Теория МВАП в сравнении с теорией предиктивного кодирования

Теория предиктивного кодирования (и шире — предиктивного процессинга) занимает одну из ключевых позиций среди теорий сознания**.** Предиктивное кодирование — это международная, академическая «теория всего» в нейронауке, поддержанная сотнями институтов по всему миру. Сравнение с МВАП.

Аннотация

В статье проводится сравнительный анализ двух теоретических подходов к пониманию механизмов работы психики и сознания: доминирующей в мировой нейронауке теории предиктивного кодирования (Predictive Coding / Free Energy Principle, К. Фристон и др.) и альтернативной инженерно-схемотехнической модели — модели волевой адаптивности психики (МВАП, В.В. Парусников, Д.Л. Петрийчук и др.). Рассмотрены исторические истоки каждой концепции, их ключевые понятия (ошибка предсказания, активный вывод, свободная энергия — в первом случае; доминанта, новизна, значимость, произвольность, сознание как «аварийный тумблер» — во втором). Показано, что предиктивное кодирование представляет собой универсальную метатеорию с мощным математическим аппаратом (вариационный Байес, принцип минимизации свободной энергии), поддержанную сотнями эмпирических исследований, но критикуемую за нефальсифицируемость и чрезмерную абстрактность. МВАП, напротив, предлагает конкретную схемотехнику психики, верифицированную действующим программным прототипом (проект Beast), и даёт чёткий функциональный ответ на вопрос о роли сознания и воли, однако остаётся малоизвестной за пределами российской научной школы и уступает в масштабе эмпирической валидации. Выделены принципиальные различия: в понимании первичности прогноза versus актуальной значимости стимула, роли сознания (побочный продукт вычислений versus аварийный механизм переучивания), природы воли (активный вывод как устранение ошибки предсказания versus произвольность как торможение автоматизма). Сделан вывод, что теории не являются взаимодополняющими в силу концептуальной несовместимости базовых принципов, однако их сравнение выявляет важные инженерные и теоретические ориентиры для будущих исследований сознания и создания сильного объяснимого искусственного интеллекта.

Ключевые слова: предиктивное кодирование, предиктивный процессинг, свободная энергия, активный вывод, МВАП, модель волевой адаптивности

психики, сознание, воля, ошибка предсказания, доминанта, новизна, объяснимый ИИ, Beast-прототип.

Работа Фристона детально прокомментирована на сайте МВАП: fornit.ru/70999.

Суть теории предиктивного кодирования

Суть теории предиктивного кодирования заключается в том, что наш мозг — это не пассивный приемник информации, а активная «машина предсказаний», которая постоянно конструирует модель окружающего мира сверху вниз (top-down) и лишь сверяет ее с сигналами от органов чувств. Мы видим, слышим и осязаем не саму реальность, а собственные ожидания о ней, скорректированные на величину ошибки прогноза.

Название происходит от английского термина Predictive Coding.

Предиктивное (от англ. *predictive* — предсказывающее) — мозг всегда работает на опережение. Он непрерывно создает внутренние гипотезы о том, что произойдет в следующий момент времени (какой звук раздастся, какое световое пятно упадет на сетчатку).

Кодирование (от англ. *coding*) — термин заимствован из теории информации и телекоммуникаций. Вместо того чтобы передавать «тяжелый» исходный сигнал целиком, система кодирует и передает по нейронным сетям только ошибку предсказания (разницу между тем, что ожидалось, и тем, что произошло). Это критически экономит вычислительные ресурсы мозга.

Вся работа психики строится на иерархическом взаимодействии двух потоков:

- Поток сверху вниз (Предсказания): Высшие отделы мозга на основе прошлого опыта генерируют гипотезу: *«Я на кухне, значит, цилиндрический предмет на столе — это кружка»*.
- Поток снизу вверх (Ошибки прогноза): Глаза видят объект. Если это кружка, то предсказание совпало с сигналом, ошибка равна нулю, и мозг тратит минимум энергии. Но если предмет оказался кошкой, возникает ошибка предсказания (prediction error).

