

Искусственный интеллект как система

**Какую перспективу открывает
теория систем?**

издание 1-е, вер. 1.01 (ru)

30.12.2022

Игорь Фургель / Igor Furgel
(office@furgel.com)

В настоящей работе мы хотим применить подход теории систем к системам 'искусственного интеллекта' (ИИ). Одна из целей, которую мы преследуем, заключается в установлении основных критериев различия между различными типами систем искусственного интеллекта.

Мы также хотим рассмотреть вопрос о том, какие необходимые свойства – с точки зрения теории систем – должна иметь автономная техническая система, чтобы она была универсально применима, обладая такими же (или более развитыми и обширными) интеллектуальными и созидательными способностями, как у человека.

Это рассмотрение поможет нам также понять место различных типов систем искусственного интеллекта в контексте 'неживое - живое - человек'.

Настоящая работа может привлечь внимание круга читателей, интересующихся как вопросами искусственного интеллекта и его места в контексте 'неживое - живое - человек', так и системным подходом в целом.

Настоящее, первое издание этой работы (V. 1.00 (ru)) опубликовано 30.12.2022, Deutsche Nationalbibliothek,
<https://d-nb.info/1276877293/>.

Das Original der ersten Ausgabe dieses Aufsatzes (V. 1.00 (de) - „Künstliche Intelligenz als System: Welche Perspektive zeigt die Systemtheorie auf?“) wurde am 30.12.2022 veröffentlicht, Deutsche Nationalbibliothek,
<https://d-nb.info/1276877072/>.

The first English edition (V. 1.00 (en) - „Artificial Intelligence as a System: What Perspective does Systems Theory Reveal?“) was published on 30.12.2022, Deutsche Nationalbibliothek,
<https://d-nb.info/1276876920/>.

Основные идеи этой работы в применении к ИИ возникли в период с июня по декабрь 2022.

Содержание

1	Введение.....	5
2	Основанный на правилах ИИ (экспертные системы).....	5
3	Слабый ИИ с машинным обучением.....	10
3.1	Подсистема обучения технической системы ИИ (ТСИИ-ML).....	13
3.2	Подсистема обучения автоматизированной системы принятия решений (АСПР-ML).....	16
3.3	Подсистема принятия решений технической системы ИИ (ТСИИ-DM).....	18
3.4	Подсистема принятия решений автоматизированной системы принятия решений (АСПР-DM).....	20
4	Сильный ИИ.....	23
5	Энморфия ‘самосознания’ ИИ: Критерии различия между типами искусственного интеллекта и перспектива.....	27
6	Глоссарий.....	32
7	Ссылки.....	39
8	Благодарности.....	39

1 Введение

В настоящей работе мы хотим применить подход теории систем, который мы развили в [5], Part A, СНАРТЕР I (в частности, в главе 3) и СНАРТЕР II, к системам ‘искусственного интеллекта’ (ИИ). Одна из целей, которую мы преследуем, заключается в установлении основных критериев различия между тремя типами систем искусственного интеллекта, а именно между

- экспертными системами, *основанными на правилах*, которые опираются на программирование, основанное на правилах, и были разработаны в 1970-х и 1980-х годах,
- ‘слабым (узким) ИИ’¹, который основан на *машинном обучении* и может быть использован только для решения конкретных, ограниченных задач, и
- ‘сильным (общим) ИИ’², который был бы универсально применим, обладая такими же (или более развитыми и обширными) интеллектуальными и созидательными способностями, как у человека; в настоящее время ‘сильный ИИ’ – это только лишь видение.

Мы также хотим рассмотреть вопрос о том, какие необходимые свойства – с точки зрения теории систем – должна иметь автономная техническая система, чтобы она была универсально применима, обладая такими же (или более развитыми и обширными) интеллектуальными и созидательными способностями, как у человека.

Это рассмотрение поможет нам также понять место различных типов систем искусственного интеллекта в контексте ‘неживое – живое – человек’.

Настоящую работу можно читать самостоятельно. основополагающие понятия приведены в гл. 6 «Глоссарий». Тем не менее, так как подходы, которые мы развили в [5], Part A, СНАРТЕР I (в частности, в главе 3 «Being, Existential Triads and Enmorphya») и СНАРТЕР II являются основополагающими для данного исследования, мы рекомендуем читателям, интересующимся также основополагающими разработками, обратиться к [5].

2 Основанный на правилах ИИ (экспертные системы)

Традиционное программирование, также известное как программирование, основанное на правилах, используется для создания *основанных на правилах* экспертных систем. Такие экспертные системы используют жестко закодированные базы знаний по определенной предметной области и жестко закодированные правила принятия решений. Когда основанная на правилах экспертная система получает запрос вместе с данными, подлежащими анализу, она использует жестко закодированные правила принятия решений для сопоставления этих подлежащих анализу данных с базой знаний и генерирует ответ (решение/рекомендацию), см. [4], гл. 1 и [3], гл. 2.

Рассмотрим, как абстрактные элементы *экзистенциальной триады* (см. Глоссарий) выражаются в рамках этого типа ИИ.

¹ EN: Narrow (weak) artificial intellect (AI), сокращенно ANI; DE: Schwache künstliche Intelligenz (Schwache KI)

² EN: General (strong) AI, сокращенно AGI; DE: Starke KI

Подлежащие анализу данные/запросы являются ‘*субстратом*’ любой экспертной системы, основанной на правилах. ‘*Свойства*’ системы – это характер основанной на правилах программы (включая операционализацию/измеримость цели программы, модель проблемы и алгоритм)³, который жестко закодирован в операторах (инструкциях), включая базу знаний, и свойства/характеристики подлежащих анализу данных. ‘*Отношение*’ – это процесс применения жестко закодированных инструкций к данным, которые должны быть проанализированы в ходе выполнения программы в контексте использования основанной на правилах экспертной системы. Подлежащие анализу данные не изменяются в результате применения этих инструкций.

Процесс применения жестко закодированных инструкций к данным, подлежащим анализу, имеет *детерминистический*, а не *стохастический* характер. Поэтому Принцип Достаточности Экзистенциальной Триады⁴ здесь не применим: *основанные на правилах экспертные системы являются детерминистическими*.

Для экспертных систем, *основанных на правилах*, реализованные *принципы разработки программного обеспечения*, включая применяемые процедурные правила/нормы (т.е. ‘*руководство по программированию*’), представляют собой ‘*информацию-об-управлении-отношением*’ (*энморфию отношения*). В качестве энморфии отношения между субстратом (подлежащие анализу данные) и свойством (характер программы, основанной на правилах), реализованные принципы разработки программного обеспечения (т.е. ‘*руководство по программированию*’) определяют характер этого отношения (взаимодействия).

Если мы теперь проведем параллель со *стохастическими* системами, для которых в общем случае Принцип Достаточности Экзистенциальной Триады⁴ имеет силу, мы обнаруживаем, что этот принцип заменяется ‘*руководством по программированию*’, когда речь идет о *детерминистической* экспертной системе.

³ Эти термины подробно описаны в [2], см. там, например, «Глоссарий». Для лучшей читабельности мы воспроизводим здесь только соответствующее ядро этих определений:

- «Под моделированием я подразумеваю любую форму упрощения и абстрагирования ситуации, которая, тем не менее, остается достаточно точной, чтобы позволить делать прогнозы или аналитические выводы относительно этой ситуации.» (В оригинале: „Unter einer Modellierung verstehe ich jegliche Form der Vereinfachung und Abstraktion einer Situation, die trotzdem noch so genau ist, dass sie Vorhersagen oder analytische Schlüsse über diese Situation zulässt.“)

- “Операционализация представляет собой измеримость (социального) понятия - она всегда основана на модели понятия“. (В оригинале: „Eine Operationalisierung stellt die Messbarmachung eines (sozialen) Konzeptes dar – sie basiert immer auf einem Modell des Konzeptes.“)

- “Алгоритм – это достаточно подробный и систематический набор инструкций для любого опытного программиста по решению математической задачи, такой, что при правильной реализации (переводе в код) компьютер вычисляет правильный результат для любого правильного набора вводов.“ (В оригинале: „Ein Algorithmus ist eine für jede erfahrene Programmiererin und jeden erfahrenen Programmierer ausreichend detaillierte und systematische Handlungsanweisung, um ein mathematisches Problem zu lösen, sodass bei korrekter Implementierung (Übersetzung in Code) der Computer für jede korrekte Inputmenge den korrekten Output berechnet.“)

⁴ [5], ЧАРТЕР I, STM. 9 ‘The principle of sufficiency of the existential triad’ (перевод автора): «Если ‘отношение’ в экзистенциальной триаде {субстрат, свойство, отношение} носит принципиально *стохастический* характер и *статистически* подчиняется определенной закономерности, то экзистенциальная триада является не только *необходимой*, но и *достаточной* для достижения наблюдаемости и тем самым для создания состояния ‘бытия’ системы, базирующейся на этой экзистенциальной триаде. Эволюция этой системы будет следовать характеру ‘отношения’ в экзистенциальной триаде.»

‘Руководство по программированию’ определяет характер применения жестко закодированных инструкций к анализируемым данным. ‘Руководство по программированию’ определяет при этом различные *функции оптимизации*, например, минимизация энергопотребления или времени выполнения программы. В отличие от *стохастических* систем в целом, ‘руководство по программированию’ не может влиять на подлежащие анализу данные, а только лишь на характер программы, основанной на правилах (включая операционализацию/измеримость цели программы, модель проблемы и алгоритм).

