



Co-evolutionary
Hybrid Intelligence

Co-Evolutionary Hybrid Intelligence and its applications

K.Krkin, Y.Shichkina

12th January 2023

AI genesis

1956

Automatic computers (autonomy)

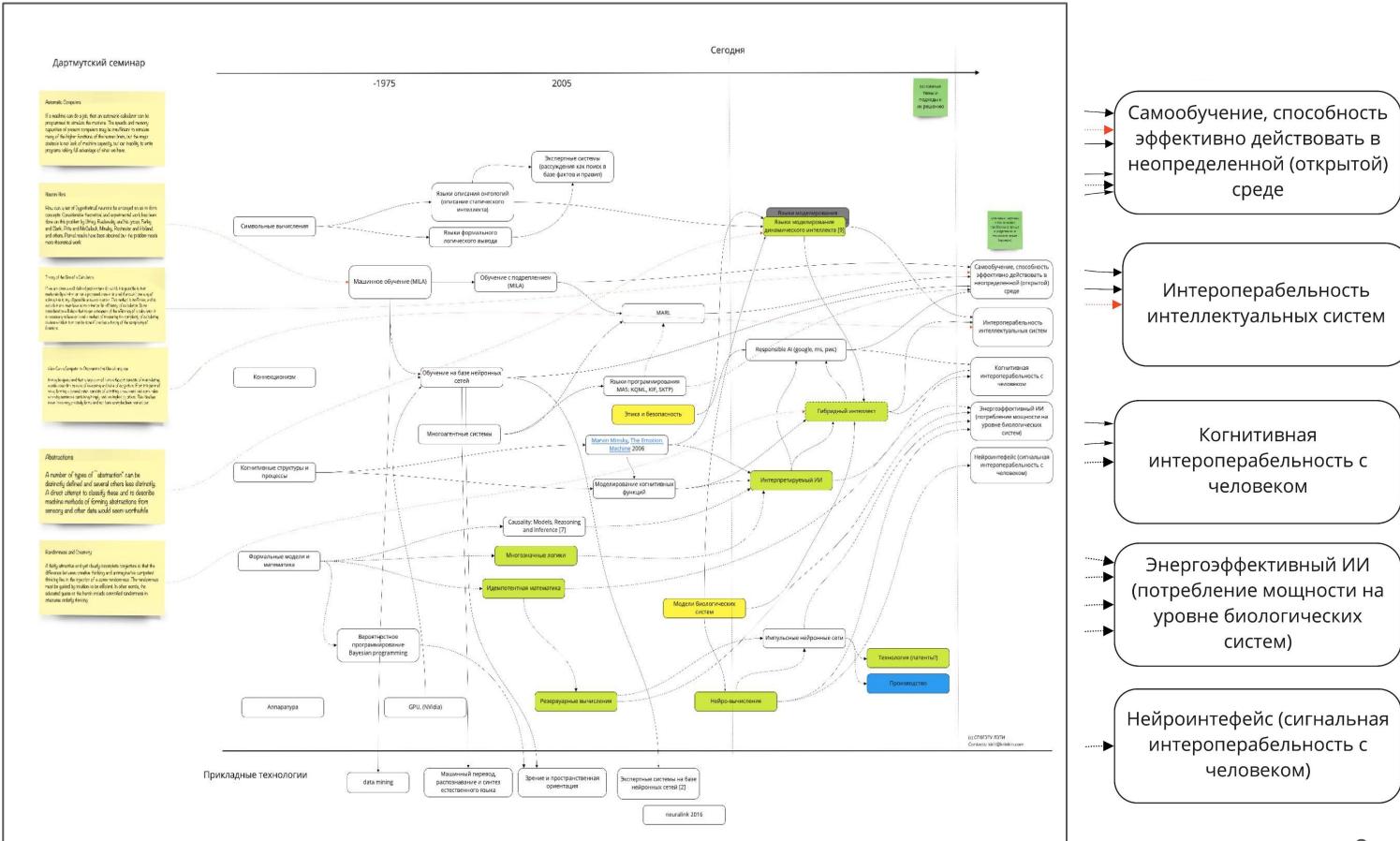
Neuron Nets

Randomness and Creativity

How can a computer be programmed to use language

Theory of the size of Calculation

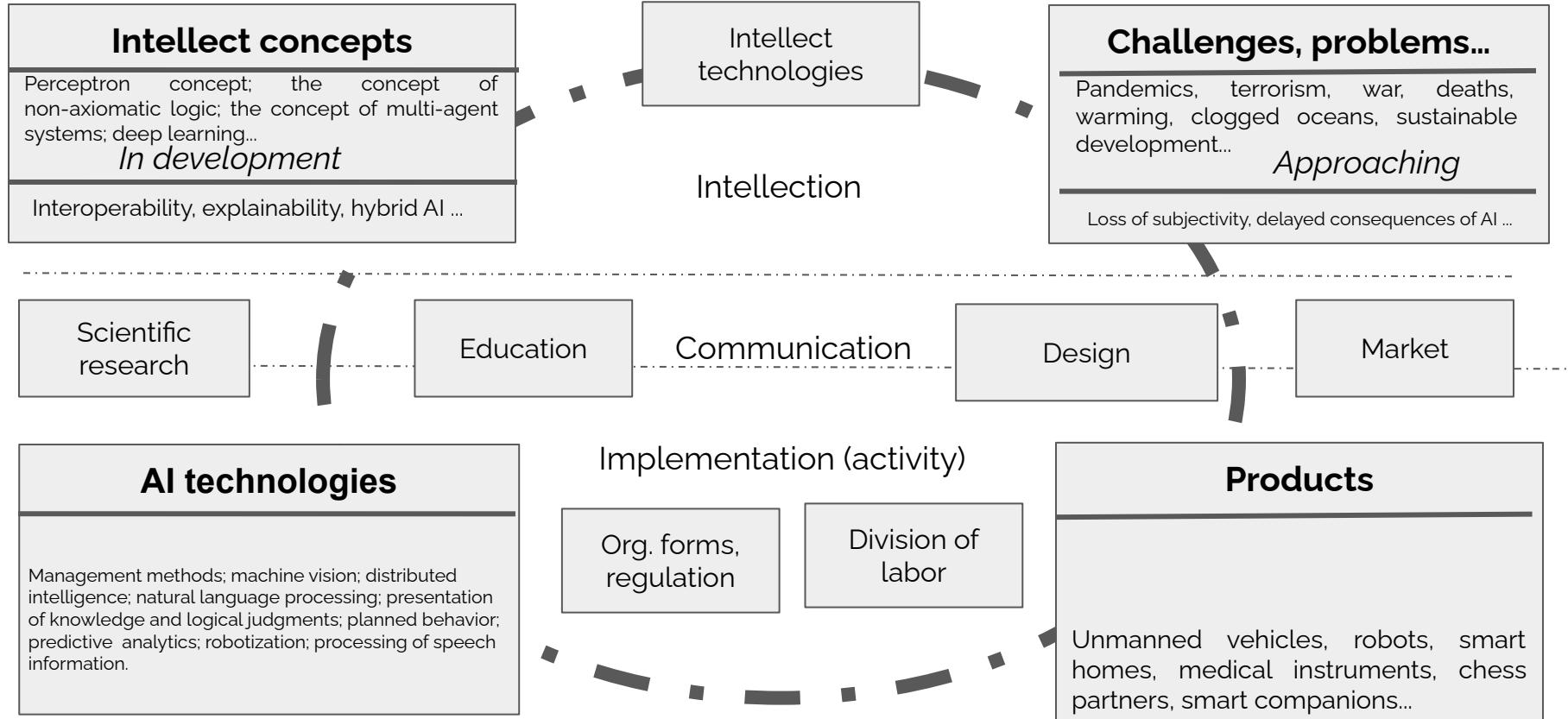
Abstractions



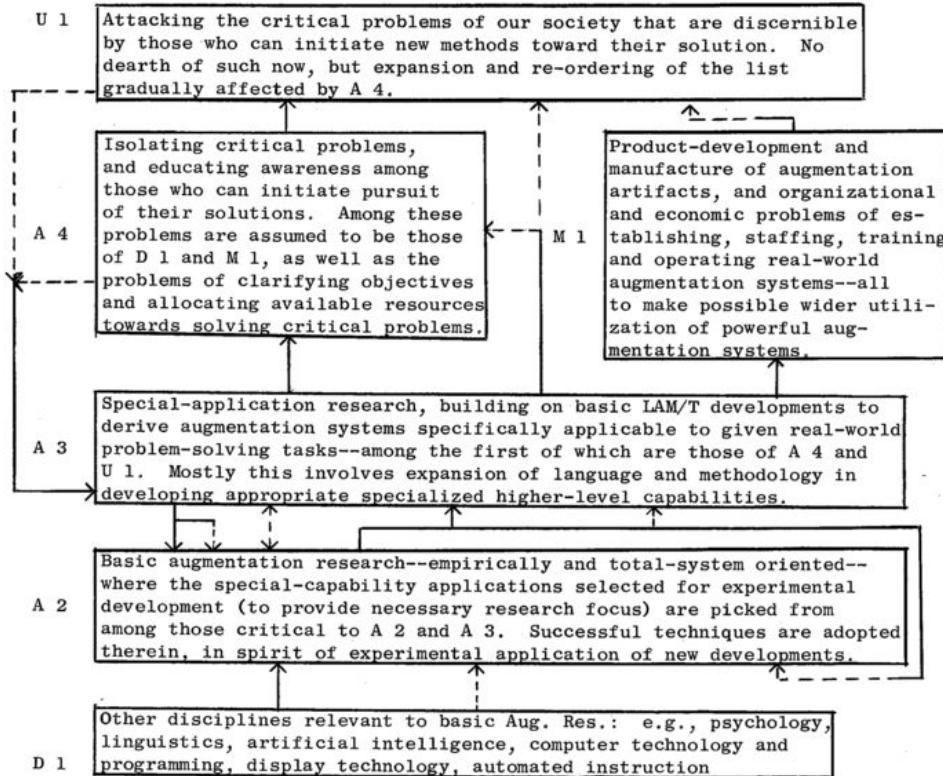
Key points for AI development

- AI is **not just a new tool** like other technologies, it can be able to improve themselves, it requires **new approaches to integration into a society**
- **Explainability of AI systems** is a fundamental for the successful human-machine collaboration and cooperation
- Expansive Machine Learning approach with exhausting computational power is not a “silver bullet”
- Mathematically proven fundamental approaches for explainability and power efficiency should be developed
- **Human - Machine (AI) integration and their co-evolution** are important for the building reliable future
- Introducing AI systems for the wide spectrum of application is essential

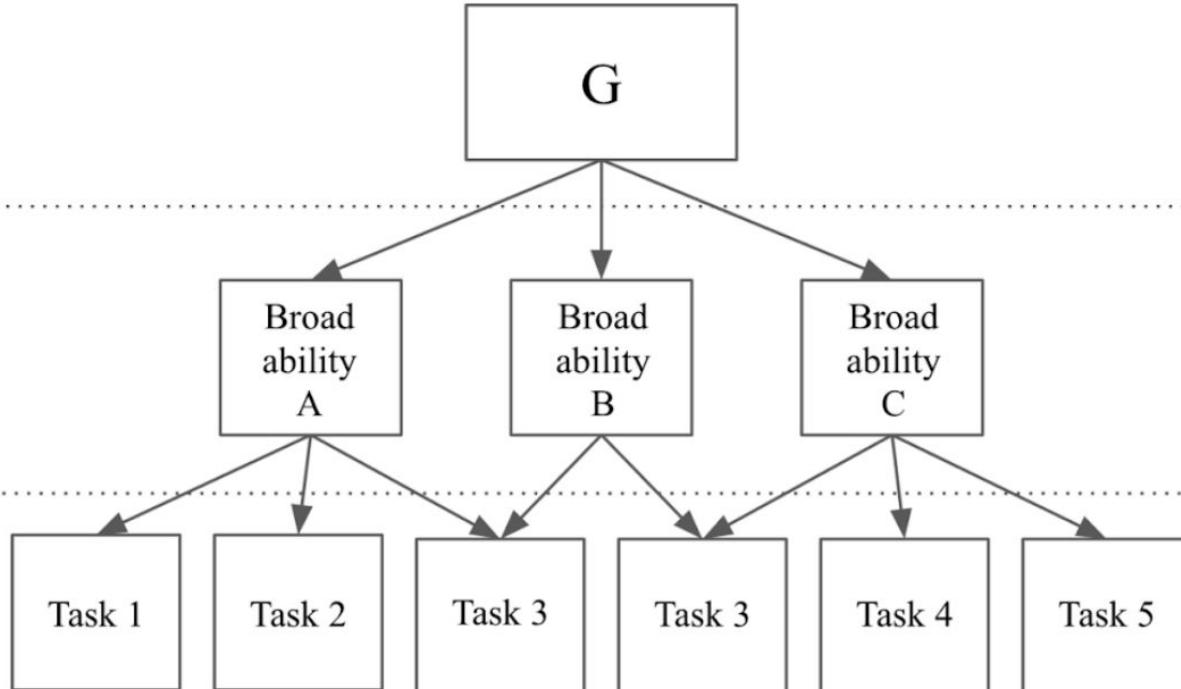
Conceptual scheme of the development of intelligence



Augmenting Human Intellect: A Conceptual Framework



Мощность интеллекта



General intelligence
Extreme generalization

Broad cognitive abilities
Broad generalization

Task-specific skills
Local generalization
(or no generalization, i.e. absence of uncertainty)

Co-Evolutionary vs Hybrid Intelligence



Computer, 2020 (DOI:10.1109/MC.2020.2996587)

Co-evolutionary hybrid intelligence

Kirill Krinkin
Alexander Popov International
Innovation Institute for AI,
Cybersecurity and Communications,
SPbETU "LETI"
St. Petersburg, Russia
kirill@krinkin.com

Yulia Shchikina
Alexander Popov International
Innovation Institute for AI,
Cybersecurity and Communications,
SPbETU "LETI"
St. Petersburg, Russia
strange.y@mail.ru

Andrey Ignatyev
MGIMO University
Center for Global IT
Cooperation (CGITC)
Moscow, Russia
