



On logic and probability synthesis in AGI
**Задачный подход, как логико-вероятностный
общий искусственный интеллект**

Evgenii Vityaev, Anton Kolonin

**Sobolev Institute of Mathematics,
Novosibirsk State University**

Проблемы глубокого обучения

1. Нейронные сети являются «черным ящиком» и не дают возможности объяснить причины принятия того или иного решения. Это делает невозможным применение нейросетей в таких областях как медицина, финансы, военные приложения, где цена ошибки слишком высока, либо объяснение решения необходимо по юридическим причинам. Например, отказ нейросетью в выдаче кредита или совершения опасной хирургической операции должно быть аргументировано юридически.
2. Во-вторых, они обладают слабой способностью к генерализации. Например, нейросеть, натренированная распознавать слонов и китов, в случае предъявления кита, выброшенного на берег, будет видеть в нем слона, а купающегося в прибое слона будет распознавать как кита.
3. Нейросети запоминают отдельные, зачастую случайные детали предъявленных в ходе обучения образцов и принимают дальнейшие решения на основе этих деталей, а не на основе полноценного обобщенного предмета. Например, замена изображения шумом может приводить к распознаванию несуществующего предмета, а замена одного пиксела на изображении к распознаванию предмета, отличного от предъявленного.
4. Нейросети не инвариантны относительно допустимых преобразований шкал – они могут принимать разные решения после преобразования единиц измерения в данных.

Альтернативные подходы к ИИ: Объяснимый ИИ (eXplainable XAI),
Агентный подход, AGI - Artificial General Intelligence,
Задачный подход.

Подходы к Искусственному Интеллекту

Объяснимый ИИ (eXplainable XAI):

Система методов, которые объясняют как и почему ИИ принимает те или иные решения.

1. Метод показывающий вклад каждого признака в полученный прогноз. В объяснении, показывается прогнозируемый класс вместе с пикселями наиболее сильно повлиявшими на результат прогноза.

2. Метод проверки влияния отдельных признаков на прогноз. Эта проверка может обеспечить понимание отдельного прогноза, но не может быть обобщена на весь класс.

Чтобы получить обобщенную информацию, можно объединить признаки по объектам класса или по всему набору данных.

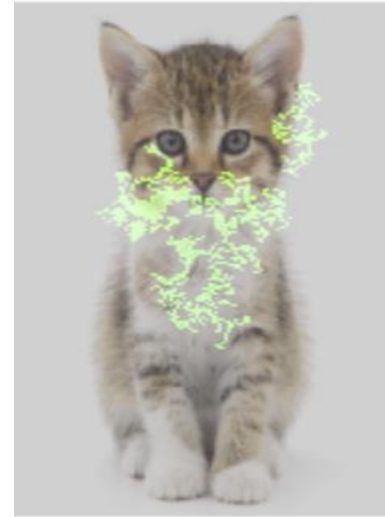
Агентный подход:

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

Различные задачи искусственного интеллекта рассматриваются как задачи взаимодействия «рационального агента» с внешней средой.

В этом подходе рассматриваются следующие агенты:

Простые рефлексивные агенты, агенты, основанные на модели мира, агенты, основанные на цели, агенты, основанные на полезности, обучающиеся агенты.



Общий Искусственный Интеллект

AGI - Artificial General Intelligence:

Глубокое обучение показало, что человеческая природа не обязательно является условием решения когнитивных задач той или иной степени сложности.

Нейронные сети не обязательно моделируют когнитивные процессы.

Соответственно, возникает понятие “общего искусственного интеллекта” (ОИИ или AGI - “artificial general intelligence”), которым, в той или иной степени потенциально может обладать, как человек или живой организм с высокоразвитой центральной нервной системой, так и абстрактная робототехническая система.

Бен Гёрцел (Ben Goertzel) - «Общий интеллект – это способность достигать сложных целей в сложных средах»,

Шейн Легге (Shane Legg) и Маркус Хуттер (Marcus Hutter) - «Интеллект измеряется способностью агента успешно действовать в широком диапазоне сред»,

Пей Ванг (Pei Wang) - «Интеллект – это способность системы адаптироваться к своей среде, работая при недостаточных знаниях и ресурсах».

AGI: «способность решать когнитивные задачи в целом, действуя целенаправленно, адаптируясь к условиям среды через обучение, минимизируя риски и оптимизируя потери на достижение своих целей».

Задачный подход. Понятие задача в основаниях математики

К.Ф. Самохвалов: «Я хочу пить» – что это значит? Нет, конечно, никакой ошибки полагать, что слова «я хочу пить» означают просто вот это, где это – определенное состояние сознания, которое я переживаю сейчас и которое я именую жаждой. Но тогда возникает новый вопрос: как ощущение жажды (хотения) связано с фактическим питьем (удовлетворением хотения)? Откуда я знаю, что удовлетворить жажду можно питьем? Содержится ли в самом переживании жажды сознание того, чем эту жажду можно удовлетворить? ... *Знать желание не означает знать желаемое, а означает способность узнать желаемое*, когда этому представится случай».

Задача определена (осмыслена) тогда и только тогда, когда у нас есть *критерий решенности задачи* – критерий проверки действительно ли предъявленное решение является решением задачи. В математических теориях таким критерием обычно считается наличие доказательства решения задачи. Но этот критерий применим только тогда, когда в рамках самой формальной системы мы имеем, как доказательство решения задачи, так и возможность убедиться средствами самой же 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.

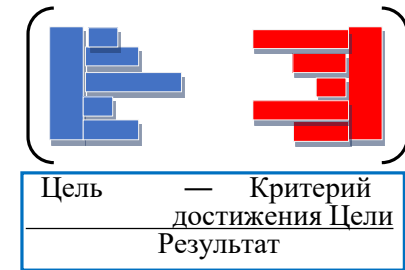
Ю.Л. Ершов, К.Ф. Самохвалов. Современная философия математики, Новосибирск, 2007.

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

Понятие «Задачи» в когнитивных науках

Обобщением понятия задача, является понятие *Цели*. Цель нельзя достичь не имея критерия её достижения, иначе всегда можно считать, что она уже достигнута. Поэтому формулировка цели всегда должна сопровождаться определением *критерия достижения цели*. Достижение цели дает определенный *результат*.

Единственной физиологической теорией, в которой достижение Цели и получение Результата рассматривается как решение мозгом ЗАДАЧИ по удовлетворению некоторой потребности, является Теория Функциональных Систем П.К Анохина. Эта теория также выявляет физиологические механизмы достижения цели и решения задачи мозгом по удовлетворению потребности.



«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.

Поэтому **необходимым** условием **целенаправленного поведения** интеллектуального агента является **целеполагание**, включающее **критерий достижения цели**.

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.

Определение цели ***парадоксально***, так как критерий достижения цели принципиально не содержит никаких знаний о том, как её достичь. *Можно определить цель, не определяя способ её достижения.* Эту парадоксальность цели назовем ***парадоксом цели***. Как будет следовать из теории функциональных систем, мозг при целенаправленном поведении постоянно разрешает парадокс цели, определяя, как и каким образом можно достичь цели.

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

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

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

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

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 целенаправленное поведение, 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. Постановка цели осуществляется возникшей потребностью, которая трансформируется в мотивационное возбуждение.

Память. Мотивационное возбуждение, на основании имеющегося опыта «извлекает из памяти» все последовательности и, соответствующие им иерархии результатов, которые могут привести к достижению цели.

Обстановочная афферентация. Из памяти извлекаются только те способы достижения цели, которые применимы к данной обстановке, т.е. учитывают поступающую обстановочную афферентацию.

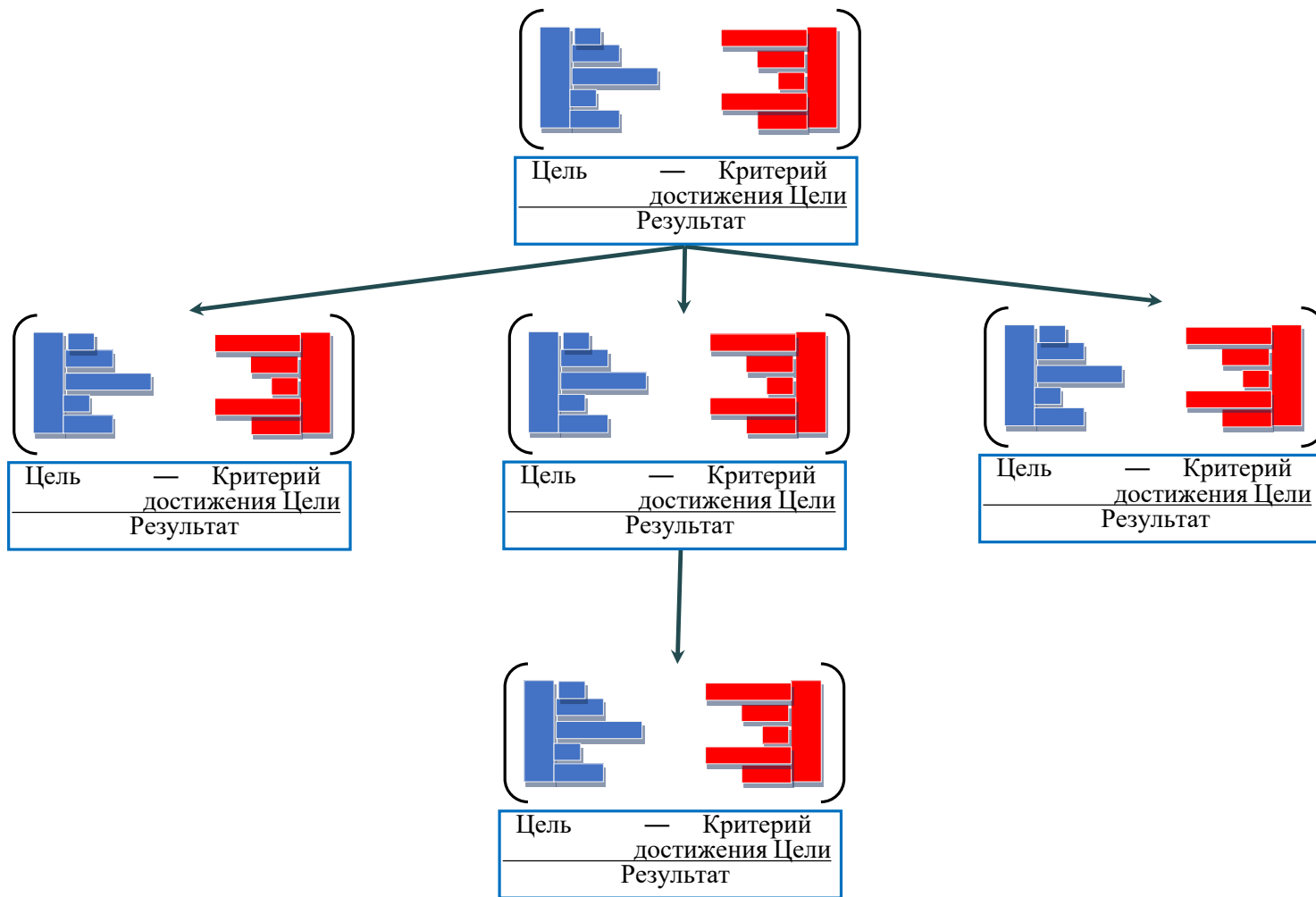
Пусковая афферентация. Пусковая афферентация также является обстановочной афферентацией, только связана со временем и местом достижения результата.

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 нужно сделать для достижения цели.

Принятие решений. На стадии афферентного синтеза мотивационным возбуждением может быть извлечено из памяти несколько способов достижения цели. На стадии принятия решения выбирается один, который формирует **конкретный план действий.**

Конкретный план действий «извлекает» из памяти также всю последовательность и иерархию функциональных систем и действий, связанных с достижением **Цели**, преобразуя **Цель в конкретную цель**, определяющую способ своего достижения.