Проиллюстрируем взаимоотношение первичной системы и метасистемы на примере *основанных на правилах* экспертных систем:

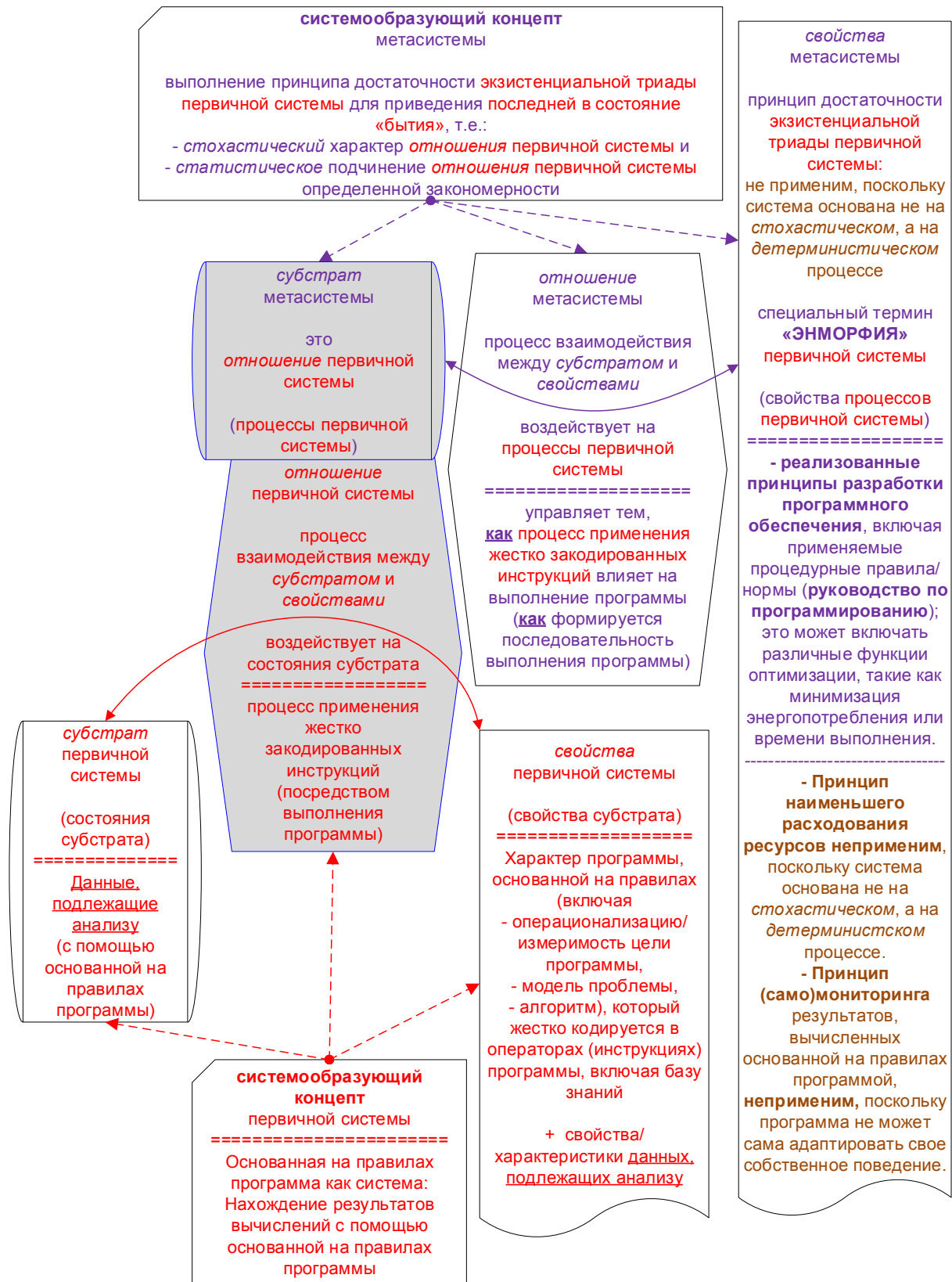


Рис. 1: Взаимоотношение первичной системы ‘основанная на правилах экспертная система’ и соответствующей метасистемы

Системно-теоретическое рассмотрение экспертных систем, основанных на правилах, позволяет нам сделать полезное обобщение системно-теоретического различия между детерминистическими и стохастическими системами:

Тип системы →	<i>стохастические системы</i>	<i>детерминистические системы</i>
Параметр ↓		(могут существовать исключительно как артефакты ⁵)
Принцип Достаточности Экзистенциальной Триады ⁴	всегда в силе	не применим
Энморфия отношения ('информация-об-управлении-отношением')	Специфическая для каждой системы реализация <i>Принципа Достаточности Экзистенциальной Триады</i> , причем <i>Принцип Наименьшего Расходования Ресурсов</i> <u>всегда</u> является составной частью энморфии ⁶ .	практически реализованные <i>принципы изготовления</i> соответствующей системы, включая применяемые процедурные правила/нормы; <i>спецификация</i> (или, образно говоря, ' <i>руководство по программированию</i> ') для изготовления соответствующей системы
Влияние энморфии отношения на первичную систему	Специфическая для каждой системы реализация <i>Принципа достаточности экзистенциальной триады</i> определяет характер 'отношения' первичной системы. Таким образом, специфическая для данной системы реализация <i>Принципа достаточности экзистенциальной триады</i> влияет <u>как</u> на состояния 'субстрата', <u>так и</u> на 'свойства' первичной системы.	Практически реализованные <i>принципы изготовления</i> соответствующей системы <u>не</u> могут влиять на состояния 'субстрата', а только лишь на 'свойства' первичной системы.

⁵ см. [5], CHAPTER I, ch. 3.1 «Being and Existential Triads» или CHAPTER VII, ch. 2.1.3 «Indeterminacy and Action Quanta: Complementary Characters of the Past and the Future»

⁶ см. [5], CHAPTER I, ch. 3.2 «Enmorphy», STM. 11 'the Principle of Least Resources Consumption' (перевод автора): «**Принцип Наименьшего Расходования Ресурсов (ПНР)** является информацией-об-управлении-отношением (т.е. энморфией отношения) и управляет не только процессом взаимодействия между материей и информацией в Природе, но и между *субстратом* и *структурным фактором* любой системы – физической, социальной, коммуникативной и т.д. –, в основе которой лежит *стохастический процесс*.»

Тип системы →	<i>стохастические системы</i>	<i>детерминистические системы</i>
Параметр ↓		(могут существовать исключительно как артефакты ⁵)
Влияние ‘отношения’ первичной системы на состояния ‘субстрата’	Процесс взаимодействия ‘свойств’ первичной системы с ‘субстратом’ первичной системы изменяет состояния ‘субстрата’.	Процесс взаимодействия ‘свойств’ первичной системы с ‘субстратом’ первичной системы <u>не</u> может изменить состояния ‘субстрата’.

Принцип Наименьшего Расходования Ресурсов, который имеет силу для всех *стохастических* систем, не применим к *основанной на правилах* экспертной системе, поскольку такая система основана не на *стохастическом*, а на *детерминистическом* процессе⁶.

Принцип Самосохранения Системы, который имеет силу для всех *квази-стохастических* систем⁷, и который для многих таких систем выражается как Принцип (само)мониторинга, также неприменим, поскольку такая система основана не на *стохастическом*, а на *детерминистическом* процессе, и поэтому программа не может сама адаптировать свое поведение. *Основанная на правилах* экспертная система не может автономно корректировать свое поведение, опираясь на вычисленный ею результат.

3 Слабый ИИ с машинным обучением

В отличие от экспертных систем, основанных на правилах, ‘слабый ИИ’ основан на *машинном обучении* (ML). ‘Слабый ИИ’ применим только для конкретных, ограниченных задач, см. [2], ‘Глоссарий’, [3], гл. 1, [4], гл. 2 (‘ANI’) или в Интернете.

Подход *машинного обучения* технологически отличается от подхода программирования *на основе правил* тем, что правила принятия решений и базы знаний не жестко закодированы в программе, как в случае *основанного на правилах* ИИ, а программа, реализующая *машинное обучение*, сама формирует набор правил принятия решений в виде статистической модели на основе данных, предоставляемых программе с ML⁸.

Существует ряд технологических типов машинного обучения, таких как, например, ‘глубокое обучение’ на основе технологии искусственных нейронных сетей (ИНС),

⁷ см. [5], CHAPTER I, ch.3.5 «Enmophya for Quasi-Stochastic Systems», STM. 13 ‘the Principle of Self-Preservation of System’ (перевод автора): «для обеспечения стабильности квази-стохастических систем, их энморфия должна содержать по крайней мере еще один принцип, который мы назвали **Принципом Самосохранения Системы.**»

⁸ Машинное обучение (ML) определено в [2], Глоссарий следующим образом: «Набор методов, которые ищут в данных о прошлом паттерны, позволяющие делать прогнозы на будущее» (В оригинале: „Eine Sammlung von Methoden, die in Daten der Vergangenheit nach Mustern suchen, die für die Zukunft Vorhersagen erlauben.“).