andrey.ignatyev.g@gmail.com

Abstract— Artificial intelligence is one of the drivers of modern technological development. The current approach to the development of intelligent systems is data centric. It has several limitations: it is fundamentally impossible to collect data for modeling complex objects and processes; training neural networks requires huge computational and energy resources; solutions are not explainable. The article discusses an alternative approach to the development of artificial intelligence systems based on the human-machine hybridization and their co-evolution.

- ability to transfer of knowledge and experience from one context (ontology) to another on the fly

It is not the purpose of this paper to find the most relevant definition of AI. Further, in the paper, we will use on the following definition of artificial intelligence.

AI is an intelligence implemented on an artificial substrate (machine), where the *Intelligence* is a measurable property of a system in the ability to acquire, accumulate, and use knowledge, based on the previous experience of solving

- *Hybrid Intelligence*

- *the combination of human and machine intelligence, augmenting human intellect and capabilities instead of replacing them and achieving goals that were unreachable by either humans or machines*

- *Co-evolutionary Hybrid Intelligence*

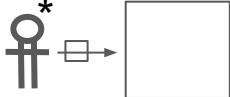
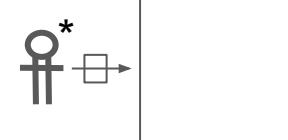
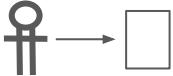
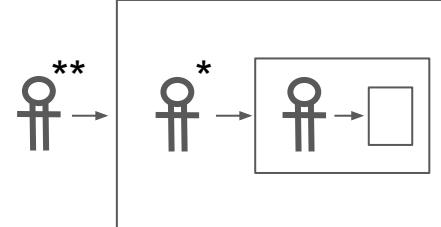
- *is a symbiosis of artificial and natural intelligence, mutually developing, teaching, and complementing each other in the process of co-evolution.*

- *The key pillars of Co-evolutionary Hybrid Intelligence are:*

- *Cognitive interoperability*
- *Human-Machine Co-evolution*
- *Reflectiveness*

K. Krinkin, Y. Shchikina and A. Ignatyev, "Co-evolutionary hybrid intelligence," 2021 5th Scientific School Dynamics of Complex Networks and their Applications (DCNA), 2021, pp. 112-115, doi: 10.1109/DCNA53427.2021.9587002.

