

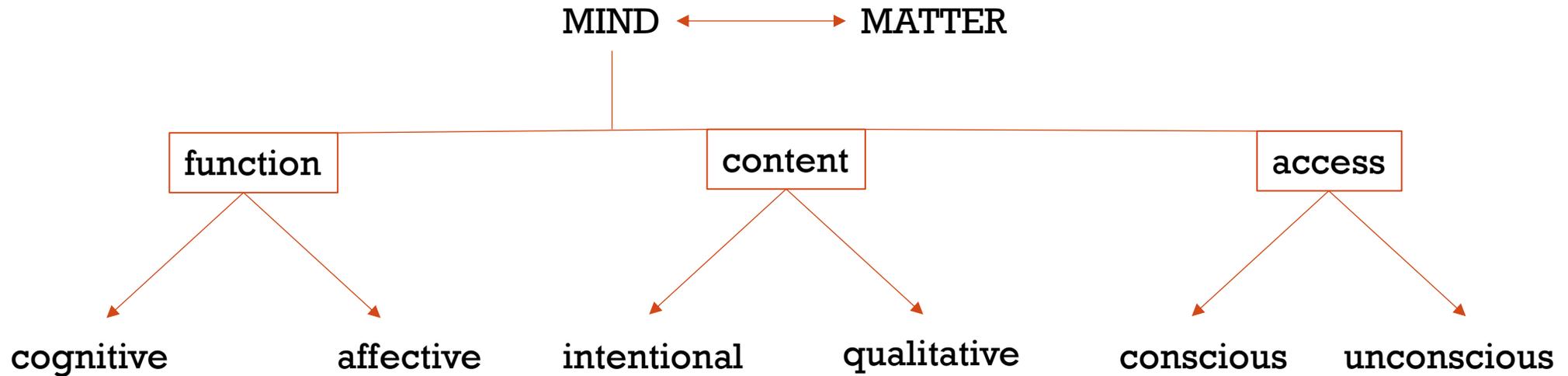
ПРЕДМЕТЫ И МЕТОДЫ ЭМПИРИЧЕСКИХ ИССЛЕДОВАНИЙ СОЗНАНИЯ

Игорь Михайлов
15 июня 2023 г.

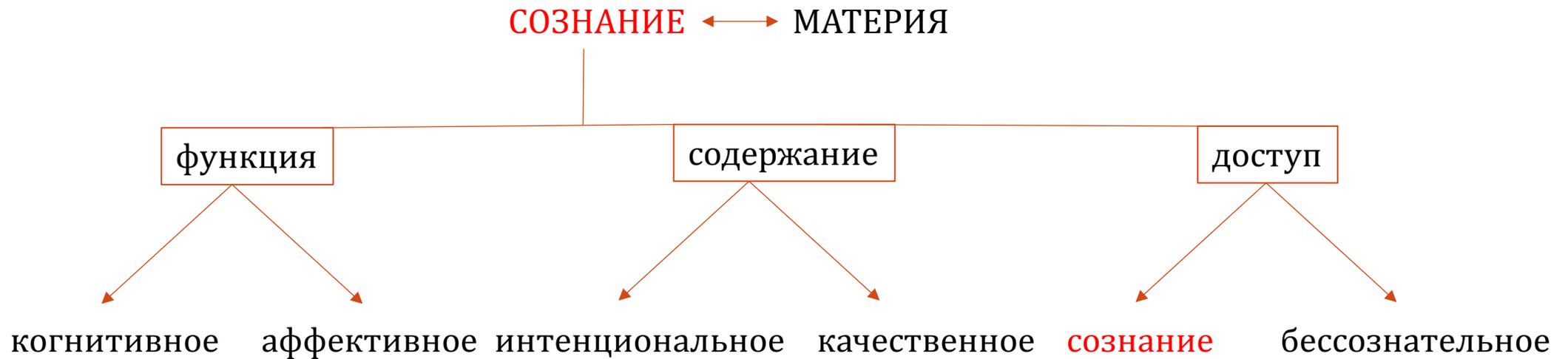
РУДН



МНОГОЗНАЧНОСТЬ И МНОГОСМЫСЛЕННОСТЬ



МНОГОЗНАЧНОСТЬ И МНОГОСМЫСЛЕННОСТЬ



КОНЦЕПТУАЛЬНОЕ И ЭМПИРИЧЕСКОЕ

Концептуальное = онтология сознания

- Может ли «сознание» быть логическим субъектом (=ουσία), или оно только приписывается подлинным сущностям (=κατηγορία)?
 - В последнем случае высказывания типа «сознание конституирует предмет» содержит категориальную ошибку.
- Это единая сущность/атрибут/акциденция, или мы одним словом обозначаем разные явления?
- Достаточно ли для «хорошей» теории сознания онтологий естественных наук?

■ **«Эмпирическое»** = теории сознания (теоретические положения в идеале должны быть фальсифицируемы)

- Каузальная структура сознания
- Необходимые и достаточные условия его проявления
- Идентификация и метрика сознания
- и т. п.



ЕЩЁ ОДИН УРОВЕНЬ

- В пространство между концептуальным и эмпирическим *может* «зайти» ещё один уровень — **вычислительный**. В его ведении вопросы:
 - Каковы формы и элементарные шаги (примитивы) процессов, порождающих сознательные состояния?
 - Как структурировано пространство состояний «сознательной» системы?
 - Описываются ли указанные параметры рациональными (дискретны) или действительными (континуальны) числами?
 - Применяются детерминистские или вероятностные правила преобразования данных?
 - Как кодируются входные данные и как интерпретируются выходные?
 - и т. п.



