

# Когнитивное компьютерное моделирование и искусственный интеллект

Александр Панов

Отдел «Интеллектуальные динамические системы и когнитивные исследования»  
Институт системного анализа  
Федеральный исследовательский центр «Информатика и управление»  
Российской академии наук

19 марта – Семинар AI@MIPT



## 1 Введение

- О лаборатории
- Психологические идеи

## 2 Модель элементов картины мира

- Знаковая картина мира
- Нейронный субстрат
- Образная компонента знака
- Каузальная сеть
- Семиотическая сеть

## 3 Модели когнитивных функций

- Операции в семиотической сети
- Планирование поведения
- Примеры
- Обучение и целеполагание

## 4 Прикладные задачи

- Многоагентная постановка
- Задача интеллектуального перемещения
- Распределение ролей в коллективе
- Обучение с подкреплением

## 5 Заключение

# ФИЦ ИУ РАН



Федеральный исследовательский центр «Информатика и управление» Российской Академии Наук — ведущая организация Российской Федерации, специализирующаяся в научных исследованиях в области информатики и управления.

## Структурные подразделения:

- Институт проблем информатики РАН (большие данные, информационная безопасность).
- Вычислительный центр РАН (машинное обучение, математическое моделирование).
- Институт системного анализа РАН (искусственный интеллект, системный анализ в экономике).

## Базовые кафедры:

- ФУПМ МФТИ («Системные исследования» и «Интеллектуальные системы»).
- ВМК МГУ («Информационной безопасности»).
- МГТУ, МИРЭА, МАИ.

# Отдел «Интеллектуальные динамические системы и когнитивные исследования»

Основные направления исследований:

- Искусственный интеллект.
- Информационный поиск и анализ текстов (компьютерная лингвистика, психолингвистика, текстовая аналитика).
- Системы принятия решений (мультиможночества).
- Медицинские экспертные системы.
- Робототехнические системы (управление, обучение, планирование).
- Компьютерное когнитивные моделирование (модели когнитивных функций, многоагентные системы).



Осипов Геннадий Семенович,  
зам. дир. ФИЦ ИУ РАН, проф.,  
д.ф.-м.н., президент Российской  
ассоциации искусственного  
интеллекта (РААИ)



# Группа когнитивного компьютерного моделирования

- **Осипов Геннадий Семенович,**  
д.ф.-м.н.,
- **Чудова Наталья**  
Владимировна, к.псих.н.,
- **Кузнецова Юлия Михайловна,**  
к.псих.н.,
- **Панов Александр, к.ф.-м.н.,**
- Киселев Глеб, асп. ИСА РАН,
- Скрыник Алексей, асп. ИСА РАН,
- Ковалев Алексей, асп. ВШЭ,
- студенты SLabAI.

Основные направления исследований:

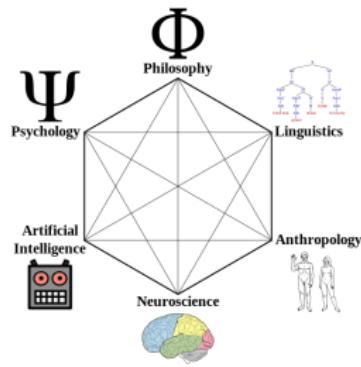
- Моделирование картины мира человека.
- Моделирование когнитивных функций человека с учетом психологических и нейрофизиологических данных.
- Планирование поведения и целеполагание.
- Обучение и логический вывод в картине мира.
- Групповое и коалиционное поведение, распределение ролей.

# Кратко о себе

**Панов Александр Игоревич, к. ф.-м. н.**

- Выпускник НГУ и МФТИ.
- Старший научный сотрудник отдела «Интеллектуальные динамические системы и когнитивные исследования» ФИЦ ИУ РАН.
- Научный сотрудник и доцент ФКН ВШЭ.
- Доцент кафедры системных исследований Московского физико-технического института (МФТИ).
- Член Российской ассоциации искусственного интеллекта (РААИ).
- Член Сообщества биологически инспирированных когнитивных архитектур (BICA Society).
- Участие в организации Международной конференции по биологически инспирированным когнитивным архитектурам (BICA-2016 — Нью-Йорк, BICA-2017 — Москва), Международной школы по биологически инспирированным когнитивным архитектурам (Fierce on BICA, Москва) и школы молодых ученых по ИИ (ISyT 2017, Санкт-Петербург).
- Член редколлегии журнала Biologically Inspired Cognitive Architectures.
- Руководитель проектов РФФИ мол\_а, мол\_а\_дк, офи\_м.
- Лауреат медали РАН для молодых ученых за 2017 год.
- Ментор студенческой лаборатории по ИИ (SLabAI).

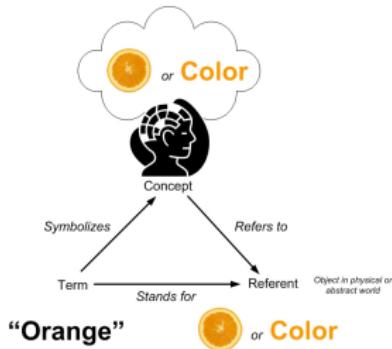
# Когнитивные науки



Когнитивная наука (лат. *cognitio* «познание») - междисциплинарное научное направление изучающее психику, разум (*mind*) человека и реализующие его процессы.

Схожие направления в Европе и Америке: **Artificial General Intelligence** и **Common model of cognition**.

# Проблема привязки символов



В оригинале Harnad — **symbol grounding problem** в 1990 г.

Классические методы искусственного интеллекта являются символыми (логика, множества правил, планирование, обучение). Однако, мышление — это больше, чем манипулирование символами.

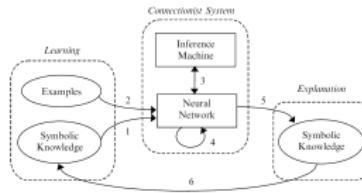
Особенно эта проблема актуальна в робототехнике (**symbol anchoring**) — система должна обучаться символам на основе своего собственного опыта.

Barsalou, Lawrence W. "Grounded cognition". *Annual review of psychology*. 2008.

Barsalou, L. W. "Perceptual symbol systems". *The Behavioral and brain sciences*. 1999.

Harnad, Stevan. "Symbol Grounding Problem". *Physica*. 1990.

# Нейросимвольные вычисления



**Основная идея:** кодирование символа путем вектора чисел и затем представление этого вектора с помощью связей в ансамбле искусственных нейронов (*embedding*).

**Основной результат:** путем введения специальных правил по распространению активности по нейронной сети реализованы некоторые простые логические схемы.

**Основной недостаток:** ограниченная интеграция обучения.

Besold, Tarek R. and Kai Uwe Kuhnberger. "Towards integrated neural-symbolic systems for human-level AI: Two research programs helping to bridge the gaps". *Biologically Inspired Cognitive Architectures*. 2015.

Garcez, Avila et al. "Neural-Symbolic Learning and Reasoning : Contributions and Challenges". *Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches: Papers from the 2015 AAAI Spring Symposium*. 2015.

Sun, Ron. *Integrating Rules and Connectionism for Robust Commonsense Reasoning*. 1994.

## Знак vs. символ

Если прямая интеграция нейронных сетей и символьной обработки работает плохо, то возможно стоит пересмотреть роль **символа** в общей схеме?

Будем считать, что у символа есть своя структура и у него есть кроме **номинативной** и другие **функциональные** роли в реализации когнитивных процессов → **знак**.

Понятие знака как элемента индивидуального знания появилось в конце XIX в.: Пирс, Фреге. Нашло свое отражение в лингвистике, логике, культурологии и **психологии**.

Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*. 2015.

Выготский, Л. С. «Мышление и речь». *Психология развития человека*. 2005.

Лотман, Ю. М. *Семиосфера*. 2000.

Пирс, Ч. С. *Начала pragmatизма*. 2000.

## Знак vs. символ

Если прямая интеграция нейронных сетей и символьной обработки работает плохо, то возможно стоит пересмотреть роль **символа** в общей схеме?

Будем считать, что у символа есть своя структура и у него есть кроме **номинативной** и другие **функциональные** роли в реализации когнитивных процессов → **знак**.

Понятие знака как элемента индивидуального знания появилось в конце XIX в.: Пирс, Фреге. Нашло свое отражение в лингвистике, логике, культурологии и **психологии**.

**Идея:** формализовать понятие знака → использовать знак в качестве основного структурного элемента системы знаний когнитивного агента → построить знаковые модели когнитивных функций.

Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*. 2015.

Выготский, Л. С. «Мышление и речь». *Психология развития человека*. 2005.

Лотман, Ю. М. *Семиосфера*. 2000.

Пирс, Ч. С. *Начала pragmatизма*. 2000.

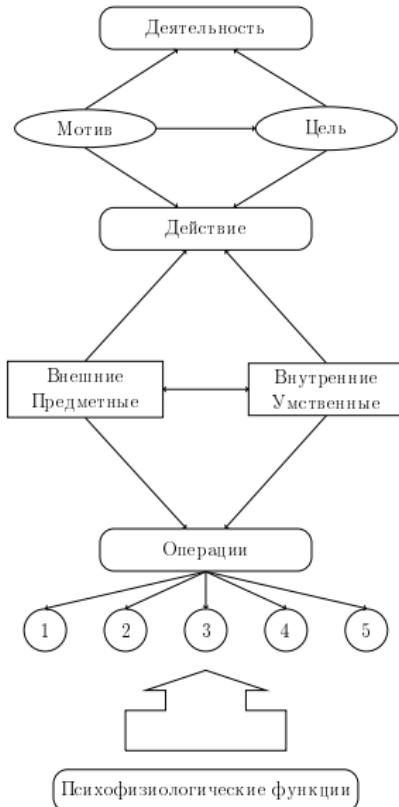
# Культурно-исторический подход Выготского



Теория происхождения и развития высших психических функций:

- **Социальная среда** — главный источник развития личности.
- Овладение культурой, способами поведения и мышления.
- Развитие когнитивных функций происходит в первую очередь через использование ребенком **«психологических орудий»**, путем овладения системой знаков-символов, таких как язык, письмо, счет.
- Внешняя деятельность, когда культурные средства имеют предметный вид, по мере отработки сворачивается (**интиериоризуется**) во внутренний план.
- На первом этапе внешней деятельности ребенок все делает в **сотрудничестве** со взрослыми, «зона ближайшего развития».
- Развитие - не ровно-постепенный, а **стадиальные** процессы.
- Сознание развивается через **диалог**: диалог ребенка со взрослым либо диалог взрослого со взрослым.

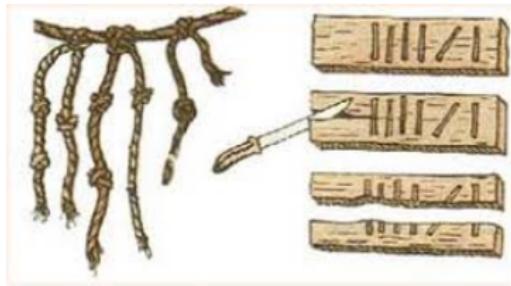
# Теория деятельности Леонтьева



## Основные положения:

- Поведение человека - это двойная иерархическая структура мотивы-цели и действия-операции.
- Деятельность – это активный, целенаправленный процесс.
- Действия человека предметны; их цели носят социальный характер.
- Сознание и деятельность неразрывно связаны.

# Знак как орудие психической деятельности



- Знак - это искусственно созданный человеком стимул, средство для управления своим поведением и поведением других.
- История развития человечества - это история развития знака: чем более развита система знаков в поколении, тем более развиты высшие психические функции.
- Знаки: наскальный рисунок, приметы, жесты, речь, ноты и т.д.

# Прикладная семиотика



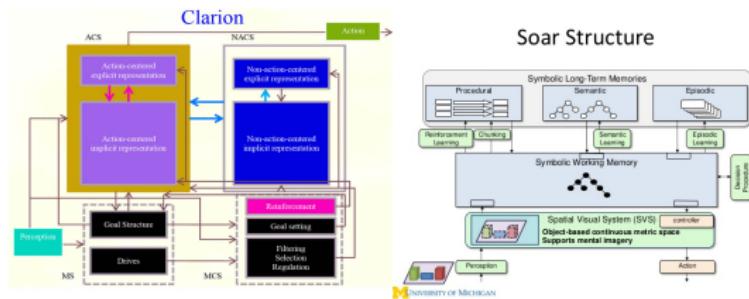
Семиотические базы знаний:

- **Именованность:** информационная единица, которая претендует на то, чтобы называться знанием, должна иметь некоторую собственную метку - имя.
- **Структурированность:** информационная единица должна обладать своей внутренней структурой.
- **Принцип матрешки:** знаки за счет связей наследования как бы вкладываются друг в друга, обеспечивая описание сущностей на различных уровнях.
- **Связность:** знаки благодаря различным отношениям объединяются в сеть.
- **Активность:** в сетях знаков становится возможной реализация принципа «активизация знаний - источник активизации процедур».
- **Рефлексивность:** появление метауровня позволяет системе рассуждать о самой себе, о характере имеющейся у нее информации об окружающем мире.

Осипов, Г. С. и Д. А. Поспелов. «Прикладная семиотика». *Новости искусственного интеллекта*. 1999.

Поспелов, Д. А. «Прикладная семиотика и искусственный интеллект». *Программные продукты и системы*. 1996.

# Когнитивные архитектуры



Недостатки современных когнитивных архитектур:

- Концептуальная нерешенность проблемы привязки символов (symbol grounding problem) - CLARION
- Отсутствие деятельностной модели поведения системы - реализация только некоторых когнитивных аспектов
- Иерархичность представления знаний (4D/RCS)
- Возможность реализации иерархического планирования
- Реализация обучения концептуальным знаниям - Cognitive Mario
- Моделирование рефлексивного поведения

Besold, Tarek R. and Kai Uwe Kuhnberger. "Towards integrated neural-symbolic systems for human-level AI: Two research programs helping to bridge the gaps". *Biologically Inspired Cognitive Architectures*. 2015.

Sun, Ron. "Autonomous generation of symbolic representations through subsymbolic activities". *Philosophical Psychology*. 2013.

# Картина мира субъекта деятельности

Картина мира субъекта деятельности - это представления субъекта о внешней среде, о своих собственных характеристиках, целях, мотивах, о других субъектах и операции (произвольные и непроизвольные), осуществляемые на основе этих представлений.

# Картина мира субъекта деятельности

Картина мира субъекта деятельности - это представления субъекта о внешней среде, о своих собственных характеристиках, целях, мотивах, о других субъектах и операции (произвольные и непроизвольные), осуществляемые на основе этих представлений.

Элементом картины мира является знак:

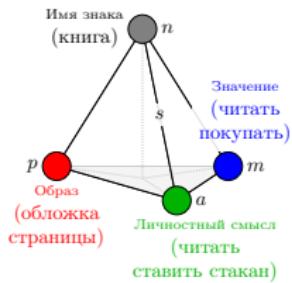
- в смысле культурно-исторического подхода Выготского-Лурии,
- выполняющий функции в соответствии с теорией деятельности Леонтьева.

