

Компьютерное когнитивное моделирование

Александр Панов

Федеральный исследовательский центр «Информатика и управление»
Российской академии наук

18 марта 2016 г.

Кратко о себе

Панов Александр Игоревич, к. ф.-м. н.

- Научный сотрудник лаборатории «Динамические интеллектуальные системы» Института системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук.
- Научный сотрудник и старший преподаватель в Высшей школе экономики.
- Ассистент в Московском физико-техническом институте.
- Член редколлегии журнала Biologically Inspired Cognitive Architectures (BICA Journal).
- Член Российской ассоциации искусственного интеллекта (РААИ).
- Член Сообщества биологически инспирированных когнитивных архитектур (BICA Society).
- Организатор Международной школы по биологически инспирированным когнитивным архитектурам (Fierces on BICA, Москва) и Международной конференции по биологически инспирированным когнитивным архитектурам (BICA-2016, Нью-Йорк).
- Член рабочей группы «Нейронет» Национальной технологической инициативы.
- Руководитель проектов РФФИ мол_а и мол_а_дк.



Научные интересы

- *Когнитивное компьютерное моделирование*: планирование поведения, модели внимания, восприятия, принятия решений и обучения, знаковые системы.
- *Многоагентные системы*: образование коалиций, распределение ролей в коллективе, целеполагание.
- *Анализ данных*: выявление причинно-следственных связей, анализ психологических и медицинских данных.
- *Распознавание изображение*: выявление объектов на сложных сценах, рекуррентные и глубокие нейронные сети.
- *Системы управления*: управление поведением, многоуровневые архитектуры, робототехника.

Подгруппа 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.
- Петров Александр
- Киселев Глеб, асп. ИСА РАН
- Скрыник Алексей, студ. РГАТУ
- Кудинов Антон, студ. ВШЭ
- Филин Дмитрий, студ. ВШЭ

Некоторые публикации

- Emel'yanov, Stanislav et al. "Multilayer cognitive architecture for UAV control". In: *Cognitive Systems Research* 39 (2016), pp. 58–72.
- Panov, Aleksandr I. and Konstantin S. Yakovlev. "Behavior and path planning for the coalition of cognitive robots in smart relocation tasks". In: *Robot Intelligence Technology and Applications 4. Advances in Intelligent Systems and Computing*. 2016, (In Press).
- Osipov, G. S., A. I. Panov, and N. V. Chudova. "Behavior Control as a Function of Consciousness. II. Synthesis of a Behavior Plan". In: *Journal of Computer and Systems Sciences International* 54.6 (2015), pp. 882–896.
- Panov, A. I., A. V. Shvets, and G. D. Volkova. "A Technique for Retrieving Cause and Effect Relationships from Optimized Fact Bases". In: *Scientific and Technical Information Processing* 42.6 (2015), pp. 420–425.
- Макаров, Д. А., А. И. Панов и К. С. Яковлев. «Архитектура многоуровневой интеллектуальной системы управления беспилотными летательными аппаратами». В: *Искусственный интеллект и принятие решений* 3 (2015), с. 18–33.
- Панов, А. И. «Представление знаний автономных агентов, планирующих согласованные перемещения». В: *Робототехника и техническая кибернетика* 4(9) (2015), с. 34–40.
- Панов, А. И., А. В. Швец и Г. Д. Волкова. «Метод извлечения причинно-следственных связей с использованием оптимизированных баз фактов». В: *Искусственный интеллект и принятие решений* 1 (2015), с. 27–34.
- Osipov, G. S., A. I. Panov, and N. V. Chudova. "Behavior control as a function of consciousness. I. World model and goal setting". In: *Journal of Computer and Systems Sciences International* 53.4 (2014), pp. 517–529.
- Panov, A. I. "Extraction of Cause-Effect Relationships from Psychological Test Data Using Logical Methods". In: *Scientific and Technical Information Processing* 41.5 (2014), pp. 275–282.