Конкретная цель называется в ТФС **«высшей мотивацией»**.

Конкретный план действий «извлекает из памяти» также всю последовательность и иерархию результатов, которые должны быть получены при выполнении плана действий. Эта последовательность и иерархия называется в ТФС **акцептором результатов действия**, являющимся **критерием достижения конкретной цели**.

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

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

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».

Осуществленный конкретный план действий подкрепляется и заносится в память.

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».

Так образом, мозгом решаются ЗАДАЧИ по удовлетворению своих потребностей.

Что такое ЗАДАЧА

Задача определена только в том случае, когда в ее формулировке присутствуют:

- указание предметной области, к которой относится задача, знания о предметной области, зафиксированные в виде ее модели, включая описание сигнатуры и структуры языка описания предметной области, набора терминов и понятий (онтологию), исходные данные, факты и знания, записанные в терминах онтологии;
- на какой запрос (вопрос), сформулированный в задаче, относящийся к предметной области, мы должны получить ответ;
- определен критерий удовлетворения запроса – определено в каком случае можно считать, что ответ на запрос получен;
- в каком контексте следует искать ответ на запрос – что мы ожидаем от полученного результата и каковы его последствия и что делать, если ответ будет отрицательным.

В задачном подходе считается, что назначением ИИ является **автоматизация решения задач**, понимая термин «автоматизация» в самом широком смысле, что бы запросы/решения формулировались в терминах **исполнимых спецификаций**.

Базовой моделью вычислений была выбрана Σ -определимость вычислений и проверка истинности Σ -формул на конструктивной модели M и ее списочной надстройке $NW(M)$.
Yu.L. Ershov, S.S. Goncharov & D.I. Sviridenko, "Semantic programming" // Information Processing 86: Proceedings of the IFIP 10th World Computer Congress, IFIP Congress Series, vol. 10, Elsevier Science, Dublin, 1986, pp. 1093–1100.*

Goncharov S.S., Sviridenko D.I. Theoretical aspects of Σ -programming. Lecture Notes in Computer Science, 1986, vol. 215, p. 169–179.

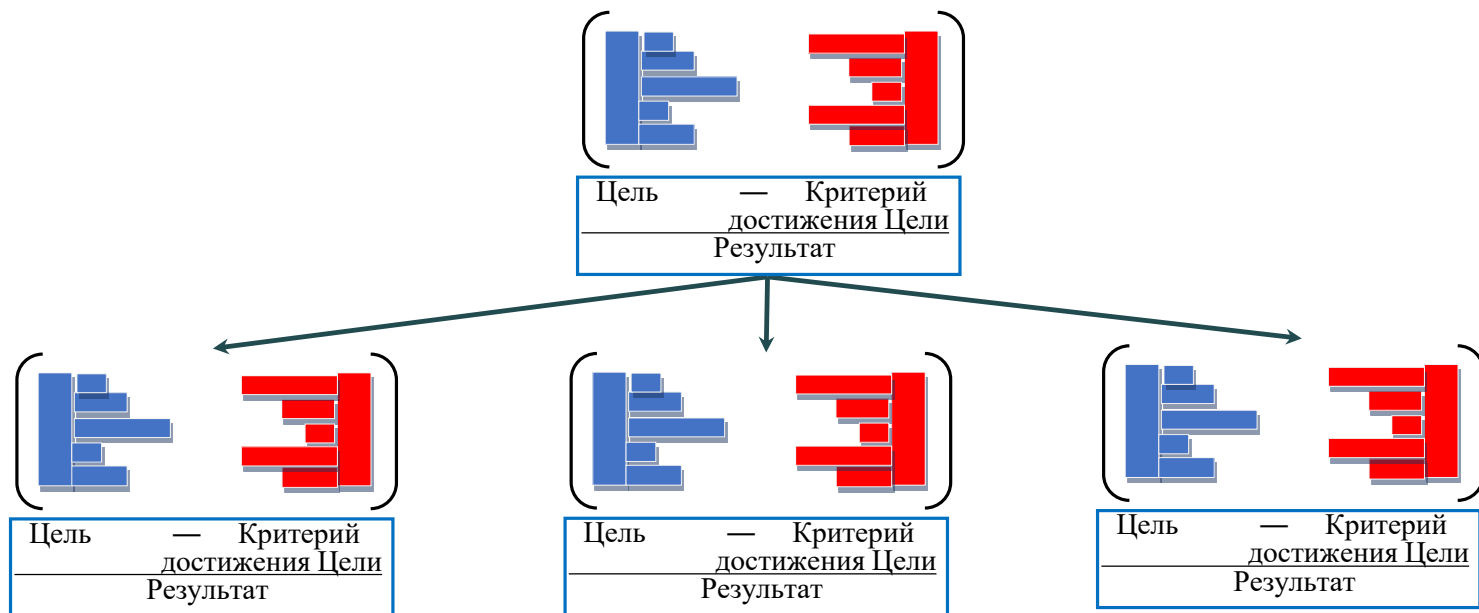
Goncharov S.S., Sviridenko D.I. Σ -programming / Transl., II. Ser., Am. Math. Soc., 1989, v.142, p.101–121.

Е. Е. Витяев, С. С. Гончаров, Д. И. Свириденко. О задачном подходе в искусственном интеллекте // Сибирский философский журнал. 2019. Т. 17, № 4.

Автоматизация решения задач – семантическое моделирование

Определение задач в терминах семантического моделирования.

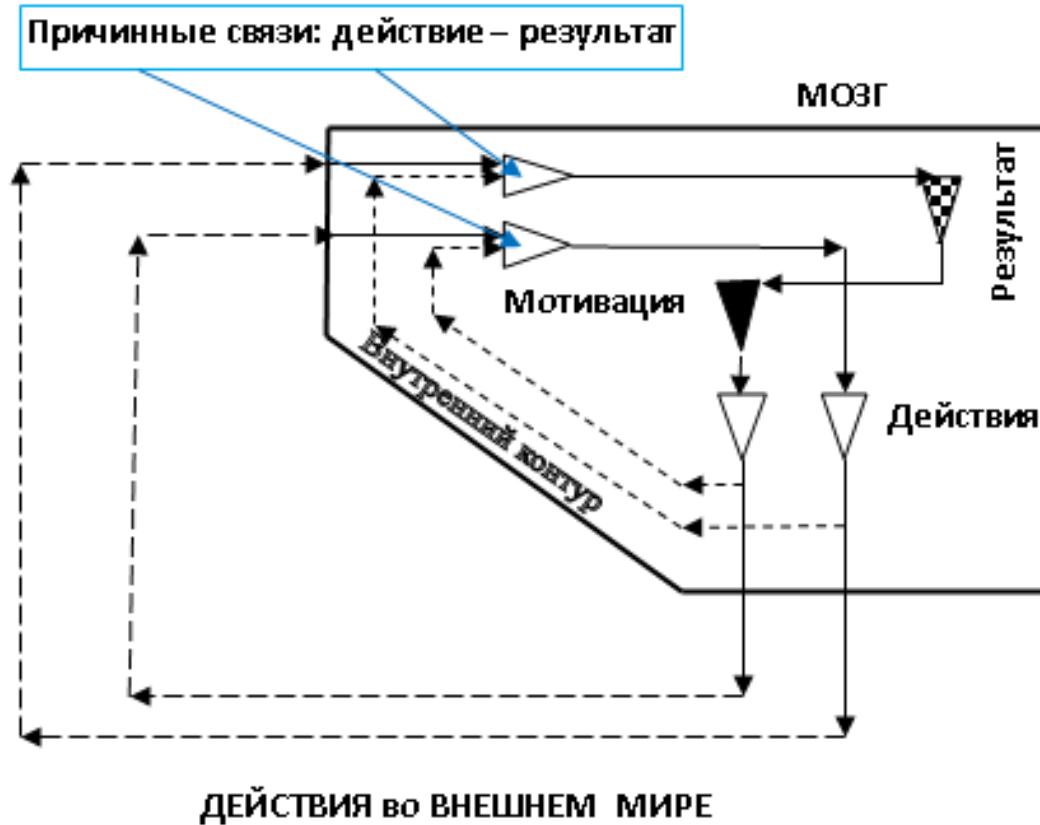
- Предполагается, что в нашем распоряжении имеется *многосортная конструктивная модель* M вместе со своей списочной надстройкой $NW(M)$, выступающая как некий базовый вычислитель.
Модель предметной области рассматриваемой задачи формулируется в сигнатуре языка исчисления предикатов этой базовой конструктивной модели M вместе с ее списочной надстройкой $NW(M)$ как набор Σ -определений, т.е. Σ -формул и Σ -термов этого языка. При этом допускаются рекурсивные схемы Σ -определений с некоторыми ограничениями на вхождения в них определяемых предикатов и термов.
- *Запрос* к модели предметной области определяется также как Σ -формула, в записи которой могут использоваться как сигнатурные конструкции базовой конструктивной модели, так и определяемые предикаты, и термы предметной области.
- Под *Результатом решения* так сформулированной задачи понимается набор констант, делающий Σ -формулу запроса при означивании ее переменных константами истинной на модели предметной области. Истинность Σ -формулы запроса, получаемой подстановкой констант вместо переменных, и есть *критерий решенности задачи*.
- Результатов решения может оказаться несколько и тогда есть возможность выбрать в некотором смысле наилучшее решение с учетом контекста задачи.



В *семантическом моделировании формализация ТФС* как теории решения мозгом задачи по удовлетворению своих потребностей формулируется так:

- *Модель* предметной области – это *опыт*, который получается в *результате обучения* и в котором осуществляется целенаправленное поведение;
- *Запрос* (Цель-ключ, обозначенный синим) к предметной области – это потребность, которую надо удовлетворить. Обозначим ее предикатом P_0 ;
- *Принятие решений* осуществляется выбором *конкретного плана действий*, на основании опыта и модели внешнего мира, приводящего к объекту a такого что $P_0(a)$.
- *Конкретный план действий* преобразует запрос-цель в *конкретный запрос*, дающий план нахождения требуемого объекта a вместе с совокупностью подзапросов-целей и иерархией соответствующих под результатов (*акцептор результатов действий*).
- *Результатом решения* задачи является объект, воспринимаемый «совокупностью афферентных раздражений» a (красный блок), на которых предикат P_0 истинен $P_0(a)$ (потребность удовлетворяется), – это и есть *критерий решения задачи*.

Причинность и Теория Функциональных Систем



«Речь идет о коллатеральных ответвлениях пирамидного тракта, отводящих ко многим нейронам “копии” тех эфферентных посылок, которые выходят на пирамидный тракт ... »
П.К.Анохин.

Принцип опережающего отражения действительности: «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.

ПРЕДСКАЗАНИЕ

Covering Law Model: Задача предсказания (объяснения) состоит в демонстрации того, что предсказываемый факт является частным случаем закона.

Есть два случая предсказания (объяснения):

1. Deductive-Nomological (D-N), объяснение основанное на факте и законах;
2. Inductive-Statistical (I-S) объяснение основанное на факте и вероятностных законах.

$\frac{L_1, \dots, L_m}{C_1, \dots, C_n} \quad \text{DEDUCTIVE-NOMOLOGICAL MODEL}$	$\frac{L_1, \dots, L_m}{C_1, \dots, C_n} \quad \text{INDUCTIVE-STATISTICAL MODEL}$
$\frac{C_1, \dots, C_n}{G}$	$\frac{C_1, \dots, C_n}{G} \quad [r]$

L_1, \dots, L_m – множество законов; C_1, \dots, C_n – множество фактов;

G – объясняемое/предсказываемое высказывание;

$L_1, \dots, L_m, C_1, \dots, C_n$ непротиворечиво и $L_1, \dots, L_m, C_1, \dots, C_n \vdash G$;

$L_1, \dots, L_m \not\models G, C_1, \dots, C_n \not\models G$; L_1, \dots, L_m законы имеют кванторы всеобщности, C_1, \dots, C_n, G – бескванторные формулы.