# Картина мира субъекта деятельности

Картина мира субъекта деятельности - это представления субъекта о внешней среде, о своих собственных характеристиках, целях, мотивах, о других субъектах и операции (произвольные и непроизвольные), осуществляемые на основе этих представлений.

Элементом картины мира является знак:

- в смысле культурно-исторического подхода Выготского-Лурии,
- выполняющий функции в соответствии с теорией деятельности Леонтьева.

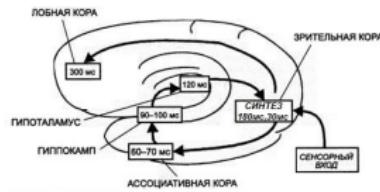
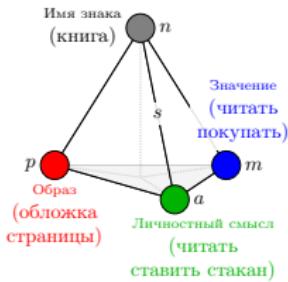


# Картина мира субъекта деятельности

Картина мира субъекта деятельности - это представления субъекта о внешней среде, о своих собственных характеристиках, целях, мотивах, о других субъектах и операции (произвольные и непроизвольные), осуществляемые на основе этих представлений.

Элементом картины мира является знак:

- в смысле культурно-исторического подхода Выготского-Лурии,
- выполняющий функции в соответствии с теорией деятельности Леонтьева.

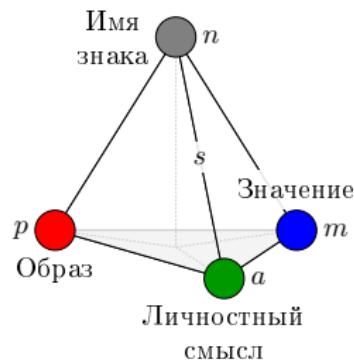


В пользу существования такой структуры свидетельствуют:

- нейрофизиологические данные (Эдельман, Иваницкий, Маунткастл и др.),
- другие психологические теории (например, трехкомпонентная модель Станович).

Оsipov, Г. С., А. И. Panov и Н. В. Чудова. «Управление поведением как функция сознания. II. Синтез плана поведения». *Известия Российской академии наук. Теория и системы управления*. 2015.  
 — . «Управление поведением как функция сознания. I. Картина мира и целеполагание». *Известия Российской академии наук. Теория и системы управления*. 2014.

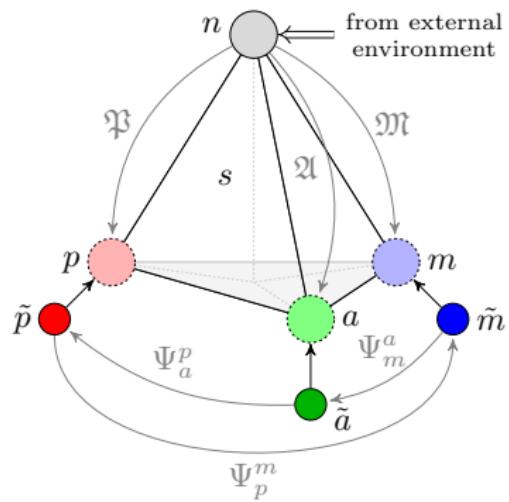
# Три образующих элемента картины мира



Представляемая сущность описывается тремя причинно-следственными (каузальными) структурами:

- **структура образа** - представление взаимосвязи внешних сигналов и внутренних характеристик субъекта (агента) - сенсо-моторное представление,
- **структура значения** - обобщенное знание о соотношениях во внешнем мире, согласованное в некоторой группе субъектов (агентов),
- **структура личностного смысла** - ситуациянная потребностно-мотивационная интерпретация знаний о соотношениях во внешней среде (значение для себя).

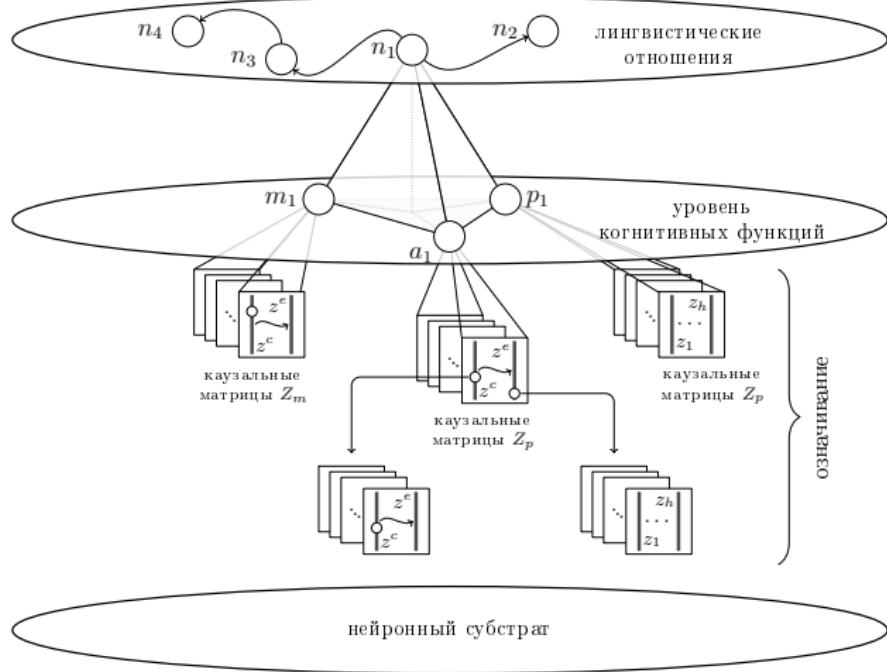
# Актуализация и формирование знака



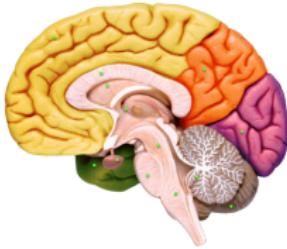
**Процесс обучения** — образование новых знаков как неподвижная точка операторов замыкания  $\Psi_p^m \Psi_m^a \Psi_a^p$ .

Реализация когнитивных функций — **актуализация** (активация) имеющихся знаков и формирование новых «сituационных» знаков — «протознаков» без конвенционального имени.

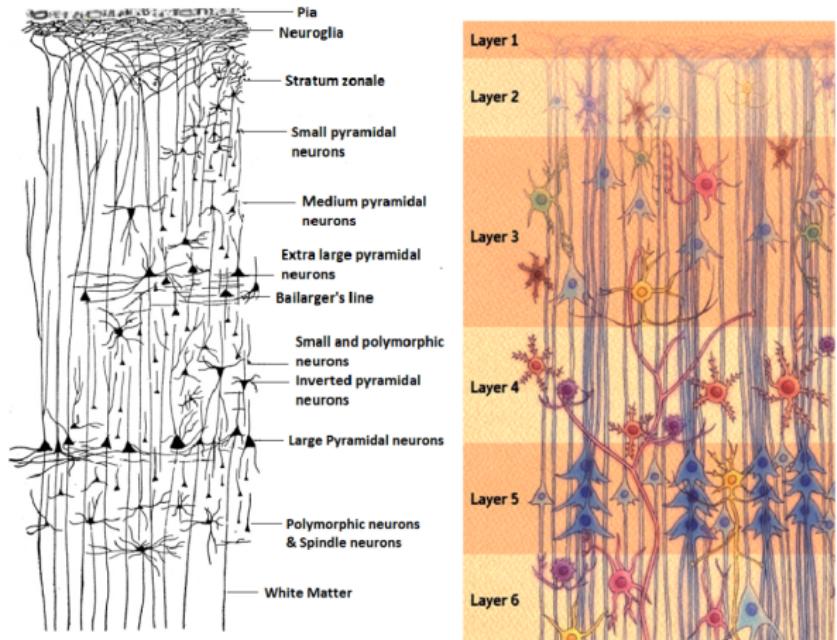
# Уровни представления



# Нейронный субстрат



## Histological Structure of the Cerebral Cortex

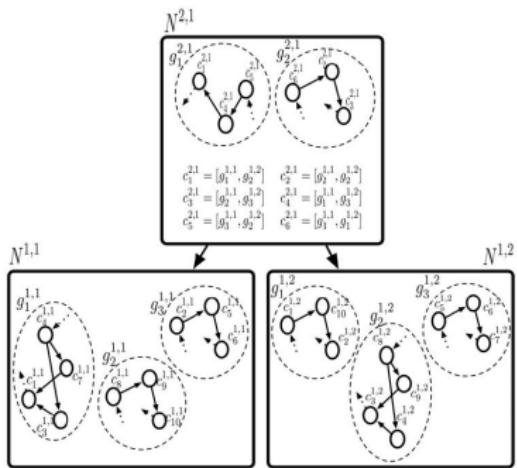


Rockland, Kathleen S. «Five points on columns». *Frontiers in neuroanatomy*. 2010.

Mountcastle, V. B. *Perceptual Neuroscience. The Cerebral Cortex*. 1998.

# Гетерархическая модель

Разработана расширенная реализация иерархической временной памяти (hierarchical temporal memory - HTM) - **гетерархическая каузальная сеть (heterarchical causal network - HCN)**.

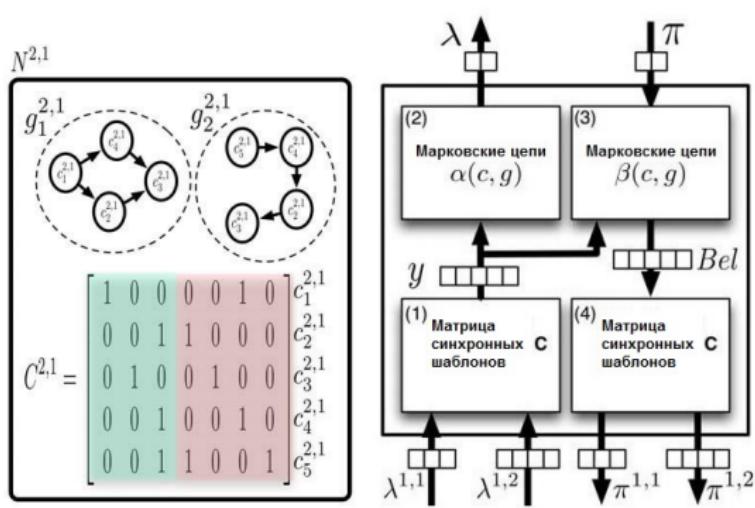


George, D. et al. "A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs". *Science*. Oct. 2017.

George, Dilip and Jeff Hawkins. "Towards a mathematical theory of cortical micro-circuits". *PLoS computational biology*. 2009.

George, Dilip. "How the Brain Might Work: a Hierarchical and Temporal Model for Learning and Recognition". PhD thesis. Stanford University, 2008, p. 191.

# Гетерархическая модель



Skrynnik, A., A. Petrov, and A. I. Panov. "Hierarchical Temporal Memory Implementation with Explicit States Extraction".

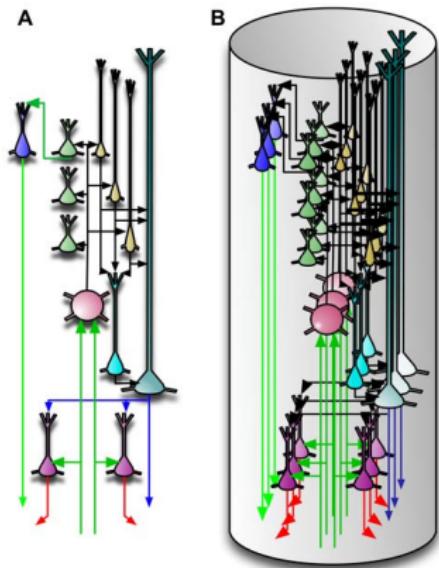
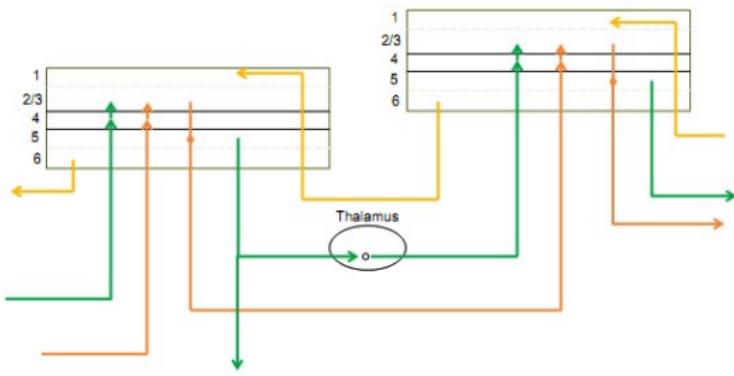
*Biologically Inspired Cognitive Architectures (BICA) for Young Scientists.* 2016.

Панов, А. И. «Алгебраические свойства операторов распознавания в моделях зрительного восприятия».

*Машинное обучение и анализ данных.* 2014.

Velichkovsky, B. M. "Hierarchy of cognition: The depths and the highs of a framework for memory research". *Memory.* 2002.

# Нейронная организация



Billaudelle, Sebastian and Subutai Ahmad. "Porting HTM Models to the Heidelberg Neuromorphic Computing Platform".  
2015.

George, Dileep and Jeff Hawkins. "Towards a mathematical theory of cortical micro-circuits". *PLoS computational biology*.  
2009.

# Модель процесса обучения

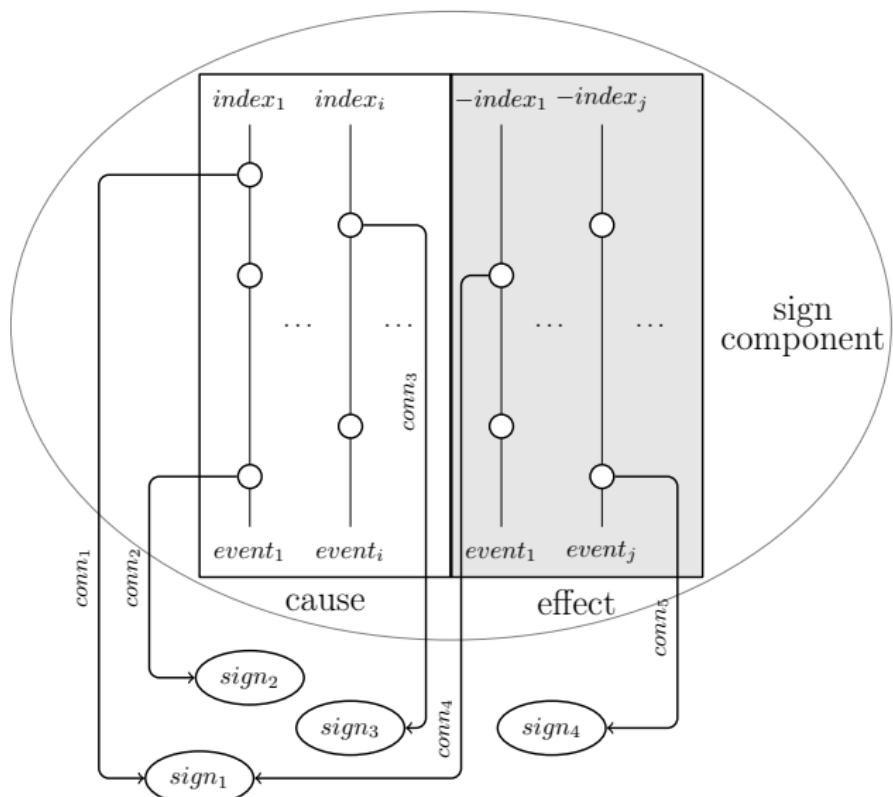
К основным принципам работы механизма обучения относятся:

- использование иерархии вычислительных узлов с восходящими и нисходящими связями,
- использование Хэббовских правил обучения,
- разделение пространственного и временного группировщиков,
- подавление второстепенной активации для формирования разреженного представления.