При возникновении ошибки у мозга есть два пути:

1. Обновить модель: Перестроить гипотезу под реальность (*«Ой, это кошка, а не кружка»*).

2. Активный вывод (Active Inference): Заставить тело совершить действие, чтобы подогнать реальность под модель (например, повернуть голову или присмотреться ближе, чтобы убрать искажение фона).

История. Концепция развивалась на стыке философии, физиологии и цифровых технологий.

- XIX век (Философский фундамент): Немецкий физик и физиолог Герман фон Гельмгольц в 1860-х годах выдвинул идею *«бессознательных умозаключений»*. Он заметил, что зрительное восприятие — это процесс логического вывода, где мозг сам достраивает картинку на основе вероятностей [1].
- Середина XX века (Отечественный след): Советский физиолог Николай Бернштейн разработал концепцию *«модели потребного будущего»*. Он доказал, что любое движение человека строится не как реакция на стимул, а как опережающее моделирование результата [2].
- 1950–1970-е годы (Технический контекст): Понятие *predictive coding* официально родилось в инженерии. Его использовали для сжатия видео- и аудиосигналов (например, в телефонии). Зачем кодировать каждый кадр фильма, если можно передавать только изменения между кадрами? Этот принцип инженеры подсмотрели у природы.
- 1999 год (Прорыв в нейронауке): Ученые Раджеш Рао и Дана Баллард опубликовали фундаментальную работу, где математически доказали, что зрительная кора человека устроена именно по принципу предиктивного кодирования [3].
- XXI век (Единая теория): Британский нейробиолог Карл Фристон расширил эту модель до глобального *«Принципа свободной энергии»*. Сегодня предиктивный процессинг считается главной метатеорией в когнитивных науках, объясняющей всё — от базового зрения до природы шизофрении, аутизма и работы искусственного интеллекта [4].

Карл Фристон (Karl Friston) сегодня безусловно является самым авторитетным, цитируемым и влиятельным ученым, развивающим эту теорию. В научном сообществе его фигура уникальна: он входит в число самых цитируемых нейробиологов мира (его h-index превышает 290, а общее число цитирований перешагнуло за 400 тысяч).

Однако его роль в этой теории имеет важные нюансы, которые разделяют ученых на два лагеря.

Почему Фристон — главный?

1. Он объединил всё в суперорганизмическую метатеорию. До Фристона предиктивное кодирование было скорее локальной моделью того, как работают зрение или слух (после работ Рао и Балларда). Фристон математически «поженил» нейробиологию с физикой и термодинамикой, создав Принцип свободной энергии (Free Energy Principle).
2. Он создал концепцию активного вывода (Active Inference). Фристон объяснил, что мозг не просто пассивно меняет свои галлюцинации под воздействием реальности. Он заставляет тело действовать (двигать глазами, перемещаться), чтобы активно менять входящие сигналы и подгонять реальность под свои ожидания [7].
3. Его авторитет непререкаем. Еще до предиктивного кодирования Фристон совершил революцию в нейровизуализации — он создал метод SPM (статистическое параметрическое картирование), с помощью которого сегодня анализируется более 90% всех снимков МРТ в мире. Поэтому, когда ученый такого масштаба занялся теорией предсказаний, к ней отнеслись максимально серьезно.

Особенности восприятия Фристона в научном мире

Несмотря на статус «рок-звезды» от нейронауки, вокруг его работ ведутся жаркие споры:

- «Фристоновский язык» невероятно сложен. Математический аппарат Фристона (опирающийся на вариационное исчисление Байеса, физику случайных динамических систем и информационные барьеры — *«одеяла Маркова»*) настолько труден для понимания, что в научном мире существует шутка: *«В мире есть всего 5 человек, которые по-настоящему понимают Фристона, и один из них — сам Фристон»*.
- Чрезмерная глобальность. Фристон утверждает, что его принципом минимизации свободной энергии можно объяснить вообще всё: от поведения одной живой клетки и колонии муравьев до эволюции видов, сознания человека и устройства Вселенной. Противники теории критикуют ее за «нефальсифицируемость» (если теория объясняет абсолютно всё, ее невозможно опровергнуть экспериментально).