ЕЩЁ ОДИН УРОВЕНЬ

Вычислительный уровень
объяснения обнаруживает
черты

← концептуальные

Создает некоторую
схематическую
контрфактуальную
«реальность» (Ср.:
схематизм чистого
рассудка у Канта).

→ эмпирические

Вычислительные
модели
принципально
фальсифицируемы
(работает/не работает,
так/не так).



ЕЩЁ ОДИН УРОВЕНЬ

- Если возможно вычислительное описание сознания,
 - задача построения научной теории сознания значительно упрощается,
 - компьютерное моделирование сознательных процессов из имитации превращается в собственно моделирование, обладающее эвристическим потенциалом,
 - полученная абстрактно-вычислительная объяснительная схема может быть экстраполирована на другие предметные области;
 - делается следующий шаг в прогрессе технологий ИИ.

Аналогия:

Электродинамика

Схемотехника

Lazebnik, Y. (2002). Can a biologist fix a radio?—Or, what I learned while studying apoptosis. *Cancer Cell* 2 (3), 179–182.
[https://doi.org/10.1016/S1535-6108\(02\)00133-2](https://doi.org/10.1016/S1535-6108(02)00133-2)



ЕЩЁ ОДИН УРОВЕНЬ

ТЕОРИИ СОЗНАНИЯ

Не-вычислительные

Квантовая теория
сознания,
различные
биологические
теории

Вынужденно
остаются де-факто
метафизическими.

Вычислительные

Рекурсивная теория
сознания, теория
предиктивного
процессинга/активного
вывода

Уже сейчас
предлагают
эмпирически
фальсифицируемые
(=научные)
описания.



КРИТЕРИИ ЭМПИРИЧЕСКИХ ТС

▪ [Doerig & Schurger & Herzog, 2021]

1. Адресуется ли ТС к парадигмальным случаям сознания?

[различение сознательного и бессознательного]

2. Свободна ли ТС от привязки к каузальной структуре на уровне имплементации?

[сети прямого распространения **vs.** рекуррентные сети]

3. Свободна ли ТС от аргумента малых и больших сетей?

[проблема панпсихизма и единства сознания]

4. Устойчива ли ТС к аргументу других систем?

[проблема объяснения схожих феноменов в разных имплементациях]



ЭМПИРИЧЕСКИЕ ТЕОРИИ СОЗНАНИЯ

- **Теория глобального рабочего пространства (Б. Баарс):** бессознательные процессы и психические состояния конкурируют за центр внимания, из которого глобальная информация транслируется по всей системе. Сознание отождествляется с этой глобальной трансляцией и, по Баарсу, является важным средством функциональной и биологической адаптации.
- **Теория интегрированной информации (Дж. Тонони, К. Кох)** предлагает измеримый показатель степени сознания в системе, обозначенный как Φ (фи), который показывает степень физической интеграции информации. По мнению Тонони, этот показатель можно измерить даже в относительно простых физических системах, что, безусловно, приводит к панпсихистским выводам.
- Обе теории можно считать «модельными», поскольку они указывают на возможные нейробиологические реализации.



ЭМПИРИЧЕСКИЕ ТЕОРИИ СОЗНАНИЯ

- Напротив, **теория «мышления высшего порядка»** Дэвида Розенталя чисто концептуальна. Согласно ей, к психическим состояниям низшего порядка относятся ощущения и восприятия. Психические состояния более высокого порядка имеют в качестве объекта любые другие психические состояния. Ментальные состояния низшего порядка становятся сознательными только тогда, когда они становятся объектом ментальных состояний высшего порядка. Эта теория имеет некоторое сходство с различными подходами, утверждающими **рекурсивный характер** индивидуального сознания.



ЭМПИРИЧЕСКИЕ ТЕОРИИ СОЗНАНИЯ

Теория предиктивного процессинга (ПП)

- ПСЭ, генеративные модели, обратная трансляция разницы, активный вывод, байесовские вычисления
- **не предполагает**, что сенсорные стимулы распространяются вверх по кортикальным иерархиям, прежде чем будут модулированы и объединены в сложные репрезентации, которые затем определяют принятие решений и управляют поведенческими реакциями (в отличие от Локка, Фодора и Марра)
- широко понятая байесовская перспектива, подразумеваемая ПП, может помочь понять так называемую метапроблему сознания, которая касается наших суждений и интуитивных представлений о сознании как таковом.



КРИТЕРИИ ЭМПИРИЧЕСКИХ ТС

NDT – Neural Darwinism Theory,
 Orch OR – Orchestrated Objective Reduction,
 IIT – Information Integration Theory,
 RPT – Recurrent Processing Theory,
 GWT – Global Workspace Theory,
 HOTT – Higher Order Thought Theory,
 PPT – Predictive Processing Theory,
 ART – Adaptive Resonance Theory,
 TLT – Thalamicocortical Loop Theory,
 NMDA – NMDA Theory,
 AST – Attention Schema Theory,
 SMT – Sensorimotor Theory,
 SCMT – Self Comes to Mind Theory.