В результате работы механизма обучения по прецедентам (без учителя) формируются так называемые **каузальные матрицы**.

Панов, А. И. и А. В. Петров. «Иерархическая времененная память как модель восприятия и её автоматное представление». *Шестая Международная конференция «Системный анализ и информационные технологии» САИТ-2015 (15-20 июня 2015 г., г. Светлогорск, Россия): Труды конференции. В 2-х т. 2015.*

# Каузальная матрица



# Алгоритм $\mathfrak{A}_{th}$ активации образа знака

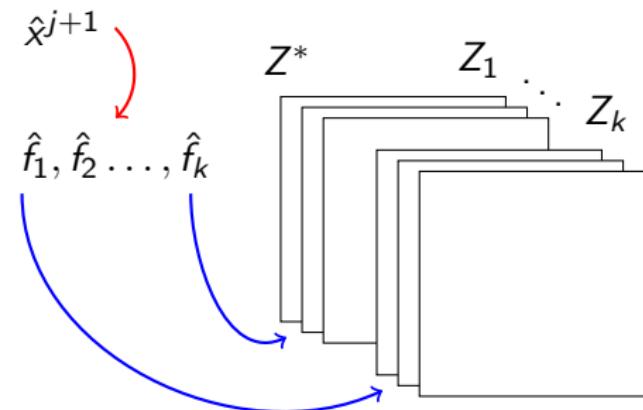
$\hat{f}_1, \hat{f}_2 \dots, \hat{f}_k$

# Алгоритм $\mathfrak{A}_{th}$ активации образа знака

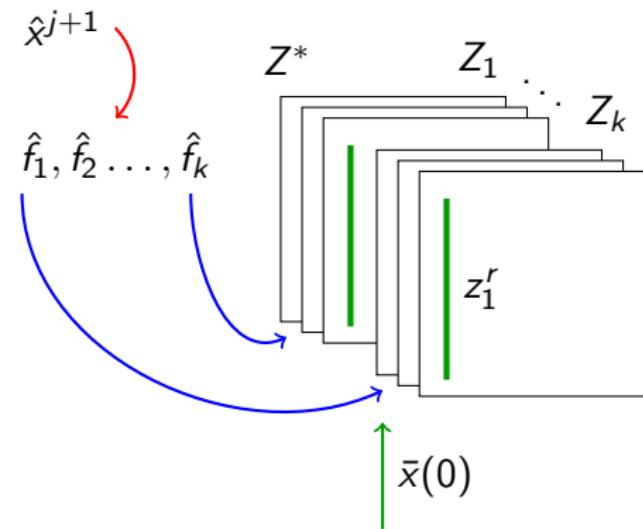
$$\hat{x}^{j+1}$$

$$\hat{f}_1, \hat{f}_2 \dots, \hat{f}_k$$

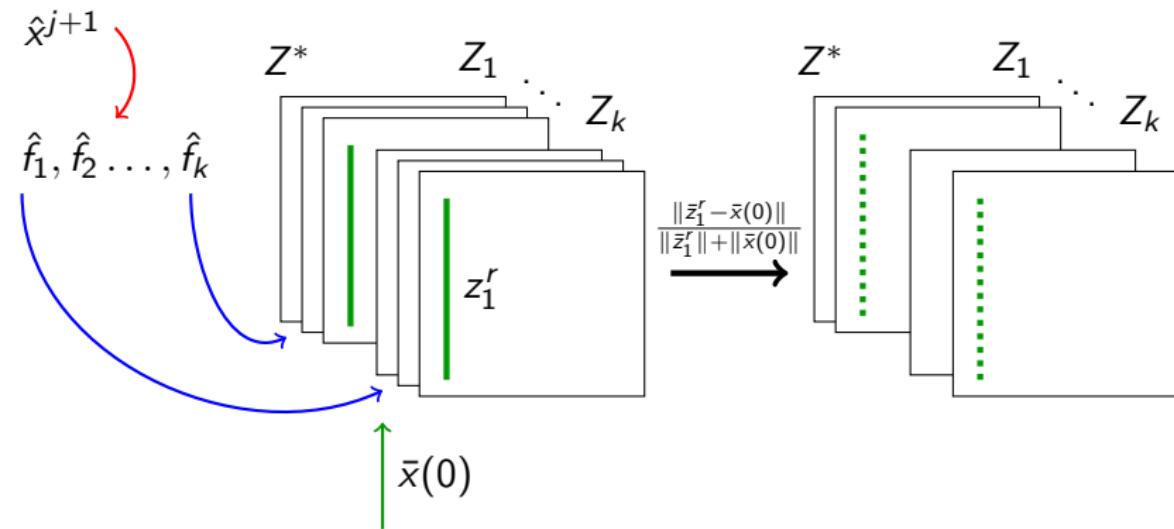
# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



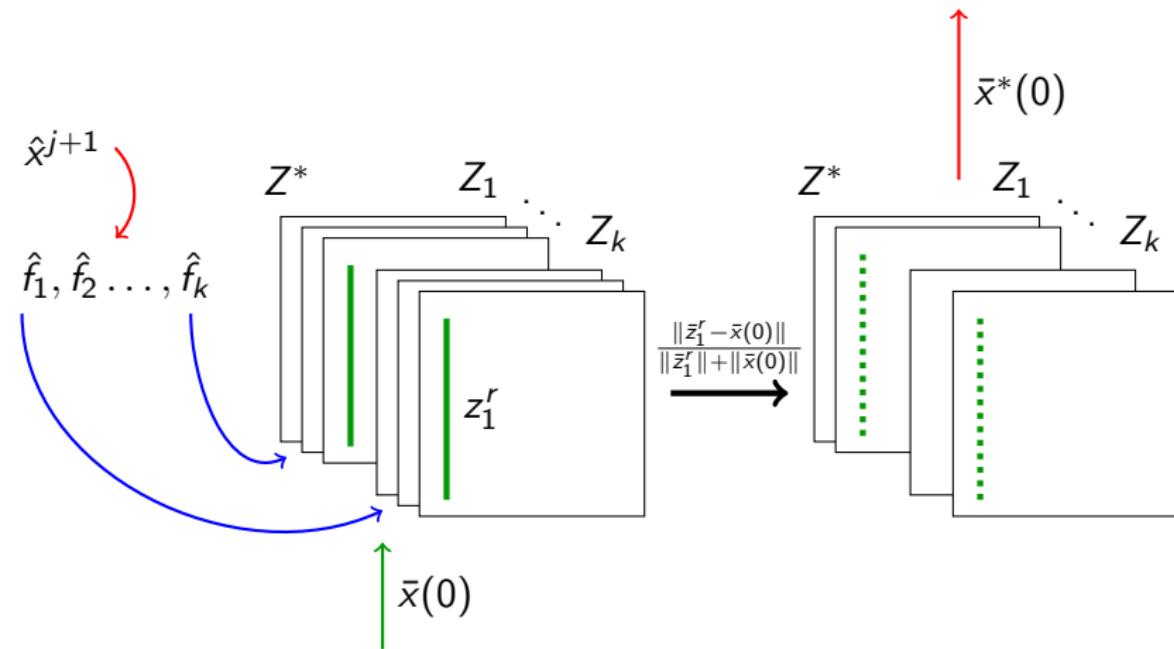
# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



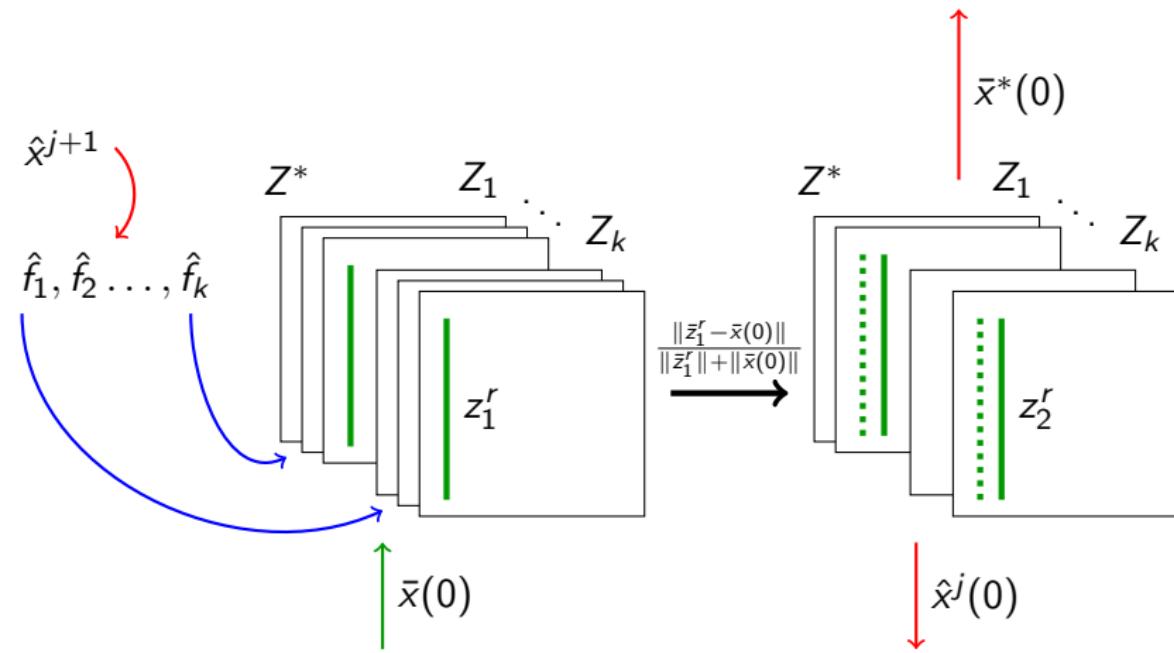
# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



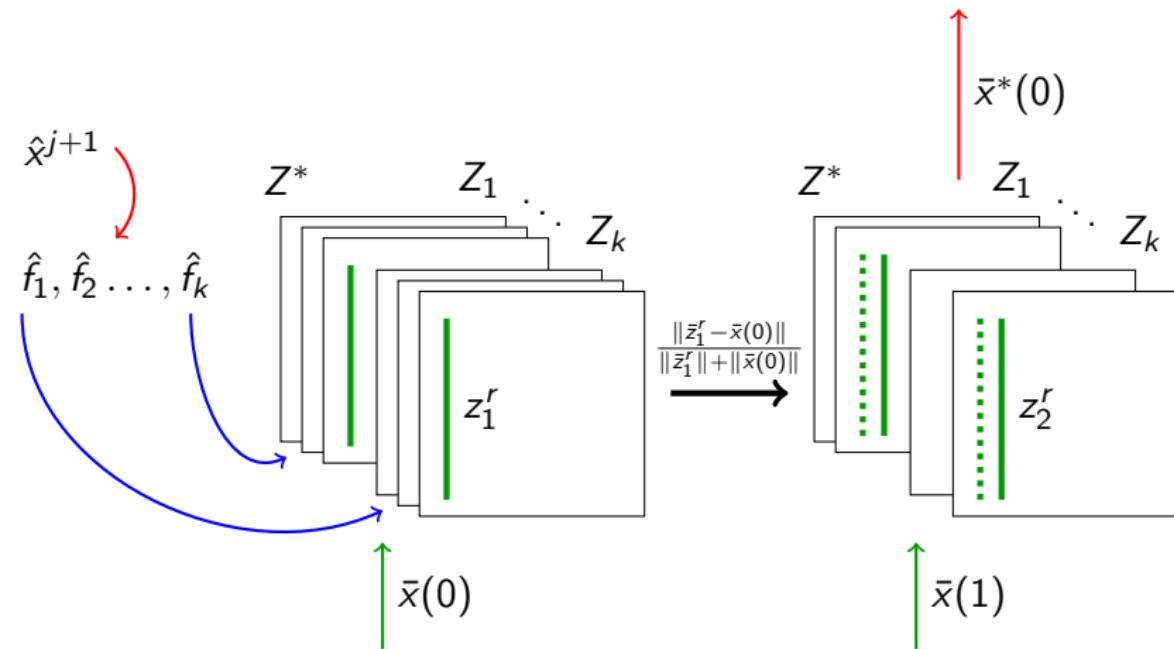
# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



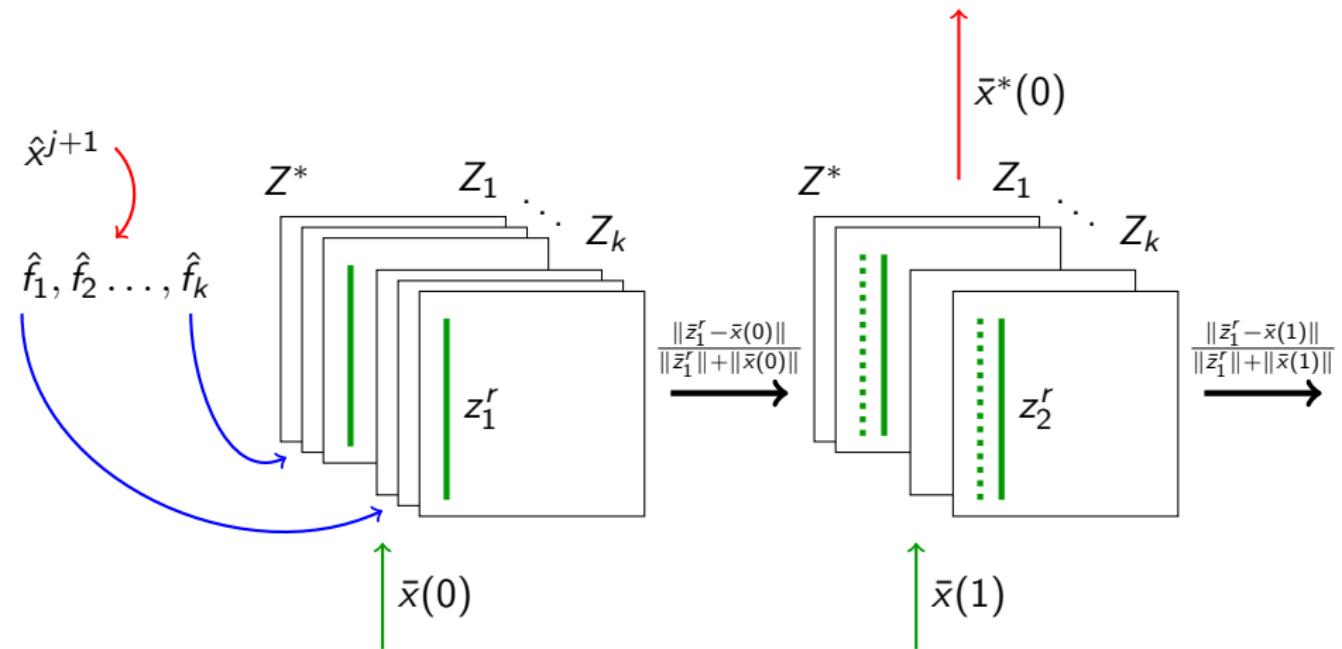
# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



