

Сознание в нейронных гиперсетях
Consciousness in neural hypernetworks
神經超網路中的意識

Анохин К.В.

академик РАН, Институт перспективных исследований мозга

МГУ им. М.В. Ломоносова, Москва, Россия

*Academician of the Russian Academy of Sciences, Institute for Advanced Brain Research,
 Moscow State University named after M.V. Lomonosov, Moscow, Russia*

俄羅斯科學院高級腦研究所院士

莫斯科國立大學以它的名字命名 MV 羅蒙諾索夫, 莫斯科, 俄羅斯

Аннотация. Речь в работе идет о нейронных гиперсетях и процессах сознания, с которыми мы работаем. Сознание - это специфический процесс в специфически организованной структуре. Это структура нейронной гиперсети, описываемой формальным языком, специфический процесс - это процесс глобальной перколяции активности в указанной сети. Статья начинается с определения сознания, далее показано, что разобраться в этой теме можно только с помощью теории. На сегодняшний день имеется большое количество теорий сознания, мы выделяем в них физико-математическое направление, разбираем некоторые из этих теорий. Далее речь идет о той теории, которую мы называем теорией нейронных гиперсетей, о некоторых принципах устройства этой структуры, так, как она представляется нейрофизиологам. Сознание представляет собой особый вид трафика в этой специфической структуре, отличающейся от известных представлений в теориях нейронных сетей.

Ключевые слова: сознание, когнитом, функциональные системы, нейрофизиология, трафик, искусственный интеллект

Annotation. The work is about neural hypernetworks with which we work. Consciousness is a specific process in a specifically organized structure. This is the structure of a neural hypernetwork, described in formal language, a specific process - the process of global percolation of activity in this network. The article begins with a definition of consciousness, then it is shown that this topic can only be understood with the help of theory. Today there are a large number of theories of consciousness, we highlight the physical and mathematical direction in them, and analyze some of these theories. Next we talk about the theory that we call the theory of neural hypernetworks, about some principles of the structure of this structure, as it appears to neurophysiologists. Consciousness represents a special kind of traffic in this specific structure, which differs from the known concepts in neural network theories.

Keywords: consciousness, cognitome, functional systems, neurophysiology, traffic, artificial intelligence

註解。 我們正在談論我們所使用的神經超網路。意識是特定組織結構中的特定過程。這是一個神經超網路的結構，用形式語言描述，一個特定的過程—這個網路中活動的全域滲透過程。文章首先對意識進行了定義，然後表明這個主題只能藉助理論來理解。今天有

大量的意識理論，我們強調其中的物理和數學方向，並分析其中的一些理論。接下來我們討論我們稱為神經超網絡理論的理論，討論這種結構的一些原理，正如神經生理學家所看到的那樣。意識代表了這種特定結構中的一種特殊的流量，它與神經網路理論中的已知概念不同。

關鍵字: 意識, 認知組, 功能係統, 神經生理學, 交通, 人工智慧

Уважаемые коллеги! Большое спасибо за приглашение выступить на этом замечательном Конгрессе, который собирает в себе специалистов как в области системной биологии и теории систем, так и математики, искусственного интеллекта. Моя область исследований лежит, как раз, на перекрестке этих дисциплин. Это нейрофизиология высших функций мозга. Так получается, что в последние годы она связана с развитием новых математических подходов. И мне хотелось бы сегодня рассказать об этой области и теории, с которой мы работаем и пригласить и попросить сотрудничества со стороны специалистов в области математики, потому что это то, что происходит сейчас в мире и то, что требуется для прогресса нашего понимания того, как мозг рождает процесс сознания. Я начну сразу с резюме, в котором суммированы основные моменты доклада. Речь будет сегодня о нейронных гиперсетях и процессах сознания, согласно теории нейронных гиперсетей, с которыми мы работаем. Сознание - это специфический процесс в специфически организованной структуре. Это структура нейронной гиперсети. По крайней мере, так она может быть описана формальным языком. А специфический процесс — это процесс глобальной перколяции активности в этой сети. И мы как нейрофизиологи заинтересованы в совместном исследовании этой структуры, органической и вместе с тем математической структуры, и этого процесса, с одной стороны физиологического, когнитивного, с другой стороны, информационного с помощью физико-математических наук. Я начну с определения сознания и скажу, почему это сегодня такая важная тема. Дальше, я перейду к тому, что разобраться в этой теме можно только с помощью теории. И что в последние годы роль математики и теоретической физики в теоретических исследованиях высших функций мозга, в том числе сознания, становится все больше и больше. Дальше, я сделаю некоторый обзор проблемной ситуации. На сегодняшний день имеется большое количество теорий сознания, я выделю в них физико-математическое направление, разобрав некоторые из этих теорий. И, наконец, в третьей части я расскажу о той теории, с которой мы работаем и которую мы называем теорией нейронных гиперсетей или гиперсетевой теорией мозга, расскажу о некоторых принципах устройства этой структуры, так, как она представляется нам нейрофизиологам. И, коротко, о том, что сознание представляет собой особый вид трафика в этой специфической структуре, отличающейся от известных нам представлений в теориях нейронных сетей.

Итак, определение сознания. Сознание — это термин, который в истории

имеет множество смыслов. Тем не менее, в последнее время в научном сообществе намечается некоторый консенсус в определении фундаментальной характеристики сознания, определяемого, как любого субъективного опыта, генерируемого или изнутри системы, или в ответ на внешнее обстоятельство, которое включает в себя как высоко порядковые мысли, планы, идеи, так и эмоциональные чувства, ощущения. Короче говоря, все, что ощущается субъектом. Наш выдающийся философ сознания Давид Израилевич Дубровский в свое время определил это охватывающим термином субъективной реальности. Это определение существенно отличается от определений и понимания сознания, которые давались ранее. Например, в большом энциклопедическом словаре сознание определялось еще 25 лет назад, как высшая форма психического отражения, которая свойственна общественно развитому человеку, связанная с речью, идеальная сторона целенаправленной деятельности. Тогда, согласно этим представлениям, сознанием может обладать только человек. Сегодня ситуация и фокус исследования сознания существенно сместились. Но прежде, чем сказать о нём, нужно сформулировать саму суть проблемы сознания. В принципе, ее очень остро в одной фразе сформулировал Иван Петрович Павлов около ста лет назад. Как материя мозга рождает субъективные явления? Давид Израилевич Дубровский дал более развернутое определение этой проблемы, сформулировав ее в двух вопросах. Как объяснить связь явлений субъективной реальности с мозговыми процессами, если первым нельзя приписывать физические свойства, а вторые ими необходимо обладают? И второй вопрос, если явлениям субъективной реальности нельзя приписывать физические свойства, то как объяснить их способность причинного действия на телесные процессы. Нужно сказать, что несмотря на то, что этой проблемой занимается огромное количество философов, написаны тысячи книг, в отношении и первого, и второго вопроса это открытая научная проблема. Значительную роль в изучении этих вопросов сегодня играет нейронаука. Теперь скажем о важности сознания, предложенного в такой фундаментальной формулировке. Если речь идет об ощущениях, субъективном опыте, любых переживаниях, в том числе эмоциональных реакциях, то мы обязаны начать рассматривать сознание как широкий природный феномен, присутствующий не только у человека, но и у других животных. И сегодня также с развитием системы искусственного интеллекта все острее встает вопрос о сознании у системы искусственного интеллекта. Ряд передовых групп в различных центрах и компаниях занимаются моделированием сегодня не просто искусственного интеллекта, а искусственного сознания. Проиллюстрировать это можно недавно состоявшимися двумя конференциями. Одна из них прошла в начале мая 2023 года и называлась Конференцией по сознанию животных (Animal Consciousness Conference 2023, Dharamshala, May 1-5). На ней собрались специалисты из дюжины стран, ведущие эксперты в области исследования сознания в

биологическом мире и консенсус, который достигается в такого рода конференциях, состоит в том, что большое количество животных, о которых мы раньше не думали, как о существах, обладающих сознанием, обладают этим феноменальным свойством, которому я выше дал определение. Это не только приматы, это не только дельфины и высшие млекопитающие с сильно развитым мозгом слоны, киты, касатки и так далее. Это птицы, это ряд видов головоногих моллюсков, такие как осьминоги, каракатицы. Я пропускаю многих других млекопитающих. Это членистоногие, пчелы, шмели, другие виды насекомых. Сознание широко распространено в биологическом мире.

Вторая конференция, которая, традиционно проходит с 1994 года, каждый год, крупная конференция, которая называется Наука о сознании, проходила в Италии в Taormina (The Science of Consciousness 2023, Taormina, May, 22-27). На ней было заслушено большое количество докладов. Вот некоторые их иллюстрации. На конференции по сознанию животных, например, была удивительная работа Мосанори Коды, опубликованная некоторое время назад (Masanori Kohda «Cleaner fish identify themselves via self-face recognition like humans: What does this mean?», Osaka Metropolitan University, Osaka, Japan), показывающая, что рыбы чистильщики способны идентифицировать самих себя и узнавать самих себя в зеркале. Таким же образом, как это делается людьми и некоторыми высшими приматами. Это удивительные данные, которые заставляют нас быть внимательными, совершенно иначе смотреть на распространение феномена сознания в естественном мире. С другой стороны, ключевая лекция на конференции в Taormina была прочитана Дэвидом Чальмесом ее название: «Способны ли большие языковые модели быть сознательными?» (David Chalmers «Could a Large Language Model be Conscious?», NYU, New York, USA). На ней он рассмотрел аргументы за и против. Нет консенсуса в этом вопросе. Есть разные точки зрения, но этот вопрос стоит. Таким образом, мы видим, насколько актуально сегодня понимание фундаментальных принципов организации сознания. И чтобы понять их недостаточно чисто экспериментальных исследований, количество фактов накопленных в нейронаучных исследованиях огромно. А они так и не привели пока к священному граалю, который был провозглашён около 30 лет назад, нобелевским лауреатом Френсисом Криком и его коллегой Кристофом Кохом - найти нейронные или нервные корреляты сознания. До сих пор нет надежного нервного коррелята сознания. И отчасти это понятно, потому что только теория позволит решить, что наблюдается в эксперименте и предсказать, какого рода корреляты следует искать. Нельзя отказываться от построения теории в начале пути. Это окажется пагубным в оценке результатов экспериментов. «Когда мы говорим, что мы понимаем совокупность явлений природы, писал Эйнштейн, мы имеем в виду, что мы нашли конструктивную теорию, которая охватывает их». И, действительно, с постепенным

разочарованием в чисто эмпирических экспериментальных исследованиях поисках тех или нервных коррелятов сознания в последние десятилетия наблюдается буквально взрыв числа теории сознаний.

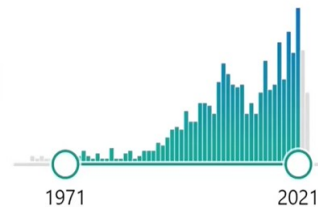


Рис.1. Рост числа публикаций по теориям сознания
 Fig. 1. Growth in the number of publications on theories of consciousness
 圖 1. 意識理論出版數量的成長

Здесь показан рост в количестве публикаций, включающих в себя ключевые слова сознания и теории, с 1971 по 2021 год, но если посмотреть просто на работы 21-22 года, обзоревающие эту область, то увидим, что появились синтетические метаанализы огромного количества теорий. Этих теорий точно больше 30 сегодня, может быть около 50. И возникают вопросы, каким образом сравнивать эти теории. Работ на эту тему достаточно много.