	CLASS I: Identify consciousness with other phenomena		CLASS II: Identify consciousness with causal structures		Class III: Identify consciousness with computational processes				CLASS IVa: Identify consciousness with biological processes		CLASS IVb: Identify consciousness with cognitive processes		
	NDT	Orch OR	IIT	RPT	GWT	HOTT	PPT	ART	TLT	NMDA	AST	SMT	SCMT
CRITERIA													
III.1: Addresses paradigm cases of consciousness													
III.2: Copes with the unfolding argument													
III.3: Copes with the small network argument													
III.4: Copes with the multiple realization argument													
SCOPE													
II.1: Addresses state, content, both or unclear (S, C, B, Unc)	Unc	B	B	B	B	B	B	B	B	S	B	B	B
II.2: Consciousness is graded, binary or unclear (G, B, Unc)	Unc	G	G	B	B	B	Unc	Unc	Unc	G	B	Unc	Unc
II.3: Consciousness is unitary, non-unitary or unclear (U, N, Unc)	U	U	U	Unc	U	Unc	U	Unc	U	Unc	Unc	Unc	Unc
II.4: Consciousness is continuous, discrete or unclear (C, D, Unc)	Unc	D	Unc	Unc	D	Unc	Unc	Unc	D	Unc	Unc	Unc	Unc
II.5: Fate of unconscious elements is clear or unclear (C, Unc)	Unc	C	C	C	C	C	C	C	C	Unc	C	C	C



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

Таблица 1. Отличительные свойства сознания
Table 1. Characteristic features of consciousness

№	Свойство сознания	I	II	III	IV	Прототипическое свойство	Специфика у сознания	Разновидность свойства
1	реальность	✓			✓	каузальность (К)	квалитативность	особый вид К
2	субъективность	✓	✓			индивидуальность (И)	квалитативность	особый вид И
3	дифференцированность		✓		✓	дифференцированность (Д)	квалитативность	особый вид Д
4	целостность	✓	✓	✓	✓	целостность (Ц)	квалитативность	особый вид Ц
5	качественность	✓	✓	✓	✓	специфичность (С)	квалитативность	особый вид С
6	интенциональность	✓	✓	✓		специфичность (С)	квалитативность	особый вид С
7	разнообразность		✓		✓	разнообразность (Р)	квалитативность	особый вид Р
8	определенность		✓		✓	определенность (О)	квалитативность	особый вид О
9	ситуативность	✓	✓			вписанность (В)	квалитативность	особый вид В
10	временная структура		✓		✓	временная структура (ВС)	квалитативность	особый вид ВС



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

О «трудной проблеме»

«Это, безусловно, не снимает вопроса о специфике конкретных механизмов интеграции, дифференциации и других свойств в сознательной системе. Это также не означает, что все разнообразие субъективных феноменов, входящих в сложную структуру сознания, должно сводиться исключительно к квалиа. Однако отмеченное обстоятельство позволяет сфокусировать теорию на задаче объяснения, в первую очередь, именно основ качественности субъективного опыта. Если теория сможет сделать это в естественно-научных понятиях, то есть надежда, что и другие специфические характеристики сознания получат свое закономерное объяснение» (с. 44).



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

4(5)-компонентная структура (исследований) сознания

«• Любой эпизод сознания развивается в структуре когнитивного агента (КТО).

- В какой-то момент времени часть когнитивных элементов этой системы активируется до надпорогового уровня (КОГДА).
- Их совместная активность создает уникальную по содержанию интеграцию – состояние сознания (ЧТО).
- Такая группа активных элементов обретает каузальный доступ в глобальное пространство других когнитивных элементов системы (ГДЕ).
- Результатом является изменение целостного состояния когнитивного агента (КТО).» (с. 47).



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

4(5)-компонентная структура (исследований) сознания

«Искомая фундаментальная теория обязана объяснить нейронаучную конкретику “кто-что-где-когда”, составляющих этот универсальный сознательный эпизод» (с. 48).



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

Циркулярная ловушка ТС:

- невозможно по-настоящему понять сущность субъективного опыта (“что”), не понимая устройство его носителя – разума (“кто”);
- невозможно по-настоящему понять устройство разума, не понимая законы его формирования в процессах обучения; (с. 54);
- невозможно по-настоящему понять процессы обучения, не понимая принципов развития нервной системы в онтогенезе;
- невозможно по-настоящему понять принципы развития нервной системы в онтогенезе, не понимая закономерности ее эволюции в филогенезе;
- невозможно по-настоящему понять закономерности эволюции нервной системы в филогенезе, не понимая роль в них нервных механизмов поведения и субъективного опыта (“что”) (с. 52).



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

Область компетенции ТС:

М1 – ментальные явления, наблюдаемые от первого лица, т.е. субъективно;

М3 – ментальные явления, наблюдаемые от третьего лица, т.е. объективно;

Н3 – нервные явления, наблюдаемые от третьего лица, т.е. объективно. (с. 52).



КРИТЕРИИ ПО К. В. АНОХИНУ

Анохин, К. В. (2021). Когнитом: в поисках фундаментальной нейронаучной теории сознания. *Журнал Высшей Нервной Деятельности Им И П Павлова*, 71(1), 39–71. <https://doi.org/10.31857/S0044467721010032>

Критерии:

1. Объяснить сознание означает создать научную теорию сознания.
2. Эта теория должна отвечать на пять вопросов “как” и “почему” в отношении отличительных свойств и ключевых ингредиентов субъективного опыта.
3. Ответы на эти вопросы должны быть совместимы с требованием выхода из “циркулярной ловушки” сознания.
4. Фундаментальная теория сознания должна охватывать своим объяснением максимальное количество феноменов М1, М3 и N3, исходя при этом из минимального числа первых принципов.
5. Первые принципы фундаментальной теории сознания должны основываться на понятиях биологического уровня, из которых должно закономерно выводиться возникновение свойств когнитивного уровня (с. 54).