Новая парадигма исследований в ИИ

- Искусственный интеллект исследует средства решения интеллектуальных задач, т.е. таких задач, которые не имеют известного алгоритма решения.
- Компьютерная аналогия перестала отвечать новым требованиям, предъявляемых к интеллектуальным задачам.
- Необходимо рассмотреть науку «Искусственный интеллект» в контексте всех когнитивных наук.
- В этом случае целью искусственного интеллекта становится построение правдоподобных гипотез (моделей) работы когнитивных функций человека, которые бы согласовывались как с экспериментальными данными когнитивных психологов, так и с нейрофизиологическими данными о строении и функционировании мозга человека.

Биологическое и психологическое правдоподобие

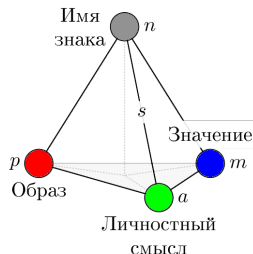
Возникает потребность в разработке новых биологически и психологически правдоподобных методов (biologically and psychologically inspired methods) моделирования когнитивных функций в искусственном интеллекте.

Пример:

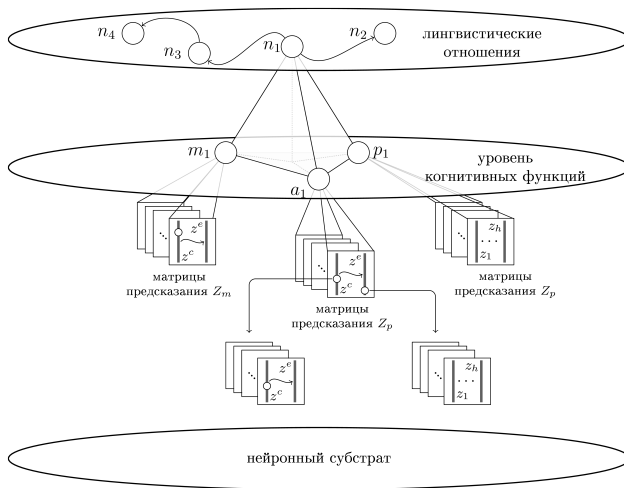
- использование *культурно-исторического подхода и психологической теории деятельности* (Лурия, Выготский, Леонтьев), как основы описания работы когнитивных функций,
- использование идей *прикладной семиотики* (Поспелов, Осипов),
- использование моделей работы *кортикальной колонки* (Маунткэсл, Хокинс) и *нейронного ансамбля* (Эйдельман, Инваницкий) в качестве нейрофизиологического «заземления» (grounding).

Знаковая картина мира

- Модель представления субъекта о действительности — картина мира.
- Элемент картины мира — знак, четырех-компонентная структура.
- Знаки опосредуют как объекты, так и процессы, ситуации, внутренние характеристики и т.п.



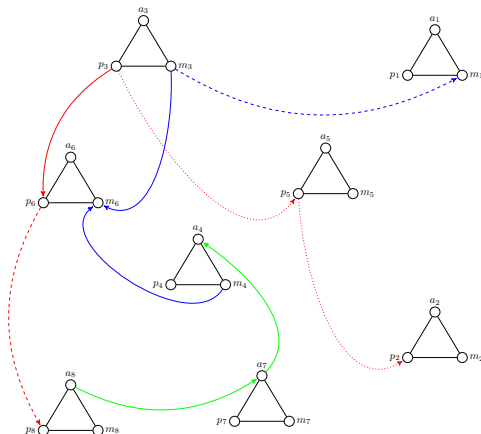
Уровни представления



Знаковая модель

- Когнитивные функции — это процессы, качественно зависимые от субстрата и воспроизводимые с помощью вычислений («заземленная» структура компонент знака).
- Вместо когнитивного солипсизма — взаимодействие с внешней средой (теория деятельности, сигналы внешней среды).
- Системный подход в задании причинно-следственных связей (пространство состояний по Тонони).
- Необходимость коллективного взаимодействия в процессе обучения и действия.