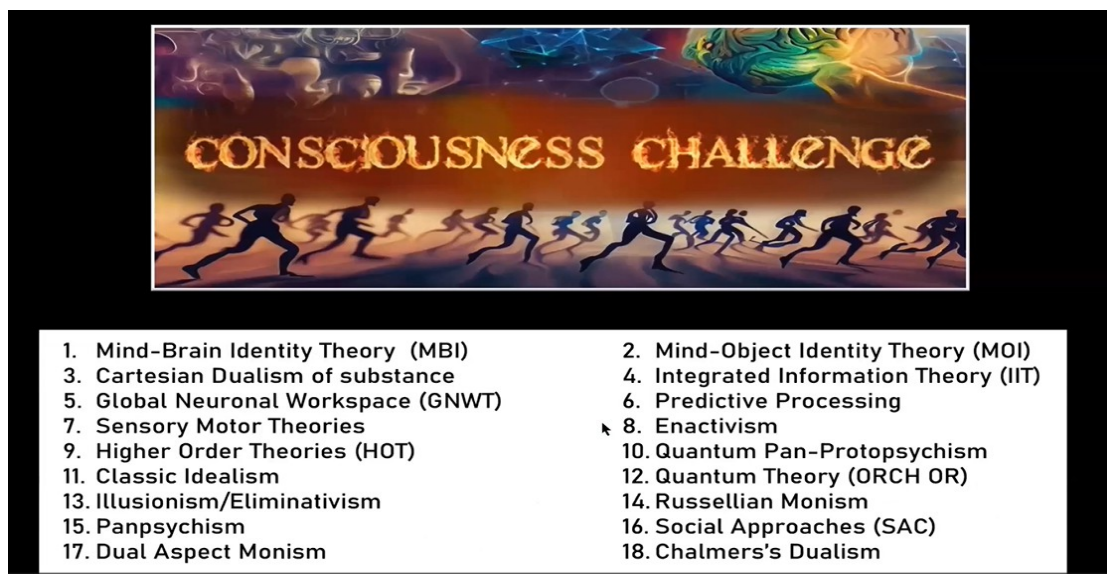


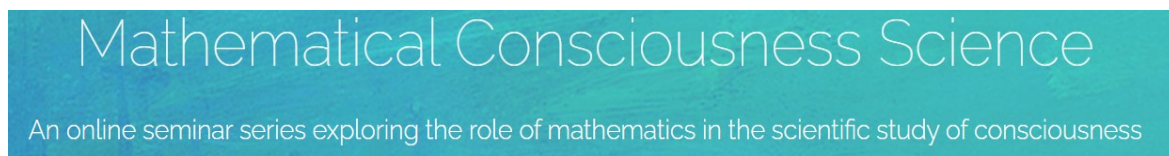
Рис.2. Ведущие теории сознания по мнению участников конференции в Taormina
 (The Science of Consciousness 2023, Taormina, May, 22-27)
 Fig. 2. Leading theories of consciousness according to conference participants in Taormina
 (The Science of Consciousness 2023, Taormina, May, 22-27)
 圖 2. 陶爾米納會議參與者認為領先的意識理論
 (《意識科學 2023》，陶爾米納，5 月 22-27 日)

На упомянутой конференции в Италии, в Taormina в мае 2023 года, был даже организован специальный конкурс - голосование всех участников конференции, которых было более 500 ученых, с выбором предпочитаемой ими теории сознания. На рисунке 2 выше перечислено только 18 теорий, есть некоторый произвол организаторов, в том, какие теории они выбрали, но суть состоит в этом. Одна из участниц докладчиков этой конференции вместе со своими коллегами год назад предложили специальную базу данных, названную Контраст, для анализа и сравнения эмпирических исследований тех или иных теорий сознания.

Ясно, что область расцветает и, может быть, через некоторое время здесь будет даже кризис перепроизводства.

Еще одна важная линия, которая намечается в последние годы, - это развитие физико-математического направления в изучении сознания. Становится понятным, что сознание, как состояние и процессы в сложных системах, требует анализа самых разных математических позиций. Это математический объект и информационный процесс в зависимости от взглядов тех или иных авторов.

Приведем названия ряда выступлений известных специалистов на семинаре «Математическая наука сознания», Mathematical Consciousness Science, который регулярно идет в сети (<https://seminar.math-consciousness.org>).



Carlo Rovelli “Understanding consciousness within the known laws of physics three small contributions”,

Karl Friston “Markov blankets and Bayesian mechanics”,

Stephen Wolfram “How Features of Our Consciousness Seem to Define Our Laws of Physics and Mathematics”,

Nao Tsuchiya “A search for an isomorphism between conscious experience and structure of information”.

Другой пример из Математического института Оксфорда, который проводит семинары по математическим моделям и основам сознания (рис.3).

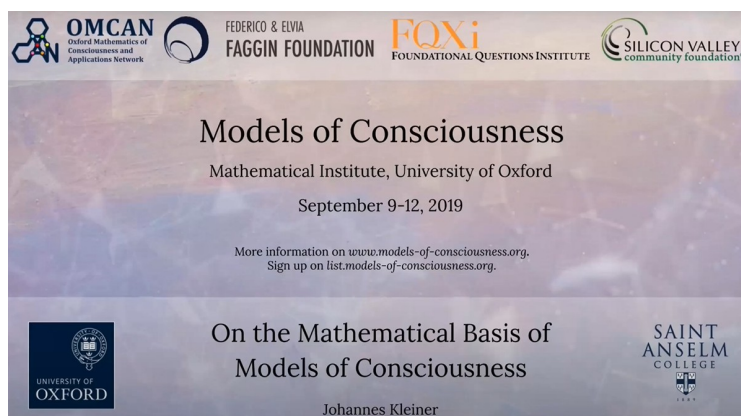


Рис.3. Семинары Математического института Оксфорда по моделированию сознания.

Fig. 3. Oxford Institute of Mathematics workshops on modeling consciousness.

圖 3. 牛津數學研究所關於建模意識的研討會。

Я мог бы приводить здесь также большое количество работ, но за неимением места ограничусь только двумя приведенными примерами. Среди теорий сознания есть, безусловно, фавориты.

В прошлом году был сделан обзор основных таких теорий (Seth & Bayne (2022) Nature Reviews Neuroscience). На схеме из обзора (рис. 4) указаны четыре наиболее крупные популярные теории сознания. Высокоуровневая теория сознания, теория глобального рабочего пространства, теории интегрированной информации и теории повторного входа.

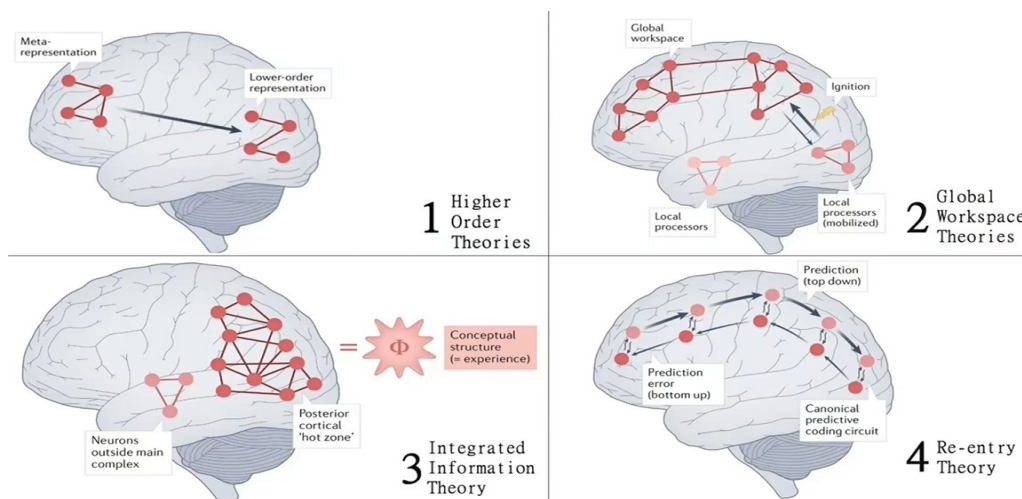


Рис.4. Популярные теории сознания.

Fig. 4. Popular theories of consciousness.

圖 4. 流行的意識理論。

Я не имею возможности рассказывать про каждую из них. Отмечу только, что по многим дискуссиям и обсуждениям, доминирующую позицию сегодня занимают теория глобального рабочего пространства, которая была предложена специалистом в области вычислительной психологии, когнитивной нейронауки Жан-Пьером Шанжё (Jean-Pierre Changeux) и теории интегрированной информации, предложенной итальянским ученым Джулио Тонони (Giulio Tononi), работающим в Америке. Сегодня ситуация продвинулась, таким образом, что был создан крупный проект очень известных ученых (см. рисунок 5), написавших протокол по соревновательному сопоставлению предсказаний двух этих теорий, теории глобального рабочего пространства и теории интегрированной информации

STUDY PROTOCOL

An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory

Lucia Melloni^{1,2†*}, Liad Mudrik^{3,4†}, Michael Pitts^{5†}, Katarina Bendtz⁶, Oscar Ferrante⁷, Urszula Gorska⁸, Rony Hirschhorn⁹, Aya Khalaf^{9,10}, Csaba Kozma⁸, Alex Lepauvre¹, Ling Liu^{11,12}, David Mazumder⁶, David Richter¹³, Hao Zhou^{14,15}, Hal Blumenfeld⁸, Melanie Boly¹⁶, David J. Chalmers¹⁷, Sasha Devore², Francis Fallon¹⁸, Floris P. de Lange¹³, Ole Jensen⁷, Gabriel Kreiman^{6,19}, Huan Luo^{11,12}, Theofanis I. Panagiotaropoulos¹⁴, Stanislas Dehaene^{14,20†}, Christof Koch^{21†}, Giulio Tononi^{8†}

PLOS ONE | <https://doi.org/10.1371/journal.pone.0268577> February 10, 2023

Рис.5. Протокол по сопоставлению теорий

Fig. 5. Protocol for comparing theories

圖 5. 理論比較協議

Сторонники этих теорий собрались, выдвинули предсказание в отношении нервных процессов, которые должны наблюдаться согласно одной и согласно другой теории, которые отличаются, договорились о протоколах. Сейчас началось широкомасштабное исследование в ряде лабораторий мира различающихся предсказаний между этими двумя лидирующими теориями.