ПРОБЛЕМА МНОЖЕСТВЕННЫХ EXPLANANDA

- Crick, F., & Koch, C. J. (1990). Towards a neurobiological theory of consciousness. *Seminars in The Neurosciences 2*, 263–275.
- «проблема сознания может в конечном счете быть решена только с помощью объяснений на нейронном уровне»;
- понятие нейронных коррелятов сознания (НКС).
- Chalmers, David J. 2000. What is a neural correlate of consciousness? In: *Neural correlates of consciousness: Empirical and conceptual questions*. The MIT Press, 17–39.
- НКС был определен как минимум нейронных механизмов, совместно необходимых для того, чтобы дать возможность любому конкретному сознательному опыту.

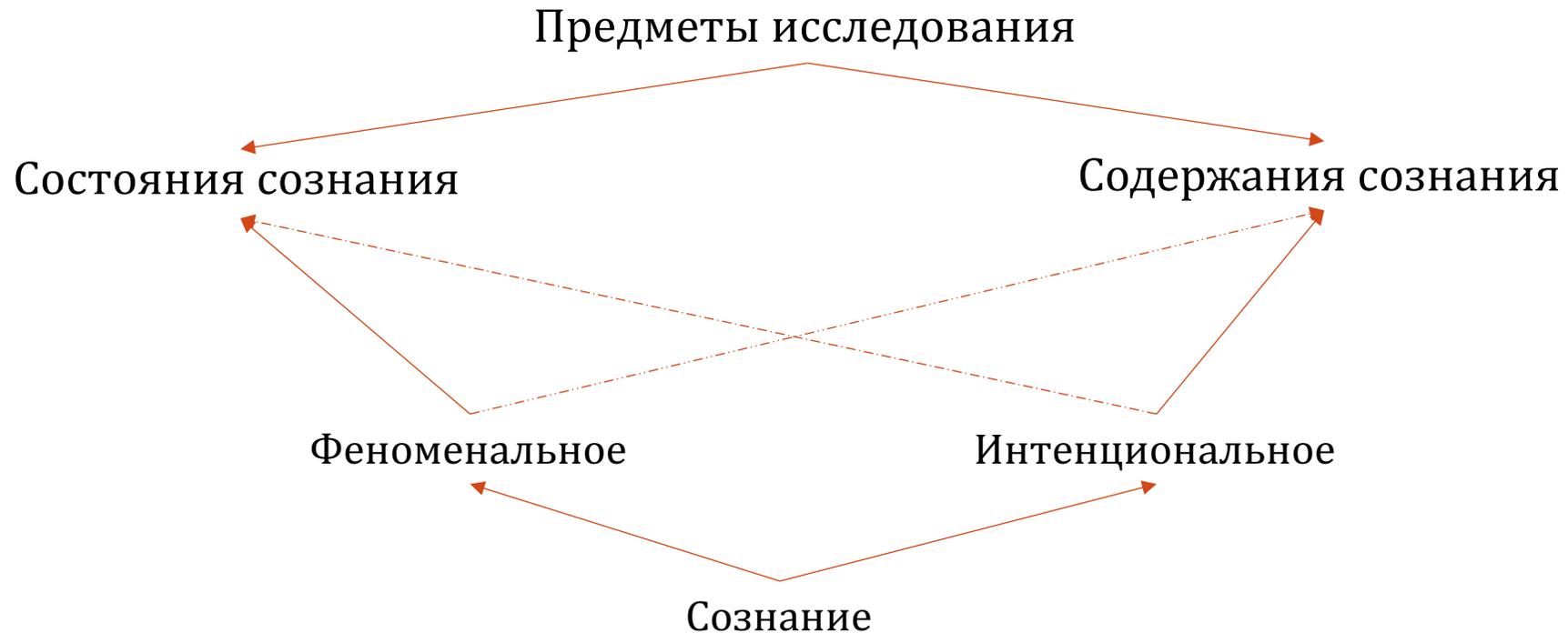


ПРОБЛЕМА МНОЖЕСТВЕННЫХ EXPLANANDA

- Vilas, M. G., Auksztulewicz, R., & Melloni, L. (2021). Active Inference as a Computational Framework for Consciousness. *Review of Philosophy and Psychology*.
<https://doi.org/10.1007/s13164-021-00579-w>
- “explanandum problem”
- отсутствие единого объяснения, т. е. того, что должна объяснять наука о сознании, и
- отсутствие консенсуса относительно того, как сознание должно быть операционализировано и измерено.
- Wiese, W. (2018). Toward a mature science of consciousness. *Frontiers in Psychology*, 9(MAY).
<https://doi.org/10.3389/FPSYG.2018.00693/FULL>.
- Сознание было описано как «набор характеристик», каждая из которых подвергается анализу с помощью различных исследовательских программ, которые сами внутренне обсуждают таксономию и экспериментальные парадигмы, предназначенные к использованию.



ПРОБЛЕМА МНОЖЕСТВЕННЫХ EXPLANANDA



ПРОБЛЕМА МНОЖЕСТВЕННЫХ EXPLANANDA

Предметы исследования



Состояния сознания

Например, исследования контрастных паттернов мозговой активности, измеренных с помощью функциональной магнитно-резонансной томографии (фМРТ), магнито/электроэнцефалографии (М/ЭЭГ), электрокортикографии (ЭКоГ), между бодрствованием и сном, между сновидения против сна без сновидений, или между бодрствованием и анестезией.