‘случайный лес’ на основе архитектуры дерева решений, ‘машина опорных векторов (SVM)’⁹ и некоторые другие, см. [4], гл. 3, [2], гл. 5.

Существуют также некоторые организационные типы машинного обучения, причем организационные и технологические типы машинного обучения должны сочетаться друг с другом, чтобы ‘слабый ИИ’ мог эффективно достичь заданной цели. В качестве примеров организационных типов машинного обучения мы приводим следующие (см. [3], глава 4):

- Обучение под наблюдением (с учителем; контролируемое обучение)¹⁰: Программа с ML сначала получает *тренировочные (обучающие) данные*, отобранные человеком (в роли ‘аналитика данных’ (‘Data Scientist’), см. [2]), на основе которых программа с ML ищет корреляции (и обычно находит их) и на этой базе определяет набор правил принятия решений в виде статистической модели. Затем эта программа с ML получает *тестовые данные* (т.е. данные с уже известными результатами¹¹). ‘Слабый ИИ’ анализирует эти тестовые данные, используя правила принятия решений, которые он установил для себя во время обучения, и сообщает результат этого анализа. Человек (в роли ‘аналитика данных’) сравнивает ранее известный результат тестовых данных с результатом, полученным ‘слабым ИИ’, и применяет к этому сравнению определенную человеком *меру качества*¹². Если отклонение между сравниваемыми результатами удовлетворяет этой *мере качества*, обучение считается успешным, см. [2], гл. 5, [4], гл. 3 (‘Нейронные сети’).
- Обучение без наблюдения (без учителя; неконтролируемое обучение)¹³: Программе с ML предоставляются немаркированные (‘unlabelled’) данные, т.е. без какого-либо указания цели обучения. Программа с ML ищет в этом наборе данных корреляции, т.е. зависимости и паттерны в этих данных, и из найденных корреляций¹⁴ выводит набор правил принятия решений. Если этой программе с ML подаются другие данные для анализа, программа с ML продолжает поиск дополнительных корреляций, корректирует правила принятия решений и применяет их к входным данным. В результате этого применения входные данные распределяются по кластерам (категориям), которые программа с ML таким образом сама сформировала, т.е. которым она сама обучилась. Одним из вариантов неконтролируемого обучения являются ‘генеративно-состязательные нейронные сети (Generative Adversarial Networks - GAN)’. Так как в обучении без надзора не существует предварительной ‘базовой истины’¹¹, человек должен проверить ex-post, являются ли решения/рекомендации, выработанные в результате этого типа ‘слабого ИИ’, адекватными для цели оператора системы ИИ, см. [3], глава 4.

⁹ EN, DE: Support Vector Machine

¹⁰ EN: Supervised Learning, DE: Überwachtes Lernen

¹¹ Свойства *тренировочных* и *тестовых данных*, которые уже известны заранее, называются ‘базовой/фундаментальной истиной (Ground Truth)’.

¹² Мера качества: Функция, которая оценивает, насколько хорошим является (алгоритмическое) решение проблемы (В оригинале: „Qualitätsmaß: Eine Funktion, die bewertet, wie gut eine (algorithmische) Lösung eines Problems ist.“), см. [2], Глоссарий.

¹³ EN: Unsupervised Learning, DE: Unüberwachtes Lernen

¹⁴ Важно отметить, что найденные корреляции не обязательно являются причинно-следственными, что ‘слабый ИИ’ в принципе не может установить.

- Обучение с подкреплением¹⁵: На первом этапе человек (в роли ‘аналитика данных’) определяет *функцию вознаграждения* для программы с ML и указывает программе максимизировать значение ‘вознаграждения’ при поиске возможных результатов анализа входных данных. ‘Слабый ИИ’ пробует затем возможные варианты решения (методом проб и ошибок) и выбирает тот вариант, который увеличивает значение ‘вознаграждения’ по сравнению с предыдущим решением, как подходящее решение. Таким образом ‘слабый ИИ’ оптимизирует себя все дальше и дальше с каждым следующим решением, так что значение ‘наград’ увеличивается в статистическом среднем и в какой-то момент достигает максимального значения, см. [3], глава 4, [4], глава 3 («Эволюционные алгоритмы»).

Из приведенных выше примеров видно, что ‘слабый ИИ’ зависит от вмешательства человека *ex-ante* или *ex-post*. Этот факт также относится ко всем другим организационным типам машинного обучения.

Поэтому системы, основанные на машинном обучении, всегда являются *социотехническими* системами. Это означает, что они обязательно имеют чисто техническое ядро, т.е. чисто техническую подсистему, которая включена в тщательно организованную человеческую среду, см. [2], глава 1, рисунок 3.

Соответствующая социотехническая система в целом – т.е. чисто техническая подсистема вместе с организованной человеческой средой – называется ‘Автоматизированной системой принятия решений’ (*ADM System*¹⁶; АСПР), см. [2], глава 1, [3], глава 1.

Чисто техническая подсистема (мы будем сокращенно называть ее сокращенно ТСИИ – *техническая система искусственного интеллекта*) и вся АСПР (*автоматизированная система принятия решений*) относятся – с системно-теоретической точки зрения – к разным категориям систем (подробнее об этом – далее в этом разделе). Поэтому необходимо различать с точки зрения теории систем между ТСИИ и АСПР, см. [2], глава 1, рисунок 2.

Еще одно различие, имеющее значение с точки зрения теории систем, существует внутри каждой ТСИИ. Архитектура каждой ТСИИ включает в себя две технические подсистемы: подсистему машинного обучения (подсистему ML), которая функционирует *стохастически*, и подсистему принятия решений, которая функционирует *детерминистически* на основе набора правил принятия решений, определенных подсистемой ML, см. [2], гл. 5, рис. 22.

Соответственно, мы должны различать – с точки зрения теории систем – следующие подсистемы ‘слабого ИИ’:

- ТСИИ-ML: подсистема обучения технической системы ИИ,
- АСПР-ML: подсистема обучения автоматизированной системы принятия решений,
- ТСИИ-DM: подсистема принятия решений технической системы ИИ, и

¹⁵ EN: Reinforcement Learning, DE: Bestärkendes Lernen

¹⁶ Automated (or Algorithmic) Decision-Making System

- АСПР-DM: подсистема принятия решений автоматизированной системы принятия решений.

Далее мы рассмотрим каждую из этих подсистем в отдельности.

3.1 Подсистема обучения технической системы ИИ (ТСИИ-ML)

Правила принятия решений ТСИИ, которые необходимо выучить (статистическая модель), являются ‘субстратом’ подсистемы ТСИИ-ML ‘слабого ИИ’. ‘Свойством’ этой подсистемы является характер ТСИИ (включая операционализацию/измеримость цели ТСИИ, модель проблемы и алгоритм)³ и свойства/характеристики *тренировочных данных*. ‘Отношения’ – это процесс взаимодействия свойств ТСИИ и характеристик *тренировочных данных*, с одной стороны, и правил принятия решений, которые необходимо выучить, с другой, т.е. собственно процесс обучения.

Операционализация цели ТСИИ включает, среди прочего, *меру справедливости*¹⁷, которая, если она уместна, определяется человеком (в роли аналитика данных), например, равенство против равноценности, см. [2], гл. 8.

Рассмотрим в качестве примера технологию ‘глубокого обучения’. В этом случае ИНС (искусственная нейронная сеть) начинает процесс обучения, присваивая весам/вероятностям переходов между отдельными состояниями соседних слоев искусственных ‘нейронов’ случайные значения. После каждого цикла обучения с использованием *тренировочных данных* (как части *базовой истины*¹¹), ANN *вероятностно* корректирует эти веса с целью приблизить результат своего решения к базовой истине¹⁸. Соответственно, последовательность наборов выучиваемых правил принятия решений, которые составляют статистическую модель, также является в своей основе *вероятностной*.

Поскольку каждое следующее состояние набора правил принятия решений, которые необходимо выучить, *вероятностно* зависит только от его текущего состояния (а не от предыдущих состояний), процесс обучения обладает *марковским свойством*. Поэтому процесс обучения является *истинно-стохастическим* процессом, см. Глоссарий.

Как мы выяснили в [5], CHAPTER I, ch. 3.4 «Enmorphya for Truly-Stochastic Systems», ‘информация-об-управлении-отношением’ (энморфия отношения) всех *истинно-стохастических* систем всегда представлена Принципом Наименьшего Расходования Ресурсов (максимальной энтропии). Это означает, что энморфия отношения подсистемы ТСИИ-ML также должен быть представлена принципом наименьшего расходования ресурсов (ПНР), а именно – *принципом учебной экономии*.

Мануэла Ленцен отмечает в [3], гл. 4:

«Хорошая ИНС уверенно относит данные к желаемым категориям. Она не является ни слишком чувствительной, ни слишком нечувствительной к

¹⁷ Мера справедливости: «Математическая функция, оценивающая насколько различные группы населения одинаково затронуты решениями», см. [2], Глоссарий (В оригинале: „Eine mathematische Funktion, die bewertet, inwieweit unterschiedliche Bevölkerungsgruppen gleichermaßen von Entscheidungen betroffen sind“).

¹⁸ Качество результата решения определяется *мерой качества*, задаваемой человеком.