Понятие гибридного интеллекта

Объект Время	t1	t2	t3
Искусственный интеллект			
Гибридный интеллект			

Определения по Gartner

Digitalization is the use of digital technologies to change a business model and provide new revenue and value-producing opportunities; it is the process of moving to a digital business.

Digital business is the creation of new business designs by blurring the digital and physical worlds.

A digital twin is a digital representation of a real-world entity or system. The implementation of a digital twin is an encapsulated software object or model that mirrors a unique physical object, process, organization, person or other abstraction. Data from multiple digital twins can be aggregated for a composite view across a number of real-world entities, such as a power plant or a city, and their related processes.

Artificial intelligence applies advanced analysis and logic-based techniques, including machine learning, to interpret events, support and automate decisions, and take actions

Intelligence definitions

- Intelligence is the ability to learn from experience and adapt to the environment (Gardner 1983; Sternberg 1984).
- Intelligence is the ability “to achieve one’s goals in life, within one’s sociocultural context.” (Sternberg, 2005).
- Intelligence is “a biopsychological potential to process information...to solve problems...” (Gardner 1999).
- Intelligences can be considered as skills that humans learn over time to adapt to their environment (Schlinger 2003).
- Intelligence is the ability of knowledge and reasoning, problem-solving, learning, communicating, perceiving, and acting (Russell and Norvig 2010)
- Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience (Darren Stevens, 2021)

AI definitions

Pei Wang,

On Defining Artificial Intelligence

Definition	The essence	Disadvantages
STRUCTURE-AI	<p>It requires an AI system to go through isomorphic states or structure changes as the brain does when they are given similar input, which will produce similar output</p>	<p>But fundamental intuition behind AI is that human intelligence is a special form of a general notion of intelligence, which have other forms</p>
BEHAVIOR-AI	<p>A computer system's verbal behaviors are indistinguishable from that of a human being</p> <p>Вербальное поведение компьютерной системы неотличимо от поведения человека</p>	<p>But to expect an AI to behave exactly like a human is too anthropocentric to non-human intelligence, since human behaviors not only depend on our intellectual competence and mechanisms, but on biological, evolutionary, and cultural factors that are unique to humans.</p>
CAPABILITY-AI	<p>The action (solution) the computer produces for a percept (problem) is similar to the action produced by a human to a similar percept</p> <p>Действие (решение), которое компьютер производит для восприятия (проблемы), аналогично действию произведенное человеком на аналогичное восприятие</p>	<p>But it leads to an iconic result: as soon as a computer system is built to solve a problem successfully, the problem is no longer "only solvable by the human mind," so does not need intelligence anymore</p>
FUNCTION-AI	<p>The function that maps a percept (input problem) into an action (output solution) in the computer is similar to that of a human. Here the function can correspond to searching, reasoning, learning, etc., and since the focus is on the functions (i.e., input-output mappings), the concrete input and output values of the two agents do not have to be similar to each other.</p>	<p>Since each function can be specified in isolation, there is little motivation to take the other functions into consideration, as this will complicate the situation, and may violate the basic assumptions shared by the researchers working on the function.</p>
PRINCIPLE-AI	<p>Based on fundamental principles that can uniformly explain the relevant phenomena.</p>	<p>Until a system built according to such a definition is widely acknowledged as intelligent, most people will not be convinced that a good definition of AI can be obtained in this way.</p>

Intelligence ecosystem (Friston's memorandum)

- Intelligence – as the capacity of systems to generate evidence for their own existence
- Active Inference
- Intellectual agent should be embodied
- Intellectual agent can be
 - Can be biological
 - Machine
 - Combination
- Belief propagation (graphs, and networks)
- Communication is a key

Cognitive science

COGNITION, BRAIN, AND CONSCIOUSNESS

INTRODUCTION TO COGNITIVE NEUROSCIENCE

Second Edition
Bernard J. Baars
Nicole M. Gage



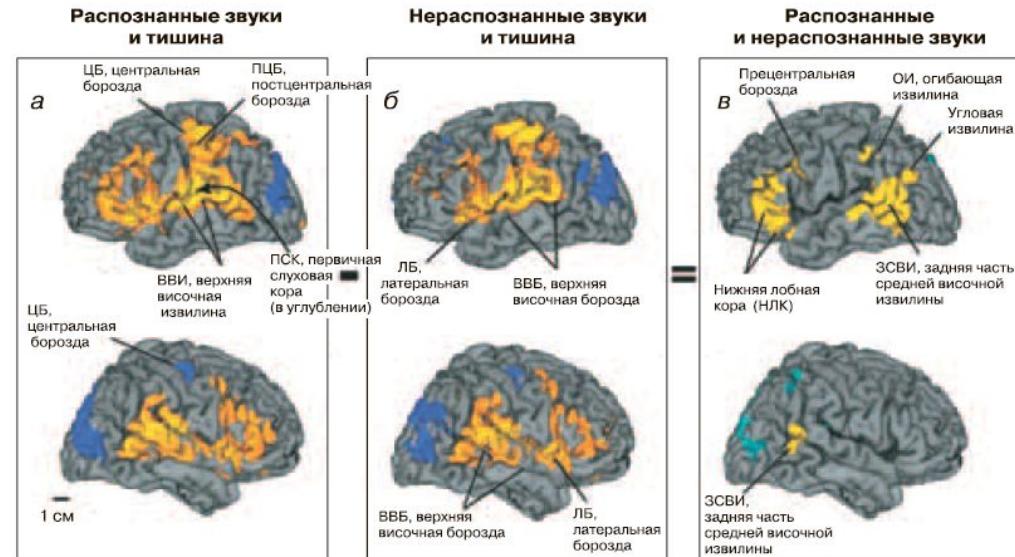
AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

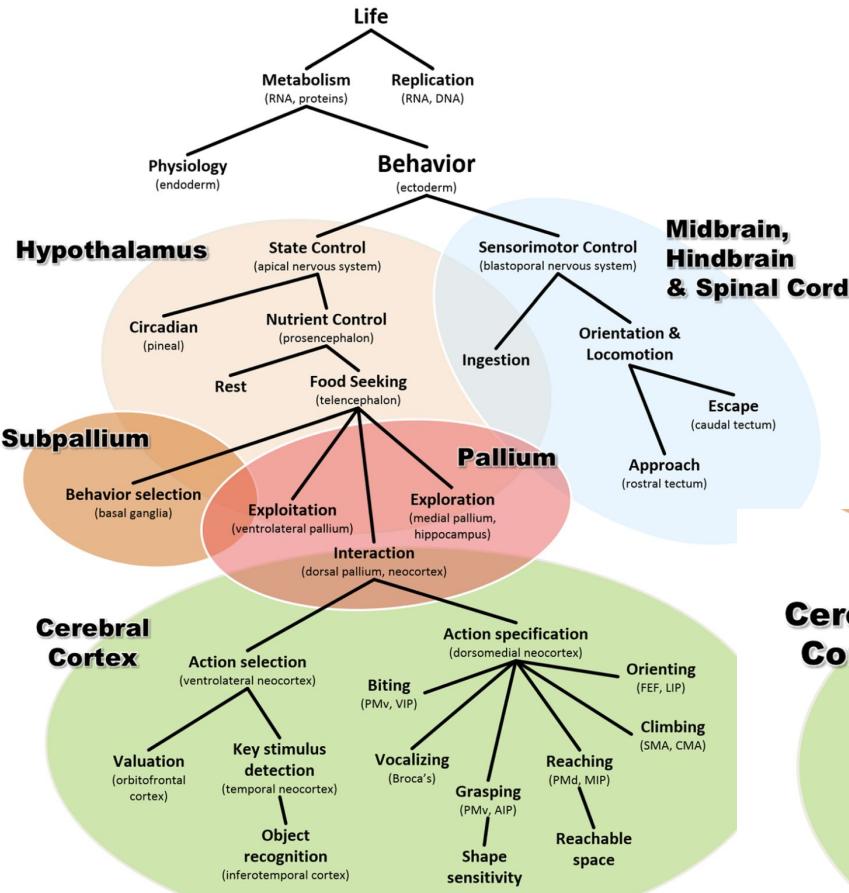
Academic Press is an imprint of Elsevier



Key point:

It makes no sense to create AI as an exact structure of the brain, because different parts of the cerebral cortex are responsible for the implementation of different cognitive functions at different age periods.