Содержания сознания

Феноменальное содержание: (1) универсальные качества опыта, (2) особенные качества конкретного опыта. Качественные исследования, отчёты от первого лица. Нейрофеноменология (Варела), ТИИ (Тонони). Пример более тонкого структурирования ФС: Haun, A., & Tononi, G. (2019). Why does space feel the way it does? Towards a principled account of spatial experience. *Entropy*, 21(12), 1160. <https://doi.org/10.3390/e21121160>



ПРОБЛЕМА МНОЖЕСТВЕННЫХ EXPLANANDA

Предметы исследования



Состояния сознания

Например, исследования контрастных паттернов мозговой активности, измеренных с помощью функциональной магнитно-резонансной томографии (фМРТ), магнито/электроэнцефалографии (М/ЭЭГ), электрокортикографии (ЭКоГ), между бодрствованием и сном, между сновидения против сна без сновидений, или между бодрствованием и анестезией.

Содержания сознания

Содержание доступа: (1) процессы, которые позволяют делать конкретную информацию доступной для использования когнитивными процессами более высокого уровня, (2) мета-когниции и мета-осознание. Методы исследования: (1) бинокулярное соперничество, (2) количественные параметры стимула (напр., длительность). Сканируются нейронные возбуждения, соответствующие разным отчётам.



INTEGRATED INFORMATION THEORY

- Haun, A., & Tononi, G. (2019). Why does space feel the way it does? Towards a principled account of spatial experience. *Entropy*, 21(12), 1160. <https://doi.org/10.3390/e21121160>
- Лучший кандидат на роль НК феноменального пространства — решётчатая структура нейронов задней коры.
- Система в некотором состоянии должна задавать максимально нередуцируемую, специфическую, композиционную, внутреннюю причинно-следственную структуру, которая состоит из различий и их отношений.



«АКТИВНЫЙ ВЫВОД» КАК ОБЩАЯ РАБОЧАЯ РАМКА

1. Принцип (минимизации) **свободной энергии** (Фристон, Фрит, Хохви, Кларк и мн. др.)

ПСЭ: все адаптивные биологические агенты стремятся свести к минимуму долговременную неожиданность (т.е. энтропию).

2. **Предиктивный процессинг** — набор гипотез, предполагающих, что мозг стремится уменьшить удивление, делая выводы о (скрытых) состояниях мира, порождающих наш сенсорный опыт, с использованием байесовских механизмов вывода.

Генеративная модель (ГМ) определяет вероятность перцептивных данных, априори предполагая «скрытые состояния» («причины»). Нейронная система использует ГМ для вычисления апостериорной вероятности скрытых состояний. Эта порождающая модель считается иерархической, где каждый уровень иерархии кодирует состояния во вложенных временных масштабах, и каждый уровень принимает в качестве наблюдений скрытые состояния уровней ниже.



«АКТИВНЫЙ ВЫВОД» КАК ОБЩАЯ РАБОЧАЯ РАМКА

Предсказанные данные сравниваются с полученными, и «ошибка предсказания» (формально эквивалентная свободной энергии) передается вверх по иерархии.

3. Активный вывод

Цель нейронной системы состоит в том, чтобы свести к минимуму эти ошибки предсказания на иерархических уровнях. Этого можно достичь двумя способами:

- делая вывод о (скрытых) состояниях мира, которые максимизируют вероятность перцептивных данных;
- или путём *активного вывода* — смены выборки, т. е. действий в мире, чтобы увеличить вероятность исполнения этих прогнозов.

Модели, предполагающие это последнее свойство, — модели активного вывода (AIM).



«АКТИВНЫЙ ВЫВОД» КАК ОБЩАЯ РАБОЧАЯ РАМКА

4. Точность предсказаний

Кроме того, организм оценивает точность своих убеждений. Эти так называемые *оценки точности* могут взвешивать влияние ошибок предсказания (функция усиления), и когда они развертываются нисходящим (сверху вниз) образом в качестве ожиданий, они рассматриваются как механизмы внимания.

В совокупности эти вычислительные принципы составляют архитектуру активного вывода, которую можно *использовать для построения теорий процессов конкретных когнитивных явлений*.



ИЕРАРХИЧЕСКИЙ БАЙЕСОВСКИЙ ВЫВОД

Теорема Байеса

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

Генеративная
модель

где

$P(H|E)$ — апостериорная вероятность H ,
 $P(H)$ — априорная вероятность H
 $P(E|H)$ — вероятность E при условии H
 $P(E)$ — априорная вероятность E

$$P(E|C) = \frac{P(C|E)P(E)}{P(C)}$$

$$P(C|N) = \frac{P(N|C)P(C)}{P(N)}$$



АІМ КАК ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ СОЗНАНИЯ

Сознание



выводы о состояниях мира, которые еще не произошли: **«утолщение времени»** (Мерло-Понти)

выводы о последствиях альтернативных действий: **контрфактуальная глубина**

Утолщение времени (temporal thickness) и **контрфактуальная глубина** — это градуированные характеристики генеративных моделей, которые позволяют делать выводы о состояниях, находящихся дальше во времени, и сравнивать большее количество «политик». Поскольку предполагается, что эти два свойства лежат в основе явлений сознания, сторонники активного вывода утверждают, что сознание, следовательно, должно быть градуированным явлением.



АИМ КАК ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ СОЗНАНИЯ

Сознание

← выводы о состояниях мира, которые еще не произошли: «**утолщение времени**» (Мерло-Понти)

→ выводы о последствиях альтернативных действий: **контрфактуальная глубина**

- Friston, K. (2018). Am I Self-Conscious? (Or Does Self-Organization Entail Self-Consciousness?). *Frontiers in Psychology*, 9(APR). <https://doi.org/10.3389/fpsyg.2018.00579>



АІМ КАК ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ СОЗНАНИЯ

Сознание

←
Содержание «доступа»

Некоторые предполагают, что содержание сознания в особенности представляет собой «убеждения» среднего уровня этой апостериорной иерархии.