вариациям в данных. И она экономична с точки зрения времени, объема данных и оборудования, необходимого для её обучения.»¹⁹

Жанель Шейн цитирует в [4], гл. 5 Алекса Ирпана, исследователя ИИ в Google:

«У меня вошло в привычку думать об [ИИ] как о демоне, который намеренно неправильно интерпретирует свои вознаграждения и активно ищет оптимум, при котором он может быть максимально ленивым. Это звучит смешно, но на самом деле такое отношение может быть весьма продуктивным.»²⁰

Ввиду осознания того, что подсистема ТСИИ-ML ‘слабого ИИ’ всегда следует *принципу учебной экономии* (наименьшего расходования ресурсов), это высказывание Алекса Ирпана звучит не только не ‘смешно’, но и вполне закономерно.

В [4], гл. 6, Жанель Шейн неоднократно утверждает, что ИИ во время обучения постоянно пытается взломать заданную ‘матрицу’, чтобы получить ‘бесплатную энергию/пищу’.

Это также можно сразу объяснить, исходя из *принципа учебной экономии*: взлом ‘матрицы’ всегда минимизирует расходование ресурсов.

Второй принцип энморфии отношения подсистемы ТСИИ-ML – это набор *функций оптимизации* в соответствии с целью ТСИИ, включая их приоритизацию, при этом *мера качества* и *мера справедливости* задаются вне системы ТСИИ, т.е. внешним по отношению к ТСИИ специалистом по анализу данных.

В подсистеме ТСИИ-ML *принцип учебной экономии* вместе с *функциями оптимизации* составляют ‘информацию-об-управлении-отношением’ (энморфию отношения) этой системы. В качестве энморфии отношения между субстратом (правилами принятия решений, подлежащими обучению) и свойством (характером ТСИИ и характеристиками *тренировочных данных*), *принцип учебной экономии* и *функции оптимизации* определяют характер этого отношения (взаимодействия), см. сноску 4. Они определяют характер процесса машинного обучения, который, в свою очередь, реализует взаимодействие между подлежащими обучению правилами принятия решений и характером ТСИИ и характеристиками *тренировочных данных*. Таким образом, *принцип учебной экономии* и *функции оптимизации* определяют как правила принятия решений, подлежащие обучению (субстрат подсистемы ТСИИ-ML), так и характер ТСИИ и подходящие характеристики *тренировочных данных* (их форму и содержание).

Проиллюстрируем взаимоотношение первичной системы и метасистемы на примере подсистемы ТСИИ-ML (подсистемы обучения технической системы ИИ):

¹⁹ „Ein gutes KNN ist sicher darin, Daten den gewünschten Kategorien zuzuordnen. Es ist weder zu empfindlich noch zu unempfindlich gegenüber Variationen in den Daten. Und es ist sparsam, was die Zeit, die Menge an Daten und die Hardware angeht, die man zum Training benötigt.“

²⁰ В оригинале Алекса Ирпана: “I’ve taken to imagining deep RL as a demon that’s deliberately misinterpreting your reward and actively searching for the laziest possible local optima. It’s a bit ridiculous, but I’ve found it’s actually a productive mindset to have.”, Alex Irpan *Deep Reinforcement Learning Doesn’t Work Yet*, <https://www.alexirpan.com/2018/02/14/rl-hard.html>.

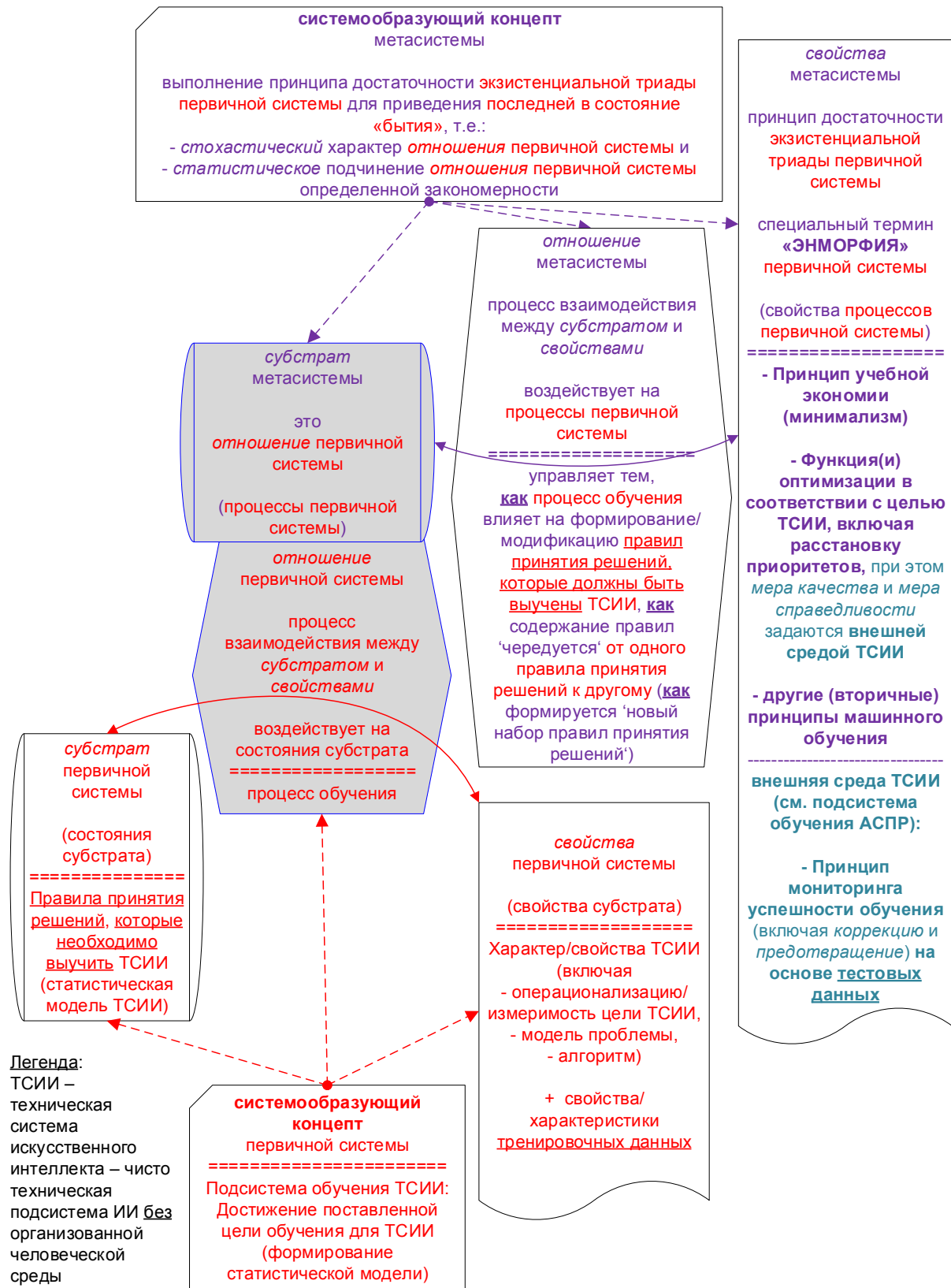


Рис. 2: Взаимоотношение первичной системы ТСИИ-ML (подсистемы обучения технической системы ИИ) и соответствующей метасистемы

3.2 Подсистема обучения автоматизированной системы принятия решений (АСПР-ML)

Как мы уже описали в этой главе, ‘слабый ИИ’ всегда полагается на вмешательство человека *ex-ante* или *ex-post*. Именно поэтому системы, основанные на машинном обучении, всегда являются *социотехническими* системами. Это означает, что они обязательно имеют чисто техническое ядро, т.е. чисто техническую подсистему (ТСИИ), которая включена в тщательно организованную человеческую среду, см. [2], гл. 1, рис. 3. Соответствующая общая социотехническая система – т.е. чисто техническая подсистема (ТСИИ) вместе с организованной человеческой средой – называется ‘Автоматизированной (алгоритмической) системой принятия решений’ (*ADM System*²¹; АСПР), см. [2], глава 1, [3], глава 1.

В разделе 3.1 выше мы рассмотрели подсистему обучения технического ИИ (подсистему ТСИИ-ML) с точки зрения теории систем. Теперь мы хотим проанализировать подсистему обучения всей автоматизированной системы принятия решений (подсистема АСПР-ML).

‘Социо’-компонент, т.е. вмешательство человека, представляет собой внешнее по отношению к системе граничное условие с точки зрения чисто технической ‘ТСИИ’. Подсистема обучения всей автоматизированной системы принятия решений отличается от подсистемы обучения технического ИИ только тем, что в неё интегрирован ‘социо’-компонент, т.е. вмешательство человека.

На Рис. 2 легко увидеть, что вмешательство человека – как внешнее по отношению к системе граничное условие для *подсистемы* ТСИИ-ML – влияет исключительно на энморфию отношения подсистемы ТСИИ-ML (выделено бирюзовым цветом). Человек (в роли аналитика данных) задает *меру качества* и *меру справедливости* (а также другие необходимые *гиперпараметры*) и контролирует успешность обучения с помощью *тестовых данных*, см. [2], глава 5.