Я должен несколько слов сказать о теории интегрированной информации просто потому, что эта теория была сформулирована как математическая теория. В разработке этой теории работ участвует большое количество математиков и специалистов в области теоретической физики. Например, Лариса Албантакис (Larissa Albantakis) специалист в области квантовой физики пришла работать в нейронауку, заниматься теорией сознания. Ну, и так далее. Несколько месяцев назад вышла четвертая версия этой теории, ИТ 4.0, которая отличается от первой, второй и третьей, это работа, которая активно развивается, и я просто отсылаю в данном случае к этой публикации [1], она

значимая. Мы тоже работаем в этом направлении и это движение в значительной степени продиктовано традициями российской физиологической школы, ставившей вопрос о физиологических, материальных основах высших функций мозга с середины 19 века, в начале Иваном Михайловичем Сеченовым, у которого были блестящие теоретические работы, в частности, «Элементы мысли», но не было экспериментальных подходов.



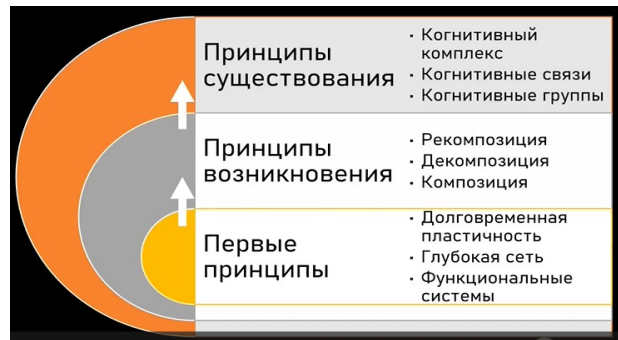
Однако, он многое сделал для теоретических моделей связей в нервной системе, так называемых, естественных групп, чисто постулируемых тогда в 19 веке, групп нейронов, связанных совместной активностью для отражения тех или иных внешних сигналов. Иван Петрович Павлов привнес в эту идею и связи между такими группами условных связей, и экспериментальный метод. Петр Кузьмич Анохин, ученик Павлова, развил Павловское представление об условных рефlekсах в системную модель работы мозга, теорию функциональных систем.

Мы пытаемся двигаться дальше по этому пути, представив себе, каким образом многие функциональные системы, выполняющие те или иные виды целенаправленной деятельности мозга, образуют совместно некоторый комплекс, составляющий сущность когнитивного агента, иначе говоря его когнитом. Мы начали работать с этой теорией около 10 лет назад, в [2-5] указаны некоторые вехи в докладах по этой теории на разных конференциях. Сейчас готовится синтетическая публикация по основам этой теории. Я коротко скажу о некоторых аспектах, которые могут иметь интерес для физико-математических подходов. Надо сразу сказать, что эта теория отличается от большинства других теорий. Во-первых, потому, что она начинается не с вопроса сознания, а с вопроса того, что такое разум, и что такое Я, которое воспринимает те или иные состояния активности. Во-вторых, эта теория начинается с первых принципов, она не постулирует существующие структуры мозга, которые обеспечивают этот высший уровень интеграции и психической деятельности, а начинает с биологических принципов, не содержащих, никакие психологические термины. Сочетание указанных принципов является начальными условиями, достаточными, чтобы система организмов дарвиновского типа, имеющая в себе сочетание этих первых принципов, начала

постепенно развивать в себе высоко порядковую структуру, являющуюся носителем разума и сознания.

Теория нейронных гиперсетей

Теория нейронных гиперсетей - алгоритмическая теория происхождения и устройства когнитивных систем, обладающих свойством разума (mind) и сознания (consciousness)



Принципы существования

Принципы существования выводятся в теории из принципов возникновения и включают три категории объектов: когнитивные нейронные группы, связи между этими группами и образуемый ими когнитивный комплекс - когнитом



Рис.6. Принципы теории нейронных гиперсетей.
Fig. 6. Principles of the theory of neural hypernetworks.
圖 6. 神經超網絡理論原理。

Из этих первых принципов теория выводит принципы возникновения этих высоко порядковых структур, а из первых принципов возникновения теория выводит то, что возникает, то есть структуру, процессы в ней и принципы существования.

Я останавливаюсь сегодня только на последней части теории, на принципах существования. Теория вводит три вида новых объектов, которые

обычно не рассматриваются в нейрофизиологическом языке и нейрофизиологическом аппарате.

Она вводит понятие когнитивных нейронных групп или **когов**, связей между этими группами (групповые связи) и образуемый этими двумя видами элементов специальный когнитивный комплекс, который теория обозначает как **когнитом**. Скажем несколько слов о каждом из этих элементов.

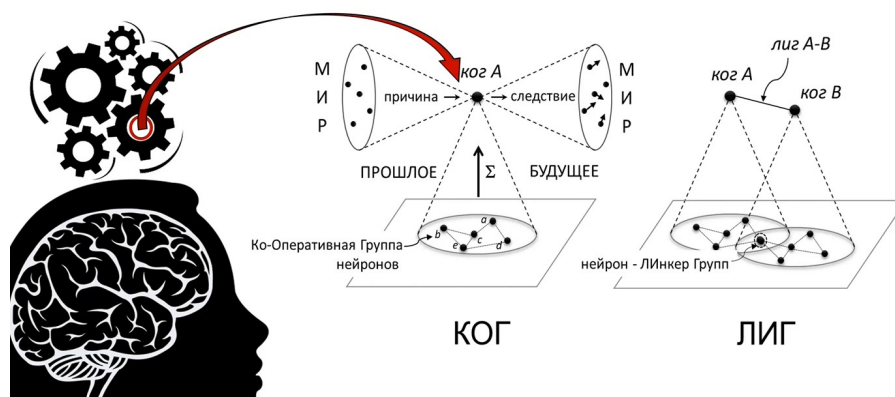


Рис. 7. COG и LIG когнитома
Fig. 7. COG and LIG cognitome
圖 7. COG 和 LIG 認知組

Когнитивная группа - это кооперативная группа нейронов, которая в своей совместной активности определяет какой-то аспект соотношения целого организма с тем или иным аспектом окружающего мира, где появление этих аспектов выступает, как причина для активации этой нейронной группы специализированных элементов, а активация этой нейронной группы способна выступать причиной для последующих изменений в поведении или ментальных состояниях других нейронных групп, определяющих динамику нервных процессов, действия и поведения организма.

Такого рода нейронные группы когнитивного уровня, вершины которых кодируют более высоко порядковые понятия, чем отдельные элементы, составляющие группу, связаны между собой линкерами, которые в теории называется лигами. Термин «ког» в теории имеет двойной смысл, с одной стороны это маленький элемент, подчиненная часть целой когнитивной системы, качественно специфического опыта, которая кодируется этой группой.

С другой стороны, это аббревиатура для физически существующей нейронной группы, активность которой и обуславливают этот элемент специфического опыта с его причинно-следственным потенциалом. Термин «лиг» линкер групп отражает специфическую особенность связей в структуре гиперсети.

В отличие от нейронной сети, где связи являются контактами между отдельными клетками, то есть аксонами, проводами, связывающими клетки в нервной системе, в гиперсети, в нейронной гиперсети, связи являются не ребрами сети, а её элементами, общими элементами между двумя группами, который будучи активированным одной группой, способен передать возбуждение (сигнал) на другую группу.

Таблица 1

Table 1

表格 1

Три фундаментальных понятия ТНГ

◆ Разум (mind) реален - это специфическая органическая и метематическая структура (К-сеть), опосредующая каузальные соотношения когнитивного агента (К-агента) со средой.

◆ Чтобы описать элементы и свойства этой специфической системы требуется введение новых фундаментальных научных понятий.

№	Понятия	Утверждения теории	Следствия
1	КОГ (COG): cognitive group (k-unit)	Разум обладает зернистой структурой, он состоит из когнитивных элементов (к-элементов). к-элементы опосредуют каузальные (=информационные) соотношения целостного когнитивного агента с его средой.	Разум <i>гранулярен</i>
2	ЛИГ (LIG): links of group (k-unit)	к-элементы образуют между собой устойчивые когнитивные связи (к-связи). к-связи отражают причинные связи элементов и процессов в среде и в соотношения когнитивного агента с ней.	Элементы разума <i>увязаны</i>
3	КОГНИТОМ (cognitome): cognitive network	к-элементы и к-связи образуют единую когнитивную сеть (к-сеть). к-сеть является носителем субъективного опыта когнитивного агента.	Разум <i>системен</i>

Таким образом, теория вносит три фундаментальные понятия, когнитивных групп, линкеров когнитивных групп и единой когнитивной сети, которая является носителем всего субъективного опыта когнитивного агента. Теория выделяет три вида элементов в такой гиперсети и важными элементами в них являются микрогруппы, возникающие на перекрытии больших групп.

Нейронная гиперсеть состоит из трех видов элементов

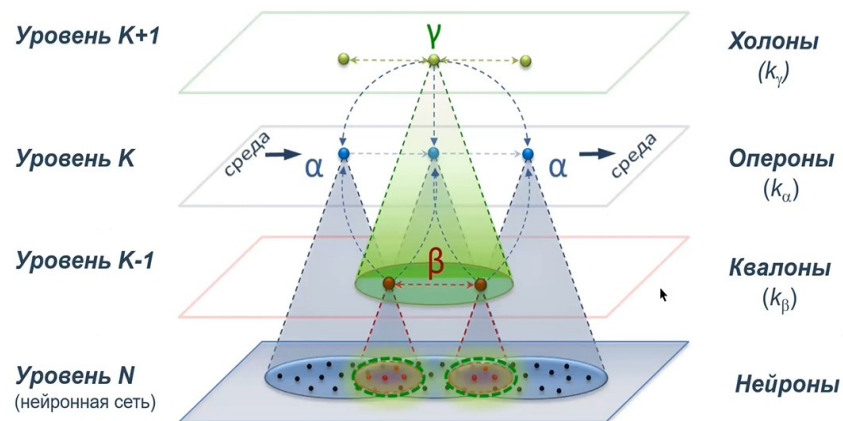


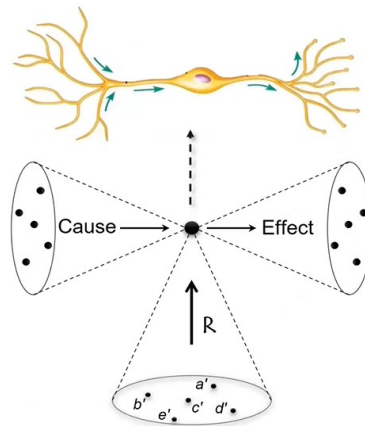
Рис.8. Нейронная гиперсеть

Fig. 8. Neural hypernetwork

圖 8. 神經超網絡

Эти микрогруппы имеют свойство того, что специализируясь в этих больших системах, нейроны способны служить как бы каналами связи между многими системами. И если речь идет о сознании, то теория говорит, что когда возникает такого рода гиперсеть со специализированными элементами, возбуждения способны перкалировать в этой сети.

Один нейрон, как минимальный когнитивный элемент



- ◆ Минимальным когнитивным элементом (КОГом) является один когнитивно специализированный нейрон
- ◆ Активность совокупности нейронов, активирующих эту клетку, является "причиной" для этого кога
- ◆ Группа нейронов, в активацию которых вносит вклад эта клетка, является эффектом этого кога

Рис.9. Минимальный когнитивный элемент нейрон

Fig. 9. Minimal cognitive element neuron

圖 9. 最小認知元神經元

Таким образом, при возбуждении одного элемента, относящегося ко многим системам, возбуждаются одновременно тысячи и десятки тысяч подобных элементов в том или ином состоянии активности мозга, возникает глобальный охват всего и элементов, и единиц опыта, накопленных субъектом, то есть его Я. Это и есть процесс осознания.

