описание результата в терминах закономерностей содержится в

Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления.
Нейроинформатика, 2008, том 3, № 1, стр. 79-107

Автоматическое обнаружение таких **результатов** позволяет обнаружить иерархию значений в соответствии с иерархией объектов и событий внешней среды.

Агенты **с необходимостью** должны уметь **обнаруживать иерархию значений** объектов и событий в целях предвосхищения развития внешней среды с учётом своих действий.

«**Каждая потребность**, даже при незначительном отклонении жизненно важной функции от оптимального для метаболизма уровня (в чём и проявляется потребность), немедленно воспринимается специальными рецепторными аппаратами» (критерием достижения цели).

Таким образом, потребность и есть цель, которая ставится перед организмом.

Ведущим возбуждением... определяющим целенаправленную деятельность является мотивационное возбуждение, формирующееся на основе доминирующей потребности

Взаимодействие результатов и целей в ТФС осуществляется несколькими способами: по «принципу доминанты», «иерархией результатов» и «моделями результатов».

По отношению к доминирующей функциональной системе остальные функциональные системы выстраиваются в иерархию по принципу «иерархии результатов».

«Так, у голодного кролика доминирует функциональная система, деятельность которой направлена на поиск пищи. В это время другие функциональные системы, определяющие, например, кровяное давление, дыхание, выделение, направлены на лучшее обеспечение доминирующей пищедобывательной функциональной системы».

Согласно П. К. Анохину, центральные механизмы функциональных систем, обеспечивающих целенаправленные поведенческие акты, имеют однотипную архитектуру.

Афферентный синтез. Начальную стадию поведенческого акта любой степени сложности составляет афферентный синтез, включающий в себя синтез **мотивационного возбуждения, памяти, обстановочной и пусковой афферентации.**

Мотивационное возбуждение. Как мы знаем, постановка цели осуществляется возникшей потребностью, которая трансформируется в мотивационное возбуждение.

Память. Мотивационное возбуждение «извлекает из памяти» также всю последовательность и иерархию результатов, которые должны быть получены для получения результата.

Пусковая афферентация. Пусковая афферентация также является обстановочной афферентацией, только связанной со временем и местом достижения результата.

Пусковая афферентация отвечает на вопрос, **где и когда** можно достичь результат.

Таким образом, на стадии афферентного синтеза в значительной степени разрешается парадокс цели и определяется, что, как и когда можно делать для достижения цели.

«Вытягивая» из памяти весь накопленный опыт, мотивационное возбуждение преобразуется в **конкретную цель**, определяющую способ своего достижения. Конкретная цель называется в ТФС **«высшей мотивацией».**

Принятие решений. На стадии афферентного синтеза мотивационным возбуждением может быть извлечено из памяти (в данной обстановке) несколько способов достижения цели. На стадии принятия решения выбирается только один из этих способов – **конкретный план действий.**

Акцептор результатов действия. Мотивационное возбуждение «извлекает из памяти» также всю последовательность и иерархию результатов, которые должны быть получены для выполнения плана действий. Эта последовательность и иерархия результатов называется в ТФС **акцептором результатов действия**.

Акцептор результатов действия представляет собой доминирующую потребность (Цель) организма, трансформированную в форме опережающего возбуждения мозга, как бы в своеобразный **комплексный рецептор** соответствующего подкрепления».

Акцептор результатов действия – одно из основных понятий ТФС, которое является *критерием достижения конкретной цели*.

Эффекторные механизмы функциональных систем. «Оценка достигнутого результата действия происходит с помощью активной *ориентировочно-исследовательской деятельности*, которая возникает во всех случаях, когда результат совершенного действия не соответствует свойствам акцептора результатов действия, т. е. при возникновении «рассогласования» в поведенческой деятельности. Благодаря включению такой реакции немедленно перестраивается афферентный синтез, принимается новое решение, строится новая программа действий».

Подкрепление. Санкционирующая стадия. «Целенаправленный поведенческий акт ... заканчивается последней санкционирующей стадией. На этой стадии при действии раздражителя, удовлетворяющего ведущую потребность (подкрепления) параметры достигнутого результата вызывают потоки обратной афферентации, которая по всем своим свойствам соответствует ранее запрограммированным свойствам подкрепляющего раздражителя в акцепторе результатов действия».