# Алгоритм $\mathfrak{A}_{th}$ активации образа знака



# Алгоритм $\mathfrak{A}_{th}$ активации образа знака

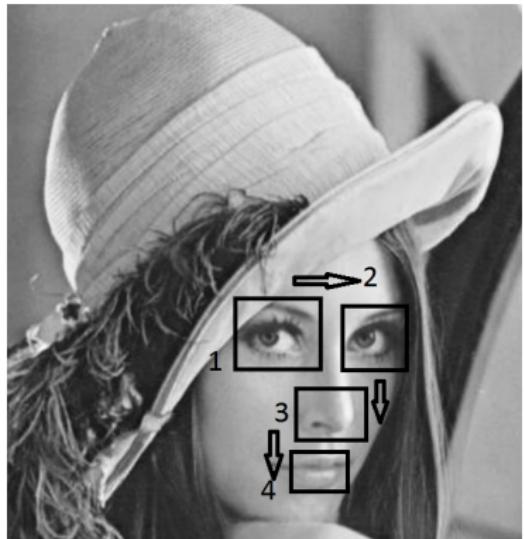
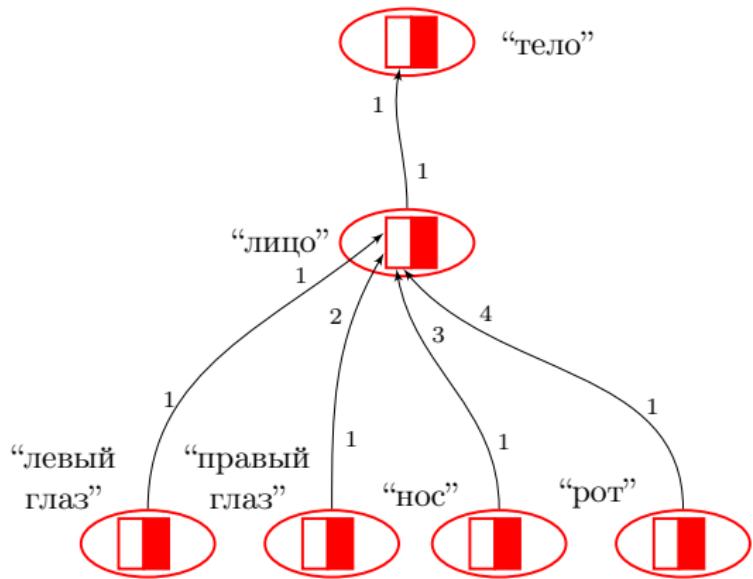


# Каузальная сеть на образах

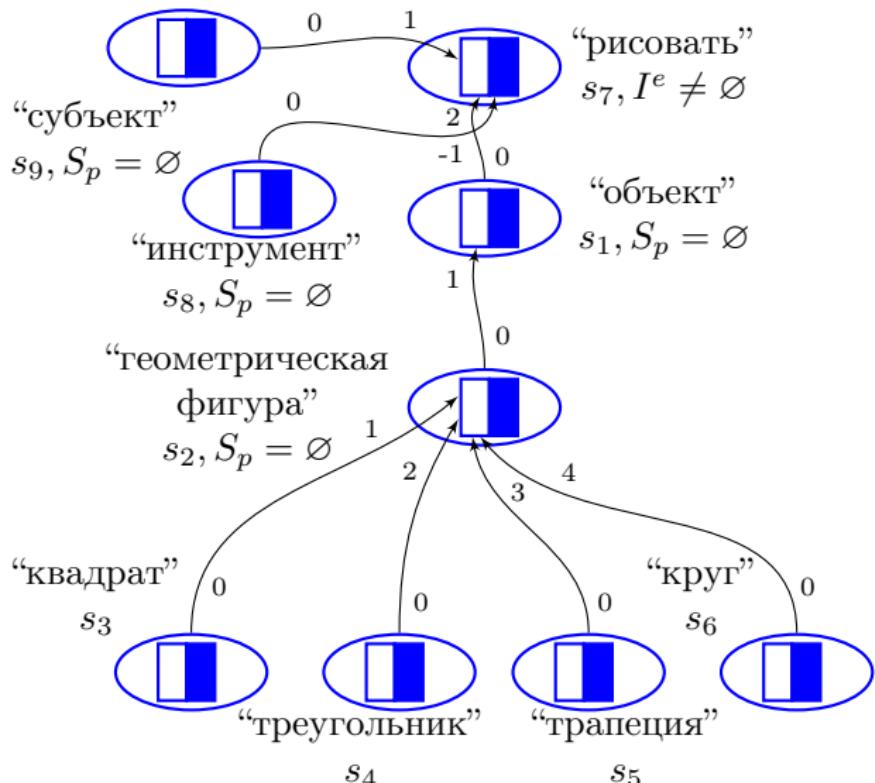
**Каузальная сеть** на множестве образов знаков  $W_p = \langle V_p, E_p \rangle$  - помеченный ориентированный граф, в котором

- каждому узлу  $v \in V_p$  ставится в соответствие кортеж каузальных матриц  $Z^p(s)$  образа некоторого знака  $s$  ( $v \rightarrow Z^p(s)$ );
- ребро  $e = (v_1, v_2)$  принадлежит множеству ребер графа  $E$ , если  $v_1 \rightarrow Z^p(s_1)$ ,  $v_2 \rightarrow Z^p(s_2)$  и  $s_1 \in S_p(s_2)$ ;
- каждому ребру графа  $e = (v_1, v_2)$ ,  $v_1 \rightarrow Z^p(s_1)$ ,  $v_2 \rightarrow Z^p(s_2)$  ставится в соответствие метка  $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3)$  - кортеж трех натуральных чисел:
  - $\epsilon_1$  - индекс исходной матрицы в кортеже  $Z^p(s_1)$ , может принимать специальное значение 0, если исходными могут служить любые матрицы из кортежа;
  - $\epsilon_2$  - индекс целевой матрицы в кортеже  $Z^p(s_2)$ , строка которой ставится в соответствие признаку  $s_1$ ;
  - $\epsilon_3$  - индекс столбца в целевой матрице, в которой в соответствующей признаку  $s_1$  строке стоит 1, может принимать положительные значения (*столбцы условий*) и отрицательные (*столбцы эффектов*).

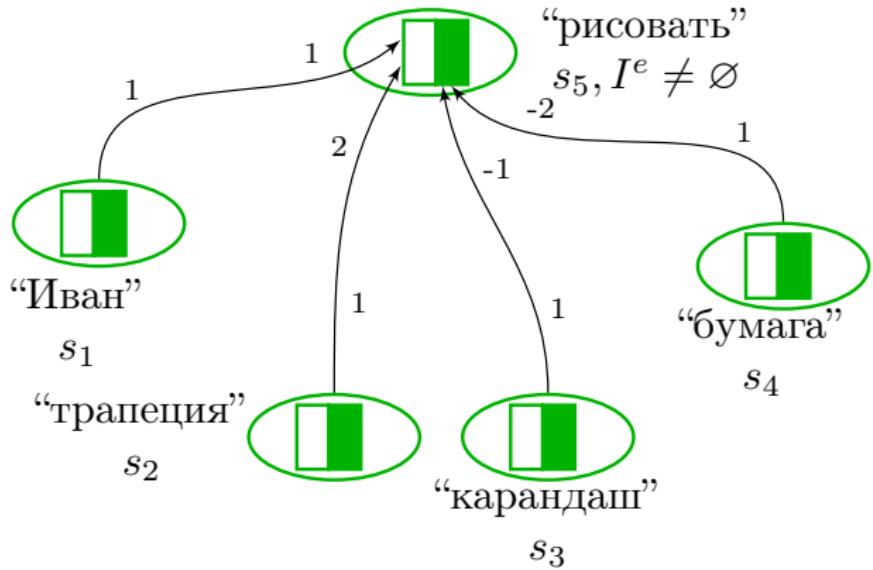
# Каузальная сеть на образах: пример



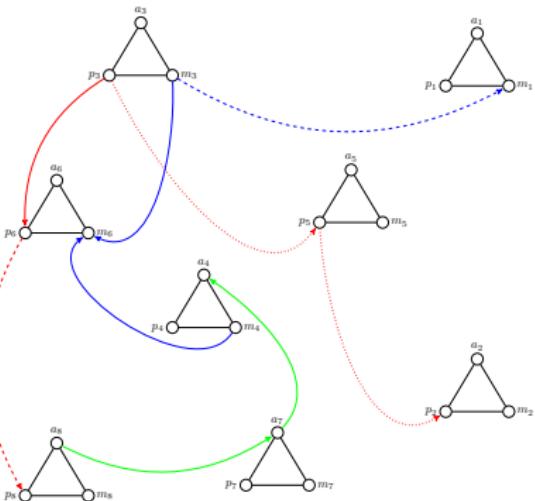
# Каузальная сеть на значениях: пример



# Каузальная сеть на личностных смыслах: пример



# Картина мира субъекта деятельности



**Знаком** будем называть четверку  $s = \langle n, p, m, a \rangle$ , где  $n$  - имя знака,  $p = Z^p$ ,  $m = Z^m$ ,  $a = Z^a$  - кортежи каузальных матриц, которые соответственно называются образом, значением и личностным смыслом знака  $s$ .

**Семиотическая сеть** - пятерка  
 $\Omega = \langle W_p, W_m, W_a, R_n, \Theta, \Phi \rangle$ , где

- $W_p, W_m, W_a$  - соответственно каузальные сети на множестве образов, значений и личностных смыслах,
- $R_n$  - семейство отношений на множестве знаков, сгенерированных на основе трех каузальных сетей, т.е.  $R_n = \{R_p, R_m, R_a\}$ ,
- $\Theta$  - семейство операций на множестве знаков,
- $\Phi$  - правила распространения активности на каузальных сетях.

Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*. 2015.

**1 Введение**

- О лаборатории
- Психологические идеи

**2 Модель элементов картины мира**

- Знаковая картина мира
- Нейронный субстрат
- Образная компонента знака
- Каузальная сеть
- Семиотическая сеть

**3 Модели когнитивных функций**

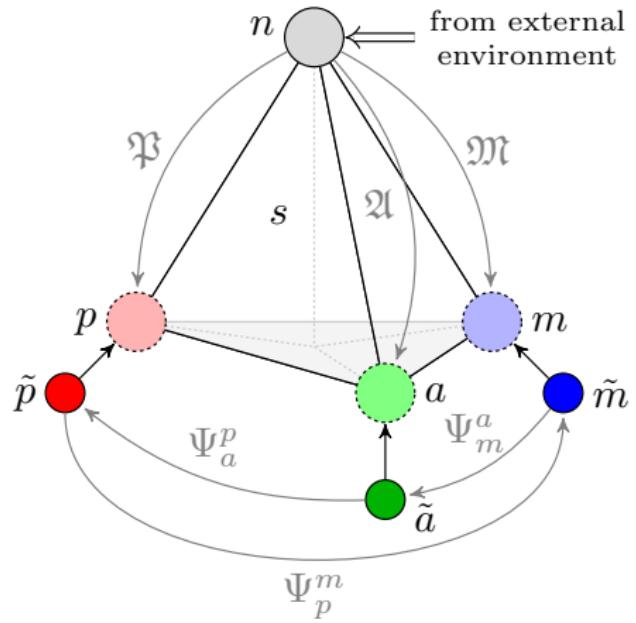
- Операции в семиотической сети
- Планирование поведения
- Примеры
- Обучение и целеполагание

**4 Прикладные задачи**

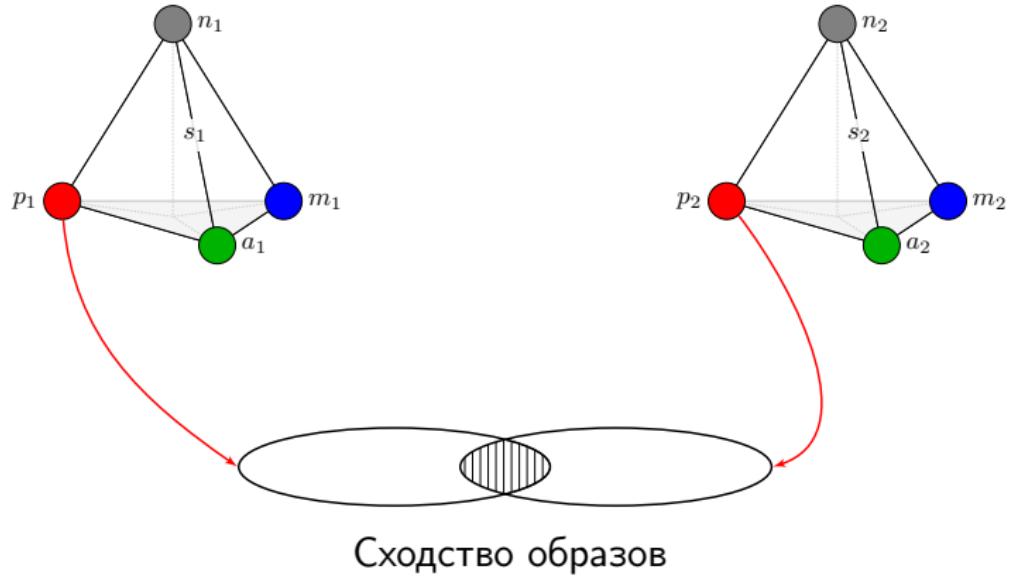
- Многоагентная постановка
- Задача интеллектуального перемещения
- Распределение ролей в коллективе
- Обучение с подкреплением