Нейронный когнитивный код

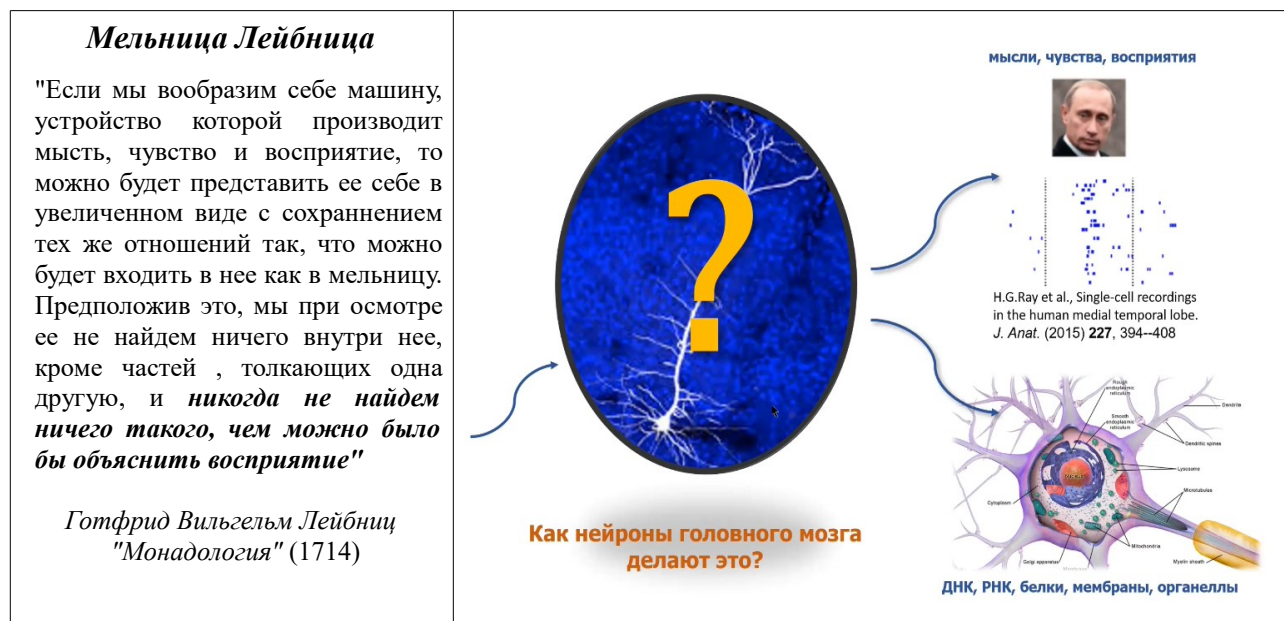


Рис. 10. Нейронный когнитивный код

Fig. 10. Neural cognitive code

圖 10. 神經認知程式碼

Мы, конечно, работаем в этой области, как нейробиологи, нейрофизиологи, нас интересует возможность и предсказания того, что минимальным, когнитивным элементом является отдельный нейрон, способный входя в разные когнитивные сообщества, нести в себе специализации общие для этой категории.

Кодирование такого когнитивного опыта отдельными клетками опускает весь процесс когнитивной системы и ее формирования на уровень отдельных клеток и молекулярных процессов в них. И это то, что нас интересует.

Суперпозиция разных функциональных систем на клетках глубоких слоёв нейронной сети

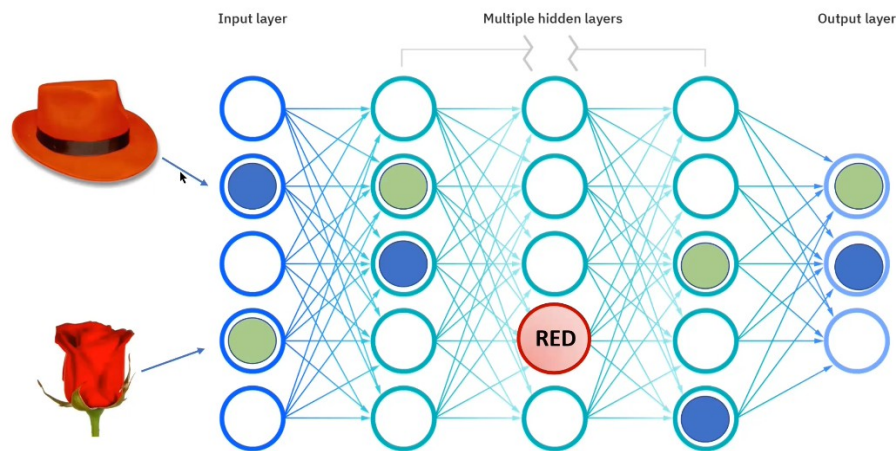
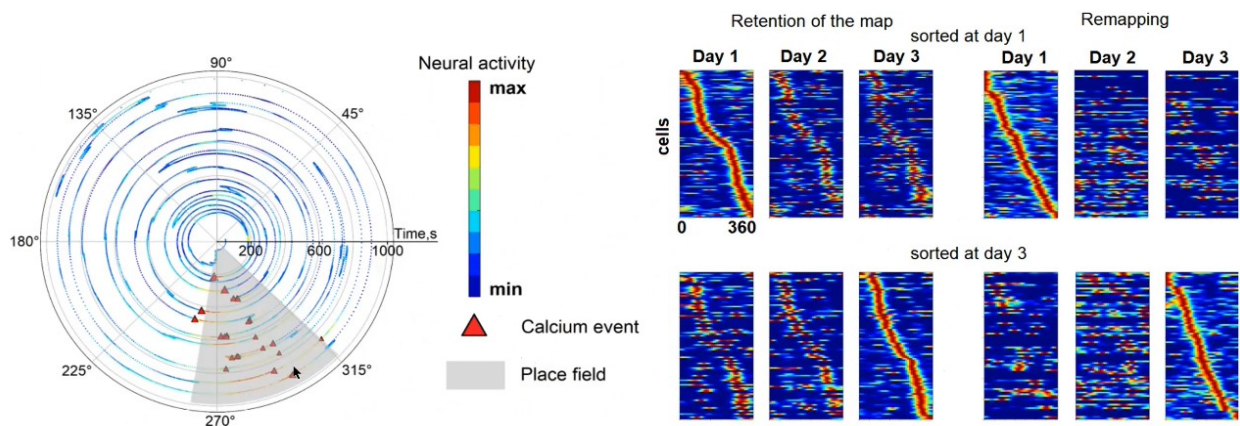


Рис. 11. Суперпозиция функциональных систем

Fig. 11. Superposition of functional systems

圖 11. 功能系統的疊加

У нас есть гипотезы, каким образом это происходит, что это происходит за счет суперпозиции разных функциональных систем, например, одной и второй, показанной действия цветами, когда общий элемент для этих двух систем может отражать некоторую активацию, связанную с тем или иным аспектом окружающего мира.



Быстрое (одномоментное?) возникновение
когнитивной специализации нейрона

Долговременное сохранение быстро возникших
когнитивных специализаций нейронов

Рис. 12. Когнитивная специализация нейрона

Fig. 12. CyCognitive neuron specialization

圖 12. 認知神經元特化

В данном случае две совершенно разные функциональные группы нейронов имеют общий аспект цвета красного цвета, который и специализирует

данный нейрон относительно этих свойств.

Мы исследуем это методами регистрации активности клеток в мозге экспериментальных животных. Мы показываем, что такого рода специализации когнитивного вида, например, специализация активности нейрона относительно одного места в среде, может происходить быстро, практически с первого раза, когда животные попадают в эту обстановку и сохраняются очень долго.

Multimodal Neurons in Artificial Neural Networks

Biological Neuron	CLIP Neuron	Previous Artificial Neuron	
Probed via depth electrodes	Neuron 244 from penultimate layer in CLIP RN50_4x	Neuron 483, generic person detector from Inception v1	
Halle Berry	Spiderman	human face	
			Photorealistic images
Responds to photos of Halle Berry and Halle Berry in costume ✓	Responds to photos of Spiderman in costume and spiders ✓ view more	Responds to faces of people ✓	
			Conceptual drawings
Responds to sketches of Halle Berry ✓	Responds to comics or drawings of Spiderman and spider-themed icons ✓ view more	Does not respond significantly to drawings of faces ✗	
			Images of text
Responds to the text "Halle Berry" ✓	Responds to the text "spider" and others ✓	Does not respond significantly to text ✗	

Рис. 13. Мультимодальные нейроны в искусственных нейронных сетях
 Fig. 13. Multimodal Neurons in Artificial Neural Networks
 圖 13. 人工神經網路中的多模態神經元

Примечательно, что в искусственных нейронных сетях выявлены такого же рода мультимодальные нейроны, это, например, работа компании OpenAI, исследовавшая специализацию клеток в глубоких слоях глубоких нейронных сетей и выявившие, что и в искусственных нейронных сетях, что ставит

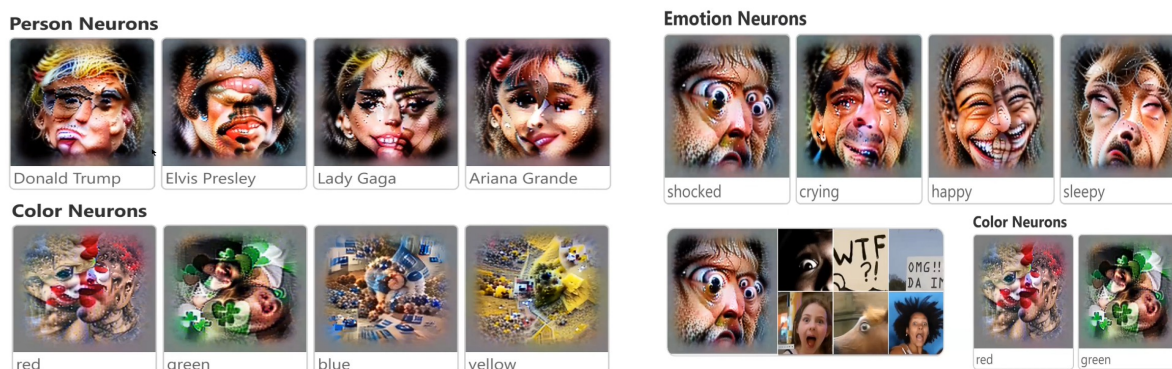


Рис. 14. Специализация нейронов

Fig. 14. Neuron specialization

圖 14. 神經元特化

серьезные глубокие вопросы, появляются нейроны, имеющие специализацию относительно отдельных личностей, отдельных цветов, отдельных эмоциональных выражений и так далее.

Ну, и последнее, о чем следует сказать, хотя нас, как нейрофизиологов, интересуют концептуальные основы и нейробиологические основы, тем не менее, стоит задача формализации этих систем и процессов в них.