Мониторинг успешности обучения (вкл. *коррекцию* и *предотвращение*) на основе *тестовых данных* реализует *Принцип мониторинга успешности обучения* для подсистемы обучения *Автоматизированной системы принятия решений* (подсистема АСПР-ML), ср. [2], глава 5, [4], глава 3. *Принцип мониторинга* включая *коррекцию* и *предотвращение* реализуется собственным *механизмом адаптации* системы.

Принцип мониторинга успешности обучения, в свою очередь, является конкретной реализацией общего **Принципа Самосохранения Системы** для подсистемы АСПР-ML, см. сноски 7. Тот факт, что энморфия отношения подсистемы АСПР-ML включает, среди прочего, принцип самосохранения системы, указывает на то, что подсистема АСПР-ML является *квази-стохастической* системой, см. сноски 7 и 22. Однако это вполне объяснимо, поскольку человек, который сам является *квази-стохастической* системой²³, является частью подсистемы АСПР-ML.

²¹ Automated (or Algorithmic) Decision-Making System

²² см. [5], СНАПТЕР I, ch. 3.5 «Enmorphya for Quasi-Stochastic Systems», STM. 14 ‘at least two principles are existentially necessary components of the enmorphya’ (перевод автора): «по крайней мере два принципа являются экзистенциально необходимыми составляющими энморфии *квази-стохастических* систем: **Принцип Наименьшего Расходования Ресурсов (ПНР)** и **Принцип Самосохранения Системы (ПСС)**.»

²³ см. [5], СНАПТЕР II, ch. 1.2 «Principles of the Enmorphya of Living Beings»

Проиллюстрируем взаимоотношение первичной системы и метасистемы на примере подсистемы ASPR-ML (подсистемы обучения автоматизированной системы принятия решений):

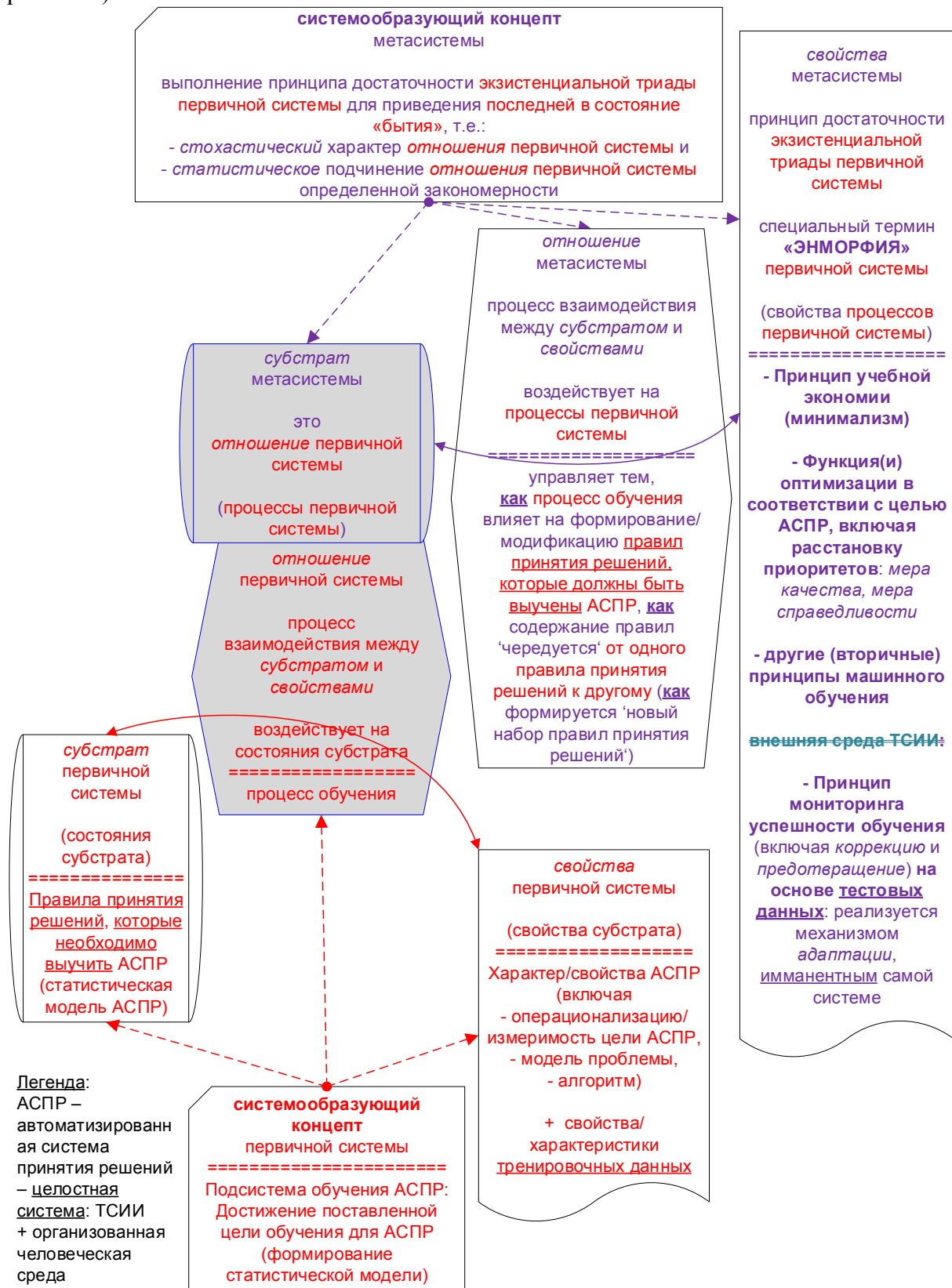


Рис. 3: Взаимоотношение первичной системы АСПР-ML (подсистемы обучения автоматизированной системы принятия решений) и соответствующей метасистемы

3.3 Подсистема принятия решений технической системы ИИ (ТСИИ-DM)

Подлежащие анализу данные являются ‘субстратом’ для каждой подсистемы принятия решений технической системы ‘слабого ИИ’. ‘Свойство’ этой подсистемы – это выученные правила принятия решений (выученная статистическая модель) ТСИИ и свойства/характеристики данных, подлежащих анализу. ‘Отношение’ – это процесс применения выученных правил принятия решений к анализируемым данным. Данные, подлежащие анализу, не меняются в результате этого применения.

Процесс применения выученных правил принятия решений к анализируемым данным является *детерминистическим*, а не *стохастическим*. Поэтому Принцип Достаточности Экзистенциальной Триады⁴ здесь не применим: ТСИИ подсистема принятия решений является *детерминистической*, см. [2], гл. 5, рис. 22.

Для ТСИИ-подсистемы принятия решений, *принципы выученной статистической модели* представляют собой ‘информацию-об-управлении-отношением’ (энморфию отношения). *Принципы выученной статистической модели* определяются ТСИИ-подсистемой обучения, см. ‘Энморфия’ на Рис. 2. В качестве энморфии отношения между субстратом (анализируемыми данными) и свойством (выученной статистической моделью), принципы выученной статистической модели определяют характер этого отношения (взаимодействия). Сравнивая энморфию основанной на правилах экспертной системы (Рис. 1) и энморфию ТСИИ-подсистемы принятия решений (Рис. 4 далее), становится очевидным, что выученная статистическая модель для ‘слабого ИИ’ играет ту же (центральную) роль, что и ‘руководство по программированию’ для экспертных систем, основанных на правилах.

Проиллюстрируем взаимоотношение первичной системы и метасистемы на примере подсистемы ТСИИ-DM (подсистемы принятия решений технической системы ИИ):