Кто еще развивает эту теорию (популяризаторы)

Поскольку статьи Фристона читать тяжело, у предиктивного кодирования появились другие ключевые лица, которые перевели эти идеи на доступный язык и сделали их популярными:

- Энди Кларк (Andy Clark) — британский философ и когнитивист. Именно он ввел в широкий оборот термин Predictive Processing (Предиктивный процессинг) как зонтичный бренд для этих идей. Его книга *«Surfing Uncertainty»* («Серфинг на волнах неопределенности») стала главным доступным учебником по этой теме [5].
- Анил Сет (Anil Seth) — известный британский нейробиолог. Автор бестселлера *«Быть собой»* (*Being You*). Он великолепно объясняет идеи Фристона через призму сознания, называя наше восприятие «контролируемой галлюцинацией» [6].

Фристон создал сложнейший математический и концептуальный фундамент предиктивного кодирования. Без него эта концепция осталась бы частным методом компьютерного зрения и обработки сигналов.

Что же описывает теория Фристона?

Фристон никогда не отказывался от того, что его теория описывает природный мозг. Изначально он создавал её именно как биологическую метатеорию. Однако сегодня его фокус внимания сильно сместился в сторону инженерии, создания нового типа ИИ и системного дизайна.

Фристон разделяет эти понятия и почему сейчас он больше говорит об ИИ:

1. Теория шире, чем просто мозг

Фристон подчеркивает, что его Принцип свободной энергии (математический фундамент предиктивного кодирования) — это не закон работы человеческого мозга, а закон физики для любых живых или самоорганизующихся систем.
С точки зрения этой теории, по принципу предиктивного кодирования живет и мыслит даже одиночная бактерия *E. Coli*, иммунная система или капля масла в специальном растворе. Мозг — это лишь один из эволюционных вариантов реализации этого принципа.

2. Мозг человека слишком «грязный» для чистой математики

Когда нейробиологи пытаются наложить строгие математические формулы Фристона на реальный биологический мозг, они сталкиваются с огромным количеством шума, эволюционных компромиссов и анатомических ограничений. Мозг развивался миллионы лет методом «проб и ошибок», в нём много хаоса. Экспериментально доказать, что каждый нейрон идеально вычисляет «Байесовскую ошибку предсказания», невероятно сложно.

3. Смещение фокуса на ИИ (Active Inference)

Понимая ограничения биологических экспериментов, Фристон в последние годы стал активно применять свои идеи в сфере искусственного интеллекта. Он даже занял пост главного научного сотрудника в технологическом стартапе *VERSES AI*, который разрабатывает нейроморфные вычислительные системы. [13]

Фристон утверждает, что современный ИИ (вроде больших языковых моделей вроде ChatGPT) зашел в тупик, потому что он работает как «пассивный процессор» данных: ему нужны терабайты текстов для обучения, и он не умеет взаимодействовать с миром напрямую.

Взамен Фристон предлагает строить ИИ на принципах Активного вывода (Active Inference):

- Такой ИИ не просто обрабатывает данные, он имеет свою цифровую «модель мира» (как мозг человека).
- Он сам решает, какие данные ему нужны, и активно исследует среду, чтобы минимизировать неопределенность.
- Ему требуется в тысячи раз меньше энергии и данных для обучения.

Итак, Фристон считает, что в технической среде ИИ эту теорию наконец-то можно реализовать в чистом, идеальном математическом виде, без биологических ограничений нашего тела. Для него ИИ — это идеальная «песочница» и практическое доказательство его правоты.

В чем специфика позиций предиктивного кодирования?

Предиктивный процессинг обладает уникальным весом, но ученые смотрят на него двояко:

1. Это великая общая теория мозга, но не обязательно сознания

Критики и многие нейробиологи (включая тех, кто работает с Фристоном) отмечают, что предиктивное кодирование — это теория работы мозга в целом, а не специализированная теория сознания.
Она идеально объясняет, как мы обрабатываем сенсорные сигналы, управляем мышцами, экономим энергию и как ИИ может обучаться. Но сам факт того, что мозг минимизирует ошибку предсказания, напрямую не отвечает на «трудную проблему сознания» (*Hard Problem of Consciousness*): *почему этот вычислительный процесс сопровождается субъективным, качественным переживанием (квалиа)?* Ведь предиктивный робот может минимизировать ошибки, оставаясь абсолютно «слепым» внутри философским зомби.