Поиск решений в семантическом моделировании может осуществляться как посредством D-N вывода, если используются точные законы L_1, \dots, L_m , так и посредством I-S вывода, если используются вероятностные законы.

В этом случае мы вместе с выводом получаем *прогноз достижения результата*.

Поиск решений в семантическом моделировании может также осуществляться некоторым перебором констант в модели, что соответствует методу «проб и ошибок» в целенаправленном поведении.

Проблема статистической двусмысленности: В индуктивном выводе получаются утверждения из которых выводятся противоречивые утверждения.

(Л1) - ‘Почти все случаи заболевания стрептококком быстро вылечиваются инъекцией пенициллина’;

(Л2) - ‘Почти всегда **устойчивая к пенициллину** стрептококковая инфекция не вылечивается после инъекции пенициллина’;

(С1) - ‘Джейн Джонс заболел стрептококковой инфекцией’;

(С2) - ‘Джейн Джонс получил инъекцию пенициллина’;

(С3) - ‘Джейн Джонс имеет устойчивую к пенициллину стрептококковую инфекцию’.

Объяснение 1		Объяснение 2	
Л1	[r]	Л2	[r]
С1,С2		С2,С3	
Е		¬Е	

$$\frac{p(G;F) = r}{F(a)} \quad [r]$$

Карл Гемпель. Требование Максимальной Специфичности (ТМС) :

Если для класса Н выполнены утверждения: $\forall x(H(x) \Rightarrow F(x))$, $H(a)$, то существует закон $p(G;H) = r'$ такой что $r = r'$.

Проблема оценок вероятностей выводимых предсказаний. Они могут быстро доходить до нуля и давать предсказания в оценкой 0, что не является предсказанием.

Эта проблема известна как проблема синтеза логики и вероятности, рассматриваемая в серии конференций Prolog (Probability+Logic), 2003, 2005, 2009, 2011, 2013, 2015, 2017, 2019.

Вывод предсказаний в логическом программировании

Предсказание формулируется как запрос G к программе,

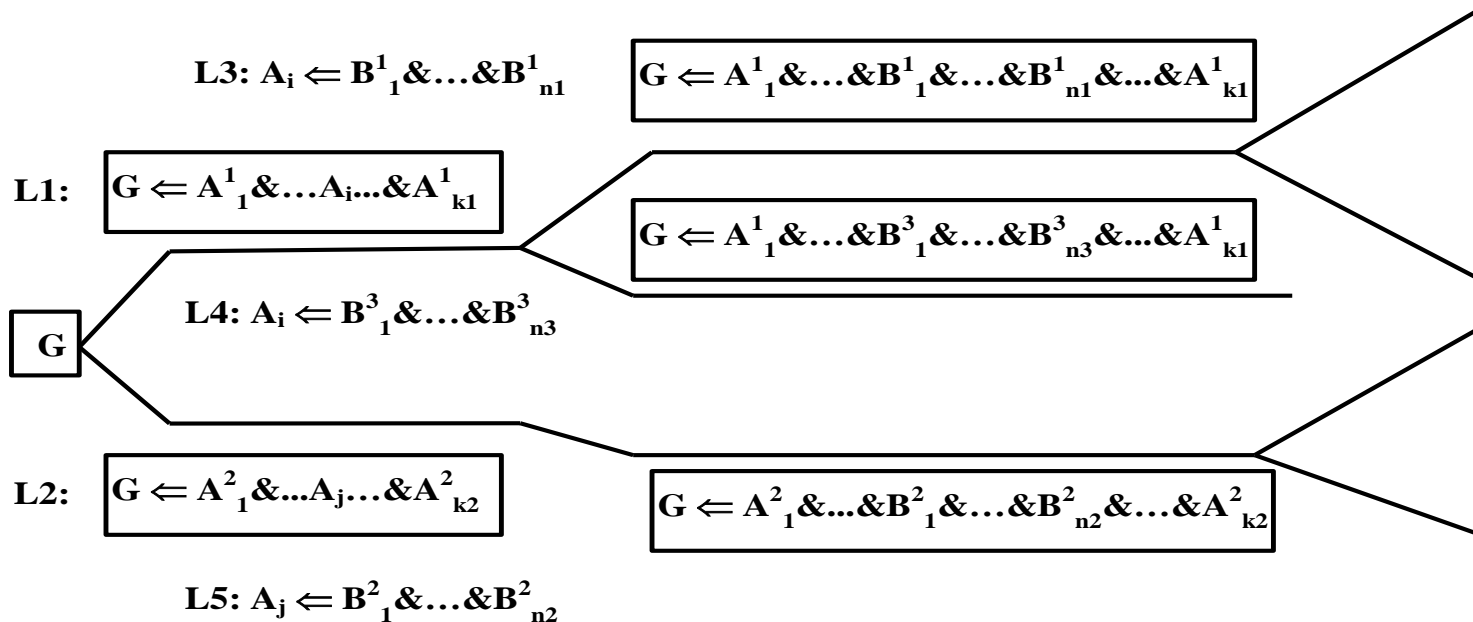
включающей (статистические) законы L_1, \dots, L_m и факты C_1, \dots, C_n

Если вывод программой запроса G успешен, то:

1. Верно, что $\{L_1, \dots, L_m, C_1, \dots, C_n\} \vdash \exists x_1, \dots, x_n G$;

2. Программа вычисляет набор термов $t_1, \dots, t_n : \{L_1, \dots, L_m, C_1, \dots, C_n\} \vdash G[x_1/t_1, \dots, x_n/t_n]$.

Процесс вычисления можно представить в виде дерева.



Предсказание с точки зрения семантического моделирования

Рассмотрим теоретико-модельную семантику логических программ, где факты являются высказываниями эмпирической системы $\mathfrak{S} = \langle \mathbf{A}; \Omega_{\mathfrak{S}} \rangle$, представляющей предметную область.

В этом случае процесс *логического вывода предсказания* можно обобщить, рассматривая логический вывод как *вычисление истинности* предсказываемого факта G , состоящее в обнаружении фактов C_1, \dots, C_n эмпирической системы \mathfrak{S} , из истинности которых по законам L_1, \dots, L_m вычисляется истинность предсказываемого высказывания.

Тогда вместе с выводом факта G , будет иметь место истинность:

$$\mathfrak{S} \models G[x_1/a_1, \dots, x_n/a_n].$$

При таком взгляде на вывод его можно обобщить, определяя новые взаимоотношения высказываний и модели. Можно рассмотреть вывод не только как проверку истинности на модели, но и как поиск фактов в модели, предсказывающих интересующее нас высказывание с максимальной вероятностью, или специфичностью.

Такой вывод будем называть семантическим.

Интеграционный проект СО РАН «Установление сходства структур реального физического мира со структурами вычислимости приобретает в современной науке столь большую значимость, что многие исследователи считают понятие вычислимости новой парадигмой философии науки».

Индуктивный вывод теории и знаний ПО

Пусть ПО представлена эмпирической системой $\mathfrak{T} = \langle A; \Omega_{\mathfrak{T}} \rangle$.

$\text{Th}(\mathfrak{T})$ – **теория ПО** – множество всех истинных на \mathfrak{T} универсальных формул.

Известно, что $\text{Th}(\mathfrak{T})$ может быть представлена как множество всех истинных на \mathfrak{T} правил вида $C = (L_1 \& \dots \& L_k \Rightarrow L_0)$, где L_j – литеры;

Правило $C = (A_1 \& \dots \& A_n \Rightarrow A_0)$ будем называть **под правилом** правила

$C' = (L_1 \& \dots \& L_k \Rightarrow A_0)$, если $\{A_1, \dots, A_n\} \subset \{L_1, \dots, L_k\}$, $0 \leq n < k$.

Законом на $\mathfrak{T} = \langle A, W \rangle$ будем называть правило $C = (L_1 \& \dots \& L_k \Rightarrow L_0)$, удовлетворяющее условиям:

1. C истинно на \mathfrak{T} ;
2. посылка правила не всегда ложна на \mathfrak{T} ;
3. каждое подправило правила C ложно на \mathfrak{T} .

L - множество всех законов на \mathfrak{T} .

Теорема. $L \vdash \text{Th}(\mathfrak{T})$.

Вероятностным законом на \mathfrak{T} будем называть правило C , условная вероятность которого строго больше условных вероятностей всех его подправил.

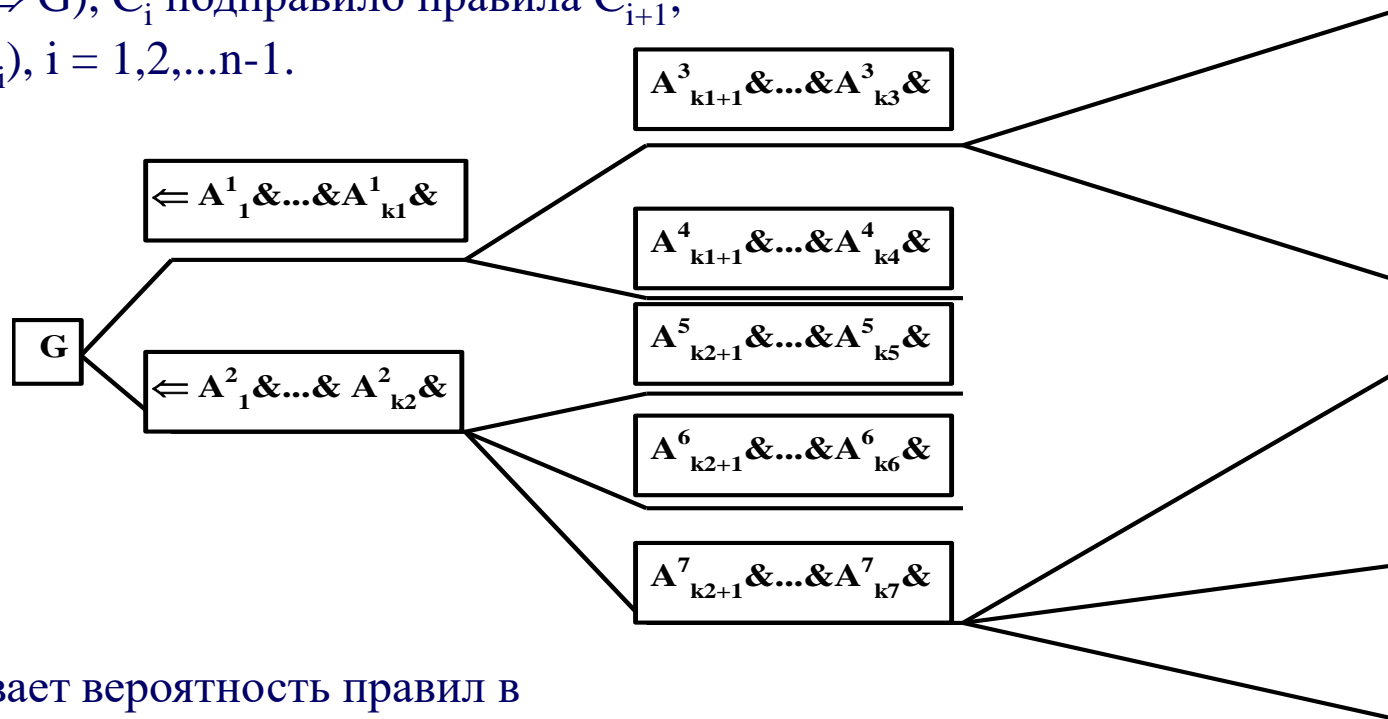
Сильнейшим Вероятностным Законом (СВЗ) будем называть вероятностный закон, не являющийся под правилом другого вероятностного закона.