→
Феноменальное содержание

Некоторые феноменологические свойства можно объяснить и иерархической природой генеративной модели.

- «(...) видение красного цвета и ощущение боли (...) сами по себе являются предполагаемыми причинами, *сконструированными* (курсив мой — И. М.), чтобы освоить (то есть наилучшим образом объяснить) сырой сенсорный поток — и иерархические махинации, которые они вызывают» (Clark, A., Friston, K., & Wilkinson, S. (2019). *Bayesian qualia: Consciousness as inference, not raw datum. Journal of Consciousness Studies*, 26(9–10), 19–33. <https://philpapers.org/rec/CLABQC>)



АИМ КАК ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ СОЗНАНИЯ

Сознание

←
Содержание «доступа»

Некоторые предполагают, что содержание сознания в особенности представляет собой «убеждения» среднего уровня этой апостериорной иерархии.

→
Феноменальное содержание

Некоторые феноменологические свойства можно объяснить и иерархической природой генеративной модели.

- «Квалиа — точно так же, как собаки и кошки — являются частью предполагаемого набора скрытых причин (то есть эмпирических гипотез), которые лучше всего объясняют и предсказывают развивающийся поток энергии через наши сенсорные поверхности» (Clark, A., Friston, K., & Wilkinson, S. (2019). *Bayesian qualia: Consciousness as inference, not raw datum. Journal of Consciousness Studies, 26(9–10), 19–33.* <https://philpapers.org/rec/CLABQC>)



АІМ КАК ВЫЧИСЛИТЕЛЬНАЯ МОДЕЛЬ СОЗНАНИЯ

Сознание

←
Содержание «доступа»

Управление «доступом»

Точно так же иерархическое развертывание точности может объяснить, почему люди способны обращать внимание на свои собственные представления о состояниях мира и манипулировать ими (т. е. **метакогнитивными способностями**). Как упоминалось ранее, взвешивание точности сверху вниз часто отождествляют с процессами внимания, когда состояния более высокого уровня модулируют надёжность и важность, приписываемые различным состояниям восприятия. Состояния **метаосознания**, когда некто осознает и сознательно манипулирует этими внутренними процессами внимания, можно рассматривать как развертывание точности на тех уровнях иерархии, которые сами задействуют механизмы модуляции внимания.



МЕХАНИСТИЧЕСКИЕ И ВЫЧИСЛИТЕЛЬНЫЕ МОДЕЛИ

- Механистическая модель:
 - точное определение *explanandum*
 - фальсифицируемое описание механизма, лежащего в основе
- Вычислительная модель
 - формализация вербальной теории средствами компьютерного кода

Сознание доступа



Модели, симулирующие
параметры известных
нейронных активаций,
сопровождающие известные
сознательные состояния.

Феноменальное сознание



Трудная задача создания
вычислительной модели
явлений, наблюдаемых
только из перспективы
первого лица.



МЕХАНИСТИЧЕСКИЕ И ВЫЧИСЛИТЕЛЬНЫЕ МОДЕЛИ

До сих пор исследования с применением АИМ имитировали:

- бинокулярное соперничество,
- маскировку,
- слепоту невнимания,
- иллюзию затухания Трокслера
- парадигмы редкости (**oddball paradigms**),
- а также парадигму рабочей памяти, использующую механизмы сознательного доступа,
- и парадигму, вызывающую галлюцинации.

Во всех исследованиях использовались относительно простые стимулы с низкой степенью экологической валидности.



МЕХАНИСТИЧЕСКИЕ И ВЫЧИСЛИТЕЛЬНЫЕ МОДЕЛИ

Особенности AIM

- использование аналогично параметризованной генеративной модели для описания эволюции скрытых состояний с активными механизмами вывода;
- концептуальная интеграция результатов различных экспериментальных задач, исследующих один и тот же аспект сознания;
- серьезный кандидат на создание теории нейронных процессов сознания;
- нужно расширить эти модели, чтобы объяснить измененные состояния сознания и структуру феноменологического опыта;
- моделировать другие нейронные маркеры, связанные с сознанием.



ОБЪЯСНИТЕЛЬНЫЕ ГИПОТЕЗЫ

Роль оценок точности и их иерархического развертывания

- содержание определяется гипотезой с наибольшей апостериорной вероятностью, на которую, в свою очередь, влияет точность, приписываемая этим убеждениям;
- увеличение неопределенности отображения между наблюдениями и скрытыми состояниями мира влияет на то, воспринимается ли содержание сознательно.



ОБЪЯСНИТЕЛЬНЫЕ ГИПОТЕЗЫ

Benrimoh, D. A., Parr, T., Vincent, P., Adams, R. A., & Friston, K. (2018). Active Inference and Auditory Hallucinations. *Computational Psychiatry*, 2(0), 183. https://doi.org/10.1162/cpsy_a_00022)

Гипотеза: активный вывод зависит от тонкого баланса между предыдущими **убеждениями (beliefs)** о ненаблюдаемых (скрытых) переменных и ощущениями, которые они вызывают.

Природа галлюцинаций: предшествующие убеждения доминируют над выводом о восприятии.

Результаты:

- были предложены вычислительные механизмы, которые могут вызвать этот дисбаланс в восприятии;
- с помощью симуляции было показано, что содержание предшествующих убеждений (и уверенность в них) зависит от убеждений о **политике (policies)** (здесь последовательности слушания и разговора)
- и от убеждений о надежности сенсорных данных.