Теория нейронных гиперсетей

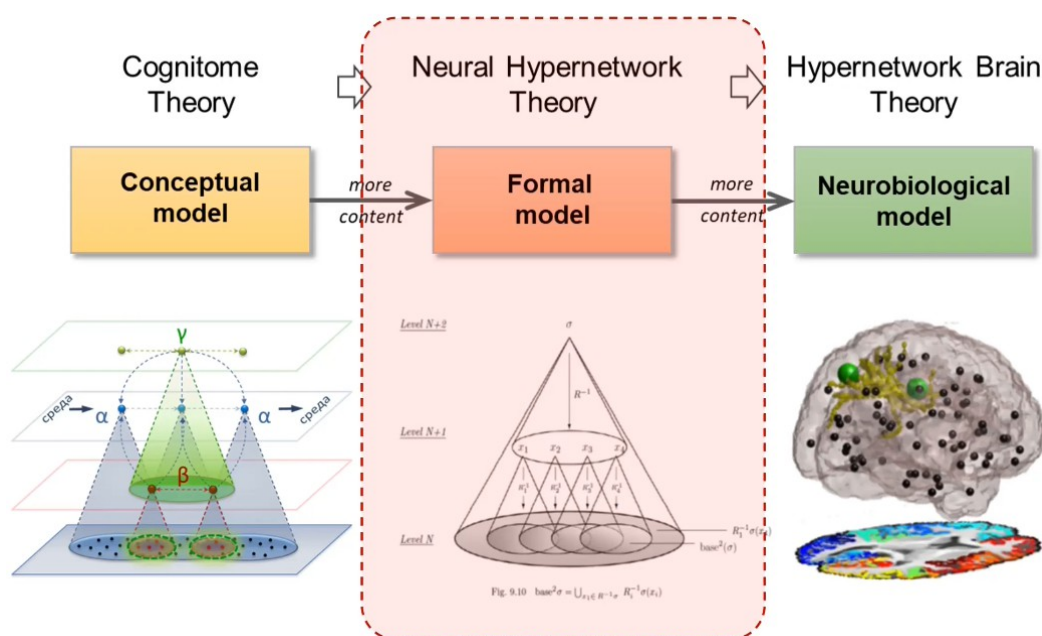
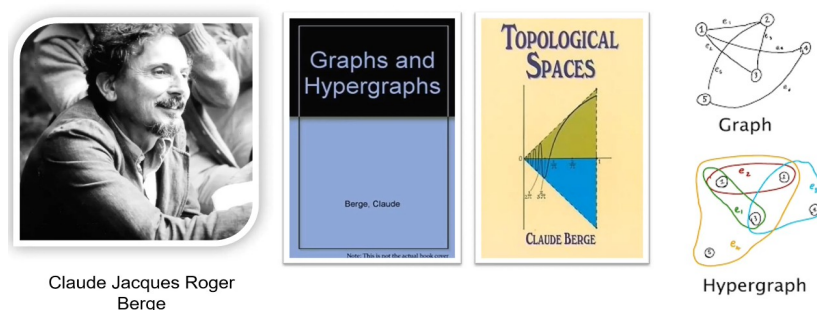


Рис. 15. Нейронные гиперсети

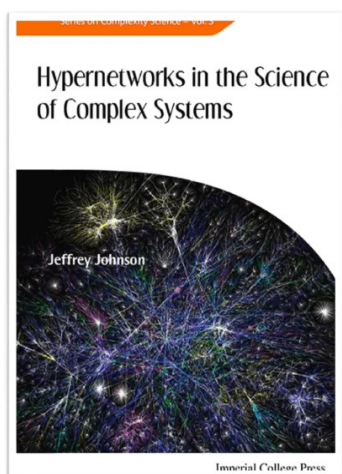
Fig. 15. Neural hypernetworks

圖 15. 神經超網絡

В начале работы над теорией в 2019 году я исходил из того, что сеть, которую мы изучаем, то есть когнитом, является когнитивной сетью, которая может соотноситься с нервной сетью, но очень быстро в течение первого года стало ясно, что это тупиковое решение: нельзя соотносить две сети, одна из которых не материальная, а вторая материальная. Необходим единый объект, который имеет две стороны, который является дуалом. Стало ясно, что сеть не подходит по ряду причин для описания такого объекта, потому что она исследует попарное отношение между элементами, а здесь я показывал это группы, объединяемые в новое свойство и, второе, что она не описывает высокопорядковых свойств. И стало ясно, что надо двигаться в сторону более сложных систем и математических моделей.



Стало ясно, что гиперграфы, введенные Бержем, также не являются подходящим понятием. И мы в 2014 году пришли к тому, что формальным описанием такого рода систем, которые мы видим, является гиперсети.



Hypernetworks are a natural extension of graphs, hypergraphs and networks

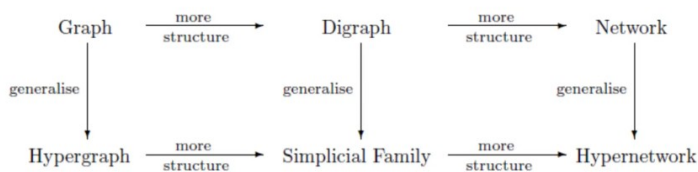


Рис. 16. Гиперсети
Fig. 16. Hypernetworks
圖 16. 超網路

В последнее время специалисты в области сети тоже пришли к этому, начиная с последних с 16-17-го года, назвав этот же объект высокопорядковыми

сетями. В обоих случаях речь идет об одном и том же и требует существенного дальнейшего исследования именно в нейронауке.

Поэтому я завершаю свой доклад первым слайдом.

Основные моменты доклада

Сознание в теории нейронных гиперсетей

1. Сознание – это специфический процесс в специфически организованной структуре
2. Структура – нейронная гиперсеть
3. Процесс – глобальная перколяция активности в ней

Мы как нейрофизиологи заинтересованы в совместном исследовании этой структуры и этого процесса с помощью подходов физико-математических наук

Институт перспективных исследований мозга МГУ
contact@brain.msu.ru

Рис. 17. Последний слайд доклада

Fig. 17. Last slide of the report

圖 17. 報告的最後一張幻燈片

Мы работаем на уровне теории высших структур и функций мозга. Мы формализуем это как понятие нейронные гиперсети или высокопорядковые сети. Мы считаем, что мозг представляет собой именно такую гиперсеть, а не низко уровневую первопорядковую нейронную сеть, так как это исследуется сегодня. И что сознание представляет собой специфический процесс в этой структуре, существующий только, когда эта структура развивается под влиянием опыта и обучения. Суть этого процесса безусловно загадочна, но мы думаем, что она может быть описана подходами перколяции, но в обоих случаях и структур, и процесса, нам совершенно очевидно, что необходимо сотрудничать со специалистами в области математики и теоретической физики и мы приглашаем к такому сотрудничеству.

Dear Colleagues! Thank you very much for the invitation to speak at this wonderful Congress, which brings together specialists in the field of systems biology and systems theory, as well as mathematics and artificial intelligence. My area of research lies precisely at the crossroads of these disciplines. This is the neurophysiology of higher brain functions. It turns out that in recent years it has been associated with the development of new mathematical approaches. And I would like today to talk about this area and the theory with which we are working and to invite and ask for cooperation from specialists in the field of mathematics, because this is what is happening now in the world and what is required for the progress of our understanding of how the brain gives birth to the process of

consciousness. I'll start right away with a summary that summarizes the main points of the report. Today we will talk about neural hypernetworks and processes of consciousness, according to the theory of neural hypernetworks with which we work. Consciousness is a specific process in a specifically organized structure. This is the structure of a neural hypernetwork. At least that's how it can be described in formal language. And the specific process is the process of global percolation of activity in this network. And we, as neurophysiologists, are interested in the joint study of this structure, an organic and at the same time mathematical structure, and this process, on the one hand physiological, cognitive, on the other hand, informational with the help of physical and mathematical sciences. I'll start by defining consciousness and why it's such an important topic today. Next, I will move on to the fact that you can understand this topic only with the help of theory. And that in recent years the role of mathematics and theoretical physics in theoretical studies of higher functions of the brain, including consciousness, is becoming greater and greater. Next, I will give some overview of the problematic situation. Today there are a large number of theories of consciousness, I will highlight the physical and mathematical direction in them, analyzing some of these theories. And finally, in the third part, I will talk about the theory with which we work and which we call the theory of neural hypernetworks or hypernetwork theory of the brain, I will talk about some principles of the structure of this structure, as it appears to us as neurophysiologists. And, briefly, that consciousness is a special type of traffic in this specific structure, which differs from the concepts known to us in the theories of neural networks.

So, the definition of consciousness. Consciousness is a term that has many meanings throughout history. However, recently there has been some consensus in the scientific community in defining the fundamental characteristics of consciousness, defined as any subjective experience generated either from within a system or in response to an external circumstance, which includes both high-order thoughts, plans, ideas, as well as emotional feelings and sensations. In short, everything that is felt by the subject. Our outstanding philosopher of consciousness, David Izrailevich Dubrovsky, once defined this as an encompassing term of subjective reality. This definition differs significantly from the definitions and understanding of consciousness that were given previously. For example, in a large encyclopedic dictionary, consciousness was defined 25 years ago as the highest form of mental reflection, which is characteristic of a socially developed person, associated with speech, the ideal side of goal-setting activity. Then, according to these ideas, only a person can have consciousness. Today, the situation and the focus of consciousness research have shifted significantly. But before talking about it, we need to formulate the very essence of the problem of consciousness. In principle, Ivan Petrovich Pavlov formulated it very sharply in one phrase about a hundred years ago. How does brain matter give rise to subjective phenomena? David Izrailevich Dubrovsky gave a more detailed definition of this problem, formulating it in two questions. How to explain the connection between the phenomena of subjective reality and brain processes, if the former cannot be attributed physical properties, and the latter necessarily possess them? And the second question is, if physical properties cannot be attributed to the phenomena of subjective reality, then how to explain their ability to cause a causal effect on bodily processes. It must be said that despite the fact that a huge number of philosophers have been studying this problem, thousands of books have been

written, in relation to both the first and second questions this is an open scientific problem. Neuroscience plays a significant role in the study of these issues today. Now let's talk about the importance of consciousness proposed in such a fundamental formulation. If we are talking about sensations, subjective experience, any experiences, including emotional reactions, then we must begin to consider consciousness as a broad natural phenomenon that is present not only in humans, but also in other animals. And today, also with the development of the artificial intelligence system, the question of consciousness in the artificial intelligence system is becoming more and more pressing. A number of advanced groups in various centers and companies are now modeling not just artificial intelligence, but artificial consciousness. This can be illustrated by two recent conferences. One of them took place in early May 2023 and was called the Animal Consciousness Conference 2023, Dharamshala, May 1-5. It brought together specialists from a dozen countries, leading experts in the field of consciousness research in the biological world, and the consensus that is reached in these kinds of conferences is that a large number of animals that we previously did not think of as beings with consciousness possess this phenomenal property, which I defined above. These are not only primates, these are not only dolphins and higher mammals with highly developed brains: elephants, whales, killer whales, and so on. These are birds, these are a number of species of cephalopods, such as octopuses and cuttlefish. I'm missing many other mammals. These are arthropods, bees, bumblebees, and other types of insects. Consciousness is widespread in the biological world.