2. Радикальный и умеренный подходы (Анил Сет vs Карл Фристон)

Чтобы превратить предиктивное кодирование именно в теорию сознания, ученые разделяются на два направления:

- Нейрорепрезентационализм (PP-NREP): Подход, который развивают такие ученые, как Анил Сет и Сайриел Пеннартц. Они утверждают, что наше субъективное сознание — это и есть содержание мультисенсорной предсказательной модели. Мы осознаем не мир, а саму виртуальную модель, которую построил мозг для минимизации «удивления» нервной системы.
- Активный вывод (PP-AI): Фристоновский взгляд, согласно которому сознание неразрывно связано с действием. Чтобы обладать сознанием, система должна не просто пассивно предсказывать картинку, но активно влиять на мир, проверять свои гипотезы через движение и осознавать себя причиной этих изменений.

Предиктивное кодирование сегодня имеет колоссальный теоретический вес. Его часто называют «главным кандидатом на звание единой теории поля в нейронауке», потому что оно способно элегантно объединить все остальные теории под одной математической крышей.

Однако в узком поле *исследований сознания* оно не является абсолютным монополистом, разделяя лидерство с квантитативной теорией IIT (Тонини) [11]и когнитивной GNWT (Деан) [12].

Сравнение с теорией МВАП

Критерий	Предиктивное кодирование (Карл Фристон и др.)	Модель волевой адаптивности психики (МВАП)
Что первично?	Прогноз и его ошибка. Мозг постоянно генерирует ожидания сверху вниз.	Актуальный образ и доминанта. Мозг реагирует на текущую значимость стимула и контекста.
Роль сознания	Опциональна. Сознание — это просто содержание наиболее точной прогностической модели (Анил Сет) или инструмент минимизации удивления.	Ключевая. Сознание необходимо строго в те моменты, когда автоматические (рефлекторные) навыки дают сбой в значимой новизне.
Понятие воли	Активный вывод (Active Inference). Воля — это предсказание движения, которое тело обязано выполнить, чтобы	Произвольность выбора. Воля — это механизм торможения привычной реакции ради оценки новой альтернативы.

	убрать ошибку.	
Аппарат теории	Высшая математика, Байесовская статистика, физика свободной энергии.	Эволюционная схемотехника, нейробиологические механизмы внимания и памяти. Формальная логика конечных автоматов.
Статус в мире	Мировой мейнстрим, фундаментальная наука.	Локальная альтернативная теория, верифицируемая через ИИ-прототипы (*Beast*).

В чем их принципиальное концептуальное различие?

1. Роль сознания: «Побочный продукт» vs «Аварийный тумблер»

- В предиктивном кодировании сознание размыто. Вычислительные циклы «прогноз-ошибка» идут на всех уровнях нервной системы постоянно (даже когда мы спим или действуем на автомате). Сознание здесь — скорее вершина айсберга, удачно собранная модель реальности.
- В МВАП сознание имеет очень жесткую функциональную задачу. Оно включается только тогда, когда привычные автоматизмы сталкиваются с непредвиденной ситуацией. Сознание в МВАП — это «аварийный режим» для переучивания и адаптации, который находит новое решение и снова засыпает, превращая действие в автомат. [

2. Понимание «Воли»

- Для Фристона (ПК) воли в бытовом понимании не существует. Если вы хотите поднять руку, ваш мозг не посылает команду «поднимись». Он генерирует *предсказание*, что рука уже поднимается. Возникает проприоцептивная ошибка (мозг думает, что рука движется, а она лежит), и мышцы сокращаются, просто чтобы ликвидировать эту ошибку. Это называется активным выводом.
- Для МВАП воля (произвольность) — это активный эволюционный механизм контроля. Это способность заблокировать привычный автоматический ответ (например, страх или привычку) с помощью механизмов внимания, оценить новизну и выдать творческое адаптивное действие.