Attention, Perception, & Psychophysics (2019) 81:2265–2287
<https://doi.org/10.3758/s13414-019-01760-1>

TIME FOR ACTION: REACHING FOR A BETTER UNDERSTANDING OF THE DYNAMICS OF COGNITION

Resynthesizing behavior through phylogenetic refinement

Paul Cisek¹

Cerebral Cortex

Cognitive processes

Basic mental processes

- Sensation
- Attention
- Perception

Complex mental operations

- Memory
- Learning
- Language use
- Problem solving
- Decision making
- Reasoning
- Intelligence

Consciousness, attention ...

OPEN ACCESS Freely available online



From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0

2014

Masafumi Oizumi^{1,2*}, Larissa Albantakis^{1*}, Giulio Tononi^{1*}

1 Department of Psychiatry, University of Wisconsin, Madison, Wisconsin, United States of America, **2** RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

Abstract

This paper presents Integrated Information Theory (IIT) of consciousness 3.0, which incorporates several advances over previous formulations. IIT starts from phenomenological axioms: information says that each experience is specific – it is what it is by how it differs from alternative experiences; integration says that it is composed of interdependent components; exclusion says that it has unique borders and are formalized into postulates that prescribe how physical mechanisms, such as neurons, interact to generate experience (phenomenology). The postulates are used to define “information about a difference” within a system, and integrated information as information that is specified by its parts. By applying the postulates both at the level of individual mechanisms, IIT arrives at an identity: an experience is a maximally irreducible set of concepts in qualia space, and the set of elements that generates it.



HYPOTHESIS AND THEORY
published: 14 November 2017
doi: 10.3389/frobt.2017.00060



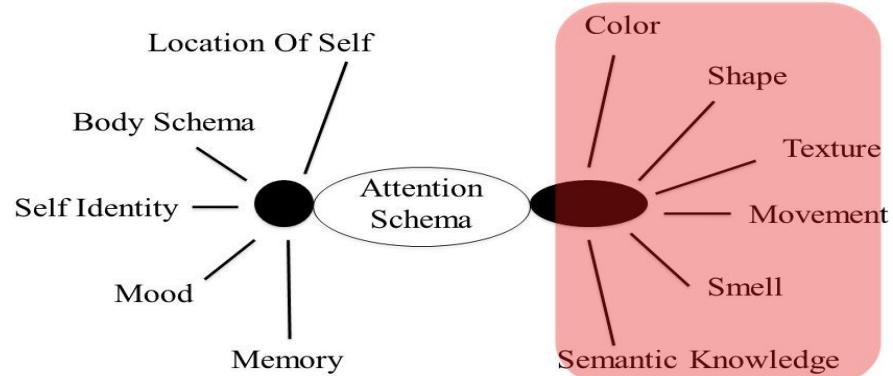
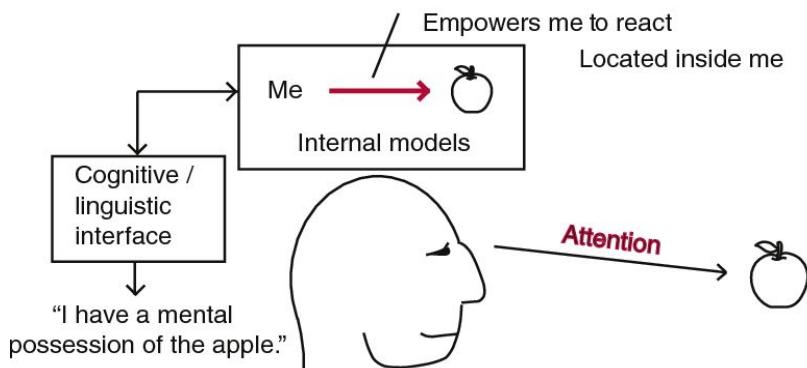
2017

The Attention Schema Theory: A Foundation for Engineering Artificial Consciousness

Michael S. A. Graziano*

Department of Psychology and Neuroscience, Princeton University, Princeton, NJ, United States

The Attention Schema Theory and Co-Modeling



The heart of the attention schema theory is that **awareness** serves as a model of attention

Integrated information theory (IIT)

EXISTENCE: Consciousness exists – it is an undeniable aspect of reality. Paraphrasing Descartes, “I experience therefore I am”.

COMPOSITION: Consciousness is compositional (structured): each experience consists of multiple aspects in various combinations. Within the same experience, one can see, for example, left and right, red and blue, a triangle and a square, a red triangle on the left, a blue square on the right, and so on.

INFORMATION: Consciousness is informative: each experience differs in its particular way from other possible experiences.

INTEGRATION: Consciousness is integrated: each experience is (strongly) irreducible to non-interdependent components. (seeing a red triangle is irreducible to seeing a triangle but no red color, plus a red patch but no triangle)

EXCLUSION: Consciousness is exclusive: each experience excludes all others – at any given time there is only one experience having its full content, rather than a superposition of multiple partial experiences; (Gestalt based)

Ключевая концепция

Коэволюционирующий гибридный интеллект (ГИ) - симбиоз искусственного и естественного интеллектов, развивающихся взаимно, обучая и дополняя друг друга в процессе коэволюции (образуют общий язык, общую онтологию, общие этические рамки).

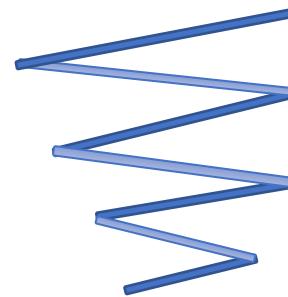
Особенности (преимущества):

- Обеспечивает **непрерывный сбор данных** о физиологических параметрах человека, его приемах работы, поведении
- Строится **индивидуализированная модель** человека, уточняется границы работоспособности
- Реализуется биологическую обратную связь, позволяя человеку управлять своим состоянием и выполнять тренировку
- Обеспечивается **когнитивная интероперабельность** (совместимость) между системами ИИ и человеком (оценка ситуации, состояния, выработка и принятие решений)

Коэволюция интеллекта (интеллектуальных агентов)

- коэволюция - основной процесс, обеспечивающий движение к сильному интеллекту
- скорость движения к сильному интеллекту зависит от когнитивной интероперабельности, то есть:
 - степени формализации когнитивных функций;
 - технологичности переноса опыта/знаний от человека машине и наоборот;
 - легкость интеграции продуктов на базе ИИ в систему разделения труда