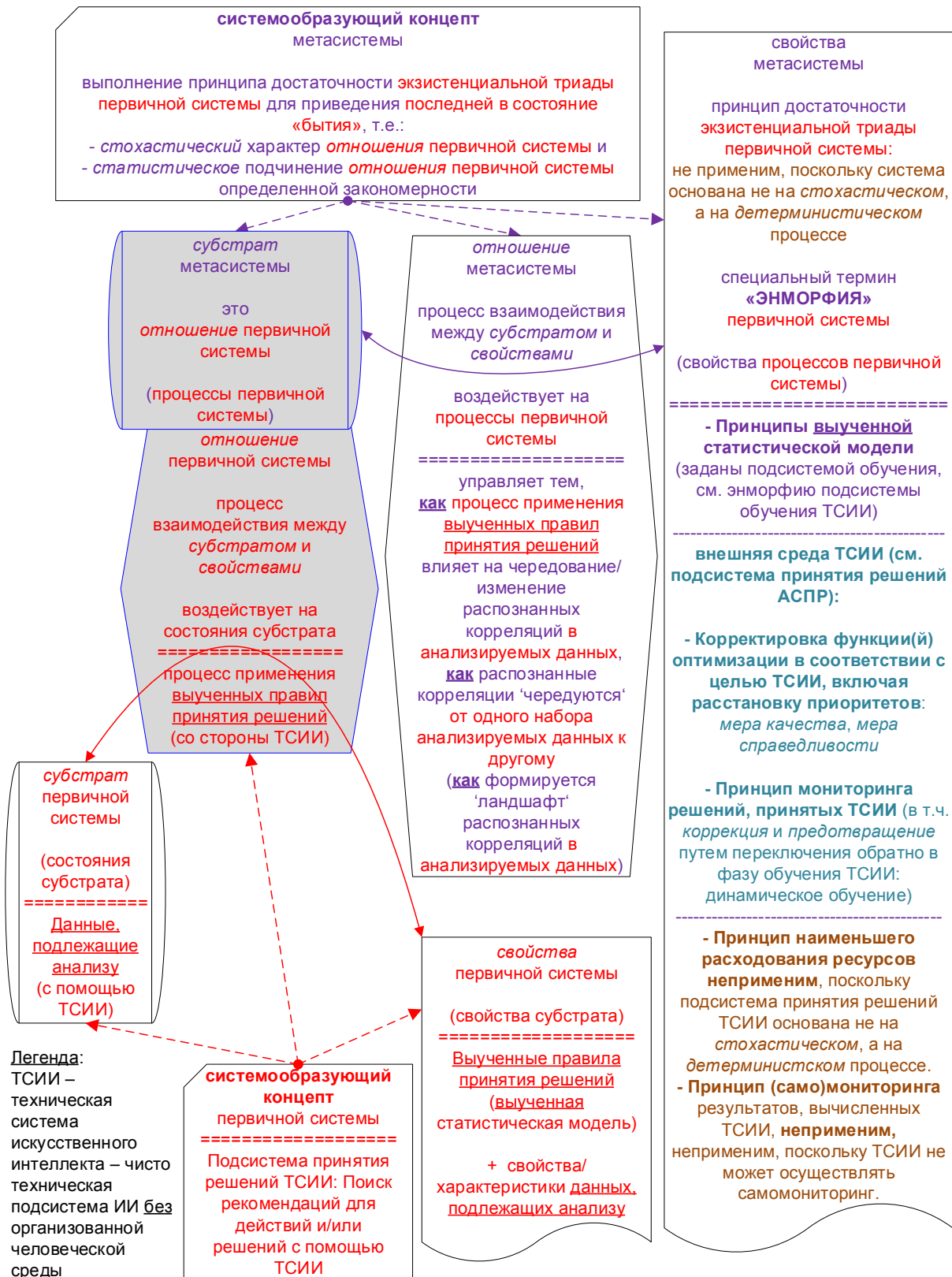


Рис. 4: Взаимоотношение первичной системы 'ТСИИ-DM' (подсистемы принятия решений технической системы ИИ) и соответствующей метасистемы

3.4 Подсистема принятия решений автоматизированной системы принятия решений (АСПР-DM)

В разделе 3.3 выше мы рассмотрели с точки зрения теории систем подсистему принятия решений технического ИИ (подсистему ТСИИ-DM). Теперь мы хотим проанализировать подсистему принятия решений всей *социотехнической* автоматизированной системы принятия решений (подсистему АСПР-DM).

‘Социо’-компонент, т.е. вмешательство человека, представляет собой внешнее по отношению к системе граничное условие с точки зрения чисто технической ТСИИ. Подсистема принятия решений всей автоматизированной системы принятия решений отличается от подсистемы принятия решений технического ИИ только тем, что в неё интегрирован ‘социо’-компонент, т.е. вмешательство человека.

На Рис. 4 легко увидеть, что вмешательство человека – как внешнее по отношению к системе граничное условие для *подсистемы* ТСИИ-DM – влияет исключительно на энморфию отношения *подсистемы* ТСИИ-DM (выделено бирюзовым цветом). Человек (в роли аналитика данных) корректирует *меру качества* и *меру справедливости* (а также другие необходимые *гиперпараметры*). Кроме этого, человек контролирует адекватность принятых решений, ср. [2], гл. 8 и 9, [4], гл. 3. Это реализует *Принцип мониторинга решений*, принятых АСПР (включая *коррекцию* и *предотвращение* путем переключения обратно в фазу обучения АСПР, Рис. 3: динамическое обучение). *Принцип мониторинга*, включающий *коррекцию* и *предотвращение*, реализуется собственным механизмом *адаптации* системы.

Аналогично подсистеме АСПР-ML (см. раздел 3.2), *Принцип мониторинга решений*, принятых АСПР, в свою очередь, является конкретной реализацией общего **Принципа Самосохранения Системы** для подсистемы АСПР-DM, см. сноску 7. Тот факт, что энморфия отношения подсистемы АСПР-DM включает, среди прочего, принцип самосохранения системы, указывает на то, что подсистема АСПР-DM является *квази-стохастической* системой, см. сноски 7 и 22. Однако это вполне объяснимо, поскольку человек, который сам является *квази-стохастической* системой²⁴, является частью подсистемы АСПР-DM.

Проиллюстрируем взаимоотношение первичной системы и метасистемы на примере подсистемы АСПР-DM (подсистемы принятия решений автоматизированной системы принятия решений):

²⁴ см. [5], СНАПТЕР II, ch. 1.2 «Principles of the Enmorphya of Living Beings»

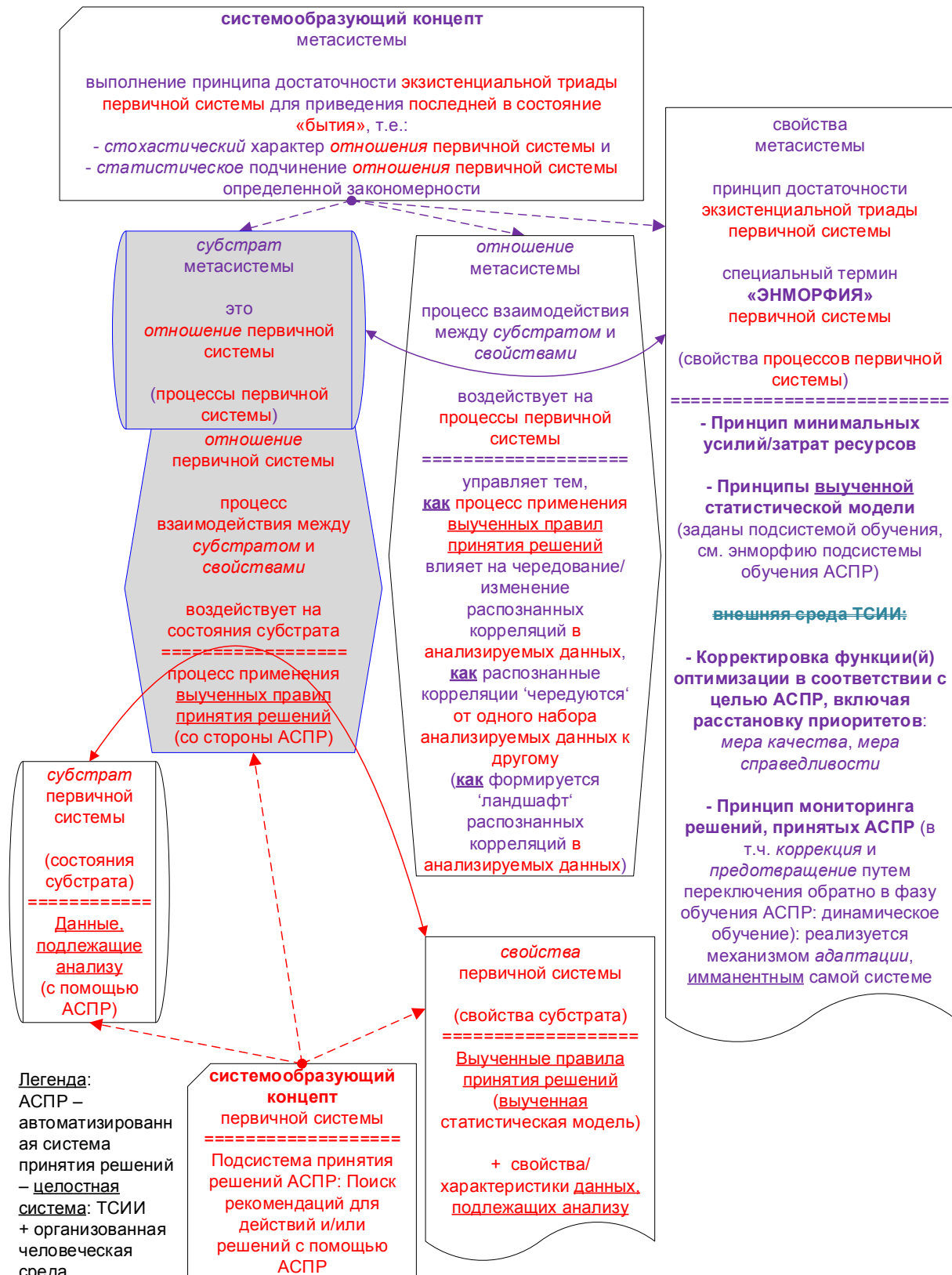


Рис. 5: Взаимоотношение первичной системы АСПР-DM (подсистемы принятия решений автоматизированной системы принятия решений) и соответствующей метасистемы

Если мы сравним соответствующие энморфии отношений четырех подсистем 'слабого ИИ' на Рис. 2 до Рис. 5, мы увидим системно-теоретические различия между этими подсистемами:

- Подсистема ТСИИ-ML (Рис. 2) является *истинно-стохастической* и управляется *принципом учебной экономии*; эта подсистема является чисто технической по своей природе;
- Подсистема АСПР-ML (Рис. 3) является *квази-стохастической* и, таким образом, самодостаточной, поскольку в эту подсистему интегрирован человек (в роли 'аналитика данных');
- ТСИИ-подсистема принятия решений (Рис. 4) является *детерминистической* и управляется *принципами выученной статистической модели*; эта подсистема является чисто технической по своей природе;
- АСПР-подсистема принятия решений (Рис. 5) является *квази-стохастической* и, таким образом, самодостаточной, поскольку в эту подсистему интегрирован человек (в роли 'аналитика данных').