**5 Заключение**

# Образование нового знака

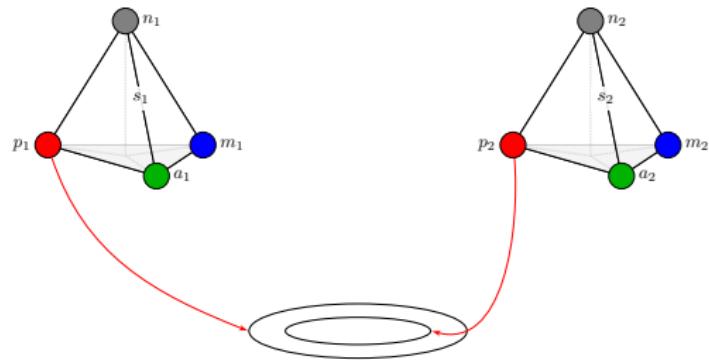


# Отношения на множестве компонент знака

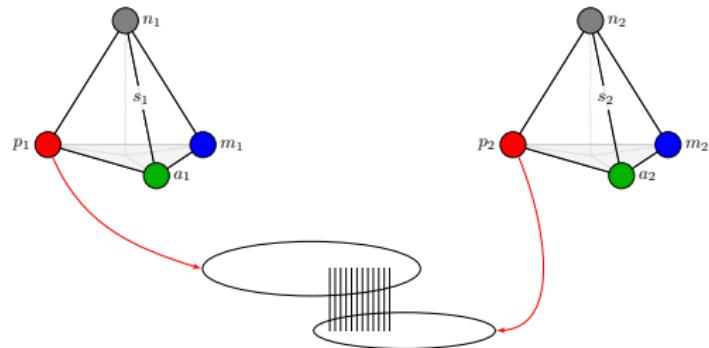


# Отношения на множестве компонент знака

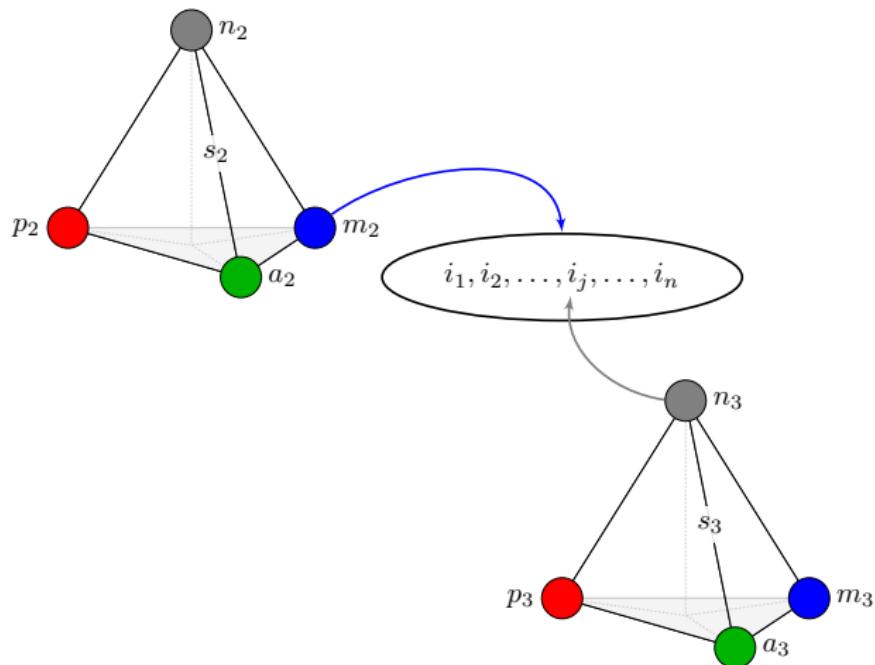
Включение образов



Противопоставление  
образов



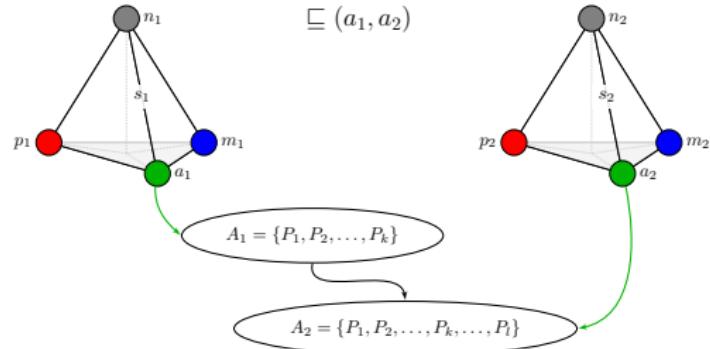
# Отношения на множестве компонент знака



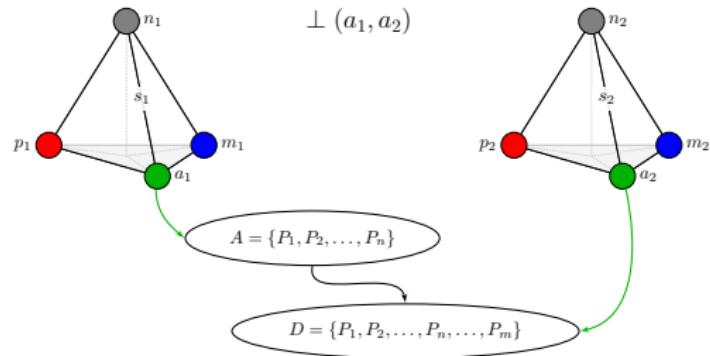
Сценарий на значениях

# Отношения на множестве компонент знака

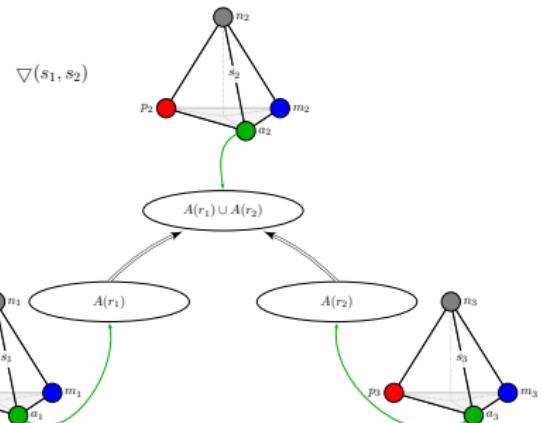
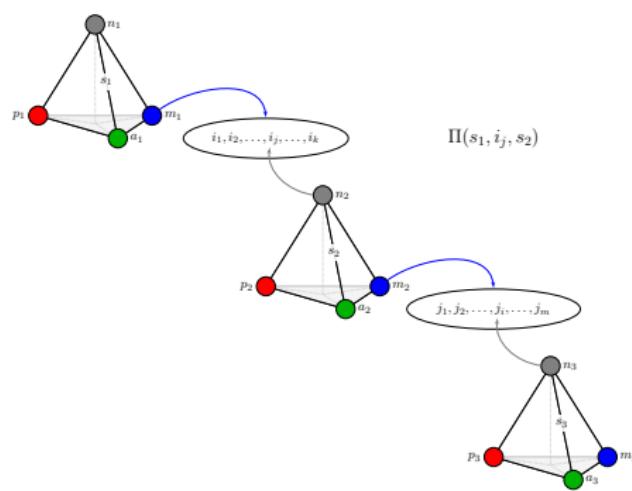
**Поглощение  
личностных смыслов**



**Противопоставление  
личностных смыслов**



# Операции на множестве компонент знака

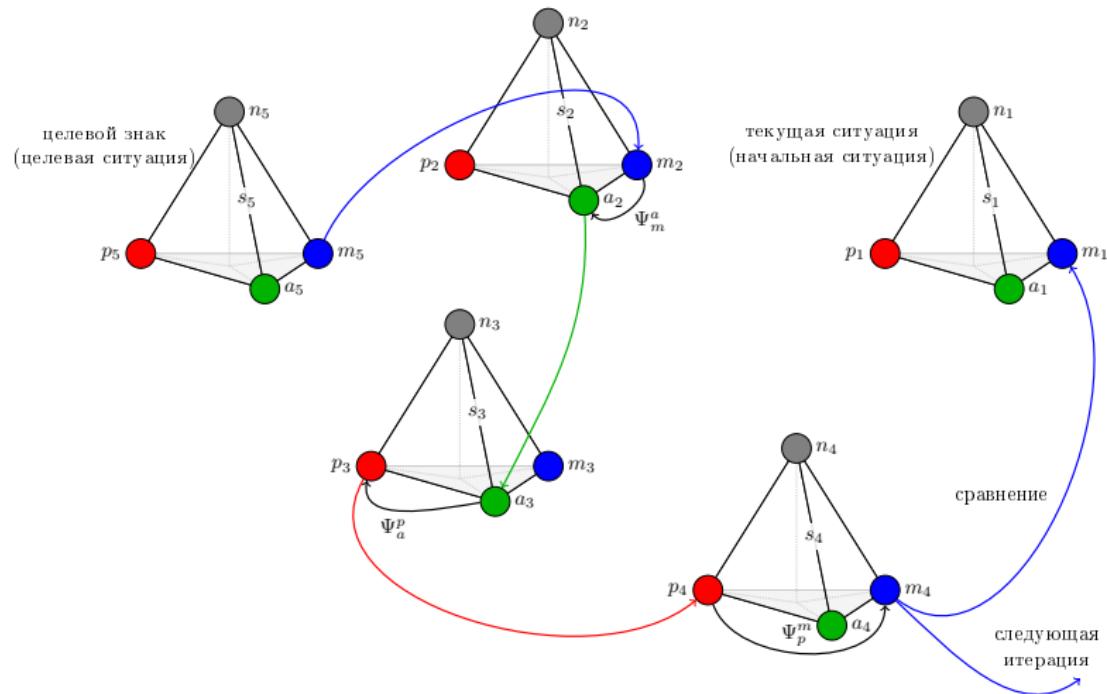


## Замыкание по значениям

## Агглютинация личностных смыслов

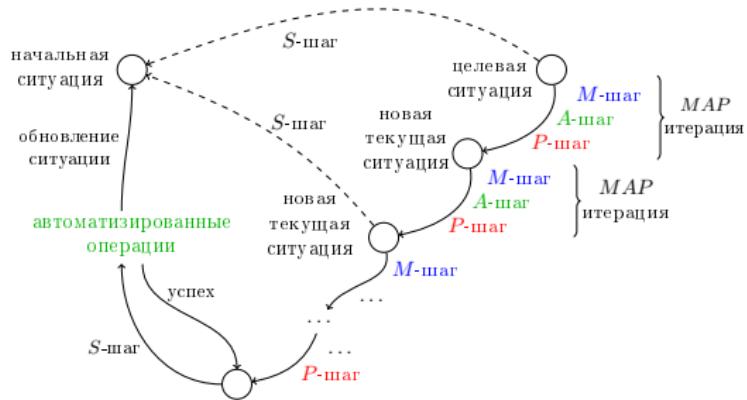
Osipov, G. S. "Sign-based representation and word model of actor". *2016 IEEE 8th International Conference on Intelligent Systems (IS)*. 2016.

# Модель функции планирование поведения



Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*. 2015.

# Алгоритм планирования поведения



Иерархический процесс планирования начинается с конченой ситуации и стремится достичь начальной ситуации.

MAP-итерация:

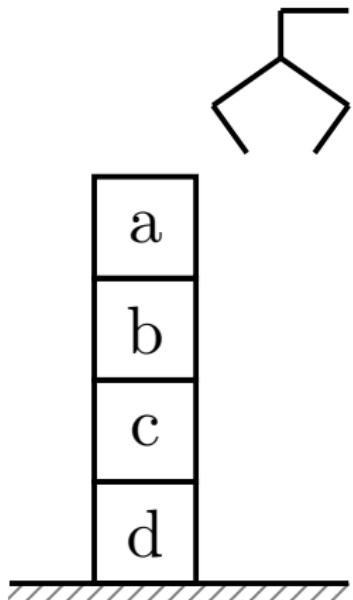
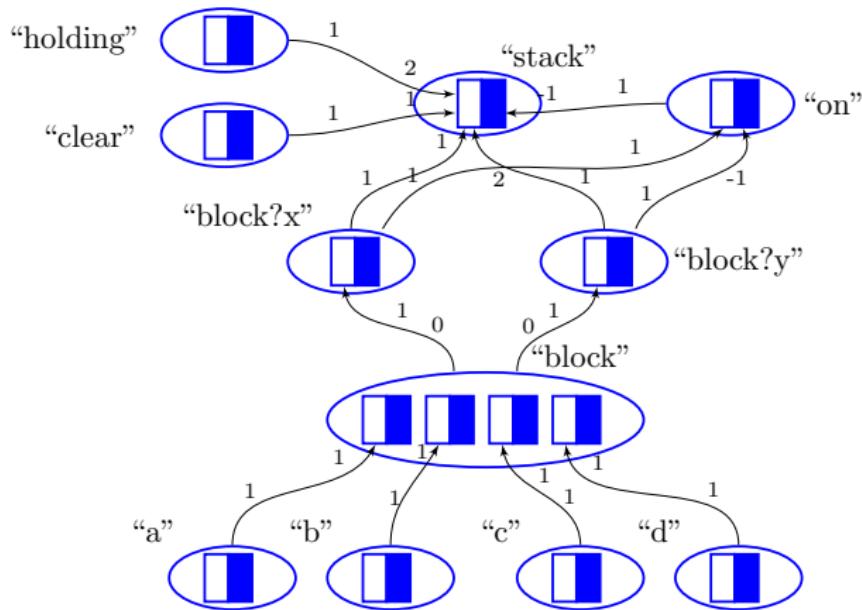
- **S-step** – поиск прецедентов выполнения действия в текущих условиях,
- **M-step** – поиск применимых действий на сети значений,
- **A-step** – генерация личностных смыслов, соответствующих найденным значениям,
- **P-step** – построение новой текущей ситуации по множеству признаков условий найденных действий.

Panov, A. I. and K. S. Yakovlev. "Behavior and path planning for the coalition of cognitive robots in smart relocation tasks". *Robot Intelligence Technology and Applications 4*. 2016.

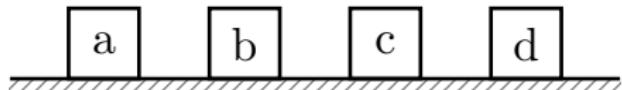
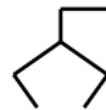
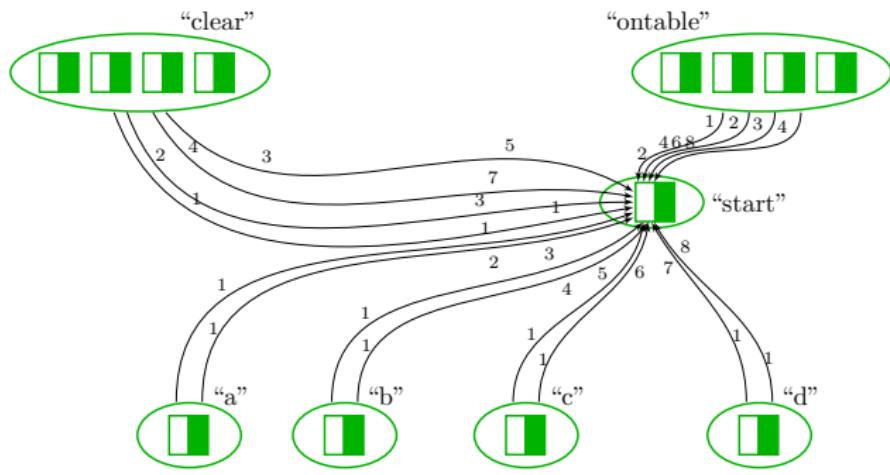
Panov, Aleksandr I. and Konstantin S. Yakovlev. "Psychologically inspired planning method for smart relocation task". *Procedia Computer Science*. 2016.

Panov, Aleksandr I. "Behavior Planning of Intelligent Agent with Sign World Model". *Biologically Inspired Cognitive Architectures*. 2017.