LP - множество всех вероятностных законов на \mathfrak{T} .

Утверждение. $L \subset LP$. Множество L дает теорию, а LP – знания о ПО.

СЕМАНТИЧЕСКИЙ ВЕРОЯТНОСТНЫЙ ВЫВОД

Семантическим Вероятностным Выводом (СВВ) называется последовательность вероятностных законов: $C_1 \supset C_2 \supset \dots \supset C_n$, где C_n – сильнейший вероятностный закон $C_i = (A_1^i \& \dots \& A_{k_i}^i \Rightarrow G)$, C_i подправило правила C_{i+1} ;
 $\text{Prob}(C_{i+1}) > \text{Prob}(C_i)$, $i = 1, 2, \dots, n-1$.



СВВ строго увеличивает вероятность правил в процессе вывода, в отличие от предсказаний в логическом программировании.

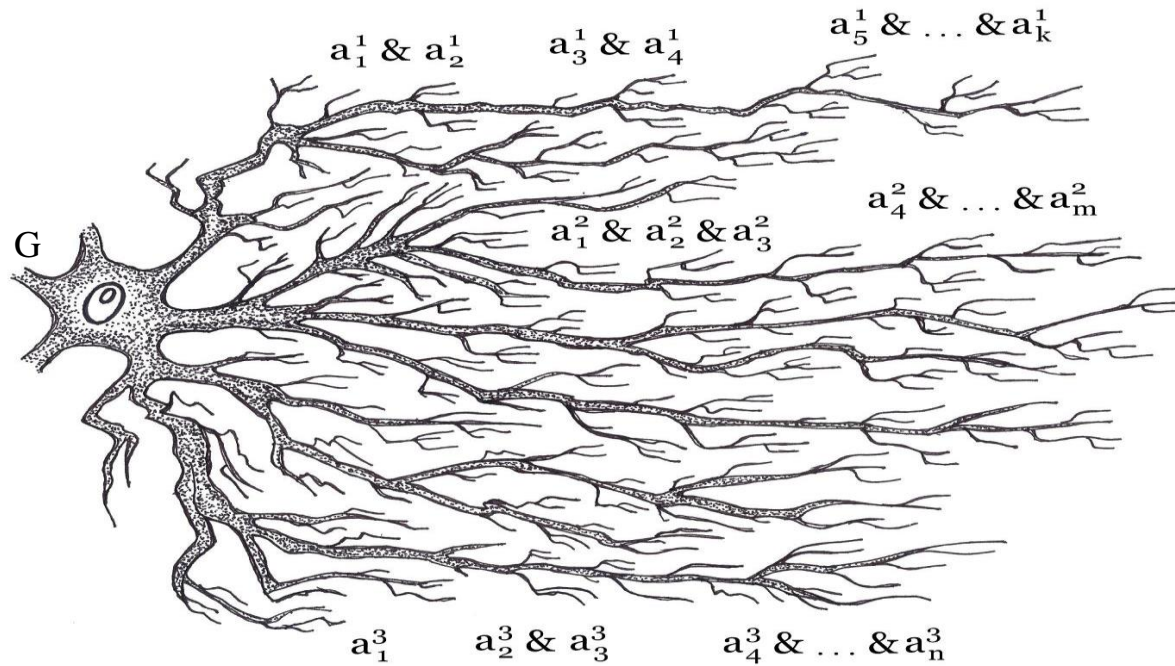
Максимально специфическим правилом МСП(G) вывода атома G назовем СВЗ-правило дерева вывода G , имеющее максимальное значение условной вероятности.

Теорема. I-S, использующий только МСП правила непротиворечив. $L \subset \text{МСП} \subset \text{LP}$.

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.

Семантический вероятностный вывод как формальная модель нейрона

Удовлетворяет правилу Хебба



В соответствии с СВВ, новые стимулы добавляются в вывод, если они увеличивают его условную вероятность. Это формирует замыкания условных связей на уровне нейрона.

Нейрон может обнаружить Максимально Специфические Правила (МСП), предсказывающие без противоречий.

Нейроны быстрее по времени отвечают на максимально вероятные условные связи.

В схеме причинных связей, объясняющих ТФС, можно предполагать такие нейроны.

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.

Переключающая функция эмоций

Семантический вероятностный вывод формализует также понятия *вероятностного прогнозирования* и *вероятности*, введенные в работах И.М. Фейгенберга и используемые в Информационной Теории Эмоций (ИТЭ) 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.

Симонов П. В. Высшая нервная деятельность человека (мотивационно-эмоциональные аспекты). М.: Наука, 1975.

Симонов П. В. Эмоциональный мозг. М.: Наука, 1981.

В ТФС принятие решений осуществляется выбором одного конкретного плана действий с учетом вероятности этого прогноза и возникающих эмоций, что является важным дополнением информационной теории эмоций.

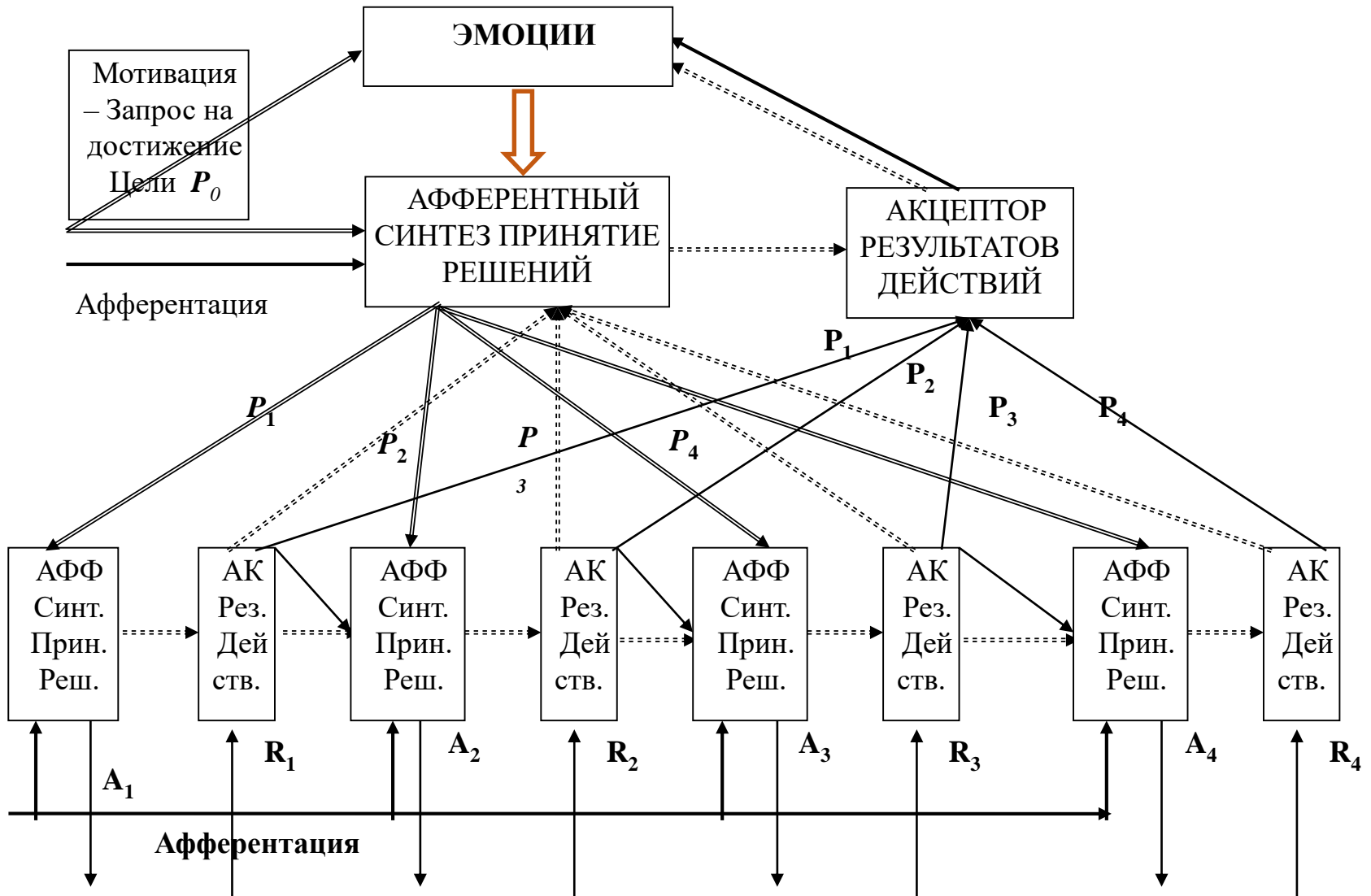
Эмоции – **необходимый** критерий выбора между различными способами достижения цели, учитывающий вероятность достижения цели, сложность и трудоёмкость, а также значимость санкционирующей афферентации от удовлетворения потребности.

Поэтому **принятие решений** и выбор конкретного способа достижения цели осуществляется **переключающей функцией эмоций**.

В Искусственном Интеллекте для интеллектуальных агентов такой критерий формулируется как **полезность**. Это самостоятельное направление исследований:

Фишберн, Питер С. Теория полезности для принятия решений // Пер. с англ. В.Н. Воробьевой и А.Я. Кируты ; Под ред. Н.Н. Воробьева. - Москва : Наука, 1978.

Переключающая функция эмоций



Многоуровневая организация движений по Н.А. Бернштейну

В работе Н.А. Бернштейна: «О построении движений» исследована многоуровневая организация движений: «Уровень организации движений В – владение телом»; «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 ...».

Определим эти уровни через цепочки действий и достигаемые ими результаты.

Введем понятие **вероятностной цели-результата (цепочки) действий** так, чтобы они могли формироваться в процессе обучения автоматически:

- 1) цель-результат увеличивает вероятность достижения конечного результата;
- 2) цель-результат обладает свойством ветвления: если в процессе действий достигнут некоторый цель-результат, то дальнейшие действия могут развиваться по-разному в соответствии с иерархией целей-результатов;
- 3) при достижении цели, результат должен фиксироваться некоторым набором признаков, фиксирующем законченность цепочки действий и возможность перехода к следующей цепочке действий;
- 4) достижение цели-результата подкрепляется эмоцией, которая возникает вследствие увеличения вероятности достижения конечной цели.

Достижение этих целей-результатов можно рассматривать как **автоматическое формирование функциональных подсистем**, достигающих этих целей-результатов.

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

Уточнение понятия Задача

Модель предметной области – это **модель внешнего мира**, которая получается в **результате** обнаружением МСП во внутреннем контуре работы мозга;

Запрос – это потребность, которую надо удовлетворить, обозначенную предикатом P_0 ;

Решением задачи является объект, ситуация, воспринимаемые «совокупностью афферентных раздражений» a , на которых предикат P_0 истинен $P_0(a)$ (удовлетворяет потребность), – это и есть **критерий решения задачи**.

Таких решений a может быть несколько, выбор конкретного решения осуществляется **переключающей функцией эмоций**.