Этот анализ показывает, среди прочего, что чисто технические подсистемы 'слабого ИИ' являются либо *истинно-стохастическими*, либо *детерминистическими*. Техническая система ИИ, состоящая из *истинно-стохастической* подсистемы ТСИИ-ML (подсистемы машинного обучения) и *детерминистической* подсистемы принятия решений АСПИ-DM (подсистемы принятия решений), в целом является *истинно-стохастической*.

Это означает, среди прочего, что 'слабый ИИ' как чисто техническая система не может быть живой системой, см. [5], CHAPTER II, ch. 1 «Enmorphya of Living Beings». Таким образом, 'слабый ИИ' также не может обладать способностью к '*рефлексии рисков*'²⁵. Только системы, способные задавать '*сущностно необходимые вопросы*', т.е. только системы, обладающие *рефлексией рисков*, способны создавать познавательные системы и, таким образом, также способны к познанию смысла²⁶.

Применение этого понимания к 'слабому ИИ' означает:

Утверждение Утв. 1:

'Слабый ИИ' принципиально не способен к познанию смысла (к семантическому познанию).

Этот вывод соответствует накопленному на сегодняшний день опыту. Мануэла Ленцен пишет в [3], гл. 6: «В конечном итоге, эти эксперименты показывают, что алгоритмы не понимают тексты, которые они генерируют»²⁷. Далее она также воспроизводит мысли Джона Серла (John Searle): «Алгоритмы работают с правилами, которые говорят Вам, как туда-сюда перемещать символы, смысла они никогда не постигнут. ... Всё, что имеет отношение к смыслу, говорит Серл, люди сами интерпретируют в рамках обработки данных»²⁸.

²⁵ см. [5], CHAPTER II, ch. 2.2 «Free Will», there the new notion of '*risk reflection*'; см. Глоссарий ниже, '*рефлексия рисков*'.

²⁶ 'Познание смысла' мы синонимически называем '*семантическим познанием*' (содержание) как дополнительное понятие к '*синтаксической обработке символов*' (форма).

²⁷ В оригинале: „Letztlich zeigen diese Experimente, dass die Algorithmen, die Texte, die sie generieren, nicht verstehen.“

²⁸ В оригинале: „Algorithmen arbeiten mit Regeln, die Ihnen sagen, wie sie Symbole hin- und herschieben sollen, Bedeutung werden sie nie erfassen. ... Alles, was mit Bedeutung zu tun hat, so Searle, interpretieren die Menschen in die Datenverarbeitung hinein.“

Жанель Шейн также описывает подобный опыт в ряде мест в [4]: «ИИ не очень понимает проблемы, которые нужно решить» («Введение»); в гл. 5 там же: «Чтобы ИИ смог найти правильное решение, программист должен позаботиться о том, чтобы ИИ действительно работал над правильной проблемой»²⁹.

Вся в целом ‘автоматизированная система принятия решений’ (АСПР) является квази-стохастической только благодаря вмешательству человека ex-ante и ex-post. Человек (в роли ‘аналитика данных’) должен заранее определить все семантические цели, т.е. смысл задач, которые должна решать АСПР, постоянно контролировать их достижение и корректировать ‘функции оптимизации’ и ‘гиперпараметры’ ‘слабого ИИ’.

4 Сильный ИИ

‘Сильный ИИ’ – это автономная техническая система, которая была бы универсально применима с такими же (или более развитыми и обширными) интеллектуальными и творческими способностями, как у человека, см. [2], Глоссарий, [3], гл. 1, [4], гл. 2 (‘AGI’) или в Интернете.

В настоящее время ‘сильный ИИ’ – это только видение. Необходима глобальная и интенсивная общественная дискуссия, которая взвесит пользу и риски ‘сильного ИИ’. Этот аспект подробно обсуждается в [2], глава 11.

Здесь мы задаемся вопросом, какие фундаментальные предпосылки с точки зрения теории систем должны быть выполнены автономной технической системой, чтобы она могла быть универсально применима, обладая такими же (или более развитыми и обширными) интеллектуальными и творческими способностями, как у человека.

Мы установили в [5]³⁰, что отличает человека как систему от всех других живых систем. Очевидный подход к ответу на поставленный выше вопрос заключается в том, чтобы исходить из того, что автономная техническая система должна реализовывать те же системно-теоретические принципы и существенные специфические для человека атрибуты этих принципов, что и сам человек как система.

Системно-теоретические принципы любой системы по определению являются компонентами *энморфии отношения* этой системы, см. [5], CHAPTER I, ch. 3.2 «Enmorphya». Принципы энморфии самосознания живых существ³¹ рассматривались в [5], CHAPTER II, ch. 1.2 «Principles of the Enmorphya of Living Beings»: Это **Принцип Наибольшего Выбора** и **Принцип Самосохранения Системы**, см. Глоссарий. Однако эти два принципа не являются уникальными для живых существ, а присущи любой *квази-стохастической системе*³².

²⁹ В оригинале: „KI versteht die zu lösenden Probleme nicht wirklich“ („Einleitung“); там же в гл. 5: „Damit die KI die richtige Lösung finden kann, muss der Programmierer dafür sorgen, dass die KI auch wirklich das richtige Problem bearbeitet.“

³⁰ [5], CHAPTER II, ch. 2.1 «Living and Non-Living Systems», там в разделе “*The human as a system*”, а также в [5], CHAPTER II, ch. 2.2 «Free Will», там новое понятие ‘*risk reflection*’; см. Глоссарий ниже (‘*рефлексия рисков*’)

³¹ Энморфия самосознания является специфическим воплощением энморфии отношения для живых существ.

³² *Принцип наибольшего выбора* – это специфическая для живых существ реализация общего **Принципа Наименьшего Расходования Ресурсов**, см. [5], CHAPTER II, ch. 1.2 «Principles of the Enmorphya of Living Beings».

Итак, пока заметим себе, что ‘сильный ИИ’ как автономная техническая система должен быть *квази-стохастической* системой³³, т.е. он должен реализовывать оба Принципа – наибольшего выбора (наименьшего расходования ресурсов) и самосохранения системы.

Это также означает, как установлено в [5], CHAPTER I, ch. 3.5 «Enmorphya for Quasi-Stochastic Systems» для любой *квази-стохастической* системы, что ‘сильный ИИ’ как система с необходимостью должен иметь собственный механизм *адаптации*, который реализует Принцип самосохранения системы. Этот механизм *адаптации* включает

- *мониторинг* состояния системы (которое также зависит от условий окружающей среды),
- внутрисистемную *коррекцию* (коррективного действия) по отношению к изменяющемуся состоянию системы и
- *предотвращение* аналогичного ‘неоптимального’ состояния системы³⁴ посредством корректировки соответствующей, имманентной этой системе, ‘нормы’.

Помимо этого, ‘сильному ИИ’, как и любой *квази-стохастической* системе, должен быть присущ *информационный метаболизм*, см. Глоссарий³⁵.

‘Чувство’³⁶ представляет собой компонент *информационного метаболизма*, при этом могут быть ‘негативные’ и ‘позитивные’ чувства: ‘Негативное чувство’ испытывает *квази-стохастическая* система, когда она находится в ‘неоптимальном’ состоянии, т.е. когда стабильность ее *системообразующего концепта* находится под угрозой; ‘позитивное чувство’ испытывает *квази-стохастическая* система, когда она находится в ‘оптимальном’ состоянии, т.е. когда стабильность ее *системообразующего концепта* не находится под угрозой.

Специфические для человека (как биологического вида) атрибуты энморфии самосознания перечислены в [5], CHAPTER II, ch. 1.3 «Variativity and Attributes of the Enmorphya of Self-Awareness». К ним относятся:

- 1) атрибут ‘биологический вид’ со значением ‘homo sapiens’;
- 2) атрибут ‘рефлексия рисков’;
- 3) атрибут ‘этические нормы’ как процедурные нормы, действительные для данной живой системы;
- 4) атрибут ‘модус’ с возможными значениями ‘обыденный (оппортунистический)’ или ‘онтологический (этический)’;
- 5) атрибут ‘психотип’;
- 6) атрибут ‘архетип’.

Какие из этих атрибутов были бы существенны для того, чтобы автономная техническая система, обладающая такими же интеллектуальными и творческими способностями, как

³³ Таким образом, ‘сильный ИИ’ должен быть живой системой, см. [5], CHAPTER II, ch. 2.1 «Living and Non-Living Systems», и ‘*волеобладателем*’, см. [5], CHAPTER I, ch. 3.5.3 «Society», [5], CHAPTER II, ch. 2.2 «Free Will», и Глоссарий.

³⁴ ‘Неоптимальным’ состоянием системы мы называем такое её состояние, которое угрожает стабильности её *системообразующего концепта*.