3. Математика против Схемотехники

- ПК строит свои модели на Байесовских вероятностях. Мозг постоянно считает, насколько вероятен тот или иной стимул.

- МВАП критикует чисто вероятностный подход. Ее создатели утверждают, что живой организм не занимается сложнейшим обсчетом абстрактных вероятностей всего вокруг. Вместо этого он оперирует цепочками «стимул — контекст — значимость — действие», усложняющимися в ходе эволюции (от простейших рефлексов до высшего волевого выбора).

Резюме: можно ли их объединить?

Предиктивное кодирование — это попытка описать мозг с точки зрения физики информационных процессов (как система выживает в хаосе). МВАП — это попытка описать психику с точки зрения эволюционной инженерии (какие конкретно блоки внимания, памяти и воли нужны, чтобы робот или животное вели себя адаптивно).

ПК занимает несопоставимо более весомые и признанные позиции в мировой науке. Однако МВАП предлагает гораздо более четкий и понятный ответ на вопрос, зачем эволюции вообще понадобилось *сознание*.

Практичность

Авторы МВАП утверждают, что в отличие от западных теорий, которые годами спорят о терминах, МВАП дает четкие ответы уровня «какой блок, куда и какой сигнал передает».

Однако, если оценивать практический путь МВАП объективно, обнаруживается серьезный разрыв между потенциалом схемы и реальностью ИТ-индустрии.

В чем сила практического подхода МВАП?

1. Наличие действующего прототипа. Теория не осталась на бумаге. Был создан программный прототип — проект Beast (и его веб-версии) [10]. Это цифровая «сущность», у которой есть виртуальные рецепторы, базовые потребности, механизмы формирования эмоций, автоматизмы и, главное, контур «сознания», активирующийся при дефиците готовых реакций.
2. Схемотехническая ясность. МВАП раскладывает психику на понятные для программиста элементы: таблицы контекстов, ветвление значимостей, триггеры новизны и доминанты внимания. Это позволяет избежать мистики в описании сознания и воли.
3. Отказ от «черного ящика». В отличие от современных нейросетей (LLM вроде ChatGPT), поведение системы на базе МВАП полностью прозрачно для разработчика. Можно пошагово отследить, почему система заблокировала один автоматизм и выбрала другую, творческую альтернативу.

Почему это (пока) не стало магистральным путем для ИИ в мире?

Несмотря на амбициозный инженерный каркас, МВАП сталкивается с рядом барьеров, которые мешают ей выйти из статуса «эксперимента энтузиастов»:

1. Проблема масштабирования (Кустарность против БигТеха)

Современный ИИ развивается по пути brute force (грубой силы): миллиарды долларов, огромные кластеры видеокарт и гигантские массивы данных. МВАП же требует кропотливого проектирования архитектуры вручную. Проект *Beast* — это во многом жестко запрограммированная (hardcoded) логическая модель. Перенести эту логику на масштаб полноценного робота, ориентирующегося в хаотичном физическом мире, силами небольшой группы исследователей технически почти невозможно.

2. Проблема первичного наполнения (Как обучить «новорожденный» ИИ?)

Система на базе МВАП начинает как младенец — у нее есть только базовые безусловные рефлексы и потребности. Чтобы она развилась в сильный интеллект, ей нужно пройти годы личного опыта адаптации, совершая ошибки и нарабатывая базу автоматизмов. У индустрии нет времени «растить» каждый процессор годами, ей нужен инструмент, который сразу умеет писать код или переводить тексты.

3. Академическая и индустриальная изоляция

МВАП практически неизвестна на английском языке и не представлена в рецензируемых мировых журналах уровня *Nature* или *Science*. Мировое сообщество разработчиков AGI (OpenAI, Google DeepMind, Anthropic) идет по другим рельсам — они пытаются вырастить адаптивность и «подобие воли» внутри огромных трансформерных сетей или через обучение с подкреплением (RL), а не собирать ее по готовой схеме.