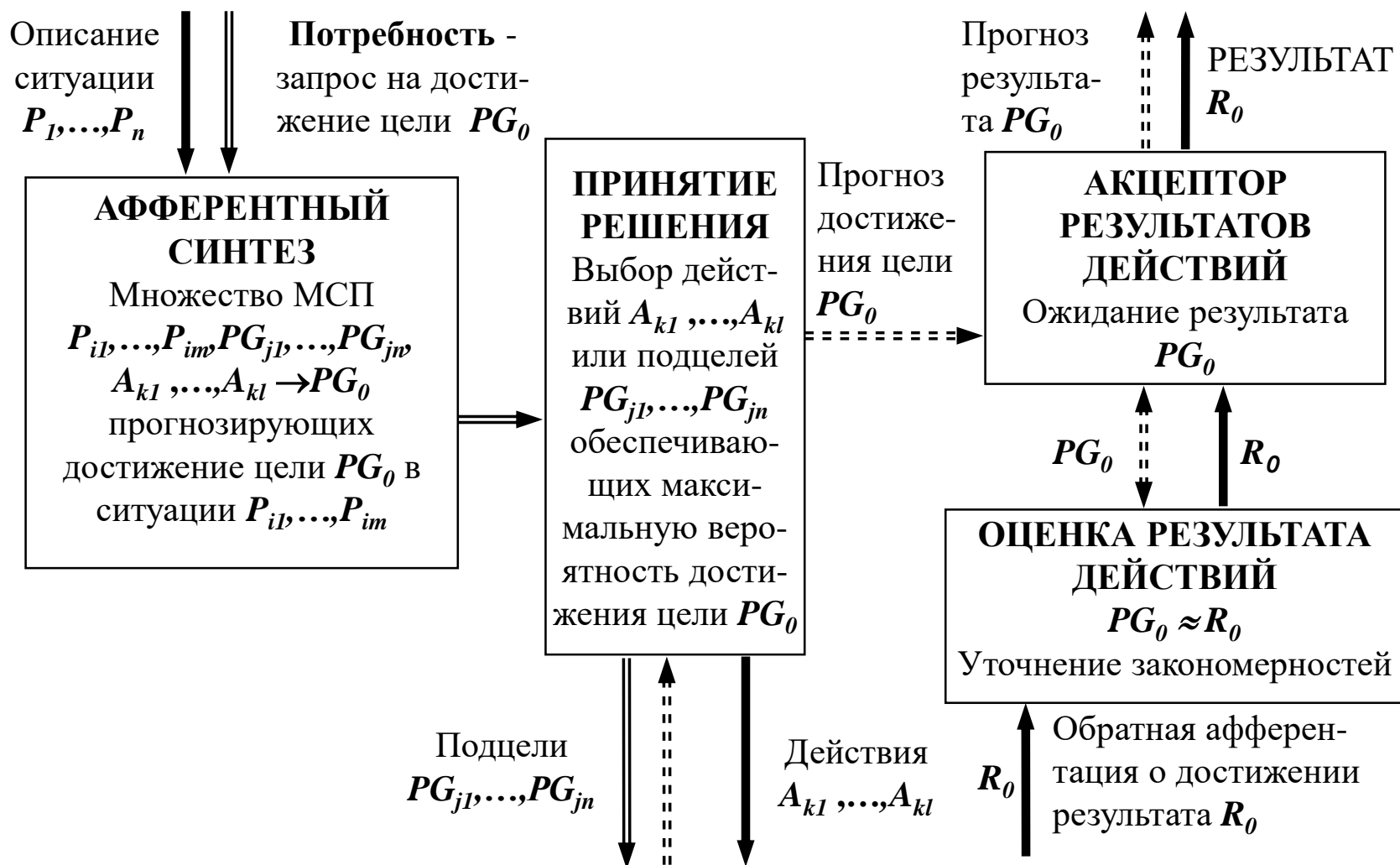
Принятие решений осуществляется выбором **конкретного плана действий** с учетом вероятностного прогноза и максимальной эмоциональной реакции:

- на основании имеющегося опыта и имеющихся МСП, предсказывающих по модели внешнего мира достижение цели (достижения объекта, ситуации a , делающие предикат P_0 истинным);
- иерархическим планированием достижение цели, используя вероятностные цели-результаты и имеющиеся МСП;
- семантическим моделированием, осуществляя предсказание достижение цели D-N и I-S выводом, используя МСП и имеющиеся знания о предметной области.

Конкретный план действий преобразует запрос-цель в **конкретный запрос-цель**, дающий способ нахождения требуемого объекта a вместе с совокупностью подзапросов-целей и иерархией целей-результатов (**акцептор результатов действий**).

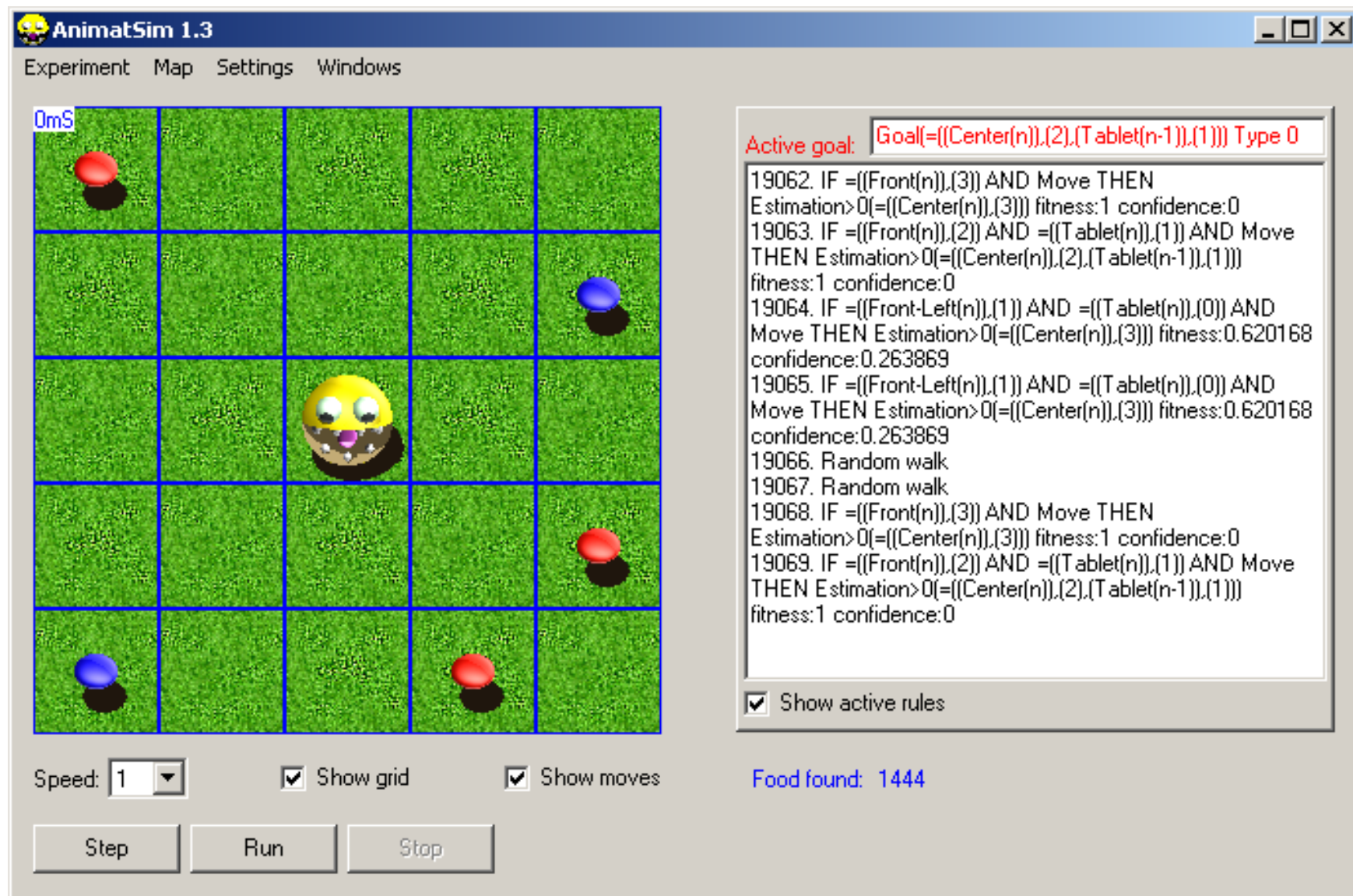
СХЕМА РЕШЕНИЯ ЗАДАЧ И РАБОТЫ ФУНКЦИОНАЛЬНЫХ СИСТЕМ В ТЕРМИНАХ МСП

$\text{Prob}(PG_0|R) = \text{Prob}(R)\text{Prob}(PG_{j1}) \dots \text{Prob}(PG_{jn})$, где $R = P_{i1}, \dots, P_{im}, PG_{j1}, \dots, PG_{jn}, A_{k1}, \dots, A_{kl} \rightarrow PG_0$



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.