ОБЪЯСНИТЕЛЬНЫЕ ГИПОТЕЗЫ

Whyte, C. J., Hohwy, J., & Smith, R. (2022). An active inference model of conscious access: How cognitive action selection reconciles the results of report and no-report paradigms. *Current Research in Neurobiology*, 3, 100036.
<https://doi.org/10.1016/j.crneur.2022.100036>

В этом интересном исследовании было показано, что

- предыдущие исследования, основанные на «когнитивных теориях» (HOT, GWS), которые приписывают ПФК роль исключительного НКС, не уделили внимание ситуации, когда от испытуемых не требовался отчёт
- в ситуации «без отчёта» активность ПФК снижается, тогда как
- «навязывание требований отчёта требует от агента целенаправленного повышения точности сообщений, передаваемых между сенсорной и префронтальной корой (а также внутри префронтальной коры) – тем самым изменяя нейронные корреляты сознания доступа» (р. 10).



ОБЪЯСНИТЕЛЬНЫЕ ГИПОТЕЗЫ

Whyte, C. J., Hohwy, J., & Smith, R. (2022). An active inference model of conscious access: How cognitive action selection reconciles the results of report and no-report paradigms. *Current Research in Neurobiology*, 3, 100036.
<https://doi.org/10.1016/j.crneur.2022.100036>

В этом интересном исследовании было показано, что

- подобно **НОТ**, **AIM** предполагает наличие более высокого, чем перцептивный, уровня репрезентации
- в отличие от **НОТ**, в **AIM** второй уровень делает выводы о временной последовательности перцепций на первом уровне = «**утолщение времени**» (ср. Кант о роли времени в схематизме чистых категорий)
- это, а также обнаружение независимых факторов состояний на втором уровне, сближает **AIM** и **GWS**.



ОБЪЯСНИТЕЛЬНЫЕ ГИПОТЕЗЫ

Whyte, C. J., Hohwy, J., & Smith, R. (2022). An active inference model of conscious access: How cognitive action selection reconciles the results of report and no-report paradigms. *Current Research in Neurobiology*, 3, 100036.
<https://doi.org/10.1016/j.crneur.2022.100036>

В этом интересном исследовании авторы заявляют, что интерпретируют «настоящую модель активного вывода как теорию *для* науки о сознании, а не как теорию *сознания*. То есть, вместо построения теории сознания *как таковой*, наш подход апеллирует к активному выводу как к общей теории функционирования мозга, которую можно использовать для построения генеративных моделей репрезентативных экспериментальных задач в исследованиях сознания» (р. 11).



ОТ СЕБЯ ДОБАВЛЮ

- «Утолщение времени», конечно, вряд ли решает трудную проблему сознания (каково это видеть красный помидор) на данном этапе, но относительную устойчивость квалиа вполне объясняет.
- Мы видим помидор постоянно красным, хотя очевидно, что частоты отражаемых его поверхностью э.-м. излучений непостоянны.
- Наличие «утолщения времени» — более высоких уровней иерархии, которые делают вероятностные выводы относительно временной последовательности предсказанных восприятий, — весьма правдоподобная гипотеза.
- Таким образом мать-природа экономит наши вычислительные и энергетические мощности, поставляя только необходимую для выживания информацию.



ОТ СЕБЯ ДОБАВЛЮ

- Правда, из этого следует, что качественные восприятия могут быть только сознательными.



ОТ СЕБЯ ДОБАВЛЮ

- А это говорит в пользу онтологического единства феноменального сознания и сознания доступа.



ГИПОТЕЗА 4Е

4Е = embodied, embedded, enactive, extended (mind).

- разум (**mind**) не находится исключительно в мозге, но распространяется на тело и даже физический мир;
- части тела и объекты внешней среды могут реализовывать когнитивные процессы и таким образом функционировать как расширения самого разума (**mind**);
- суть и цель разума (**mind**) — в обслуживании действий;
- разум (**mind**) охватывает все уровни познания, включая физический уровень;
- Ф. Варела, Э. Томпсон и Э. Рош «Воплощённый разум: когнитивная наука и человеческий опыт» (1991);
- Энди Кларк и Дэвид Чалмерс «Расширенный разум» (1998).

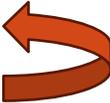


ГИПОТЕЗА 4Е

Методологический вопрос:

Является ли тело и окружающая среда *необходимыми* (с каузальной т. зр.) условиями и элементами когнитивной системы?

Предлагаю комплексный мысленный эксперимент.