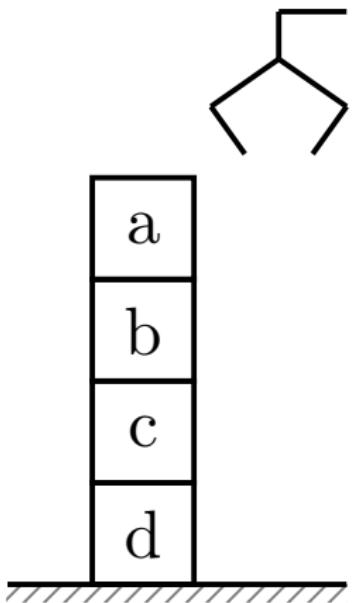
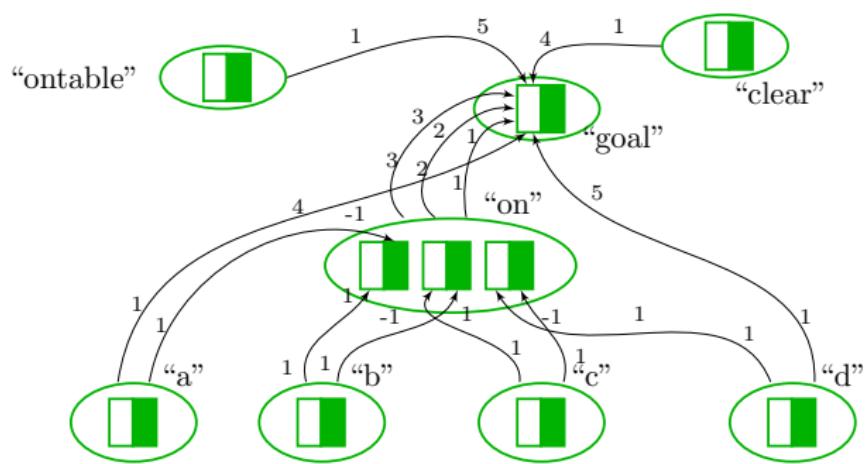
# Пример: фрагмент сети на значениях



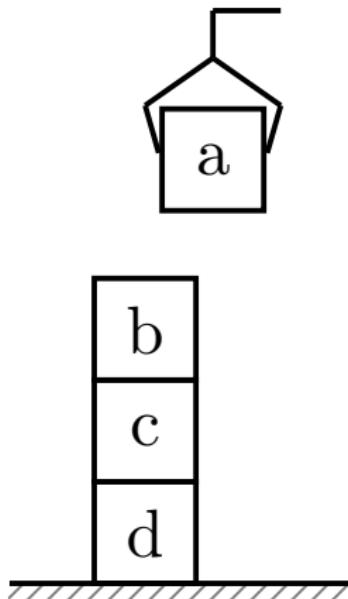
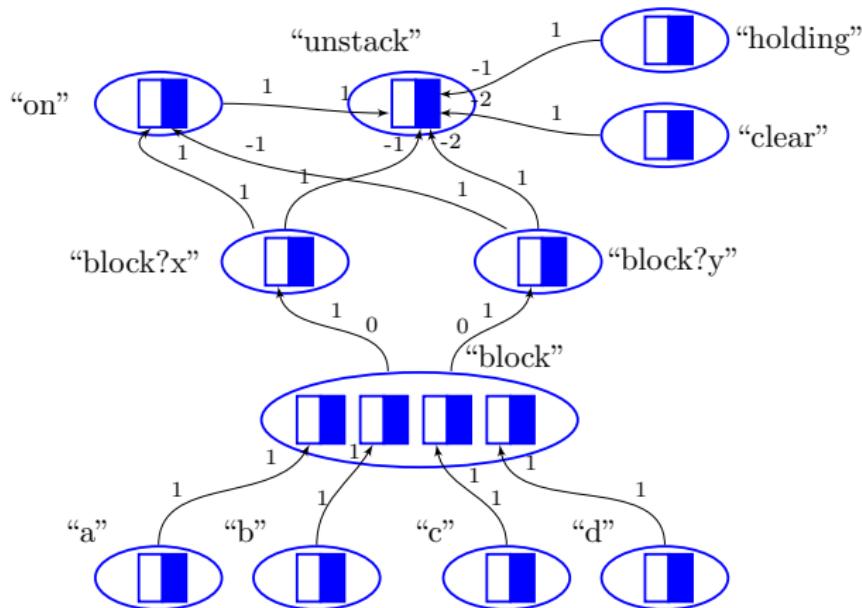
# Пример: сеть на смыслах - начальная ситуация



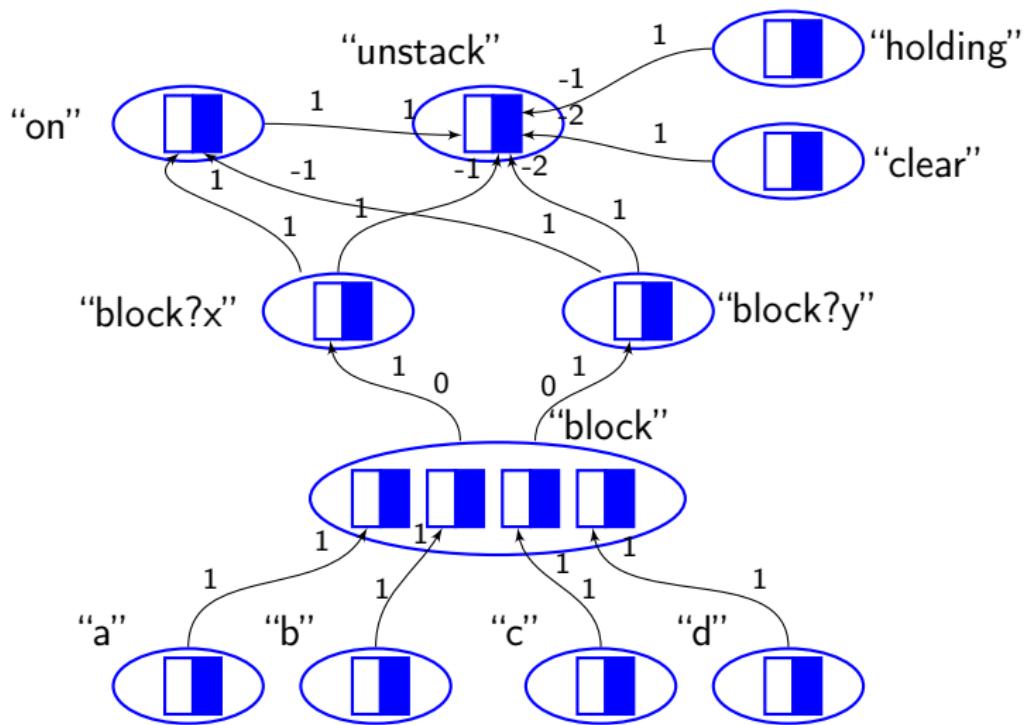
# Пример: сеть на смыслах - целевая ситуация