Искусственный интеллект

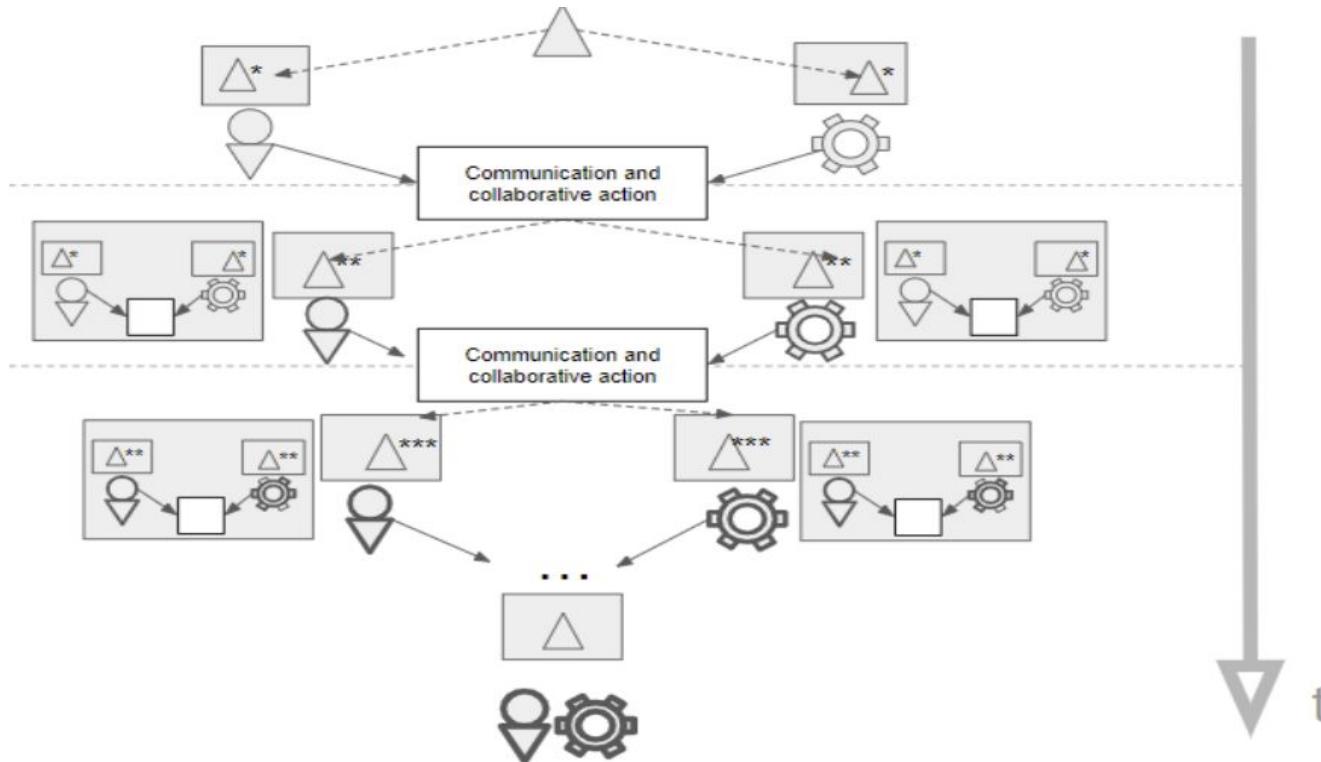


Естественный интеллект

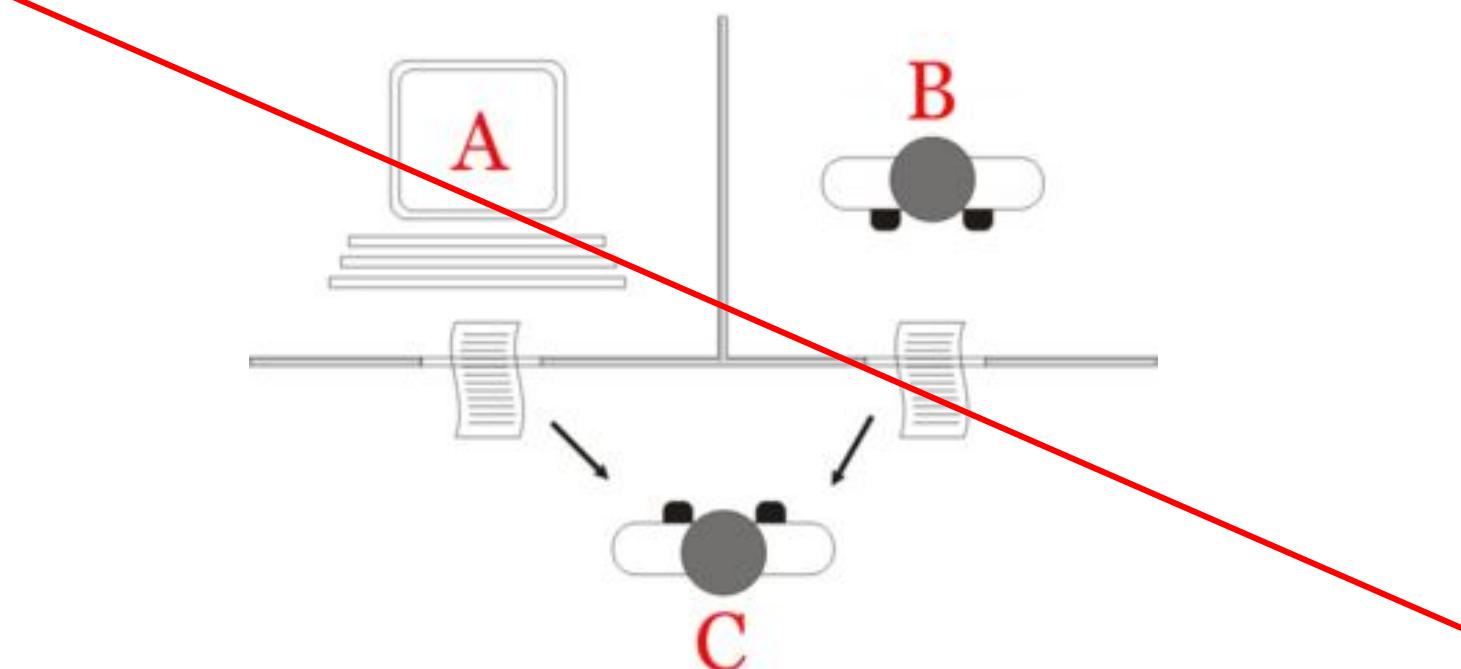
Коэволюционирующий гибридный интеллект (ГИ) -- симбиоз искусственного и естественного интеллектов, развивающихся взаимно, обучая и дополняя друг друга в процессе коэволюции.