The second conference, which traditionally takes place every year since 1994, a major conference called the Science of Consciousness, was held in Italy in Taormina (The Science of Consciousness 2023, Taormina, May, 22-27). A large number of reports were heard there. Here are some of their illustrations. At the conference on animal consciousness, for example, there was an amazing paper by Masanori Kohda, published some time ago (Masanori Kohda "Cleaner fish identify themselves via self-face recognition like humans: What does this mean?", Osaka Metropolitan University, Osaka, Japan), showing that cleaner fish are able to identify themselves and recognize themselves in the mirror. In the same way as is done by humans and some great apes. This is amazing data that forces us to be attentive, to look completely differently at the spread of the phenomenon of consciousness in the natural world. On the other hand, the key lecture at the Taormina conference was given by David Chalmers, its title: "Are Large Language Models Capable of Being Conscious?" (David Chalmers "Could a Large Language Model be Conscious?", NYU, New York, USA). There he reviewed the arguments for and against. There is no consensus on this issue. There are different points of view, but this question stands. Thus, we see how relevant it is today to understand the fundamental principles of the organization of consciousness. And to understand them, purely experimental studies are not enough; the amount of facts accumulated in neuroscientific research is enormous. But they have not yet led to the holy grail, which was proclaimed about 30 years ago by Nobel laureate Francis Crick and his colleague Christoph Koch - to find the neural or neural correlates of consciousness. There is still no reliable neural correlate of consciousness. And this is partly understandable, because only theory will allow us to decide what is observed in experiment and predict what kind of correlates should be looked for. You cannot refuse to build a theory at the beginning of your journey. This will prove detrimental to the evaluation of the

experimental results. “When we say that we understand the totality of natural phenomena,” Einstein wrote, we mean that we have found a constructive theory that embraces them.” And, indeed, with the gradual disappointment in purely empirical experimental studies in the search for certain neural correlates of consciousness, in recent decades there has been a literally explosion in the number of theories of consciousness.

Fig. 1. Growth in the number of publications on theories of consciousness

This shows an increase in the number of publications that include the keywords consciousness and theory from 1971 to 2021, but if you look simply at the 21-22 years of work covering this field, you will see that synthetic meta-analyses of a huge number of theories have appeared. There are definitely more than 30 of these theories today, maybe about 50. And questions arise about how to compare these theories. There are quite a lot of works on this topic.

Fig. 2. Leading theories of consciousness according to conference participants in Taormina (The Science of Consciousness 2023, Taormina, May, 22-27)

At the mentioned conference in Italy, in Taormina in May 2023, a special competition was even organized - a vote of all conference participants, of whom there were more than 500 scientists, with the choice of their preferred theory of consciousness. In Figure 2 above, only 18 theories are listed, there is some arbitrariness of the organizers in what theories they chose, but this is the essence. One of the speakers at this conference, together with her colleagues, a year ago proposed a special database, called Contrast, for analyzing and comparing empirical studies of certain theories of consciousness.

It is clear that the region is flourishing and, perhaps, after some time there will even be a crisis of overproduction.

Another important line that has been emerging in recent years is the development of the physical and mathematical direction in the study of consciousness. It becomes clear that consciousness, as a state and processes in complex systems, requires analysis of a variety of mathematical positions. This is a mathematical object and information process, depending on the views of certain authors.

Here are the names of a number of presentations by famous experts at the seminar “Mathematical Consciousness Science”, which is regularly held online (<https://seminar.math-consciousness.org>).

Carlo Rovelli “Understanding consciousness within the known laws of physics three small contributions”,

Karl Friston “Markov blankets and Bayesian mechanics”,

Stephen Wolfram “How Features of Our Consciousness Seem to Define Our

Lows of Physics and Mathematics”,

Nao Tsuchiya “A search for an isomorphism between conscious experience and structure of information”.

Another example is from the Oxford Institute of Mathematics, which runs workshops on mathematical models and the foundations of consciousness (Fig. 3).

Fig. 3. Oxford Institute of Mathematics workshops on modeling consciousness.

I could also cite a large number of works here, but due to lack of space I will limit myself to only two examples given. Among theories of consciousness there are, of course, favorites.

The main such theories were reviewed last year (Seth & Bayne (2022) Nature Reviews Neuroscience). The diagram from the review (Fig. 4) shows the four largest popular theories of consciousness. High-level theory of consciousness, global workspace theory, integrated information theory, and reentry theory.

Fig. 4. Popular theories of consciousness.

I don't have the opportunity to talk about each of them. I will only note that in many discussions and debates, the dominant position today is occupied by the theory of the global workspace, which was proposed by a specialist in the field of computational psychology and cognitive neuroscience Jean-Pierre Changeux and the theory of integrated information proposed by the Italian scientist Giulio Tononi (Giulio Tononi), working in America. Today the situation has advanced in such a way that a large project has been created by very famous scientists (see Figure 5), who have written a protocol for competitively comparing the predictions of these two theories, global workspace theory and integrated information theory

Fig. 5. Protocol for comparing theories

Proponents of these theories gathered, made predictions regarding the neural processes that should be observed according to one theory and according to the other, which differ, and agreed on protocols. A large-scale study has now begun in a number of laboratories around the world regarding the differing predictions between these two leading theories.

I must say a few words about integrated information theory simply because this theory was formulated as a mathematical theory. A large number of mathematicians and specialists in the field of theoretical physics are involved in the development of this theory of work. For example, Larissa Albantakis, a specialist in the field of quantum physics, came to work in neuroscience, to study the theory of consciousness. Well, and so on. A few months ago, the fourth version of this theory, IIT 4.0, was

released, which differs from the first, second and third, this is work that is actively being developed, and I simply refer in this case to this publication [1], it is significant. We are also working in this direction, and this movement is largely dictated by the traditions of the Russian physiological school, which raised the question of the physiological, material foundations of the higher functions of the brain since the mid-19th century, at the beginning of Ivan Mikhailovich Sechenov, who had brilliant theoretical works, in particular, "Elements of Thought", but there were no experimental approaches.

However, he did a lot for theoretical models of connections in the nervous system, the so-called natural groups, purely postulated then in the 19th century, groups of neurons connected by joint activity to reflect certain external signals. Ivan Petrovich Pavlov brought to this idea both the connections between such groups of conditional connections and the experimental method. Pyotr Kuzmich Anokhin, Pavlov's student, developed Pavlov's idea of conditioned reflexes into a systemic model of brain function, the theory of functional systems.

We are trying to move further along this path by imagining how many functional systems that perform certain types of purposeful brain activity together form a certain complex that constitutes the essence of a cognitive agent, in other words, its cognitome. We started working with this theory about 10 years ago; in [2-5] some milestones are indicated in reports on this theory at various conferences. A synthetic publication on the basics of this theory is currently being prepared. I will briefly talk about some aspects that may be of interest for physical and mathematical approaches. It must be said right away that this theory is different from most other theories. Firstly, because it begins not with the question of consciousness, but with the question of what the mind is, and what the Self is, which perceives certain states of activity. Secondly, this theory begins from first principles; it does not postulate existing brain structures that provide this highest level of integration and mental activity, but begins from biological principles that do not contain any psychological terms. The combination of these principles is the initial conditions sufficient for a system of organisms of the Darwinian type, which has a combination of these first principles, to gradually begin to develop within itself a high-order structure that is the bearer of intelligence and consciousness.

Fig. 6. Principles of the theory of neural hypernetworks.

From these first principles, the theory deduces the principles of the emergence of these high-order structures, and from the first principles of emergence, the theory deduces what arises, that is, the structure, the processes in it and the principles of existence.

Today I dwell only on the last part of the theory, on the principles of existence. The theory introduces three types of new objects that are not usually considered in neurophysiological language and neurophysiological apparatus.

She introduces the concept of cognitive neural groups or cogs, connections between these groups (group connections) and a special cognitive complex formed by these two types of elements, which the theory designates as a cognitome. Let's say a few words about each of these elements.

Fig. 7. COG and LIG cognitome

A cognitive group is a cooperative group of neurons, which, in its joint activity, determines some aspect of the relationship of the whole organism with one or another aspect of the surrounding world, where the appearance of these aspects acts as a reason for the activation of this neural group of specialized elements, and the activation of this neural group can act as a cause for subsequent changes in behavior or mental states of other neural groups that determine the dynamics of nervous processes, actions and behavior of the body.

These kinds of cognitive-level neural groups, the vertices of which encode higher-order concepts than the individual elements that make up the group, are interconnected by linkers, which in theory are called leagues. The term “cog” in theory has a double meaning: on the one hand, it is a small element, a subordinate part of an entire cognitive system, a qualitatively specific experience, which is encoded by this group.

On the other hand, it is an abbreviation for a physically existing neural group, the activity of which determines this element of a specific experience with its cause-and-effect potential. The term “lig” linker groups reflects the specific feature of the connections in the structure of the hypernet.

Unlike a neural network, where connections are contacts between individual cells, that is, axons, wires connecting cells in the nervous system, in a hypernetwork, in a neural hypernetwork, connections are not edges of the network, but its elements, common elements between two groups, which being activated by one group, it is able to transmit excitation (signal) to another group.

Three fundamental concepts of TNG

The mind is real - it is a specific organic and methematic structure (K-network) that mediates the causal relationships of the cognitive agent (K-agent) with the environment.

To describe the elements and properties of this specific system requires the introduction of new fundamental scientific concepts.

Table 1

Thus, the theory introduces three fundamental concepts, cognitive groups, linkers of cognitive groups and a single cognitive network, which is the bearer of all the subjective experience of a cognitive agent. The theory identifies three types of elements in such a hypernetwork, and the important elements in them are

microgroups that arise on the overlap of large groups.

Fig. 8. Neural hypernetwork

These microgroups have the property that, specializing in these large systems, neurons are able to serve as communication channels between many systems. And if we are talking about consciousness, then the theory says that when this kind of hypernetwork with specialized elements arises, excitations are capable of percolating in this network.

Fig. 9. Minimal cognitive element neuron

Thus, when one element related to many systems is excited, thousands and tens of thousands of similar elements are simultaneously excited in one or another state of brain activity, a global coverage of all elements and units of experience accumulated by the subject, that is, his Self, arises. there is a process of awareness.

Fig. 10. Neural cognitive code

We, of course, work in this area, as neurobiologists, neurophysiologists, we are interested in the possibility and predictions that the minimum cognitive element is a single neuron, capable of entering different cognitive communities and carrying specializations common to this category.

The encoding of such cognitive experience by individual cells brings the entire process of the cognitive system and its formation down to the level of individual cells and the molecular processes in them. And this is what interests us.

Fig. 11. Superposition of functional systems

We have hypotheses about how this happens, that this occurs due to the superposition of different functional systems, for example, one and the other,

actions shown by colors when a common element for these two systems may reflect some activation associated with one or another aspect of the surrounding world.