ПРОГРАММА, МОДЕЛИРУЮЩАЯ АНИМАТ И СРЕДУ ЕГО ОБИТАНИЯ



ОПИСАНИЕ АНИМАТА

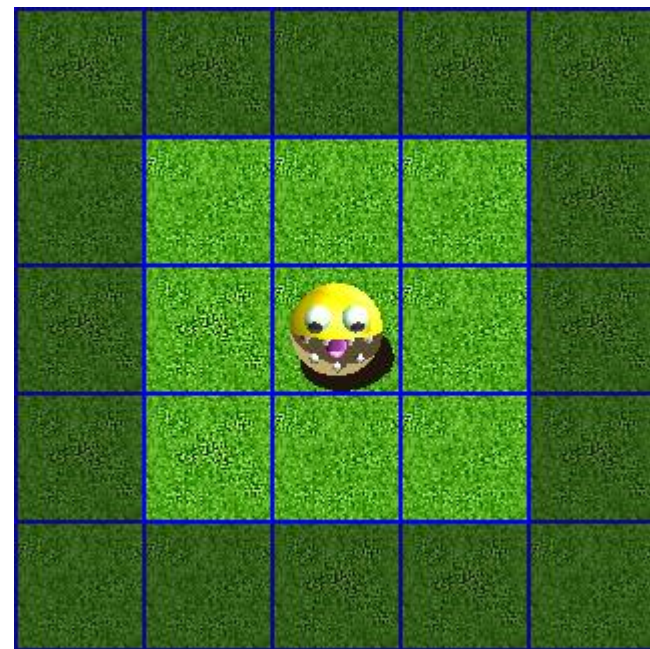
Действия

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

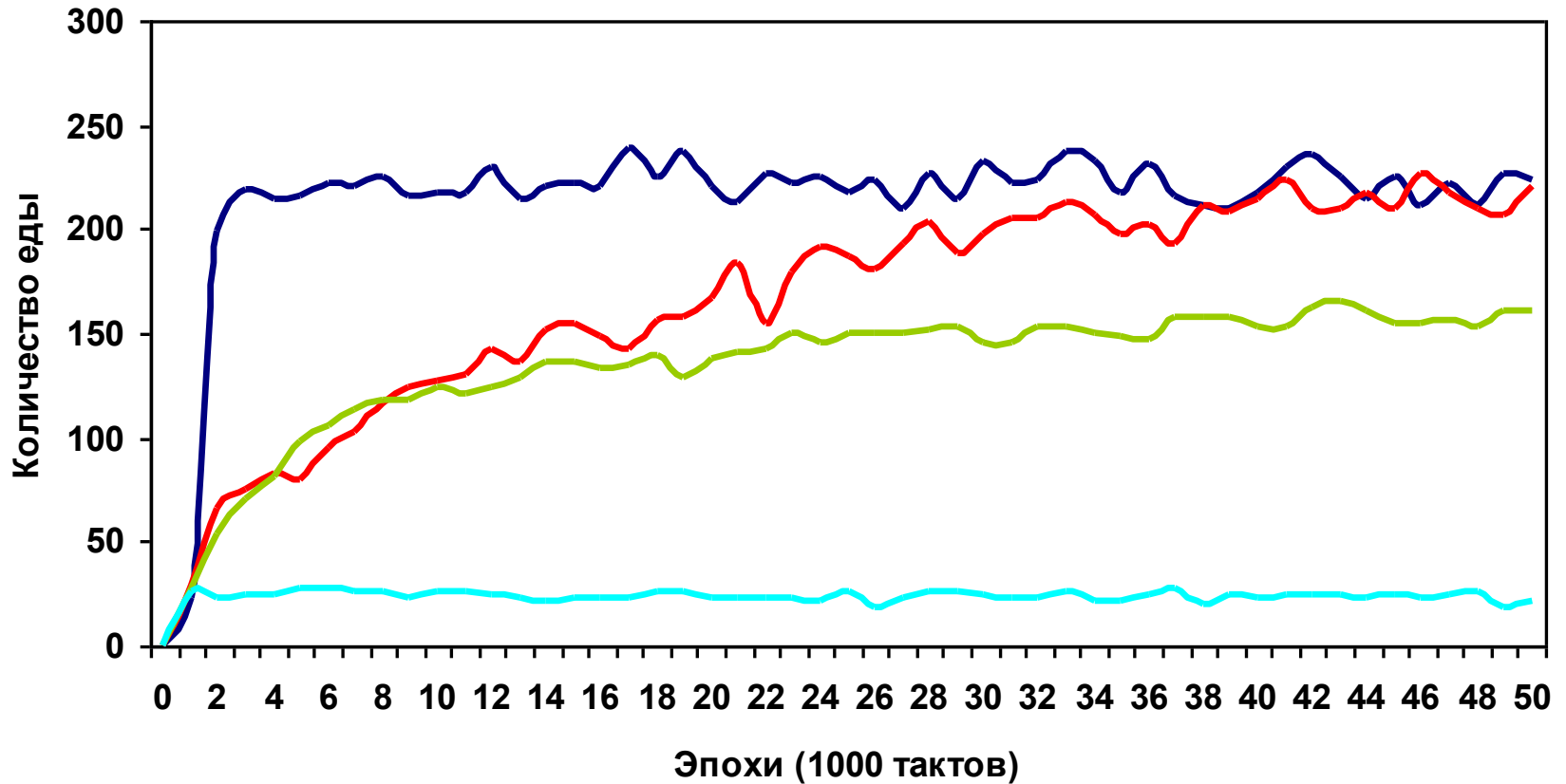
$\langle s = \text{«пусто»} \rangle$, $\langle s = \text{«препятствие»} \rangle$, $\langle s = \text{«еда»} \rangle$, $\langle s = \text{«таблетка»} \rangle$, где s – окружающая клетка, и $\langle \text{«есть таблетка»} = \text{«да»} \rangle$.

Основная цель

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

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

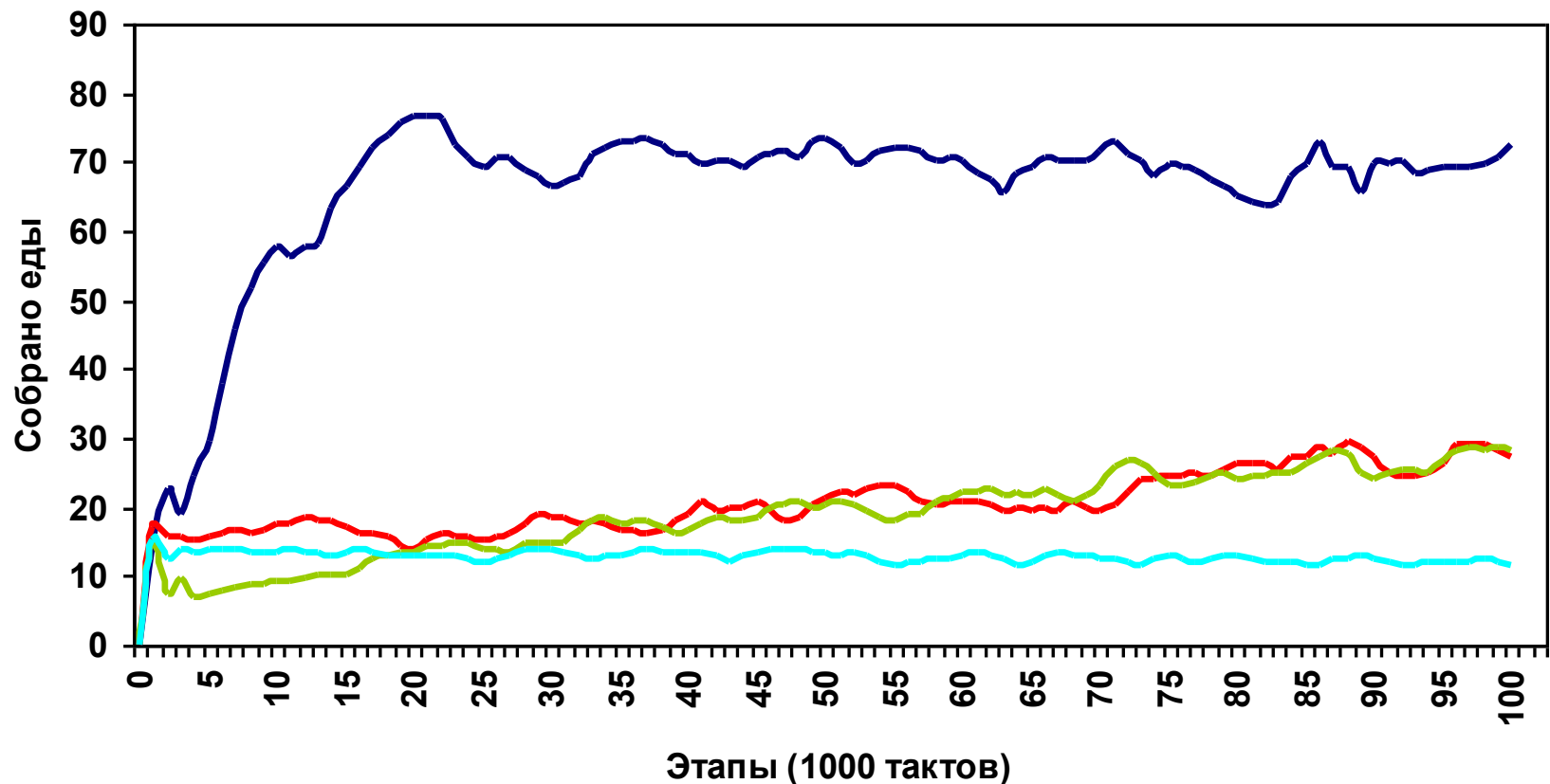
РЕЗУЛЬТАТЫ СРАВНЕНИЯ НА ПРИМЕРЕ РЕШЕНИЯ КЛАССИЧЕСКОЙ ЗАДАЧИ ФУРАЖИРОВАНИЯ



— 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



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

ПОДХОДЫ К МОДЕЛИРОВАНИЮ МЫШЛЕНИЯ



НАУКИ об
ИСКУССТВЕННОМ



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.

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.

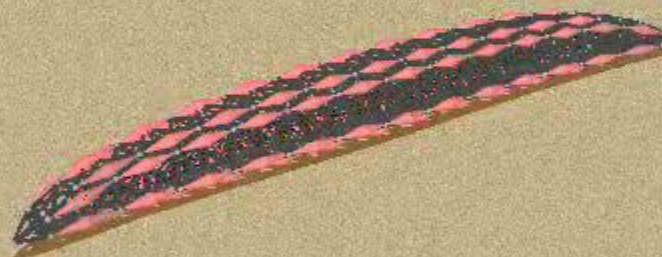
Evgenii Vityaev. Consciousness as a
logically consistent and prognostic model of
reality // Cognitive Systems Research, 2019
Elsevier, 59 (2020), 231–246.

Витяев Е.Е. Сознание - логически
непротиворечивая прогностическая
модель реальности // Нелинейная
динамика в когнитивных исследованиях
(Труды V Всероссийской конференции по
когнитивным наукам), Нижний Новгород,
ИФП РАН, 2017, с. 67-70.

Physics Simulation Engine 1.0 beta, Palyanov Andrey 2007-2010



Cyber Elegans 2.087 beta, build 30.12.2010 (c) Palyanov Andrey, Khayrulin Serge
A.P. Ershov Institute of Informatics Systems SB RAS, Novosibirsk. e-mail: palya
F1 = help show / hide



Time: 86



Semantic Mouse

Round 40

08:18:16 DEBUG: goal < GOAL_CLOSER >

rule < GOAL_SOMEWHERE_AHEAD_RIGHT > p=0.999999999999992 a=move

goal < GOAL_SOMEWHERE_AHEAD_RIGHT >

rule < GOAL_SOMEWHERE_BEHIND_RIGHT > p=1.0 a=turn right

goal < GOAL_SOMEWHERE_BEHIND_RIGHT >

rule < OBSTACLE_LEFT GOAL_SOMEWHERE_AHEAD_RIGHT > p=0.75 a=move

rule < GOAL_SOMEWHERE_AHEAD_LEFT > p=1.0 a=turn left

goal < GOAL_SOMEWHERE_AHEAD_LEFT >

rule < GOAL_SOMEWHERE_BEHIND_LEFT > p=1.0 a=turn left

rule < GOAL_SOMEWHERE_AHEAD_RIGHT > p=1.0 a=turn right

rule < GOAL_SOMEWHERE_AHEAD > p=0.9999999999999948 a=move

goal < GOAL_SOMEWHERE_AHEAD >

rule < GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

goal < GOAL_SOMEWHERE_RIGHT >

rule < GOAL_SOMEWHERE_AHEAD_RIGHT > p=0.8 a=move

rule < GOAL_SOMEWHERE_AHEAD > p=1.0 a=turn left

goal < GOAL_SOMEWHERE_AHEAD >

rule < GOAL_RIGHT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < OBSTACLE_LEFT GOAL_SOMEWHERE_AHEAD > p=1.0 a=move

rule < OBSTACLE_AHEAD GOAL_RIGHT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < GOAL_SOMEWHERE_LEFT > p=1.0 a=turn left

rule < OBSTACLE_AHEAD OBSTACLE_LEFT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < OBSTACLE_LEFT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < OBSTACLE_LEFT GOAL_RIGHT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < OBSTACLE_RIGHT GOAL_SOMEWHERE_AHEAD > p=1.0 a=move

rule < GOAL_SOMEWHERE_AHEAD > p=1.0 a=move

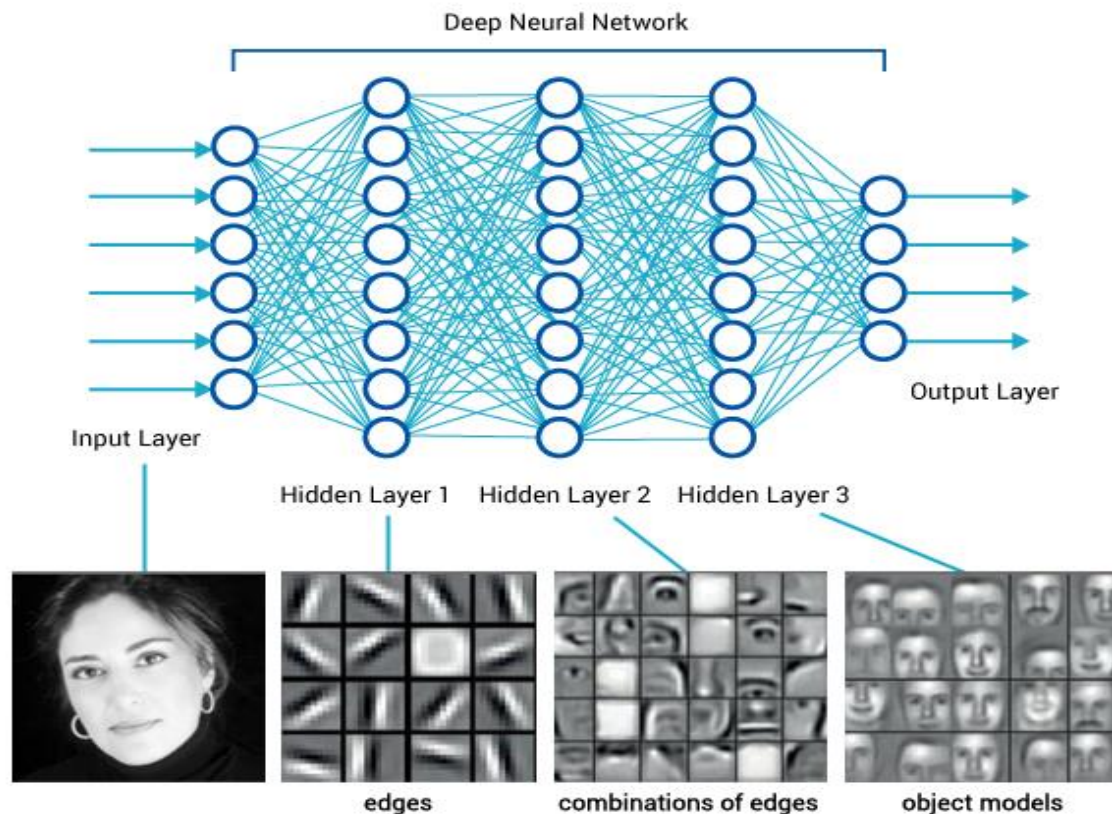
rule < OBSTACLE_AHEAD OBSTACLE_LEFT GOAL_RIGHT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn right

rule < GOAL_SOMEWHERE_BEHIND > p=1.0 a=turn right

goal < GOAL_SOMEWHERE_BEHIND >

rule < OBSTACLE_AHEAD OBSTACLE_LEFT GOAL_RIGHT GOAL_SOMEWHERE_RIGHT > p=1.0 a=turn left