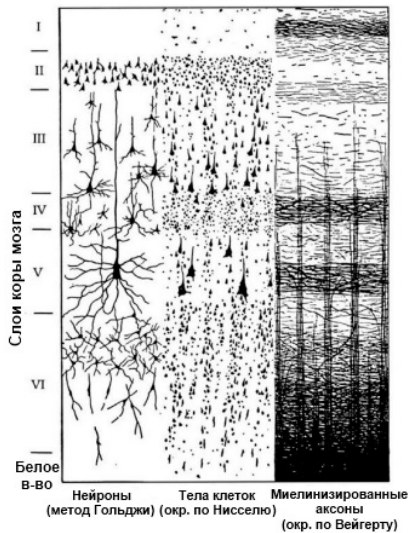
Картина мира субъекта деятельности



Семиотическая сеть
 $H = \langle H_P, H_A, H_M \rangle$, где

- $H_P = \langle 2^P, \mathfrak{R}_P \rangle$ — семантическая сеть на множестве образов знаков,
- $H_A = \langle 2^A, \mathfrak{R}_A \rangle$ — семантическая сеть на множестве значений знаков,
- $H_M = \langle 2^M, \mathfrak{R}_M \rangle$ — семантическая сеть на множестве личностных смыслов знаков.

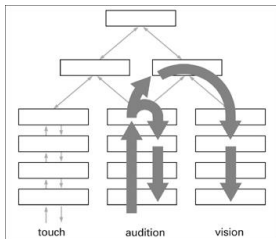
Нейронный субстрат



Основные свойства модели и используемые упрощения

Принимается следующие гипотезы:

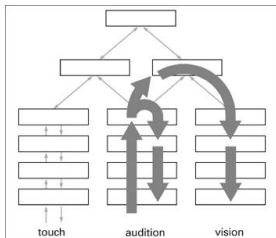
- неокортекс состоит из зон (регионов), состоящих в свою очередь из колонок и имеющих одинаковое строение на всех участках коры;
- колонки в регионе объединены латеральными связями;
- таламус формирует последовательности паттернов за счет задержки возбуждения/торможения.



Основные свойства модели и используемые упрощения

Основные свойства:

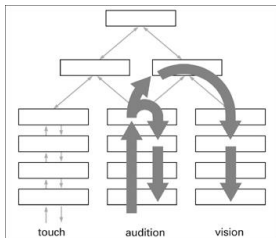
- хранение последовательности паттернов в инвариантной форме,
- воспроизведение паттернов автоассоциативно,
- хранение паттернов в иерархической системе,
- использование обратной связи для предсказания поступающей на данный уровень иерархии информации.



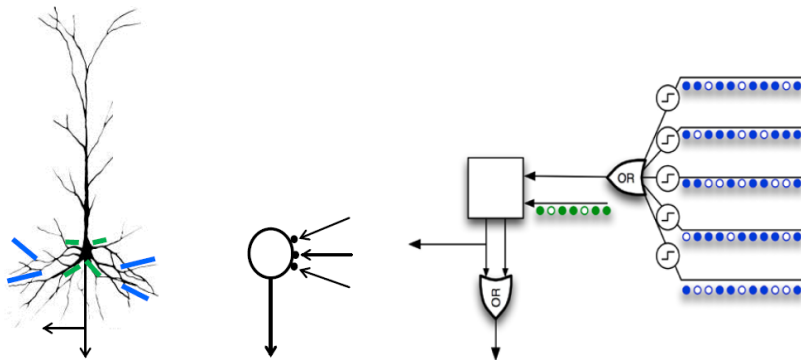
Основные свойства модели и используемые упрощения

Упрощения:

- дискретность во времени,
- простейшая строгая иерархия со связями только между ближайшими уровнями,
- гипотеза одинаковой длительности распознаваемых явлений в рамках одного региона,
- пороговая модель принятия решений в случае неопределенности результата распознавания,
- подавление непредвиденного сигнала,
- отсутствие моторной составляющей обратной связи.

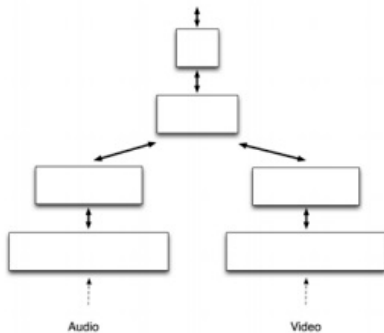
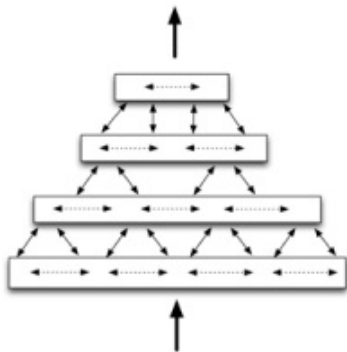