Резюме

МВАП действительно предлагает практический путь, а не просто теоретизирует, и проект *Beast* — прямое тому доказательство. Эта модель показывает, как можно собрать сильный интеллект «на кончике пера» через эволюционную схемотехнику.

Но на сегодняшний день этот путь остается альтернативной, непризнанной тропой. Чтобы МВАП стала реальным прикладным стандартом, ей необходимы миллионные инвестиции и перевод на уровень современных методов программирования, чтобы доказать превосходство над классическими нейросетями в реальных задачах.

Объективное (со стороны) сравнение теоретических и прагматических потенциалов Predictive Coding / Free Energy Principle (Фристон и др.) и МВАП (Парусников, Петрийчук и др.)

1. Обоснованность (эмпирическая и теоретическая обоснованность)

Predictive Coding (PC) / Free Energy Principle (FEP):

- **Сильные стороны**: Мощная математическая база (вариационный Байес, иерархическая байесовская inference). Хорошо объясняет и предсказывает множество феноменов: repetition suppression, mismatch negativity (MMN), precision weighting, активное восприятие. Поддерживается сотнями экспериментов в сенсорной нейронауке, нейровизуализации и моделировании поведения. SPM-метод Фристона — золотой стандарт анализа fMRI.
- **Слабые стороны**: FEP часто критикуют как **нефальсифицируемый** (unfalsifiable) или тавтологичный: почти любое адаптивное поведение можно постфактум описать как минимизацию free energy. Эмпирическая поддержка PC умеренная — многие результаты совместимы и с альтернативными feedforward-моделями. Сложно строго протестировать биологическую реализацию точных байесовских вычислений в «грязном» мозге.

МВАП (Модель Волевой Адаптивности Психики):

- **Сильные стороны**: Основана на обобщении классических советских/российских работ (Соколов, Виноградова и др.) по ориентировочному рефлексу, доминанте, значимости и новизне. Предлагает **конкретную схемотехнику** (блоки, потоки, триггеры), верифицированную через работающий прототип Beast (открытый код). Чётко разграничивает уровни (рефлексы → автоматизмы → сознание как аварийный механизм) [8], [9].
- **Слабые стороны**: Минимальная международная видимость и цитируемость. Публикации преимущественно в российских/локальных журналах. Эмпирическая база — в основном ретроспективное объяснение известных данных + демонстрация на прототипе. Отсутствуют крупные независимые валидационные исследования в mainstream-нейронауке.

Вердикт: PC/FEP значительно сильнее обоснована в глобальном научном сообществе и имеет больше эмпирических коррелятов. МВАП выигрывает в **конкретности и инженерной проверяемости** на уровне прототипа, но уступает в широте и независимой верификации.

2. Целостность описания (единство теории, охват феноменов)

PC/FEP:

- Стремится к **универсальной метатеории** — от бактерий и клеток до сознания, психопатологии, эволюции и ИИ. Отлично объединяет perception, action, learning и homeostasis под одним принципом (минимизация surprise/ free energy). Хорошо объясняет иллюзии, психические расстройства (например, шизофрения как нарушение precision weighting).

- Проблема: чрезмерная абстрактность и «всё объясняет» → риск потери специфичности (особенно для «трудной проблемы» сознания/квалиа).

МВАП:

- Фокусируется на **индивидуальной адаптивности** как схемотехнической системе: от гомеостаза и рефлексов до волевого сознания как механизма разрешения новизны/конфликта. Чётко описывает функциональную роль сознания («аварийный тумблер»), волю (торможение + выбор), память, эмоции. Решает «трудную проблему» через субъективность как свойство работы конкретных контуров.
- Проблема: менее универсальна (меньше внимания физике/математике низкоуровневых процессов), но более прагматична для психологии и инженерии.

Вердикт: PC/FEP целостнее как **фундаментальная теория** (единый принцип для всего живого). МВАП целостнее как **инженерная модель психики** — даёт понятную «схему» с чёткими блоками и ролями.