Fig. 12. CyCognitive neuron specialization

In this case, two completely different functional groups of neurons have a common color aspect of red, which specializes this neuron with respect to these properties.

We study this by recording cell activity in the brains of experimental animals. We show that this kind of specialization of a cognitive type, for example, the specialization of neuron activity relative to one location in the environment, can occur quickly, almost from the first time animals enter this environment, and persist

for a very long time.

Fig. 13. Multimodal Neurons in Artificial Neural Networks

It is noteworthy that the same kind of multimodal neurons have been identified in artificial neural networks; this is, for example, the work of the OpenAI company, which studied the specialization of cells in the deep layers of deep neural networks and revealed that in artificial neural networks, which puts

Fig. 14. Neuron specialization

serious deep questions, neurons appear having specializations regarding particular personalities, particular colors, particular emotional expressions, and so on. Well, the last thing that should be said is that although we, as neurophysiologists, are interested in the conceptual foundations and neurobiological foundations, nevertheless, we face the task of formalizing these systems and the processes in them.

Fig. 15. Neural hypernetworks

When I started working on the theory in 2019, I assumed that the network we were studying, the cognitome, was a cognitive network that could be correlated with a neural network, but very quickly within the first year it became clear that this was a dead end: it is impossible to correlate two networks, one of which is not material, and the second is material. What is needed is a single object that has two sides, which is a dual. It became clear that the network is not suitable for a number of reasons for describing such an object, because it explores the pairwise relationship between elements, and here I showed these groups being combined into a new property and, secondly, that it does not describe high-order properties. And it became clear that we need to move towards more complex systems and mathematical models.

It became clear that hypergraphs, as introduced by Berge, were also not a suitable concept. And in 2014 we came to the conclusion that the formal description of this kind of systems that we see is hypernetworks.

Fig. 16. Hypernetworks

Recently, network specialists have also come to this conclusion, starting from the last 16-17 years, calling the same object high-order networks. In both cases we are talking about the same thing and requires significant further research in neuroscience.

Therefore, I conclude my report with the first slide.

Fig. 17. Last slide of the report

We work at the level of theory of higher structures and functions of the brain. We formalize this as the concept of neural hypernetworks or high-order networks. We believe that the brain is just such a hypernetwork, rather than a low-level first-order neural network, as is being researched today. And that consciousness is a specific process in this structure, existing only when this structure develops under the influence of experience and learning. The essence of this process is certainly mysterious, but we think that it can be described by percolation approaches, but in both cases of structures and process, it is quite clear to us that it is necessary to cooperate with specialists in the field of mathematics and theoretical physics and we invite such cooperation.

親愛的同事們！非常感謝您受邀在這次精彩的大會上發言，這次大會匯集了系統生物學和系統理論以及數學和人工智慧領域的專家。我的研究領域恰恰位於這些學科的十字路口。這是高級大腦功能的神經生理學。事實證明近年來它與新數學方法的發展有關。今天我想談談這個領域和我們正在研究的理論，並邀請和請求數學領域專家的合作，因為這就是世界上正在發生的事情，也是進步所需要的我們對大腦如何產生意識過程的理解。我將立即開始總結報告的要點。今天，我們將根據我們所研究的神經超網絡理論來討論神經超網絡和意識過程。意識是特定組織結構中的特定過程。這就是神經超網絡的結構。至少用正式語言可以這麼描述。而具體的過程就是這個網路中活動的全域滲透的過程。身為神經生理學家，我們對這種結構的共同研究感興趣，這種結構是一種有機的、同時也是數學的結構，而這個過程一方面是生理的、認知的，另一方面是在物理和認知的幫助下提供資訊的。數學科學。我將首先定義意識以及為什麼它是今天如此重要的主題。接下來，我將繼續說明您只能藉助理論來理解這個主題。近年來，數學和理論物理學在包括意識在內的大腦高級功能的理論研究中的作用變得越來越大。接下來，我將概述一下有問題的情況。今天有大量的意識理論，我將突出其中的物理和數學方向，分析其中的一些理論。最後，在第三部分中，我將討論我們所使用的理論，我們稱之為神經超網絡理論或大腦超網絡理論，我將討論這種結構的一些原理，因為它看起來對於我們作為神經生理學家來說。簡而言之，意識是這種特定結構中的一種特殊類型的流量，它與我們在神經網路理論中已知的概念不同。

所以，意識的定義。意識是一個在歷史上有多種意義的術語。然而，最近科學界在定義意識的基本特徵方面達成了一些共識，意識被定義為系統內部或響應外部環境而產生的任何主觀體驗，其中包括高階思想、計劃、想法，以及情感感受和感覺。簡而言之，就是主體所感受到的一切。我們傑出的意識

哲學家大衛·伊茲萊耶維奇·杜布羅夫斯基（David Izrailevich Dubrovsky）曾將其定義為主觀現實的包羅萬象的術語。這個定義與之前給出的意識的定義和理解有很大不同。例如，在一本大型百科全書詞典中，意識在 25 年前被定義為心理反映的最高形式，這是社會發展的人的特徵，與言語相關，是目標設定活動的理想方面。那麼，根據這些想法，只有人才擁有意識。如今，意識研究的情況和重點已經發生了重大變化。但在討論它之前，我們需要先闡明意識問題的本質。原則上，伊凡彼得羅維奇巴甫洛夫在大約一百年前用一句話非常尖銳地表述了這一點。大腦物質如何產生主觀現象？David Izrailevich Dubrovsky 對這個問題給出了更詳細的定義，將其表述為兩個問題。如果主觀現實現象與大腦過程不能歸因於物理屬性，而後者必然具有物理屬性，那麼如何解釋主觀現實現象與大腦過程之間的關聯？第二個問題是，如果物理特性不能歸因於主觀現實的現象，那麼如何解釋它們對身體過程造成因果影響的能力。必須說，儘管有大量的哲學家一直在研究這個問題，已經寫了數千本書，但就第一個問題和第二個問題而言，這是一個開放的科學問題。神經科學在當今這些問題的研究中發揮著重要作用。現在讓我們來談談在這樣一個基本表述中提出的意識的重要性。如果我們談論的是感覺、主觀經驗、任何經驗，包括情緒反應，那麼我們必須開始將意識視為一種廣泛的自然現象，它不僅存在於人類身上，也存在於其他動物身上。而今天，同樣隨著人工智慧系統的發展，人工智慧系統的意識問題變得越來越迫切。各個中心和公司的許多先進團隊現在不僅對人工智慧進行建模，而且對人工意識進行建模。最近的兩次會議可以說明這一點。其中一場於 2023 年 5 月初舉行，名為“2023 年動物意識會議”，於 5 月 1 日至 5 日在達蘭薩拉舉行。它匯集了來自十幾個國家的專家，生物界意識研究領域的頂尖專家，這類會議達成的共識是，大量我們以前不認為是生命體的動物意識擁有我上面定義的這種現象屬性。這些不僅是靈長類動物，不只是海豚，還有大腦高度發達的高等哺乳動物：大象、鯨魚、虎鯨等等。這些是鳥類，是多種頭足類動物，例如章魚和墨魚。我錯過了許多其他哺乳動物。這些是節肢動物、蜜蜂、大黃蜂和其他類型的昆蟲。意識在生物世界中廣泛存在。

第二次會議自 1994 年起每年舉行一次，這是一個名為“意識科學”的重要會議，在義大利陶爾米納舉行（“意識科學 2023”，陶爾米納，5 月 22-27 日）。那裡聽到了大量的報告。這是他們的一些插圖。例如，在動物意識會議上，前一段時間發表了 Mosanori Kohda 的一篇令人驚嘆的論文（Masanori Kohda“清潔魚像人類一樣通過自我面部識別來識別自己：這是什麼意思？”，大阪都立大學，日本大阪），表明清潔魚能夠識別自己並在鏡子中認出自己。就像人類和一些類人猿所做的那樣。這是令人驚嘆的數據，迫使我們集中注意力，以完全不同的方式看待自然世界中意識現象的傳播。另一方面，陶爾

米納會議的主題演講由 David Chalmes 做，題目是：“大型語言模型能夠有意識嗎？”（David Chalmers“大型語言模型可以有意識嗎？”，紐約大學，紐約，美國）。在那裡，他回顧了支持和反對的論點。在這個問題上尚未達成共識。有不同的觀點，但這個問題是成立的。因此，我們看到今天理解意識組織的基本原則是多麼重要。要理解它們，純粹的實驗研究是不夠的；神經科學研究中累積的事實數量是巨大的。但它們尚未實現諾貝爾獎得主弗朗西斯·克里克和他的同事克里斯托夫·科赫大約 30 年前宣布的聖杯——找到意識的神經或神經關聯。意識仍然沒有可靠的神經關聯。這在某種程度上是可以理解的，因為只有理論才能讓我們決定在實驗中觀察到什麼，並預測應該尋找什麼樣的相關性。你不能拒絕在旅程開始時建立一個理論。這將不利於實驗結果的評估。愛因斯坦寫道：“當我們說我們理解了自然現象的整體時，我們的意思是我們已經找到了包含它們的建設性理論。”事實上，隨著尋找意識的某些神經相關性的純粹經驗實驗研究逐漸令人失望，近幾十年來，意識理論的數量出現了真正的爆炸性增長。

圖 1. 意識理論出版數量的成長

這表明從 1971 年到 2021 年，包含關鍵字「意識」和「理論」的出版物數量有所增加，但如果你簡單地看一下涵蓋該領域的 21-22 年的工作，你會發現大量的綜合薈萃分析的理論已經出現。如今這些理論肯定超過 30 種，也許大約 50 種。如何比較這些理論的問題就出現了。關於這個主題有很多作品。

圖 2. 陶爾米納會議參與者認為領先的意識理論
（《意識科學 2023》，陶爾米納，5 月 22-27 日）

在 2023 年 5 月在義大利陶爾米納舉行的上述會議上，甚至舉辦了一場特別競賽——所有會議參與者（其中有 500 多名科學家）進行投票，選擇他們喜歡的意識理論。上圖 2 只列出了 18 種理論，組織者選擇的理論有一定的隨意性，但這就是本質。這次會議的一位發言人與她的同事一年前提出了一個名為「對比」的特殊資料庫，用於分析和比較某些意識理論的實證研究。

顯然，該地區正在蓬勃發展，也許一段時間後甚至會出現生產過剩的危機。

近年來出現的另一條重要路線是意識研究中物理和數學方向的發展。很明顯，意識作為複雜系統中的一種狀態和過程，需要對各種數學位置進行分析這是一個數學對象和資訊過程，取決於某些作者的觀點。