Формальная модель нейрона

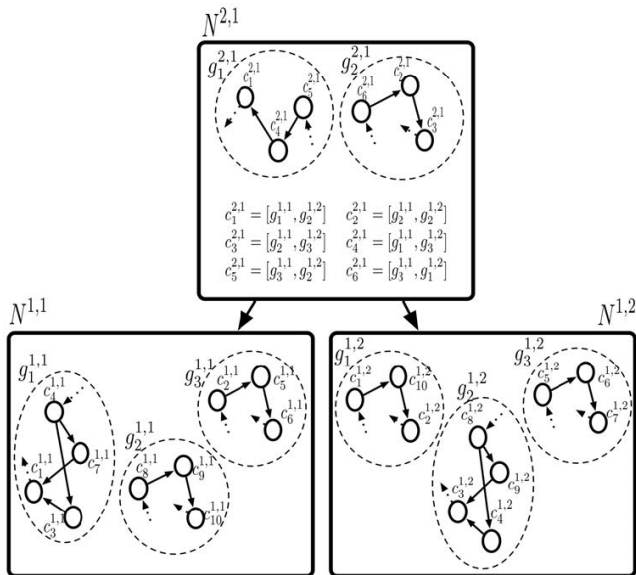


- Проксимальный дендритный сегмент — прямая активация.
- Дистальные дендритные сегменты — латеральный вход и состояние предсказания.

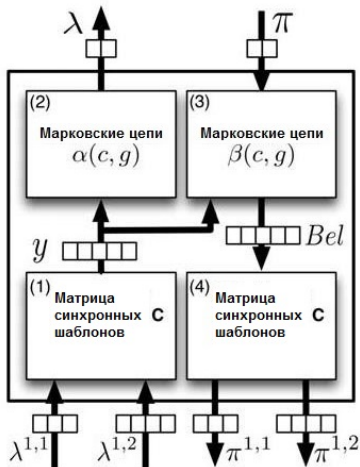
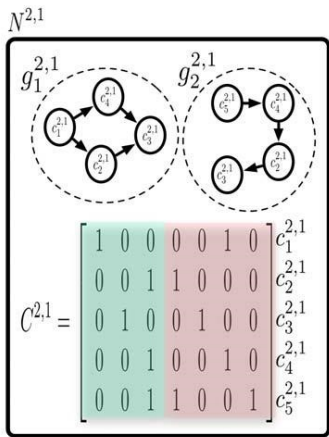
Иерархическая организация нейронов



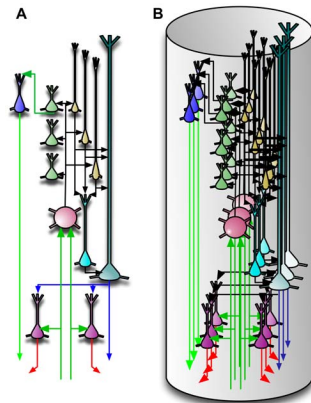
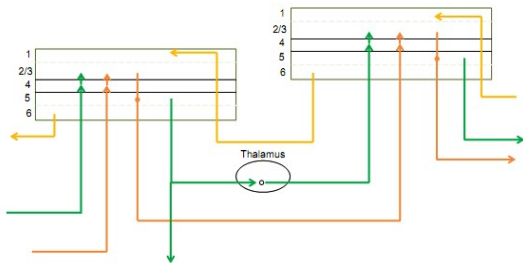
Иерархическая модель



Иерархическая модель



Послойная организация



Образная компонента знака

При окончании процесса обучения синапсы определяют как вертикальные связи между узлами, так и горизонтальные связи в рамках одного узла.

Далее будет рассмотрена автоматная модель процесса восприятия, на основе которой будут определены образная компонента знака.

Каждому узлу соответствует набор матриц предсказания, которые формируются в результате процесса обучения по алгоритму НТМ.

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).

Спасибо за внимание!

ФИЦ ИУ РАН, лаб. «Динамические интеллектуальные системы»,
pan@isa.ru