3. Практические возможности исследований и разработок

PC/FEP:

- **Исследования**: Огромный объём — тысячи статей, конференции, гранты. Используется в нейровизуализации, психиатрии, computational neuroscience.
- **Разработки (ИИ/робототехника)**: Влияет на Active Inference frameworks, neuromorphic computing (VERSES AI и др.). Потенциал в data-efficient learning, но реализация сложна из-за математической тяжести и проблем масштабирования в реальном мире. Современный mainstream ИИ (трансформеры + RL) идёт параллельно, заимствуя отдельные идеи.

МВАП:

- **Исследования**: Позволяет моделировать и тестировать конкретные механизмы (внимание, воля, сознание) в контролируемой среде. Прототип Beast демонстрирует прозрачность (traceable поведение, а не чёрный ящик).
- **Разработки**: Высокий инженерный потенциал — схемотехника легко переводима в код, правила, архитектуры. Подходит для explainable AI, роботов с ограниченными ресурсами, образовательных/терапевтических систем. Главный плюс: уже есть рабочий прототип. Минус: масштабирование требует значительных инвестиций и интеграции с современными ML-инструментами.

Вердикт по прагматике:

- PC/FEP — мощный генератор гипотез и научного престижа, но сложна для прямой инженерной реализации «из коробки».
- МВАП — более практична для создания **прозрачных, адаптивных, волевых агентов** здесь и сейчас, особенно в задачах, где важны explainability и низкие вычислительные затраты. Однако нуждается в большем финансировании и интеграции с mainstream-технологиями.

Общий вывод

Predictive Coding / FEP — это доминирующая **фундаментальная парадигма** мировой нейронауки: математически строгая, влиятельная, с богатой эмпирикой, но страдающая от чрезмерной абстрактности и проблем фальсифицируемости. Она лучше объясняет «как всё работает в принципе».

МВАП — это **инженерно-схемотехническая альтернатива**, сильная в конкретности, функциональной роли сознания/воли и практической верифицируемости через код. Она предлагает ясный путь к сильному, объяснимому ИИ, но пока остаётся нишевой (в основном российским) разработкой.

Обе теории дополняют друг друга: PC/FEP даёт широкий математический и биологический контекст, МВАП — детальную «монтажную схему» для реализации. Их синтез (или конкуренция) мог бы значительно продвинуть как понимание сознания, так и разработку AGI. Ни одна не является окончательной истиной — это конкурирующие полезные модели реальности.

Оценка Predictive Coding с точки зрения МВАП

В наиболее амбициозной формулировке Фристон утверждает, что все самоорганизующиеся системы, сохраняющие свою идентичность во времени, могут быть описаны через единый математический аппарат. Это включает организмы, клетки, мозг и даже некоторые небиологические системы.

Основная идея такова:

- Живая система должна сохранять свои состояния в узком диапазоне, иначе она распадётся или погибнет.
- Для этого она должна каким-то образом «сопротивляться» энтропии среды.
- Из определённых предпосылок о статистическом существовании устойчивой системы выводится, что её динамика может быть описана как минимизация некоторой величины — вариационной свободной энергии.

В этой логике даже отдельная клетка является кандидатом на такое описание: мембрана играет роль границы, отделяющей внутренние состояния от внешних, а обмен веществ поддерживает клетку в жизнеспособном состоянии. Вариационная свободная энергия — это своего рода мера того, насколько текущее состояние системы совместимо с её продолжением существования. Вариационная свободная энергия — это математическая величина, измеряющая одновременно:

1. насколько хорошо внутренняя модель объясняет наблюдения;
2. насколько неопределённа эта модель.

В некотором смысле свободная энергия = ошибка объяснения + неопределённость.

Один случай: глюкоза упала до критического значения, но кислород еще раньше достиг критического состояния. Второй случай: глюкоза несколько понизилась, нужно бы поесть, но появилось нечто настолько инертное (новое и важное), что состояние глюкозы по сравнению игнорируется. Как может быть величина "меры" учитывать разные состояния гомеостата и еще субъективную оценку важности для выбора? Как одно число может одновременно учитывать кислород, глюкозу, температуру, угрозы, любопытство, цели и ещё расставлять между ними приоритеты?