Co-modeling and common ontology

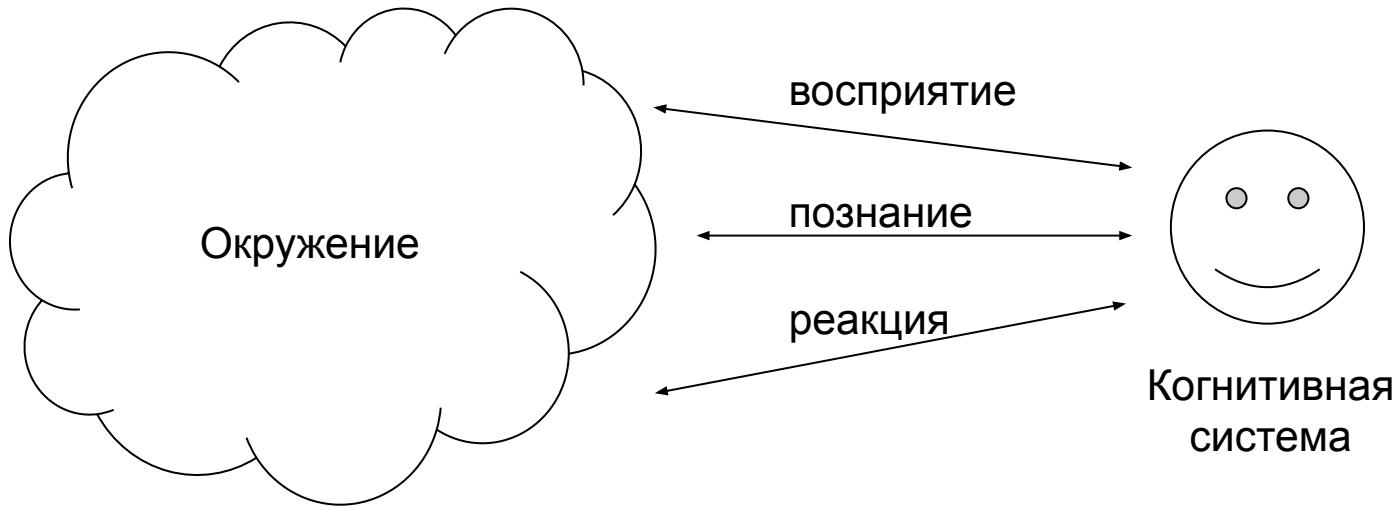


Turing test reconsidering



Когнитивная архитектура

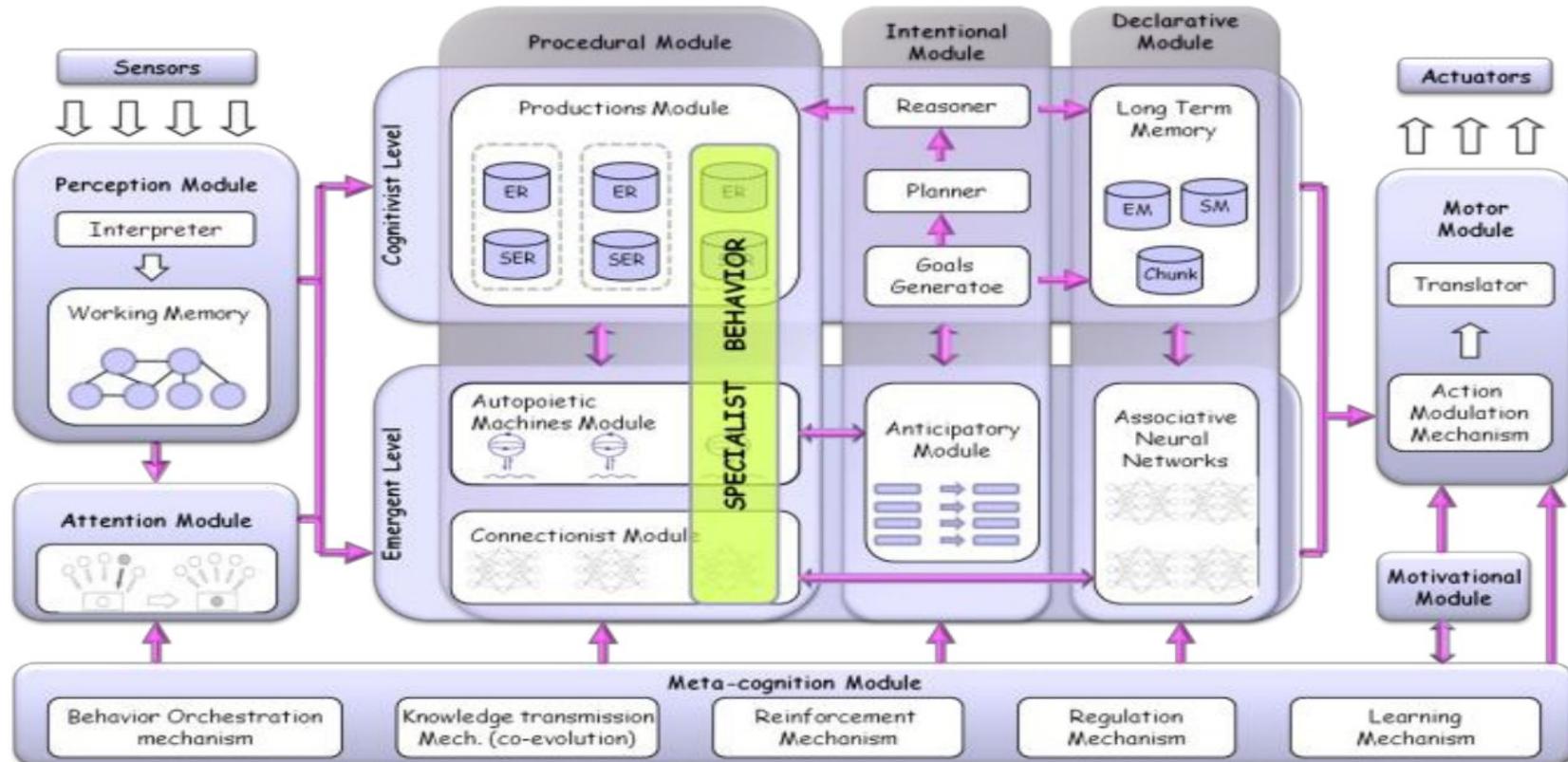
Когнитивная архитектура (или архитектура когнитивной системы) – это сочетание элементов, обеспечивающих восприятие, познание и реакцию



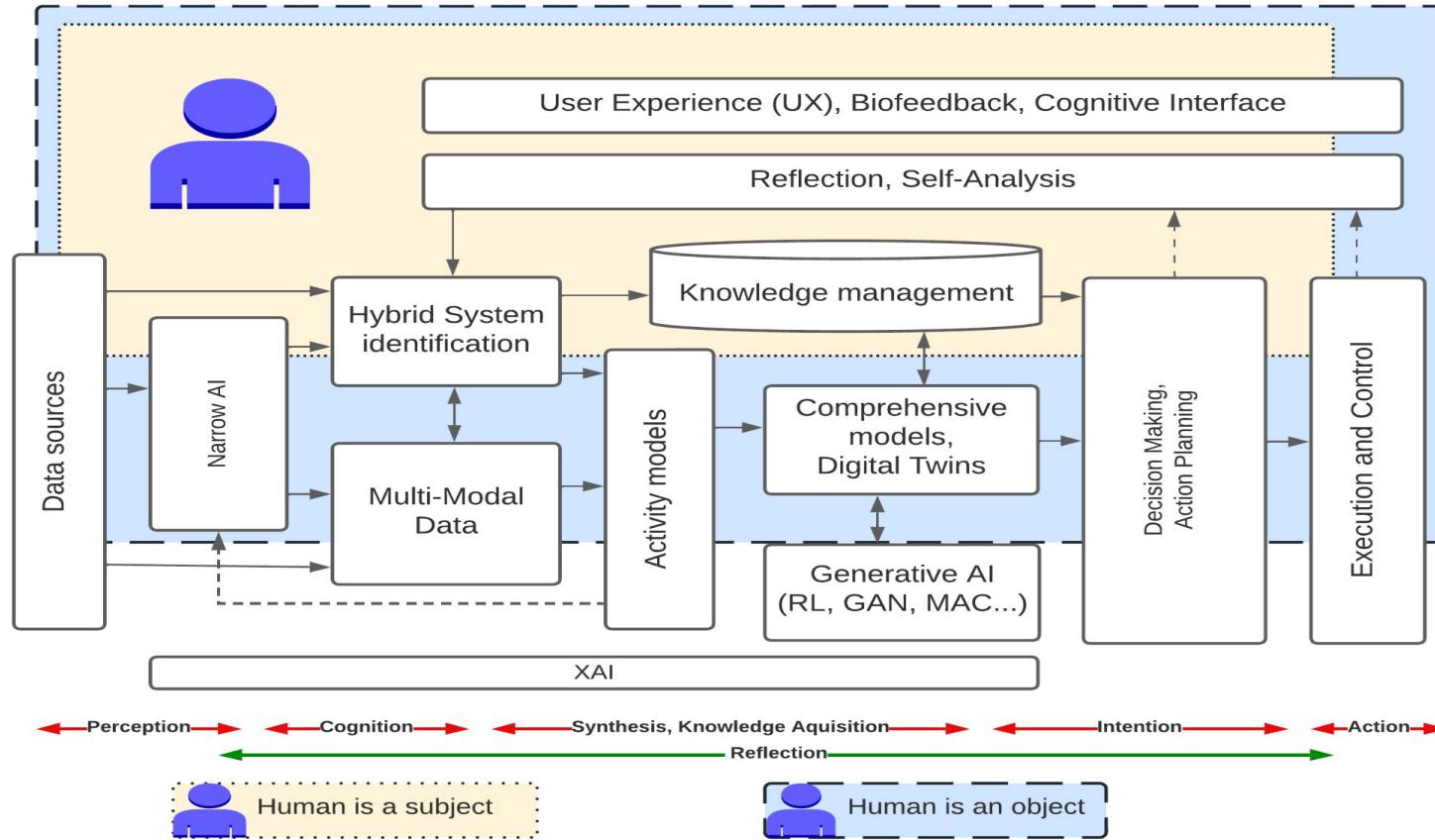
Проблемы существующих искусственных когнитивных систем:

- на этапе восприятия: зависят от отчуждённых данных от человека (а эти данные либо невозможно, либо крайне трудоёмко получить)
- на этапе познания: не могут пользоваться опытом из решения “соседних” задач; не могут моделировать сложные объекты (например, человека)

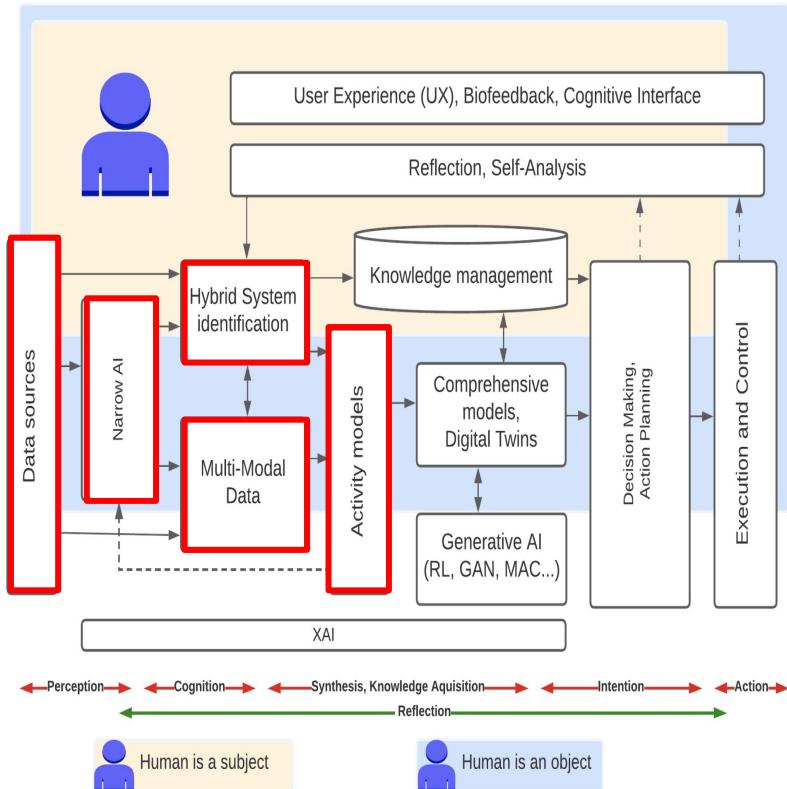
Perception-Cognition-Decision-Action Pipeline