Deep learning



Дж. Гибсон «Экологическая теория восприятия»: Мир устроен иерархически: «мелкие элементы содержатся в более крупных ... поэтому я ввожу для него специальный термин: **встроенность**». «... 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».

«Естественные» понятия в когнитивных науках

В работах Eleanor Rosch были сформулированы принципы категоризации «естественных» категорий:

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 теорией** категоризации (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, используя Байесовские сети. Однако Байесовские сети не поддерживают циклов и поэтому не могут моделировать циклические причинные связи.

Нами предполагается *новый математический аппарат – вероятностное обобщение формальных понятий* для формализации «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 порождает причинные модели.

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».

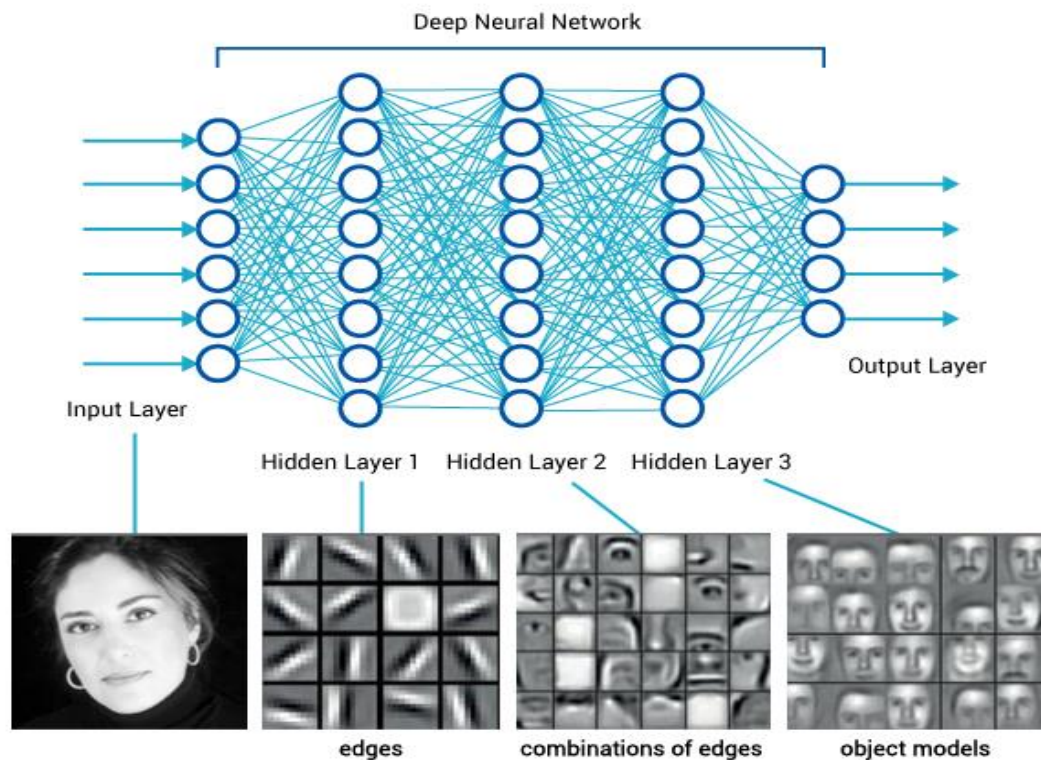
Этапы вероятностного обобщения формальных понятий :

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 “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

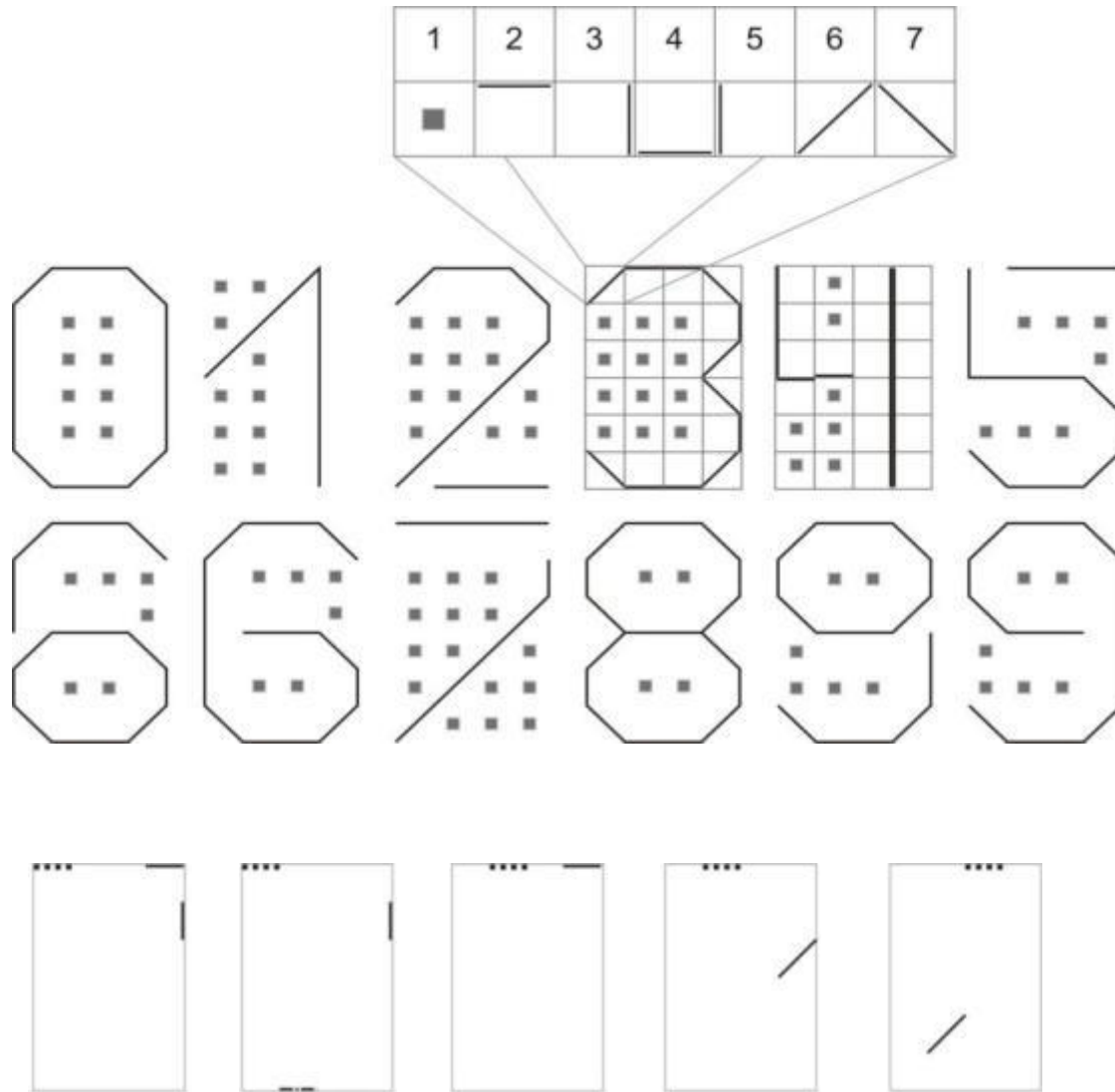
E. E. Vityaev, V. V. Martinovich. Probabilistic Formal Concepts with Negation // A. Voronkov, I. Virbitskaite (Eds.): PCI 2014, LNCS 8974, 2015, pp.385-399.

Probabilistic explainable deep learning



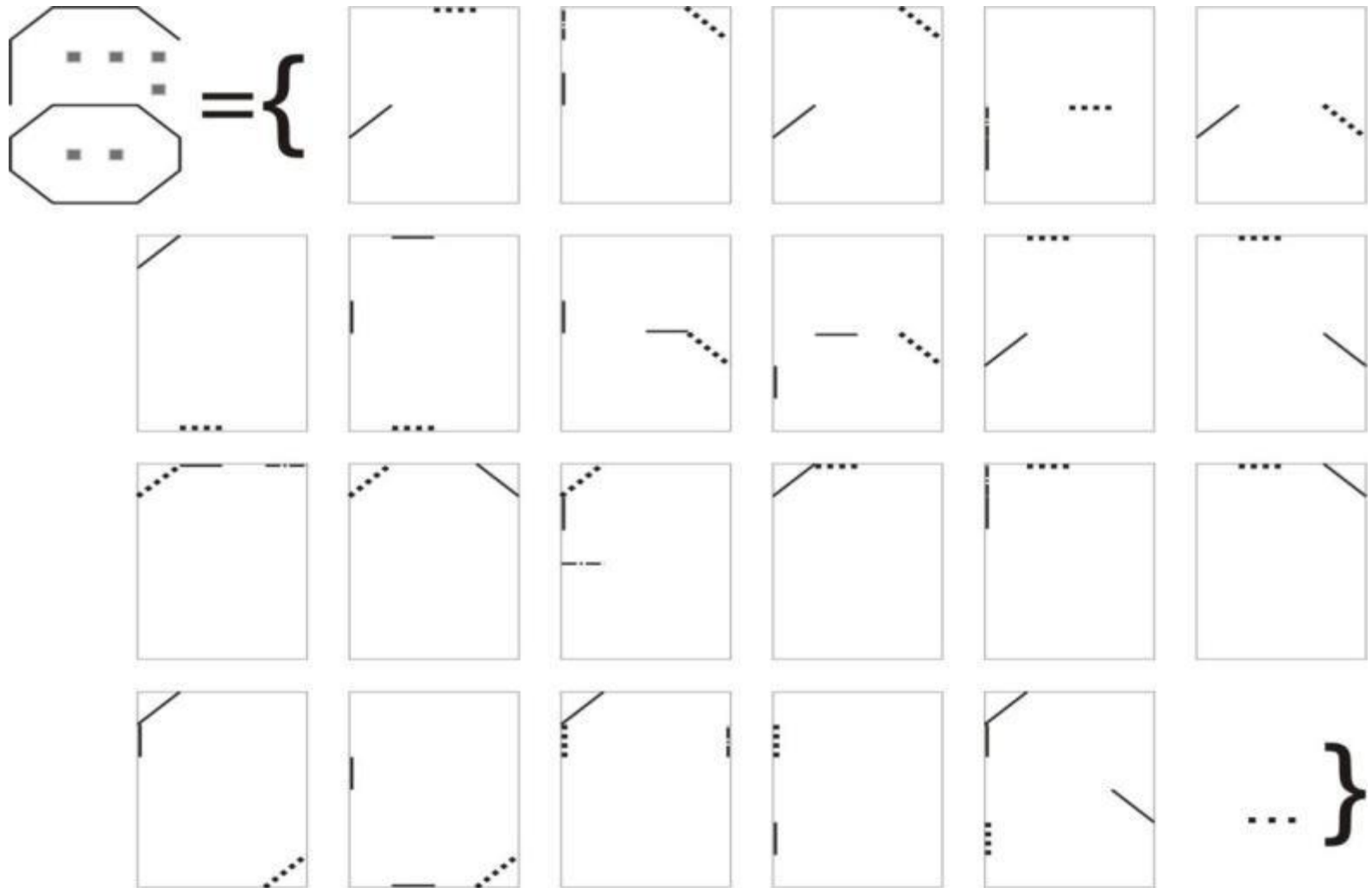
- автоматически обнаруживает иерархию вторичных признаков;
- максимально точно и непротиворечиво формирует классы;
- непрерывно во времени и в пространстве делает предсказания одних свойств по другим и отслеживает их подтверждение;
- Прозрочно и объяснимо.