Ответ Фристона примерно такой: сама свободная энергия не содержит готовую шкалу важности. Важность задаётся моделью организма. Он рассчитывает отклонения жизненных параметров как ошибку предсказания и поведение следует величине самой большой из таких ошибок.

Но почему уход жизненного параметра трактуется как ошибка предсказания? Такое состояние может быть исходным, когда организм очнется от комы или глубокого сна, это вовсе никакая не ошибка. Но по Фистону ошибка предсказания возникает только тогда, когда наблюдаемое состояние расходится с тем, что предполагает текущая внутренняя модель. Поэтому современные версии FEP опираются не только на ошибки предсказания, но и на систему предпочтительных состояний, т.е. все больше погружаются в попытки объяснений, чтобы сохранить декларируемый принцип.

На рефлекторном уровне нет никаких таких предсказаний, все совершается по принципу рефлекторной реакции в контексте, который заготовлен для случая восстановления нормы жизненного параметра. При этом есть только конкуренция, какой их контекстов реагирования активировать преимущественно, что решает все проблемы на допсихическом уровне индивидуальной адаптивности.

Сторонник FEP ответил бы, что вы используете слишком узкое понимание предсказания. Он сказал бы: Любая система управления, поддерживающая целевое состояние, уже неявно содержит модель желаемого состояния. И, таким образом, спор выход на уровень спора о словах. Если каждую уставку назвать предсказанием, каждую ошибку регулирования назвать ошибкой предсказания, а каждый контур обратной связи назвать инференсом, то возникает риск, что теория становится универсально применимой ценой потери различий между механизмами.

Теория МВАП полностью обходится без таких семантических переопределений, используя прямые понятия и принципы индивидуальной адаптивности всех живых существ с гомеостатической регуляцией, вплоть до человека и даже искусственных живых существ.

Таким образом, с точки зрения теории МВАП между ней и теорией предиктивного кодирования нет ничего общего и нельзя сказать, что одна дополняет другую.

Список литературы

1. Гельмгольц Г. фон. *Физиологическая оптика* / пер. с нем. — М.: ЛКИ, 2008. (оригинал: Helmholtz H. von. Handbuch der physiologischen Optik. — Leipzig: Voss, 1867).
2. Бернштейн Н.А. *Очерки по физиологии движений и физиологии активности*. — М.: Медицина, 1966.
3. Rao R.P.N., Ballard D.H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects // *Nature Neuroscience*. — 1999. — Vol. 2, No 1. — P. 79–87.
4. Friston K. The free-energy principle: a unified brain theory? // *Nature Reviews Neuroscience*. — 2010. — Vol. 11, No 2. — P. 127–138.
5. Clark A. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. — Oxford: Oxford University Press, 2016.
6. Seth A.K. *Being You: A New Science of Consciousness*. — London: Faber & Faber, 2021.
7. Friston K., FitzGerald T., Rigoli F. et al. Active inference and learning // *Neuroscience & Biobehavioral Reviews*. — 2016. — Vol. 68. — P. 862–879.
8. Соколов Е.Н. *Очерки о работе мозга*. — М.: Изд-во МГУ, 1990.
9. Виноградова О.С. *Ориентировочный рефлекс и его нейрофизиологические механизмы*. — М.: Наука, 1961.
10. Парусников В.В., Петрийчук Д.Л. Модель волевой адаптивности психики (МВАП): схемотехника и прототип // *Искусственные общества*. — 2018. — Т. 13, № 1–2. — URL: fornit.ru (дата обращения: 11.06.2026).
11. Тонини Г., Кох К. Сознание как интегрированная информация // *В мире науки*. — 2015. — № 4. — С. 30–39. (оригинал: Tononi G., Koch C.

Consciousness: here, there and everywhere? // *Philosophical Transactions of the Royal Society B*. — 2015. — Vol. 370, No 1668).

12. Деан С. *Сознание и мозг: как мозг кодирует мысли* / пер. с англ. — М.: АСТ, 2018. (оригинал: Dehaene S. Consciousness and the Brain. — NY: Viking, 2014).
13. VERSES AI. Active Inference for next-generation AI. — URL: verses.ai (дата обращения: 11.06.2026).