COGNITIVE ARCHITECTURE FOR CO-EVOLUTIONARY HYBRID INTELLIGENCE

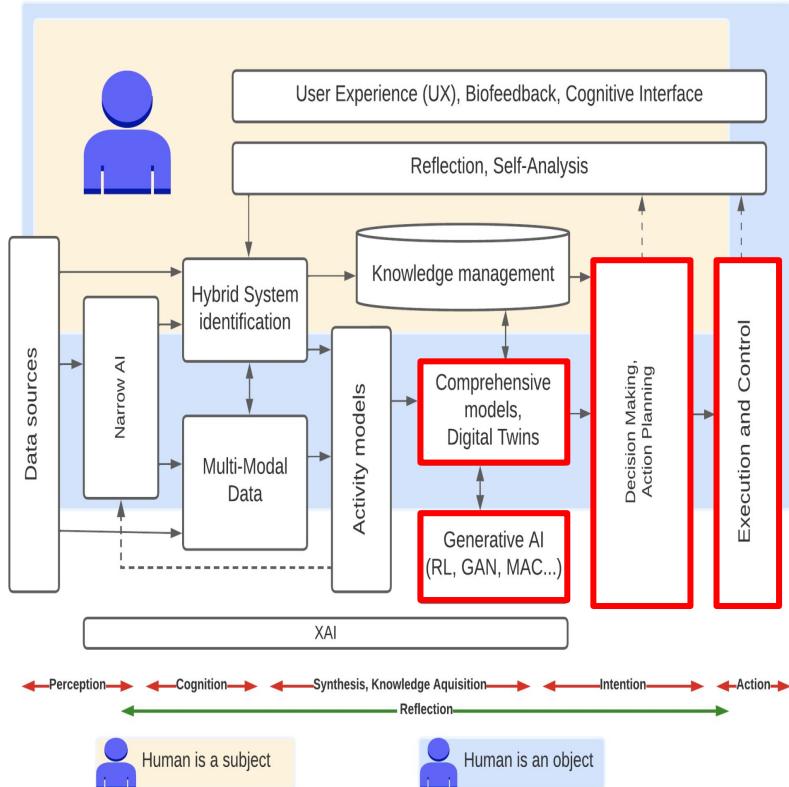


Cognitive Architecture: perception and cognition



- **Data Sources.** Various primary data sources (sensors) which receive information about the control object and about the parameters of a human
- **Narrow AI.** Data processing techniques including signal cleaning, primary pattern recognition, classification, and approximation.
- **Multi-Modal Data.** A generalized data model, containing signals from independent sources, reduced to a single time and space
- **Hybrid System Identification.** One of the key building blocks defining a model of a hybrid system as such. In order to optimize an intelligent system, you must have a model.
- **Activity Models.** Recognition the patterns of human actions including the decision-making cycle

Cognitive Architecture: knowledge extraction, intention, action



- **Digital Twins.** Full information about the hybrid system's operation during its lifetime and a set of methods for automatically identifying trends and predicting states.
- **Generative AI.** Algorithms and methods for generating decisions and directional search for options on a set of hybrid system parameters
- **Decision Making and Action Planning.** Scenario planning for management in the short-term and long-term planning horizons.
- **Execution and Control.** Execution of action scenarios

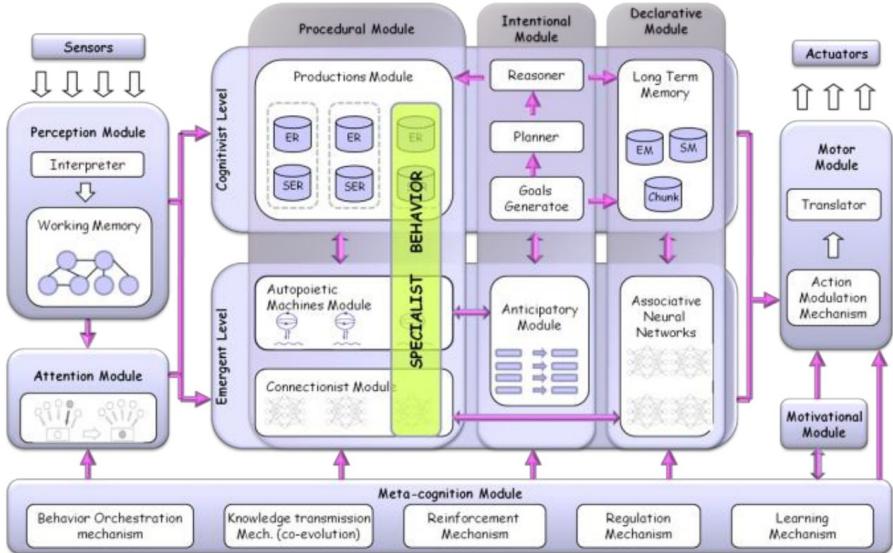
Applications examples

- A new view at automated system
- Data (knowledge management)
 - fMRI
 - eeg
- Human state control
 - functional control
- Personalized medicine
- Personal Environment automation (definition)
- Behavioral analysis and education

Кризис применимости нейросетевых решений ИИ

- Отсутствие человека в контуре решения задач – ключевая проблема применения современных решений на базе ИИ
- Недостатки искусственных нейронных сетей:
 - Сильная зависимость от данных: количества, качества, несмещенности
 - Неуниверсальность: при изменениях состава(структуры) входных данных требуется полный цикл переобучения
 - Отсутствие возможности по интерпретации результатов работы
 - Высокие требования к вычислительным ресурсам
 - Низкая энергоэффективность
 - Отсутствие теоретических основ и (практических примеров) для построения композитных решений для нейровычислений
 - Неспособность решать простейшие когнитивные задачи (обобщение, перенос опыта)

Предпосылки для разработки



O. J. Romero Lopez and A. de Antonio Jimenez, "Hybridization of cognitive models using evolutionary strategies," 2009 IEEE Congress on Evolutionary Computation, 2009, pp. 3213-3218, doi: 10.1109/CEC.2009.4983351.

AFOSR-3223

Summary Report

AUGMENTING HUMAN INTELLECT: A CONCEPTUAL FRAMEWORK

Prepared for:

DIRECTOR OF INFORMATION SCIENCES
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
WASHINGTON 25, D.C.