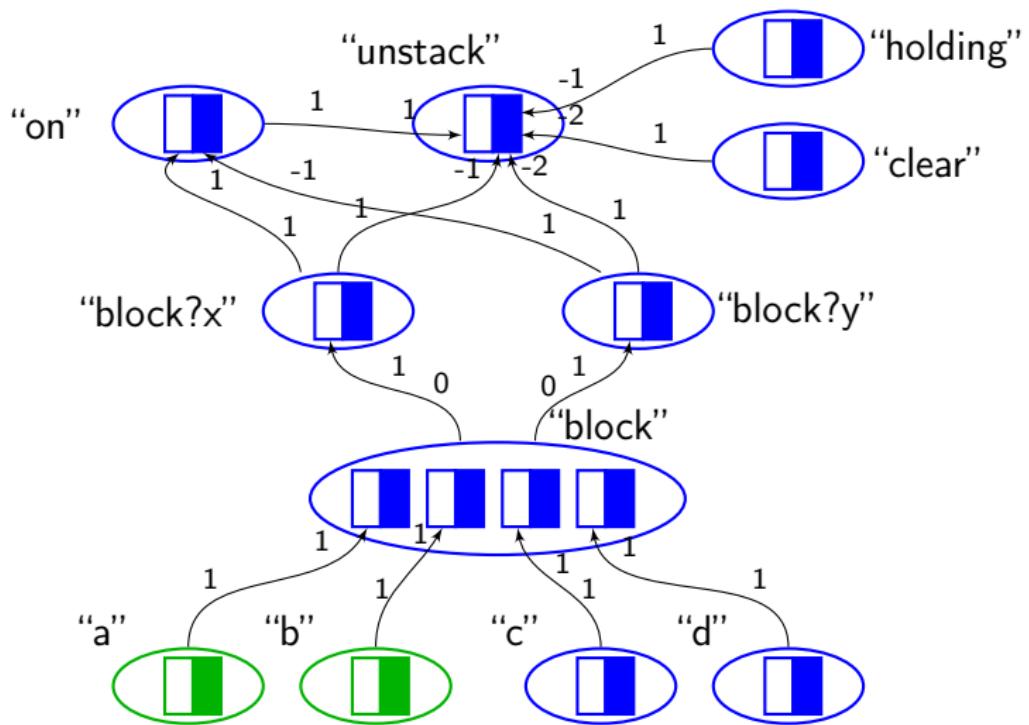
# Пример: фрагмент сети на значениях



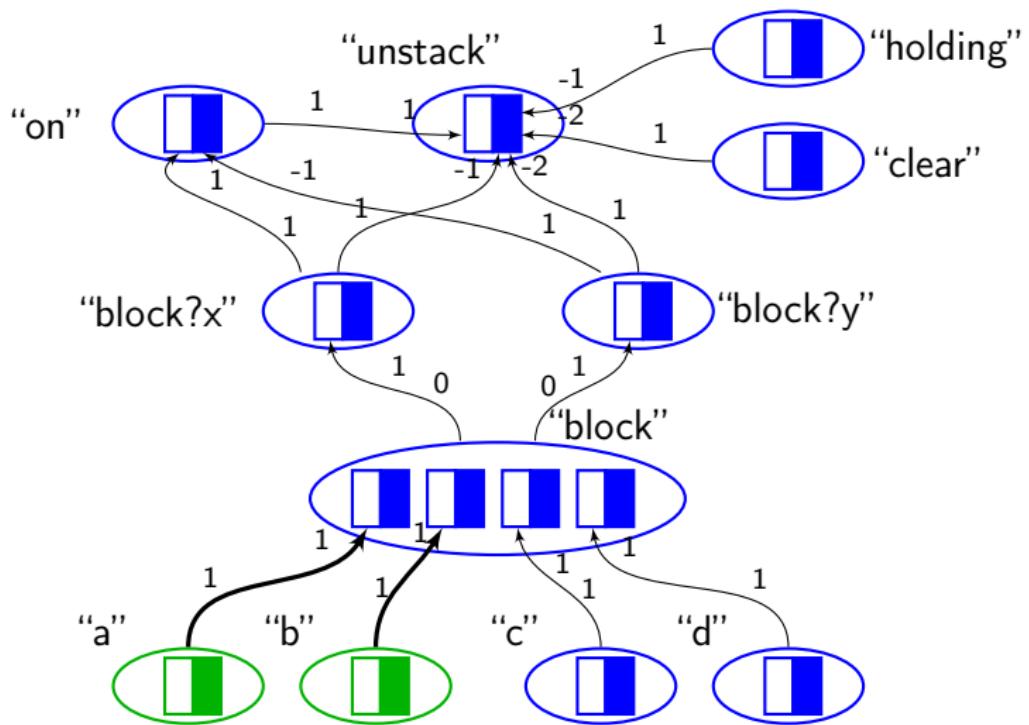
# Пример: генерация личностного смысла



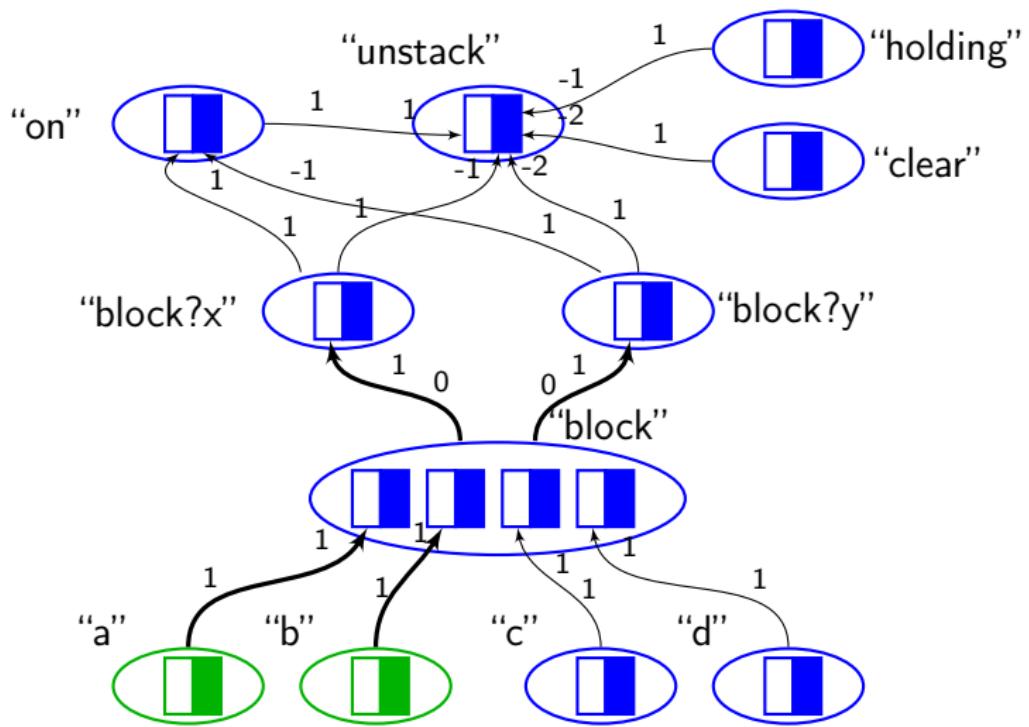
# Пример: генерация личностного смысла



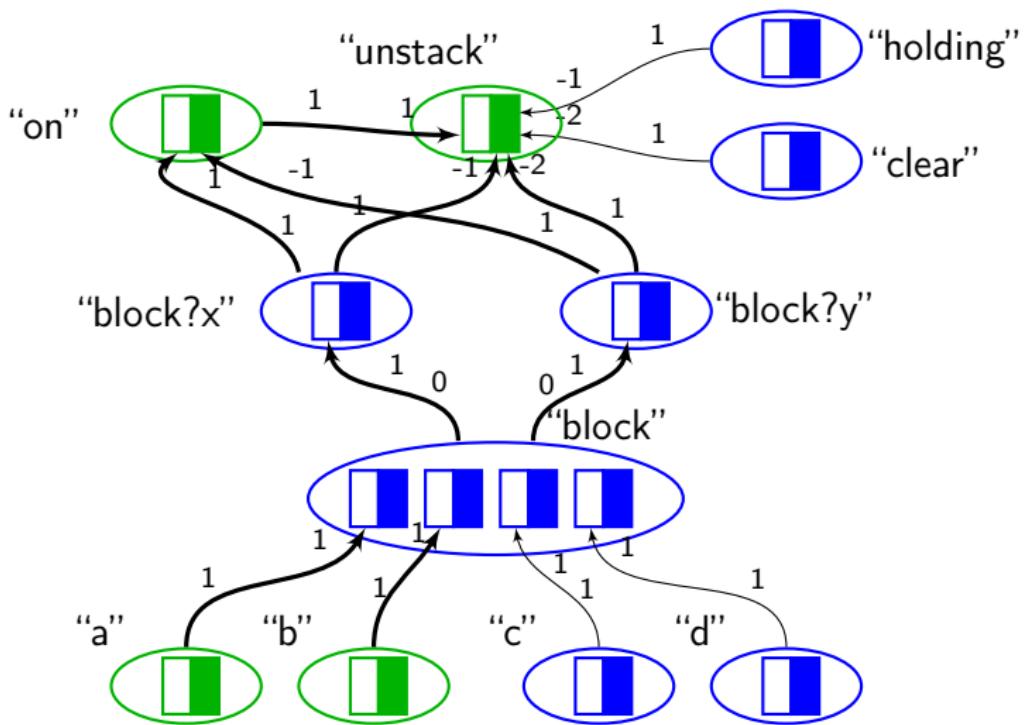
# Пример: генерация личностного смысла



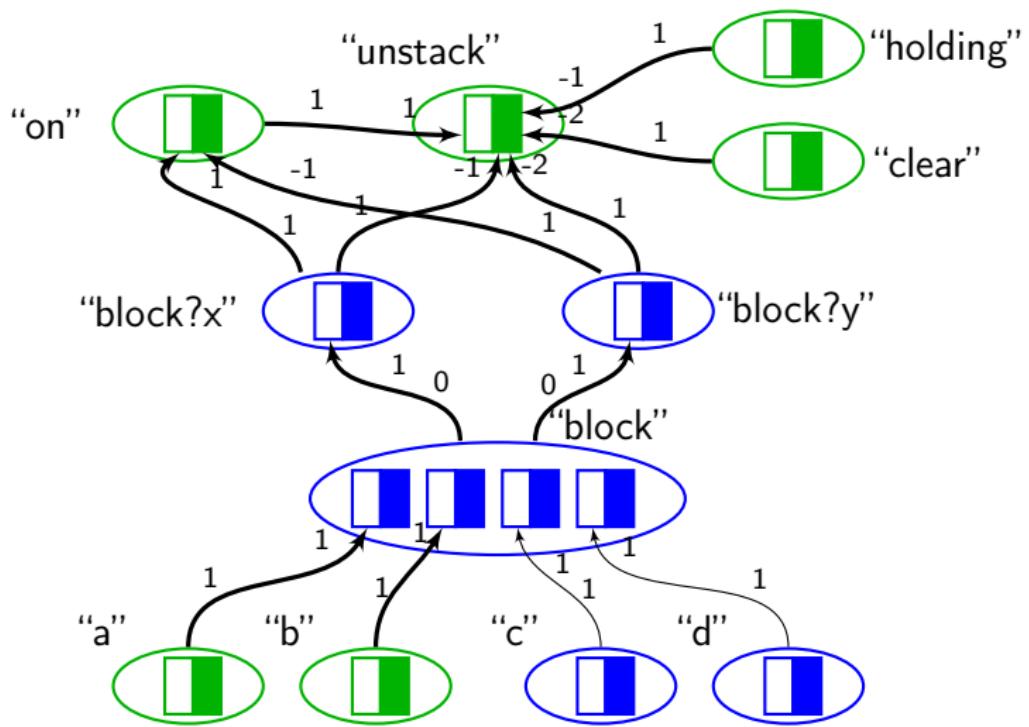
# Пример: генерация личностного смысла



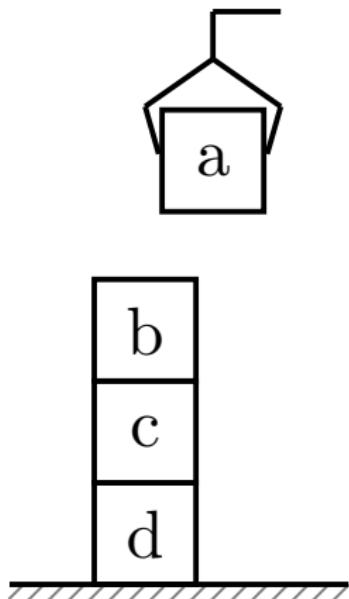
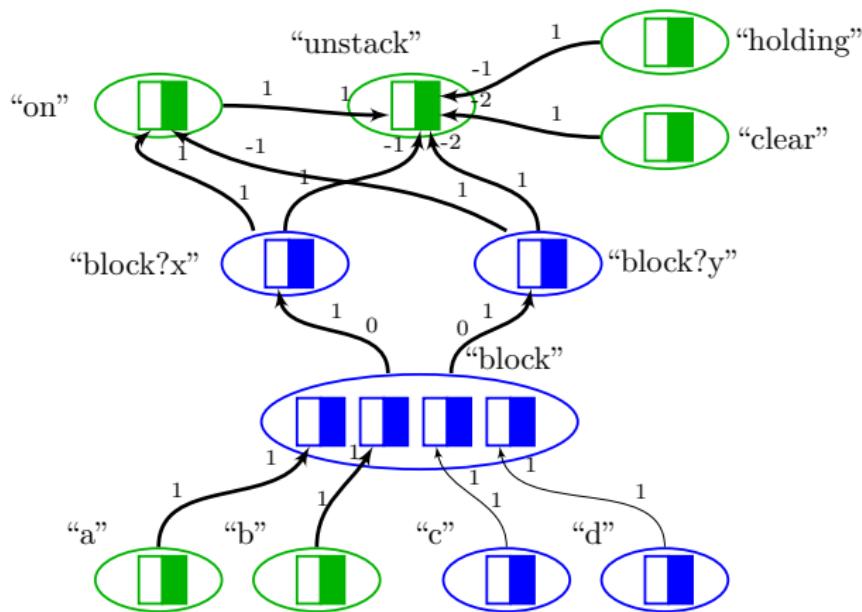
# Пример: генерация личностного смысла



# Пример: генерация личностного смысла

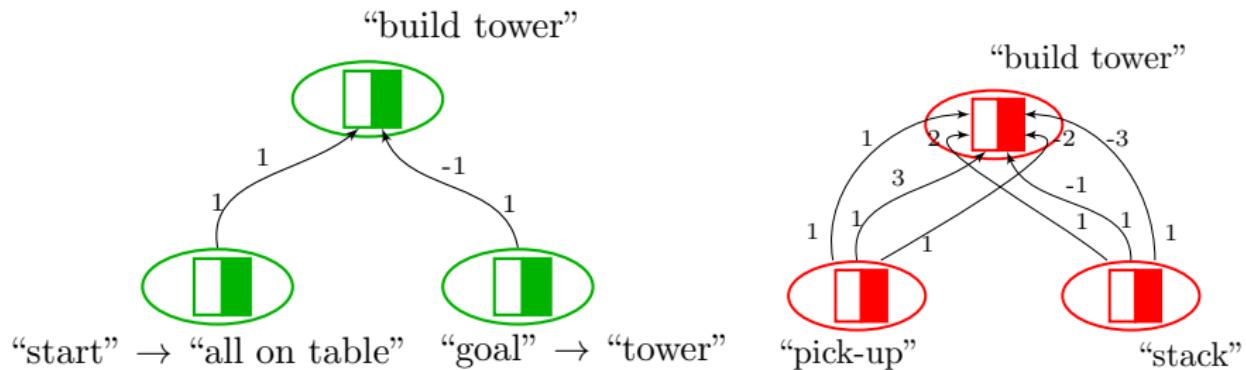


# Пример: текущая ситуация

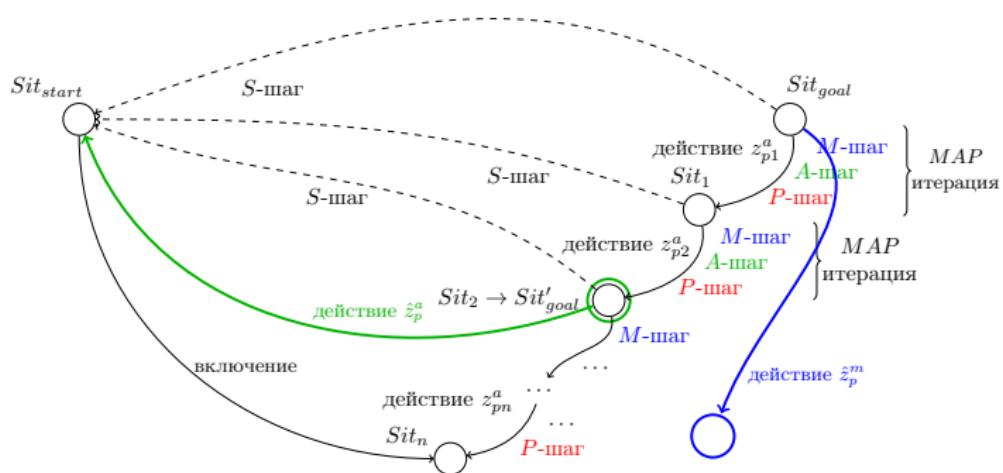


# Обучение в процессе планирования

Образование нового правила и сохранение ситуаций - образование новых каузальных матриц



# Этап целеполагания



Панов, А. И. «Целеполагание и синтез плана поведения когнитивным агентом». Искусственный интеллект и принятие решений. 2018.

Roberts, Mark et al. “Iterative Goal Refinement for Robotics”. Working Notes of the Planning and Robotics Workshop at ICAPS. 2014.

Samsonovich, Alexei V. “Goal reasoning as a general form of metacognition in BICA”. Biologically Inspired Cognitive Architectures. 2014.

## Особенности постановки задачи

Рассматривается случай группового взаимодействия автономных технических объектов (агентов), в котором:

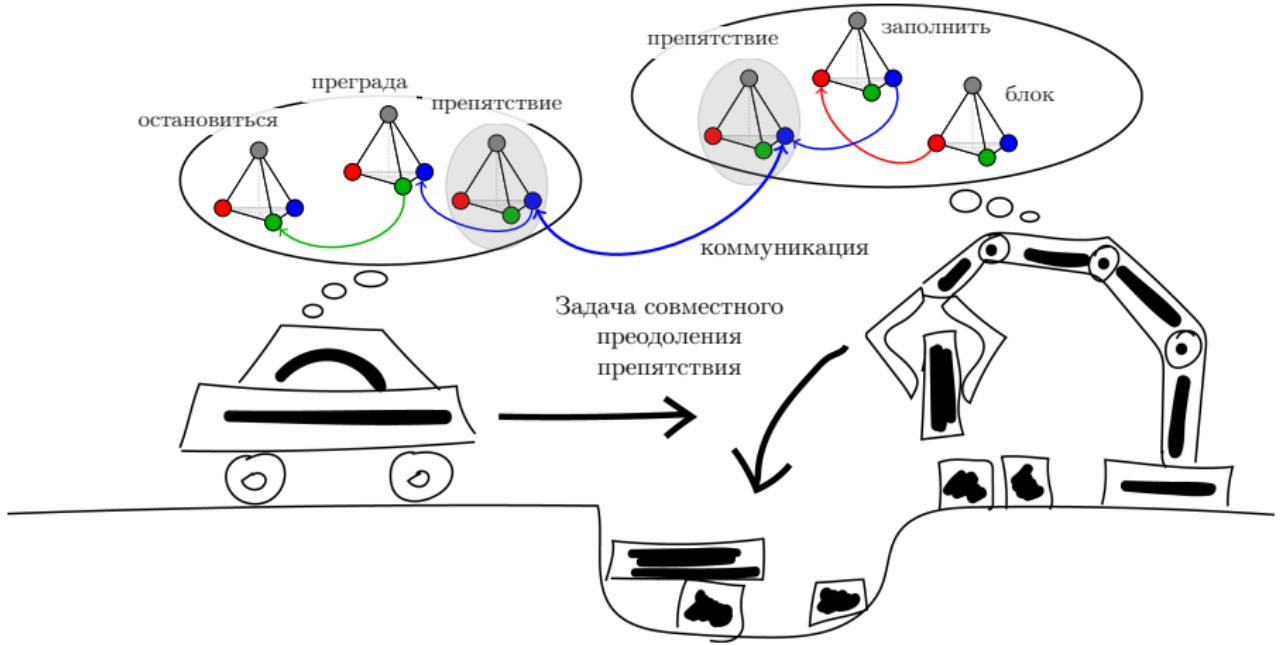
- агенты решают общую задачу (имеют общую цель высшего уровня),
- агенты действуют независимо друг от друга (децентрализованное управление), в т.ч. могут ставить индивидуальные подцели и достигать их,
- агенты обладают различными характеристиками, как техническими, так и когнитивными, т.е. разными стратегиями поведения,
- агенты обладают различными картинами мира,
- агенты действуют в меняющейся среде.

# Требования к представлению знаний

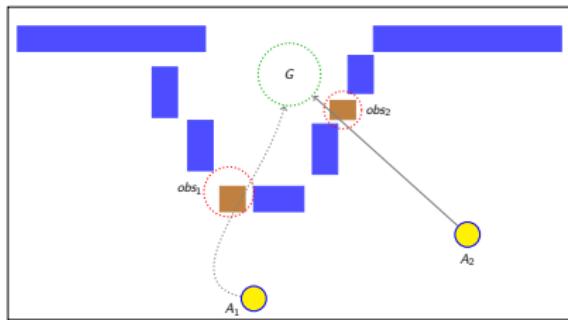
На представление пространственных и временных знаний в задаче согласованного перемещения с такими особенностями налагается ряд ограничений:

- необходимость поддержки некоторого протокола коммуникации, разделение знаний на коммуницируемые и некоммуницируемые (личные),
- необходимость выделения компоненты знания, не зависящей от индивидуальных (личных) характеристик агента,
- требование к наличию механизма связывания реальных объектов внешней среды и процедур их распознавания с символным коммуницируемым представлением (*symbol grounding problem*),
- поддержка механизмов пополнения картины мира (обучение и абстрагирование).

# Практические задачи



# Задача интеллектуального перемещения



## Задача

Целевая область не достижима некоторым агентом самостоятельно (с использованием только методов планирования траектории).