“Естественные” понятия как вероятностные формальные понятия



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



$$\neg \text{ / } = \text{ / }$$

Regularities use an attribute negation

Интеллектуальный анализ данных с точки зрения задачного подхода

Обнаружение знаний на данных модели (ПО)

Понятие *задачи* включает указание предметной области и знаний о предметной области, зафиксированные в виде ее *модели*.

Рассмотрим проблему *индуктивного вывода новых знаний на данных модели с помощью методов интеллектуальный анализ данных*.

Из данных – свойств, характеристик и атрибутов можно извлечь только те знания, которые интерпретируемы в онтологии ПО.

Сами по себе числовые значения величин знаний не содержат.

Смысл величин содержится в их интерпретации: 5 метров, 5 литров, 5 кг.

Знания о данных задаются шкалами величин – эмпирическими системами

$\mathfrak{S} = \langle \mathbf{A}; \Omega_{\mathfrak{S}} \rangle$, где

\mathbf{A} – множество значений величины;

$\Omega_{\mathfrak{S}}$ - множество отношений и операций, интерпретируемых в онтологии ПО.

Шкала наименований: $\langle \mathbf{A}; =, \approx \rangle$, Шкала порядка $\langle \mathbf{A}; =, <, > \rangle$,

Шкала интервалов $\langle \mathbf{A} \times \mathbf{A}; =, <, > \rangle$, заданная на интервалах $\langle a, b \rangle$,

Шкала отношений $\langle \mathbf{A}; =, <, >, +, -, \cdot, / \rangle$

Шкалы определяются с точностью до *допустимых преобразований шкал*.

Онтология методов Data Mining (DM) и Machine Learning (ML)

Знания получаемые методами Data Mining и Machine Learning интерпетируемы в терминах их онтологии, а не в терминах онтологии ПО.

1-st International Workshop on Philosophies and Methodologies for Knowledge discovery. 22-26 August 2005, Copenhagen, Denmark.

Evgenii Vityaev, Keith Rennolls. Guest Editorial. Philosophies and methodologies for knowledge discovery // *Intelligent Data Analysis*. Special issue on “Philosophies and Methodologies for Knowledge Discovery and Intelligent Data Analysis” eds. Keith Rennolls, Evgenii Vityaev. v.12(2), IOS Press, 2008.

- 1) типов данных, с которыми работает метод;
- 2) априорные классы гипотез, которые проверяет метод на данных.

Для получения знаний методами DM и ML необходимо, чтобы онтология метода интерпетировалась в онтологии ПО:

- 1) типы данных метода, должны интерпетироваться в онтологии ПО;
- 2) априорные классы гипотез должны интерпетироваться в онтологии ПО – метод должен использовать только интерпетируемые в онтологии ПО математические отношения и операции.

Если условия 1-2 не выполнены, то метод не будет инвариантен относительно выбора единиц измерения и допустимых преобразований шкал.

Онтологический (семантический) подход к обнаружению знаний на данных

Онтологический подход к извлечению знаний ПО состоит в том, чтобы:

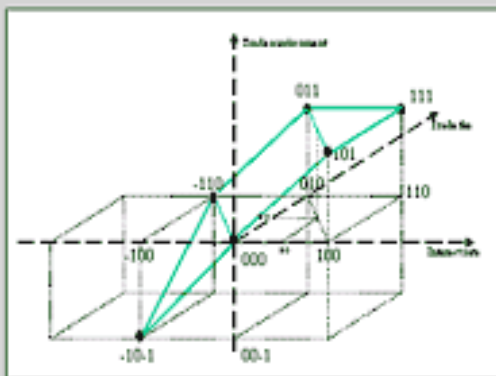
1. определить онтологию данных, используя онтологию предметной области, шкалы величин и теорию измерений;
2. представить данные в онтологии ПО в рамках модели ПО;
3. задать интересующий нас класс гипотез H , который должен задаваться экспертом в предметной области, для которого эти знания нужны.
4. обнаружить знания проверяя гипотезы из H на этих данных, а также:
 - a) теорию ПО;
 - b) все правила с **максимальными условными вероятностями**;
 - c) все **максимально специфические знания**, позволяющие делать **непротиворечивые предсказания**;
5. Разработана система “Discovery”, реализующая семантический вероятностный вывод и обнаруживающая различные классы гипотез H , $L \subset MSR \subset LP$.

Модель предметной области в семантическом моделировании может быть пополнена правилами $L \subset MSR \subset LP$, дающими теорию ПО и вероятностные знания о ПО.

DATA MINING IN FINANCE

Advances in Relational
and Hybrid Methods

by
Boris Kovalerchuk
Evgenii Vityaev



Kluwer Academic Publishers

Е.Е. Витяев



ИЗВЛЕЧЕНИЕ ЗНАНИЙ ИЗ ДАННЫХ
КОМПЬЮТЕРНОЕ ПОЗНАНИЕ
МОДЕЛИ КОГНИТИВНЫХ ПРОЦЕССОВ



ПРИЛОЖЕНИЕ В ФИНАНСОВОМ ПРОГНОЗИРОВАНИИ

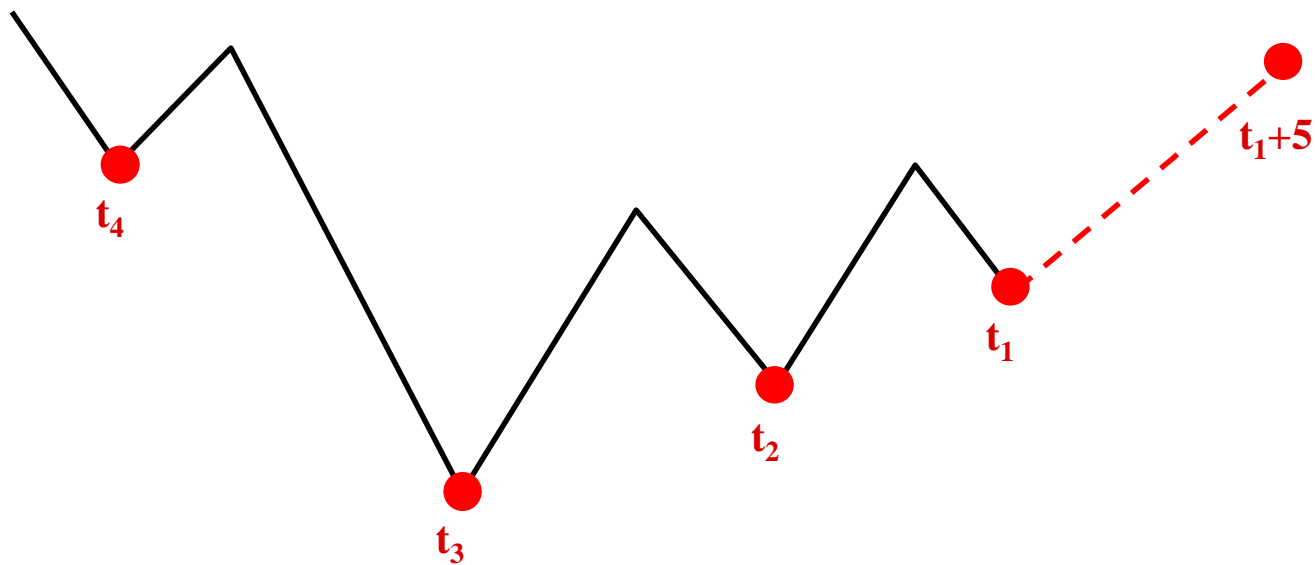
structure1	structure2	structure3	structure4	weekday	week
			forecast for	Friday	forecast week
		forecast for		Thursday	forecast week
		up	up	Wednesday	forecast week
	forecast for	up		Tuesday	forecast week
forecast for	up			Monday	forecast week
up			current day	Friday	current week
				Thursday	current week
		current day	down	Wednesday	current week
current day	current day	down		Tuesday	current week
	down	anchor2	anchor2	Monday	current week
	down anchor1	down		Friday	one week ago
				Thursday	one week ago
				Wednesday	one week ago
				Tuesday	one week ago
		anchor1		Monday	one week ago
			up	Friday	two weeks ago
				Thursday	two weeks ago
				Wednesday	two weeks ago
				Tuesday	two weeks ago
anchor2			anchor1	Monday	two weeks ago
up				Friday	three weeks ago
anchor1				Thursday	three weeks ago
				Wednesday	three weeks ago
				Tuesday	three weeks ago
				Monday	three weeks ago
training 0.74	training 0.72	training 0.7	training 0.7		
testing 0.78	testing 0.73	testing 0.71	testing 0.82		

Обнаружение фигур технического анализа для финансового прогноза

Пример правила:

ДЛЯ ЛЮБОГО t_1 СУЩЕСТВУЮТ t_2, t_3, t_4 , такие что
ЕСЛИ t_1, t_2, t_3, t_4 – локальные минимумы
И $c(t_3) < c(t_2)$ И $c(t_2) < c(t_1)$ И $c(t_1) < c(t_4)$ ТОГДА $c(t_1) < c(t_1+5)$

Фигура, описываемая правилом:



Разработка диагностической системы рака груди

Diagnostic rule	F-criterion for features		total significance of F-criterion			Accuracy of diagnosis for test cases (%)
			0.01	0.05	0.1	
IF <u>NUMBER of calcifications per cm²</u> is between 10 and 20 AND <u>VOLUME > 5 cm³</u> THEN <u>Malignant</u>	NUM VOL	0.0029 0.0040	+	+	+	93.3
IF <u>TOTAL number of calcifications >30</u> AND <u>VOLUME > 5 cm³</u> AND <u>DENSITY of calcifications is moderate</u> THEN <u>Malignant</u>	TOT VOL DEN	0.0229 0.0124 0.0325	- - -	+ + +	+ + +	100.0
IF <u>VARIATION in shape of calcifications</u> is marked AND <u>NUMBER of calcifications</u> is between 10 and 20 AND <u>IRREGULARITY in shape of calcifications</u> is moderate THEN <u>Malignant</u>	VAR NUM IRR	0.0044 0.0039 0.0254	+ + -	+ + +	+ + +	100.0
IF <u>variation in SIZE of calcifications</u> is moderate AND <u>Variation in SHAPE of calcifications</u> is mild AND <u>IRREGULARITY in shape of calcifications</u> is mild THEN <u>Benign</u>	SIZE SHAPE IRR	0.0150 0.0114 0.0878	- - -	+ + -	+ + +	92.86



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>