以下是定期在線上舉辦的「數學意識科學」研討會上著名專家的一些演講的名稱（<https://seminar.math-thought.org>）。

卡羅·羅維利（Carlo Rovelli）“在物理學已知低點內理解意識的三個小貢

獻”，

Karl Friston“馬可夫毛毯和波耶森力學”，

史蒂芬·沃爾夫勒姆“我們意識的特徵如何定義我們的物理和數學水平”，

Nao Tsuchiya「尋找意識經驗與資訊結構之間的同構」。

另一個例子來自牛津數學研究所，該研究所舉辦數學模型和意識基礎研討會（圖 3）。

圖 3. 牛津數學研究所關於建模意識的研討會。

我還可以在這裡引用大量的作品，但由於篇幅有限，我只能舉出兩個例子當然，在意識理論中，也有一些最受歡迎的理論。

去年對此類主要理論進行了回顧（Seth & Bayne (2022) *Nature Reviews Neuroscience*）。評論中的圖表（圖 4）顯示了四種最受歡迎的意識理論。高階意識理論、全局工作空間理論、綜合資訊理論、再入理論。

圖 4. 流行的意識理論。

我沒有機會談論他們每一個人。我只想指出，在許多討論和辯論中，今天佔據主導地位的是全局工作空間理論，該理論是由計算心理學和認知神經科學領域的專家 Jean-Pierre Changeux 和綜合資訊理論提出的。由在美國工作的義大利科學家朱利奧·托諾尼（Giulio Tononi）提出。如今，情況已經發展到這樣的程度：非常著名的科學家創建了一個大型專案（見圖 5），他們編寫了一個協議，用於競爭性地比較這兩種理論（全局工作空間理論和整合資訊理論）的預測

圖 5. 理論比較協議

這些理論的支持者聚集在一起，對應該根據一種理論和另一種不同的理論觀察的神經過程做出預測，並就協議達成一致。世界各地的許多實驗室現已開始針對這兩種主要理論之間的不同預測進行大規模研究。

我必須對綜合資訊理論說幾句話，因為這個理論被表述為數學理論。大量數學家和理論物理領域的專家參與了這個工作理論的發展。例如，量子物理學領域的專家拉里薩·阿爾班塔基斯(Larissa Albantakis)開始從事神經科學工作，研究意識理論。嗯，等等。幾個月前，這個理論的第四個版本，IIT 4.0，發布了，它與第一個、第二個和第三個不同，這是一個正在積極開發的工作，我在這裡簡單地參考這個出版品[1]，意義重大。我們也在朝這個方向努力，這項運動很大程度上是由俄羅斯生理學派的傳統決定的，自 19 世紀中葉以來，俄羅斯生理學派就提出了大腦高級功能的生理和物質基礎的問題。伊万·米哈伊洛

維奇·謝切諾夫（Ivan Mikhailovich Sechenov）有出色的理論著作，特別是《思想要素》，但沒有實驗方法。

然而，他為神經系統連接的理論模型做了很多工作，即所謂的自然群體，純粹是在 19 世紀假設的，神經元群體透過關節活動連接起來以反映某些外部訊號。伊凡·彼得羅維奇·巴甫洛夫（Ivan Petrovich Pavlov）將這些條件連結組之間的聯繫和實驗方法引入了這個想法。巴甫洛夫的學生彼得·庫茲米奇·阿諾欣（Pyotr Kuzmich Anokhin）將巴甫洛夫的條件反射思想發展成為大腦功能的系統模型，即功能係統理論。

我們試著沿著這條道路進一步前進，想像有多少個功能係統共同執行某些類型的有目的的大腦活動，形成某種複合體，該複合體構成了認知主體的本質換句話說，就是它的認知體。我們大約 10 年前開始研究這個理論；在 [2-5] 中，在各種會議上關於這個理論的報告中指出了一些里程碑。目前正在準備一份關於該理論基礎知識的綜合出版物。我將簡要討論物理和數學方法可能感興趣的一些方面。必須立即指出，該理論與大多數其他理論不同。首先，因為它不是從意識的問題開始，而是從什麼是心靈、什麼是感知某些活動狀態的自我的問題開始。其次，該理論從第一原理開始；它並不假設現有的大腦結構能夠提供最高水平的整合和心理活動，而是從不包含任何心理學術語的生物學原理開始。這些原則的結合是達爾文類型的有機體系統充分的初始條件，該系統具有這些基本原則的組合，逐漸開始在其自身內部發展出承載智慧和意識的高階結構。

圖 6. 神經超網絡理論原理。

從這些第一原理，理論推導出這些高階結構出現的原理，從出現的第一原理，理論推導出出現的東西，即結構、其中的過程和存在的原理。

今天我只談理論的最後一部分，也就是存在的原則。該理論引入了神經生理學語言和神經生理學裝置中通常不考慮的三種新對象。

她介紹了認知神經群體或齒輪的概念、這些群體之間的聯繫（群體聯繫）以及由這兩類元素形成的特殊認知複合體，該理論將其稱為認知組。讓我們對每個元素簡單說一下。

圖 7. COG 和 LIG 認知組

認知群體是一組合作的神經元，在其聯合活動中，決定了整個有機體與周圍世界的一個或另一個方面的關係的某些方面，其中這些方面的出現充當激活的原因這個由特殊元素組成的神經群的激活，可以導致其他神經群的行為或精神狀態發生後續變化，從而決定身體的神經過程、行動和行為的動態。

這些認知級神經組的頂點編碼比構成該組的各個元素更高階的概念，它們透過連結器互連，理論上稱為聯盟。「齒輪」一詞在理論上具有雙重意義：一方面，它是一個小元素，是整個認知系統的從屬部分，是由該群體編碼的定性特定體驗。

另一方面，它是物理上存在的神經群的縮寫，其活動決定了特定體驗的這一元素及其因果潛力。術語「lig」連接基團反映了超網路結構中連接的特定特徵。

與神經網路不同，神經網路中的連接是單一細胞之間的接觸，即軸突，連接神經系統中細胞的電線，在超網路中，在神經超網路中，連接不是網路的邊緣，而是它的元素，之間共同元素兩組，被一組激活後，能夠將激勵（信號）傳輸到另一組。

TNG 的三個基本概念

心靈是真實的一它是一種特定的有機和數學結構（K-網絡），調節認知主體（K-主體）與環境的因果關係。

為了描述這個特定系統的元素和特性，需要引入新的基本科學概念。

表格 1

因此，該理論引入了三個基本概念：認知群體、認知群體的連結者和單一認知網絡，認知網絡承載著認知主體的所有主觀經驗。該理論確定了這種超網絡中的三類元素，其中重要的元素是在大群體重疊時出現的微群體。

圖 8. 神經超網絡

這些微群具有這樣的特性，即專門針對這些大系統的神經元能夠充當許多系統之間的通訊通道。如果我們談論的是意識，那麼理論表明，當這種具有專門元素的超網絡出現時，興奮能夠在該網絡中滲透。

圖 9. 最小認知元神經元

因此，當與許多系統相關的一個元素被激發時，成千上萬個類似的元素在一種或另一種大腦活動狀態下同時被激發，這是受試者積累的所有元素和經驗單元的全域覆蓋，即他的自我出現了，有一個覺知的過程。

圖 10. 神經認知程式碼

當然，我們作為神經生物學家、神經生理學家在這一領域工作，我們對最小認知元素是單一神經元的可能性和預測感興趣，它能夠進入不同的認知社區

並具有該類別共有的專業化。

個別細胞對此認知經驗的編碼將認知系統及其形成的整個過程降低到個別細胞及其分子過程的層次。這就是我們感興趣的地方。

圖 11. 功能系統的疊加

我們對這是如何發生的有假設，這是由於不同功能系統的疊加而發生的，例如，一個和另一個，

當這兩個系統的共同元素可能反映與周圍世界的一個或另一個方面相關的某些激活時，用顏色顯示的動作。

圖 12. 認知神經元特化

在這種情況下，兩個完全不同的神經元功能組具有共同的紅色顏色，這使得該神經元在這些屬性方面具有特殊性。

我們透過記錄實驗動物大腦中的細胞活動來研究這一點。我們表明，這種認知類型的專門化，例如，神經元活動相對於環境中某個位置的專門化，可以很快發生，幾乎從動物第一次進入這個環境時就可以發生，並且持續很長時間。

圖 13. 人工神經網路中的多模態神經元

值得注意的是，在人工神經網路中已經發現了相同類型的多模態神經元；例如，OpenAI 公司的工作，該公司研究了深度神經網路深層細胞的特化，並揭示了在人工神經網路中神經網絡，它把

圖 14. 神經元特化

對於嚴肅的深層問題，神經元似乎在個體性格、個體顏色、個體情緒表達等方面具有專門化。好吧，最後要說的是，儘管我們作為神經生理學家對概念基礎和神經生物學基礎感興趣，但我們面臨著將這些系統及其過程形式化的任務。

圖 15. 神經超網絡

當我在 2019 年開始研究這個理論時，我假設我們正在研究的網絡（認知組）是一個可以與神經網絡相關的認知網絡，但很快在第一年就發現這是一個死的網路 end：不可能關聯兩個網絡，其中一個是非物質的，第二個是物質的。

我們需要的是一個有兩個面的單一物體，它是對偶的。很明顯，由於多種原因，該網絡不適合描述這樣的對象，因為它探索了元素之間的成對關係，在這

裡我展示了這些組被組合成一個新的屬性，其次，它不描述高階屬性。很明顯，我們需要轉向更複雜的系統和數學模型。

很明顯，伯格引入的超圖也不是適當的概念。2014年，我們得出的結論是，我們所看到的此類系統的正式描述是超網路。

圖 16. 超網路

最近，網路專家也得出了這個結論，從最近16-17年開始，將同一物件稱為高階網路。在這兩種情況下，我們談論的是同一件事，並且需要在神經科學方面進行大量的進一步研究。

因此，我以第一張投影片結束我的報告。

圖 17. 報告的最後一張幻燈片

我們在大腦高階結構和功能的理論層面上工作。我們將其形式化為神經超網路或高階網路的概念。我們相信大腦就是這樣一個超網路，而不是像今天所研究的那樣是一個低階的一階神經網絡。而意識是這個結構中的特定過程，只有當這個結構在經驗和學習的影響下發展時才存在。這個過程的本質固然是神秘的，但我們認為它可以用滲透方法來描述，但無論是結構還是過程我們都非常清楚，有必要與數學和數學領域的專家合作。理論物理學和我們邀請這樣的合作。

Литература

References

參考書目

1. Albantakis L., Barbosa L., Findlay G., Grasso M., Haum A.M., Marshall W., Mayner W., Zaeemzadeh A., Boly M., Juel B.E., Sasai S., Fujii K., David I., Hendren J., Lang J.P., Tononi G. Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms, arXiv:2212.14787v1[q-bio.NC] 30 Dec 2022.
2. Anokhin K.V. Cognitome: theory of realized degrees of freedom of the brain, Fifth International Conference on Cognitive Science, Kaliningrad, 2012.
3. Anokhin K.V. Cognitome: in search of a general theory of cognitive science, Sixth International Conference on Cognitive Science, Kaliningrad, 2014.
4. Anokhin K.V. Connectome and Cognit: bridging the gap between brain and mind, Seventh International Conference on Cognitive Science, Svetlogorsk, 2016.
5. Anokhin K.V. Neural code: in search of cellular principles of coding cognitive information, Baltic Forum. Neuroscience, artificial intelligence, complex systems, Kaliningrad, 2021.