Следующий слайд 

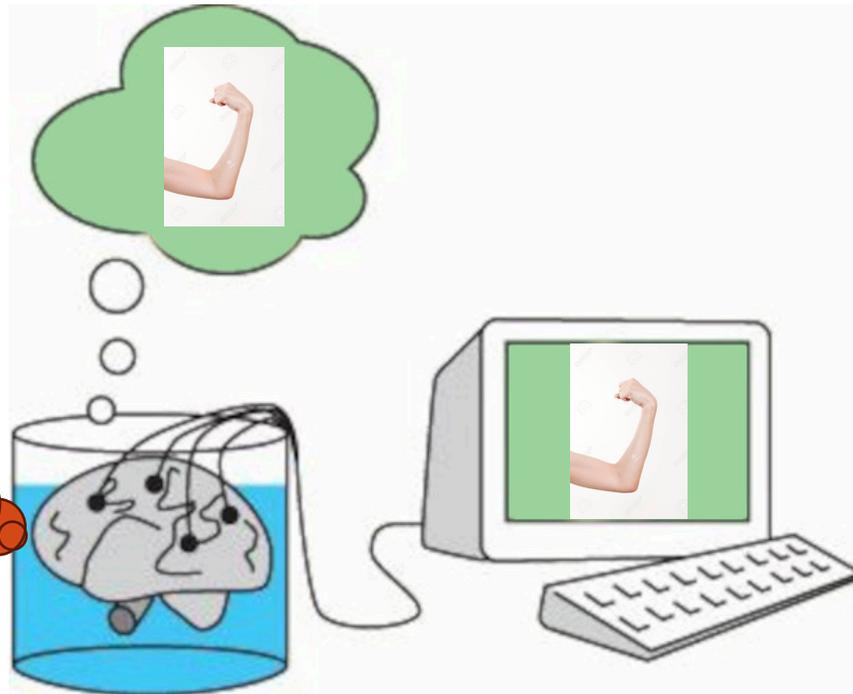


МОЗГ В БОЧКЕ...



МОЗГ В БОЧКЕ, МУР...

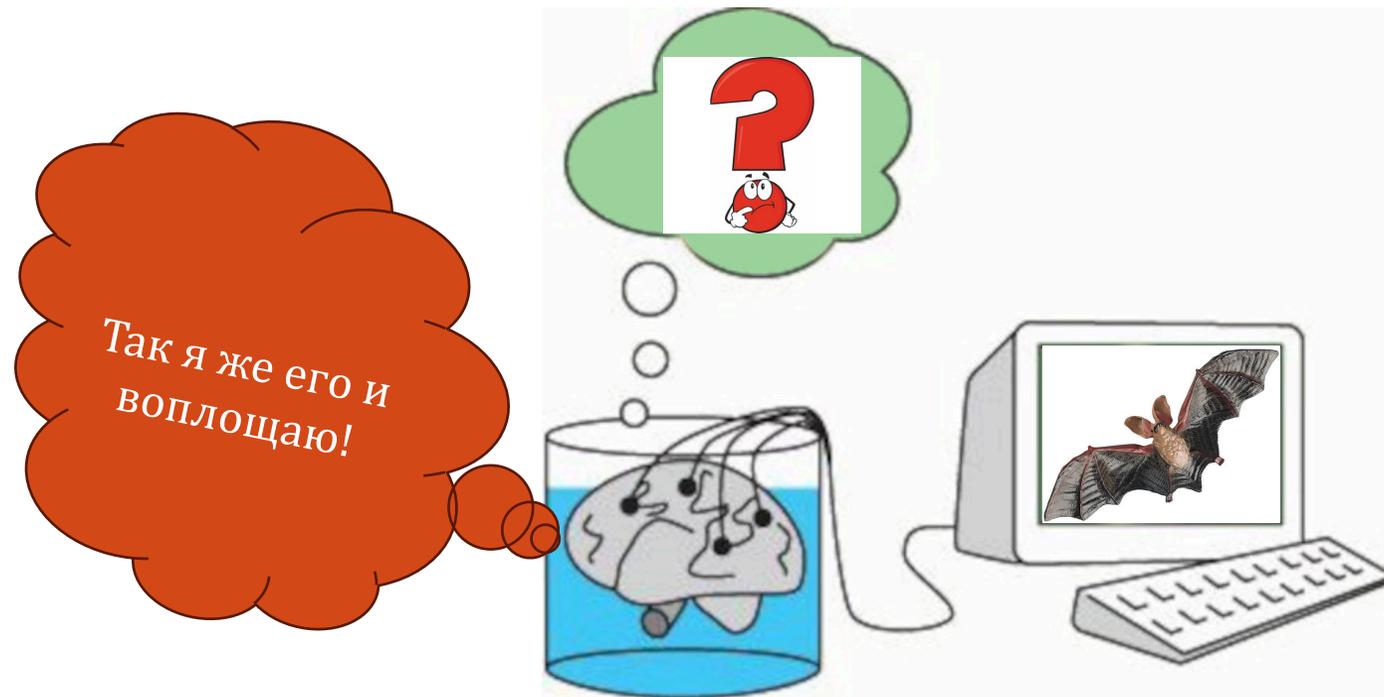
А где моё
воплощённое
познание?



Мур Дж. Э. Опровержение идеализма // Историко-философский ежегодник. М.: Наука, 1987. С. 242–265



МОЗГ В БОЧКЕ, МУР И НАГЕЛЬ



Nagel T. What is it Like to Be a Bat? // The Philosophical Review. – 1974. – Vol. 83, № 4. – P. 435–450.

- Механизмы репрезентации видоспецифичны.
- Они развиваются эволюционно...
- ...в корреляции с телесной организацией.



ФИНАЛЬНЫЕ СООБРАЖЕНИЯ

- Метод прямого сопоставления феноменальных (интроспективных) представлений с нейрофизиологическими данными пока успеха не принёс.
- Причина — «починка радио биологом»: слепой поиск корреляций между нечёткими и неverified концептами и неструктурированным потоком эмпирических данных.
- Недостающее звено — «инженерный уровень»: функциональные схемы, составленные из конечного множества типовых элементов и интерпретируемые как в терминах теоретических положений («законов»), так и в терминах эмпирических описаний (помним о Канте).
- Кандидат на это место, достойный рассмотрения, — вычислительные алгоритмы (как по-оккамовски простые описания сложных мультистабильных саморегулирующихся систем).
- Учитывая общепризнанную теперь необходимость учёта обучения и эволюции, эти алгоритмы скорее всего должны быть вероятностными.



НА ЭТОМ ВСЁ.

Спасибо за внимание!

ifmikhailov@gmail.com