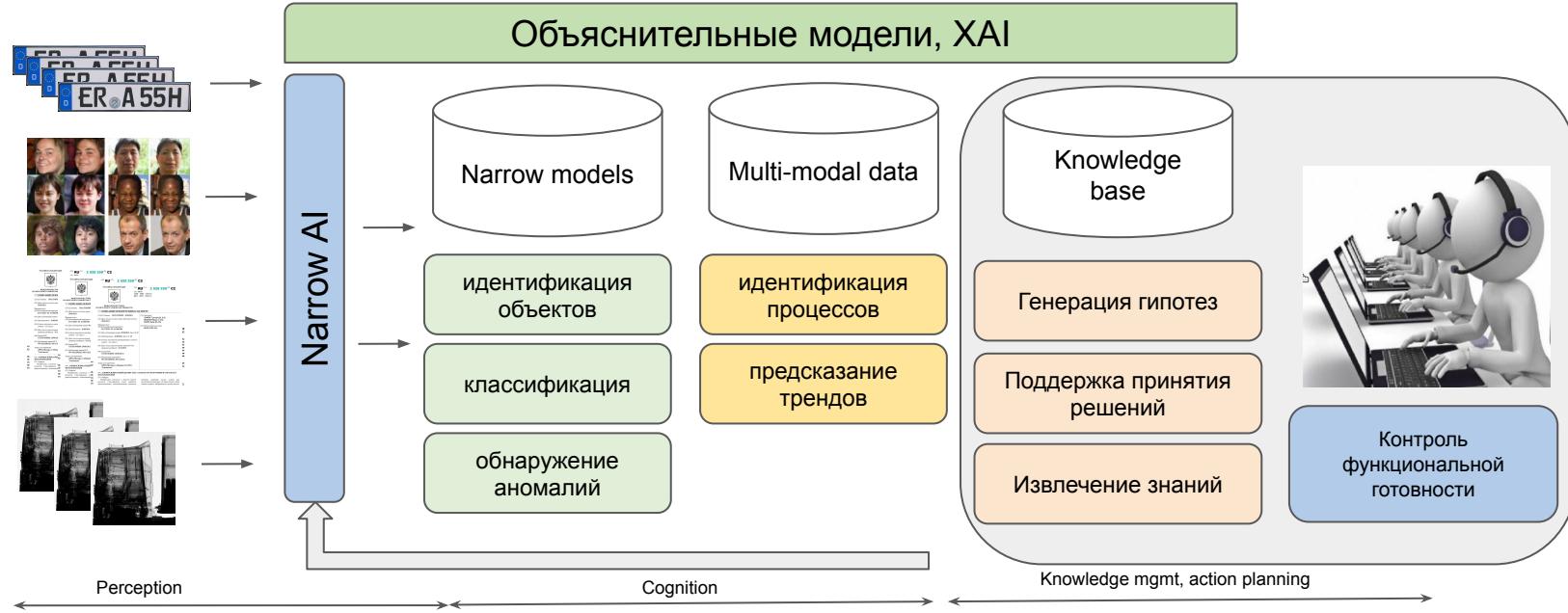
CONTRACT AF 49(638)-1024

By: D. C. Engelbart

[ГОСТ 34.003-90] автоматизированная система; АС: Система, состоящая из персонала и комплекса средств автоматизации его деятельности, реализующая информационную технологию выполнения установленных функций

Пример реализации архитектуры

Цифровой двойник,
гибридная система



Сложный объект
управления

Комплексная система оценки функциональной готовности организма человека к выполнению трудовых (и иных) функций

Области применения

- Операторы сложных технических систем
- Службы МЧС
- Операторы беспилотного транспорта



Функциональные требования

- **Неинвазивный** сбор первичных данных
- Оценка психологического и физиологического состояния организма
- Построение **индивидуального** аватара (цифрового двойника) человека
- Генерация **индивидуальных** режимов (календарей) нагрузки и отдыха
- Управление (совместно с пользователем) целями по состоянию, трекинг и учет параметров физической активности, сна, питания, интеграция с системами управления здоровьем
- Возможность подключения дополнительных источников информации

ПРОТОТИП СИСТЕМЫ КОНТРОЛЯ БДИТЕЛЬНОСТИ ЧЕЛОВЕКА НА ОСНОВЕ ТЕХНОЛОГИИ НЕИНВАЗИВНОЙ РЕГИСТРАЦИИ ПУЛЬСОВОЙ ВОЛНЫ

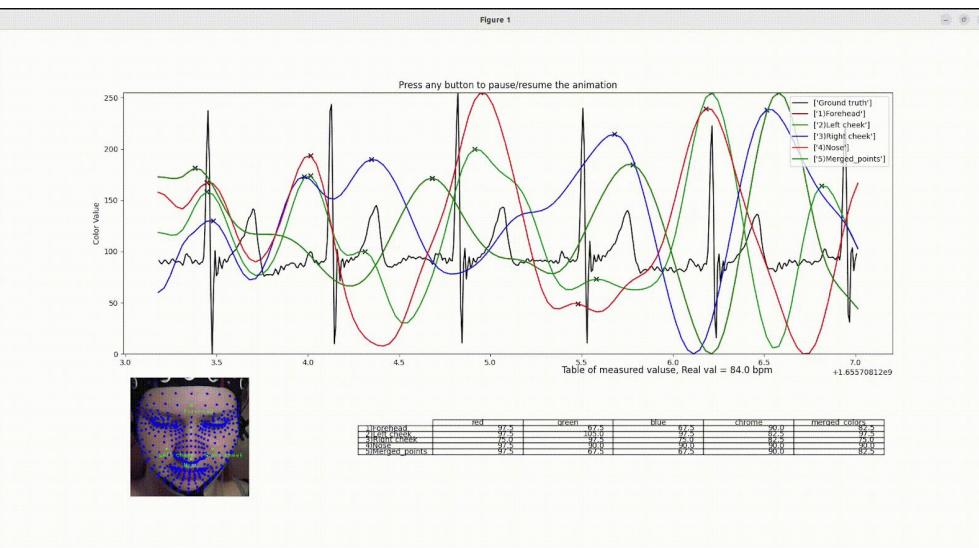


Функциональные возможности

- Фиксация изменения цвета кожи человека с помощью RGB видео камеры
- Построение пульсовой волны человека
- Определение индивидуальных особенностей пульсовой волны человека
- Определение числа классов пульсовой волны, характерных определенному виду деятельности человека.
- Определение соответствия класса типовой волны и вида деятельности
- Определение нормативных и критических поведений пульсовой волны в каждом классе.
- Контроль бдительности человека

Когнитивные возможности

- Обмен информацией и знаниями с пользователем
- Объяснимость решений
- Выработка решений на основе опыта



ТЕХНОЛОГИЯ АНАЛИЗА ЭПИЛЕПТИМОРФНОЙ АКТИВНОСТИ

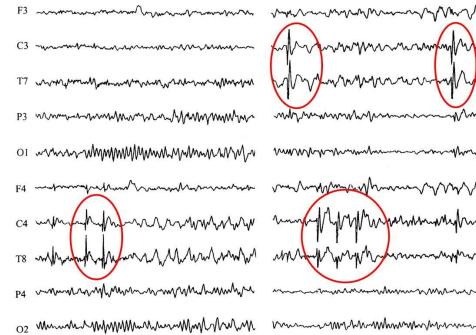
Основные факты

- Эпилепсия – хроническое неинфекционное заболевание головного мозга, поражающее людей в любом возрасте.
- Эпилепсия является одним из самых распространенных неврологических заболеваний в мире, которым страдает порядка 50 млн человек.
- Почти 80% страдающих эпилепсией проживают в странах с низким и средним уровнем дохода.
- По оценкам, до 70% людей с эпилепсией могут жить без приступов болезни при условии обеспечения надлежащей диагностики и лечения.
- Риск преждевременной смерти у больных эпилепсией почти в три раза превышает средний показатель по популяции.

Функциональные возможности

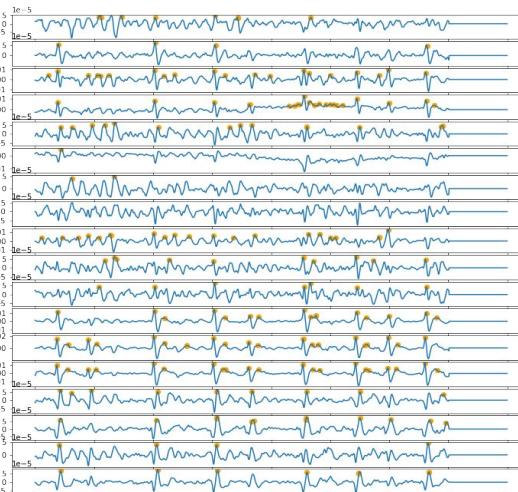
- визуализация ЭЭГ-сигнала;
- выявление пиков сигнала;
- кластеризации пиков на спайки и не-спайки;
- классификации различных паттернов спайков;
- определение по классу спайков формы эпилепсии;
- предсказание эпилептических приступов;

Пример эпилептиформной активности на ЭЭГ



Спайк (англ. spike) — это пароксизмальный (т. е. внезапно появляющийся) электрический разряд на ЭЭГ, четко отличающийся от фоновой активности. Может свидетельствовать о нарушениях работы мозга.

Автоматическая предразметка



Contacts

Kirill Krinkin, kirill@krinkin.com

Yulia Shichkina, shichkina@co-evolution.ai

Etu.ai

co-evolution.ai