## Решение

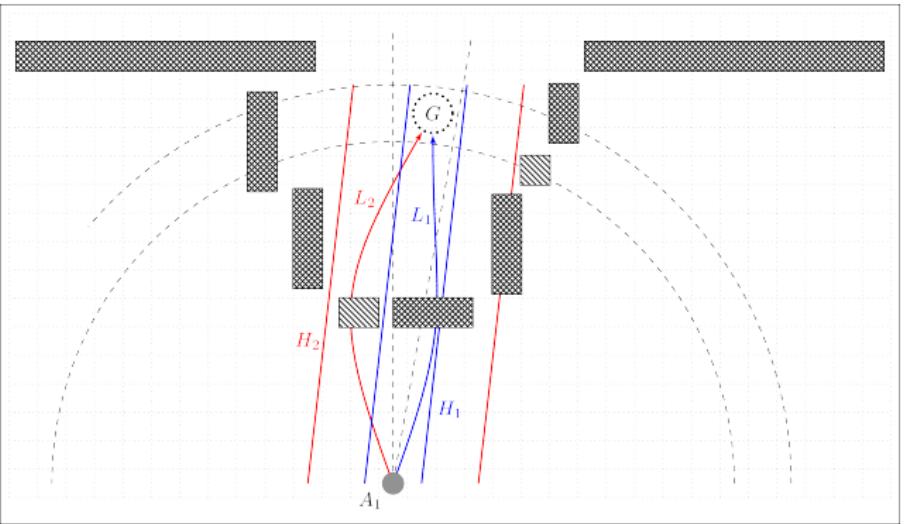
Агенты должны поддерживать коммуникацию и модифицировать свои собственные планы с учетом коалиционных подзадач.

## Особенности:

- Меняющаяся внешняя среда.
- Различные типы препятствий (некоторые могут быть разрушены).
- Агенты обладают различной функциональностью.
- Общая пространственная цель (ВСЕ агенты должны достичь определенной области на карте).

# Распределение ролей при решении задачи

# Представление пространственных знаний



Киселев, Г. А. и А. И. Панов. «Знаковый подход к задаче распределения ролей в коалиции когнитивных агентов». *Труды СПИИ РАН*. 2018.

Панов, А. И. и К. С. Яковлев. «Взаимодействие стратегического и тактического планирования поведения коалиций агентов в динамической среде». *Искусственный интеллект и принятие решений*. 2016.

Панов, А. И. «Представление знаний автономных агентов, планирующих согласованные перемещения». *Робототехника и техническая кибернетика*. 2015.

## Представление действий по перемещению

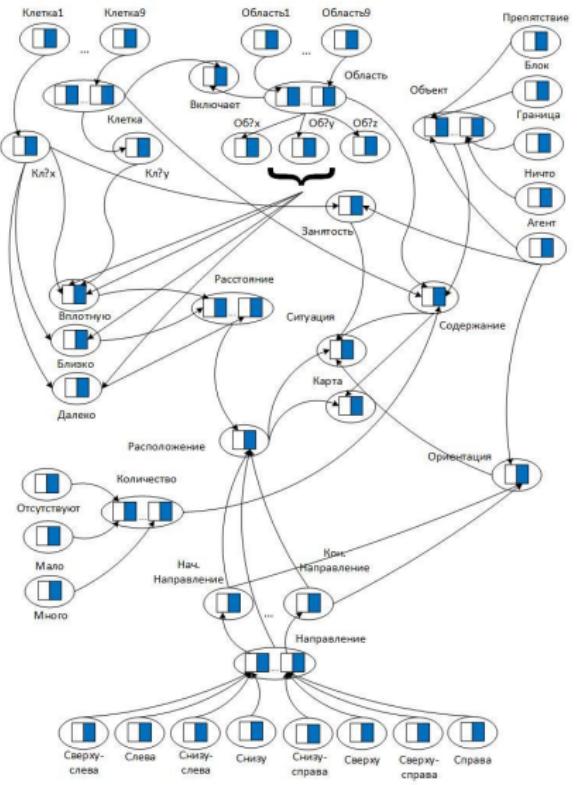
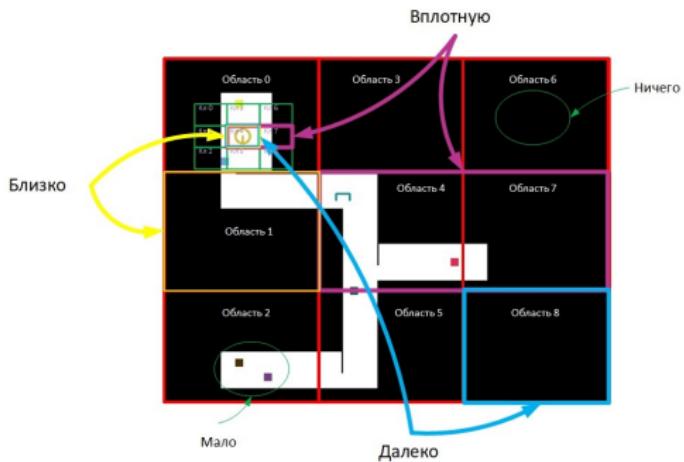
Действия по перемещению — знаки  $s_t$  (признаки  $f_t$ ,  $t$  — тип перемещения), которым соответствуют каузальные матрицы типа  $Z_t$ , состоящие из трёх столбцов

$$z_1 = (I_x, I), z_2 = (I_y, d_u, E), z_3 = (I_y, I, t_v),$$

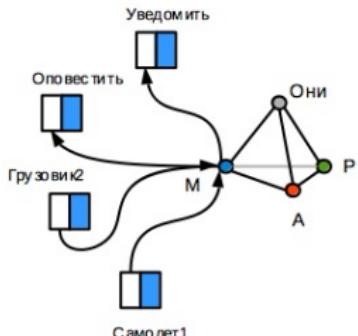
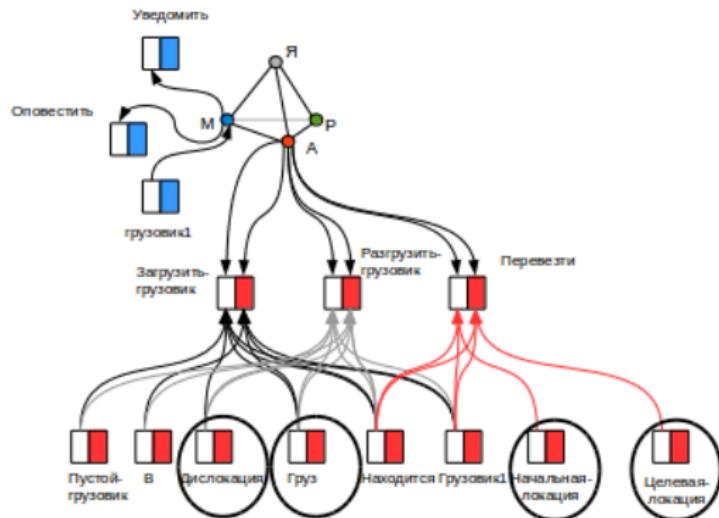
где

- $I_x$ ,  $I_y$  — признаки, соответствующие категории расстояния в пространственной логике (например, вплотную, близко, далеко и др.),
- $d_u$  — признак, соответствующий категории направления в пространственной логике (например, впереди, слева и др.),
- $t_v$  — признак, соответствующий категории времени во временной логике (например, скоро, в будущем и др.),
- $I$  — признак присутствия самого агента,
- $E$  — признак отсутствия препятствия.

# Представление пространственных знаний

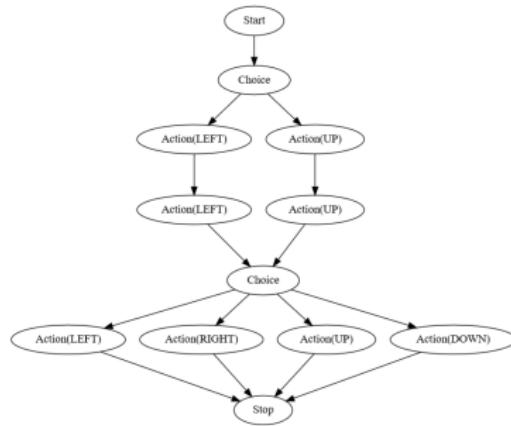
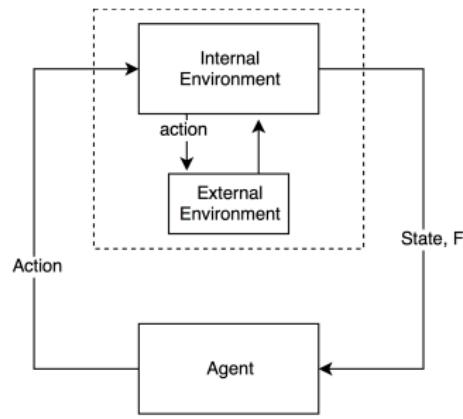


# Знаки Я и Они в алгоритме планирования МАР



Kiselev, Gleb A. and Aleksandr I. Panov. "Synthesis of the Behavior Plan for Group of Robots with Sign Based World Model".  
*Interactive Collaborative Robotics*. 2017.

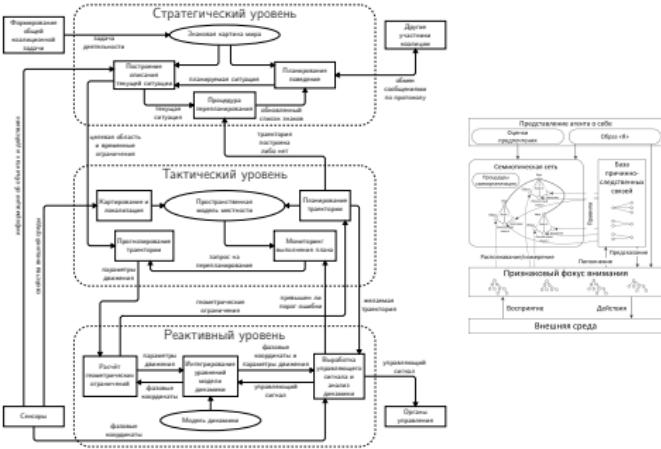
# Подкрепление для формирования каузальных матриц



- Иерархическое обучение с подкреплением для формирования представления о новых действиях.
- Чередование процессов абстрагирования действий и абстрагирования состояний среды.
- Автоматическое формирование иерархии операций.

## Применение для решения интеллектуальных задач

- Моделирование внимания.
  - Образование нового знания (концепта).
  - Планирование поведения.
  - Построение картины мира субъекта на основе текстов.
  - Генерация сообщений на основе картин мира определенного типа (виртуальные ассистенты).
  - Построение многоуровневых архитектур управления.



Emel'yanov, S., D. Makarov, A. I. Panov, and K. Yakovlev. "Multilayer cognitive architecture for UAV control". *Cognitive Systems Research*. 2016.

Макаров, Д. А., А. И. Панов и К. С. Яковлев. «Архитектура многоуровневой интеллектуальной системы управления беспилотными летательными аппаратами». *Искусственный интеллект и принятие решений*. 2015.

Зубарев, Д. В., Д. А. Макаров, А. И. Панов и К. С. Яковлев. «Принципы построения многоуровневых архитектур систем управления беспилотными летательными аппаратами». *Авиакосмическое приборостроение*. 2013.

# Список публикаций

- Panov, Aleksandr I. "Behavior Planning of Intelligent Agent with Sign World Model". *Biologically Inspired Cognitive Architectures*. 2017.
- Осипов, Г. С. и А. И. Панов.** «Отношения и операции в знаковой картине мира субъекта поведения». *Искусственный интеллект и принятие решений*. 2017.
- Emel'yanov, S., D. Makarov, A. I. Panov, and K. Yakovlev. "Multilayer cognitive architecture for UAV control". *Cognitive Systems Research*. 2016.
- Osipov, G. S. "Sign-based representation and word model of actor". *2016 IEEE 8th International Conference on Intelligent Systems (IS)*. 2016.
- Panov, A. I. and K. S. Yakovlev. "Behavior and path planning for the coalition of cognitive robots in smart relocation tasks". *Robot Intelligence Technology and Applications 4*. 2016.
- Panov, Aleksandr I. and Konstantin S. Yakovlev. "Psychologically inspired planning method for smart relocation task". *Procedia Computer Science*. 2016.
- Панов, А. И. и К. С. Яковлев.** «Взаимодействие стратегического и тактического планирования поведения коалиций агентов в динамической среде». *Искусственный интеллект и принятие решений*. 2016.
- Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*. 2015.
- Макаров, Д. А., А. И. Панов и К. С. Яковлев.** «Архитектура многоуровневой интеллектуальной системы управления беспилотными летательными аппаратами». *Искусственный интеллект и принятие решений*. 2015.
- Осипов, Г. С., А. И. Панов и Н. В. Чудова.** «Управление поведением как функция сознания. II. Синтез плана поведения». *Известия Российской академии наук. Теория и системы управления*. 2015.
- Панов, А. И.** «Представление знаний автономных агентов, планирующих согласованные перемещения». *Робототехника и техническая кибернетика*. 2015.
- Осипов, Г. С., А. И. Панов и Н. В. Чудова.** «Управление поведением как функция сознания. I. Картина мира и целеполагание». *Известия Российской академии наук. Теория и системы управления*. 2014.
- Панов, А. И.** «Алгебраические свойства операторов распознавания в моделях зрительного восприятия». *Машинное обучение и анализ данных*. 2014.

# Направления исследований и будущие работы

- Реализация элементов логического вывода в знаковой картине мира.
- Развитие кортикоморфных алгоритмов обучения.
- Разработка иерархических методов обучения с подкреплением и интеграция их в алгоритмы семейства MAP.
- Моделирование рефлексивного поведения, например, добавление этапа выбора эвристики при планирования поведения (reflectiveMAP).
- Исследование мотивационно-потребностного аспекта личного смысла для более тонкого регулирования поведения.
- Подробная разработка алгоритмов коммуникации и согласования знаний.

Приглашаем студентов и аспирантов для участия в научной работе по направлениям работы отдела:

- курсовая и проектная работа в базовой кафедре МФТИ,
- аспирантура в ФИЦ ИУ РАН,
- стажировки (в т.ч. и оплачиваемые) в SLabAI.

# Анонсы

- XVI Национальная конференция по искусственному интеллекту (КИИ 2018) — крупнейшее академическое мероприятие в области ИИ в России (24-27 сентября, подача - 14 апреля).



 КИИ-2018

НАЦИОНАЛЬНАЯ КОНФЕРЕНЦИЯ ПО ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ

[ГЛАВНАЯ](#) [ОРГАНИЗАТОРЫ](#) [ПОДАЧА СТАТЕЙ](#) [МЕСТО ПРОВЕДЕНИЯ](#) [ФЕН](#) [Q](#)

- Пятая всероссийская научная конференция молодых ученых с международным участием «Информатика, Управление и Системный Анализ» (ИУСА-2018) (6-8 июня).



 ИУСА-2018

ИНФОРМАТИКА, УПРАВЛЕНИЕ И СИСТЕМНЫЙ АНАЛИЗ

[О КОНФЕРЕНЦИИ](#) [ОРГАНИЗАТОРЫ](#) [ЛЕКТОРЫ](#) [ПОДАЧА СТАТЕЙ](#) [ОСНОВНЫЕ ДАТЫ](#) [ФЕН](#) [Q](#)

- Презентация кафедры «Системных исследований» — 20 марта, 17:50 в 202 нового корпуса.

# Спасибо за внимание!

Дополнительно лекции на Постнауке:

Идеи нейрофизиологии в ИИ — <https://postnauka.ru/video/81699>

Когнитивная робототехника — <https://postnauka.ru/video/77709>

apanov@hse.ru

Репозиторий — <https://github.com/cog-isa>