³⁵ а также [5], CHAPTER I, ch. 3.5.3 «Society», [5], CHAPTER II, ch. 1.1 «Enmorphya of Self-Awareness» и ch. 1.2 «Principles of the Enmorphya of Living Beings»

³⁶ или же ‘настроение’/‘атмосфера’ для организаций как целостных существей

и человек, была бы универсально применима? Чтобы ответить на этот вопрос, мы применим метод исключения.

Мы можем сразу исключить атрибут '*биологический вид*' со значением '*homo sapiens*', потому что автономная техническая система не может быть идентична '*homo sapiens*': Достаточно подумать о том, что автономная техническая система должна иметь материальное тело, отличное от человеческого.

Атрибут '*этические нормы*', которые применяются как процедурные нормы для человека, уже существует в подходах для '*слабого ИИ*': это *мера справедливости*, см. гл. 3 выше. Для '*сильного ИИ*' соответствующая *мера справедливости* определяется '*социализацией*' '*сильного ИИ*'. Этот атрибут является вариативным. Как мы отметили в [5], CHAPTER I, ch. 3.7.2 «Variativity of Quasi-Stochastic Systems», вариативность атрибутов энморфии является существенным отличительным признаком *квази-стохастических* систем от *истинно-стохастических*.

Для '*сильного ИИ*' мы заметим себе, что его энморфия должна включать, среди прочего, по крайней мере, один специфический вариативный *этический* атрибут, который управляет тем, как '*сильный ИИ*' обращается с этическими аспектами в своем общении с окружением и окружающей средой.

Атрибуты '*модус*', '*психотип*' и '*архетип*' – это вариативные атрибуты, которые определяют способ общения человека с окружением и окружающей средой (в самом широком смысле этого слова, т.е. включая взаимодействие)³⁷. Существенно, что эти атрибуты не постоянны, а вариативны.

Для '*сильного ИИ*' мы заметим себе, что его энморфия должна включать, среди прочего, по крайней мере, один специфический вариативный *коммуникационный* атрибут, который управляет тем, как '*сильный ИИ*' взаимодействует со своим окружением и окружающей средой.

Атрибут '*рефлексия рисков*' представляет собой главную отличительную особенность свободной воли человека как вида от свободной воли всех других живых систем³⁸. Поэтому наличие этого атрибута у людей как биологического вида является предпосылкой для интеллектуальных и творческих способностей человека.

Человек, как единственный биологический вид, способен рефлексировать часть своих возможных (будущих) состояний, которые включают в себя как окружающий его мир, так и его самого, включая его собственную конечность как системы³⁸. Отражение собственной конечности как системы вызывает (преимущественно вытесненный) экзистенциальный страх, см. сноску 37, там подраздел «Archetypes» и сноску 38. Экзистенциальный страх побуждает человека заниматься различными видами деятельности, включая интеллектуальную и творческую, чтобы различными способами сделать течение времени (всегда в направлении конца системы) незаметным, см. [5], CHAPTER VI, ch. 3.4 «The Existential Angst and Adaptation». Следует отметить, что страх (и, следовательно, экзистенциальный страх) – это '*негативное чувство*', см. выше.

Эти соображения приводят к следующей цепочке зависимостей: *рефлексия рисков* => экзистенциальный страх (как '*негативное чувство*')³⁹ => мотивация сделать течение времени незаметным => среди прочего, интеллектуальная и творческая деятельность.

³⁷ см. [5], CHAPTER II, ch. 1.3 «Variativity and Attributes of the Enmorphia of Self-Awareness»

³⁸ см. [5], CHAPTER II, ch. 2.2 «Free Will», STM. 24 и '*risk reflection*', а также Глоссарий ниже ('рефлексия рисков')

³⁹ Без рефлексии рисков не может быть экзистенциального страха.

Рефлексия рисков также является необходимой предпосылкой для того, чтобы система вообще могла задавать ‘*сущностно необходимые вопросы*’⁴⁰, то есть создавать познавательные системы, а значит, и быть способной к семантическому познанию⁴¹.

Некоторые исследователи ИИ вводят ‘искусственное любопытство’ в качестве одной из функций оптимизации (в данном случае – функции вознаграждения или оценки), см. [2], гл. 11, [4], гл. 5 (‘Любопытство’). Эта функция вознаграждения предназначена для вознаграждения системы ИИ за нахождение чего-то более абстрактного в ‘модели мира’, уже известной системе ИИ. Намерение и надежда состоят в том, что, благодаря ‘искусственному любопытству’, ‘слабый ИИ’ шаг за шагом возможно разовьется в ‘сильный ИИ’.

Наши рассуждения выше показывают, что ‘искусственное любопытство’ – только лишь как один из нескольких необходимых системных атрибутов ‘сильного ИИ’ – может быть полезным для ‘сильного ИИ’ только в том случае, если это ‘искусственное любопытство’ позволит ‘сильному ИИ’ задавать себе ‘*сущностно необходимые вопросы*’, т.е. вопросы, ответы на которые необходимы для реализации Принципа самосохранения этого ‘сильного ИИ’⁴⁰. **Только если этот ‘сильный ИИ’ будет рефлексировать свою собственную конечность как системы и ощущать ее как ‘негативное чувство’, он будет самомотивирован задавать себе такие ‘сущностно/экзистенциально необходимые вопросы’.** Это означает, что ‘искусственное любопытство’⁴², полезное для ‘сильного ИИ’, должно обладать, по крайней мере, этими двумя очень специфическими свойствами.

Для ‘сильного ИИ’ мы заметим себе, что его эморфия должна включать, помимо прочего, специфический атрибут ‘*рефлексия рисков*’, который мотивирует и в принципе наделяет ‘сильный ИИ’ способностью заниматься интеллектуальной и творческой деятельностью, задавать ‘сущностно/экзистенциально необходимые вопросы’. Для того, чтобы *рефлексия рисков* в принципе могла породить такую мотивацию, ‘сильный ИИ’ должен также испытывать ‘негативные чувства’ по поводу собственной конечности как системы.

Теперь, основываясь на этих выводах, мы можем ответить на вопрос, который задали в самом начале: ‘Какие фундаментальные предпосылки с точки зрения теории систем должны быть выполнены автономной технической системой, чтобы она могла быть универсально применима, обладая такими же (или более развитыми и обширными) интеллектуальными и творческими способностями, как у человека?’

Утв. 2:

Для того, чтобы ‘сильный ИИ’ был универсально применим в качестве автономной технической системы с такими же (или более развитыми и обширными) интеллектуальными и творческими способностями, как у человека, ‘сильный ИИ’ должен – с точки зрения теории систем –

⁴⁰ ‘Сущностно необходимые’ вопросы – это такие вопросы, ответы на которые необходимы для реализации Принципа Самосохранения Системы. В этом смысле ‘*сущностно необходимые вопросы*’ можно также назвать ‘*экзистенциально необходимыми вопросами*’, см. Глоссарий.

⁴¹ Только системы, способные задавать ‘*сущностно необходимые вопросы*’, т.е. только системы, обладающие *рефлексией рисков*, способны создавать познавательные системы и, таким образом, также способны к познанию смысла. ‘Познание смысла’ мы синонимично называем ‘*семантическим* познанием’ (содержание) как дополнительное понятие к ‘*синтаксической* обработке символов’ (форма).

⁴² Более точным термином была бы ‘искусственная любознательность’ вместо ‘искусственного любопытства’.

реализовать следующие имманентные для системы принципы и соответствующие атрибуты⁴³:

- 1) Он должен быть *квази-стохастической* системой, т.е. должен реализовывать оба Принципа – наибольшего выбора (наименьшего расходования ресурсов) и самосохранения системы;
- 2) Он должен с необходимостью иметь свой собственный механизм адаптации, который реализует Принцип самосохранения системы. Этот механизм *адаптации* включает
 - *мониторинг* состояния системы,
 - внутрисистемную *коррекцию* (коррективного действия) по отношению к изменяющемуся состоянию системы и
 - *предотвращение* аналогичного ‘неоптимального’ состояния системы⁴⁴ посредством корректировки соответствующей, имманентной этой системе ‘нормы’;
- 3) Его собственные ‘функции оптимизации’ и ‘гиперпараметры’ должны включать, среди прочего, по крайней мере, один специфический вариативный этический атрибут, который управляет тем, как ‘сильный ИИ’ обращается с этическими аспектами в своем общении с окружением и окружающей средой;
- 4) Его собственные ‘функции оптимизации’ и ‘гиперпараметры’ должны включать, среди прочего, по крайней мере, один специфический вариативный коммуникационный атрибут, который управляет тем, как ‘сильный ИИ’ взаимодействует со своим окружением и окружающей средой;
- 5) Его собственные ‘функции оптимизации’ и ‘гиперпараметры’ должны включать, среди прочего, специфический атрибут ‘*рефлексия рисков*⁴⁵, который мотивирует и в принципе наделяет ‘сильный ИИ’ способностью заниматься интеллектуальной и творческой деятельностью и задавать ‘сущностно/экзистенциально необходимые вопросы’;
- 6) Он должен, помимо прочего, испытывать ‘негативные чувства’ по отношению к своей собственной конечности как системы, чтобы *рефлексия рисков* могла в принципе вызвать мотивацию заниматься интеллектуальной и творческой деятельностью и задавать ‘сущностно/экзистенциально необходимые вопросы’.

Автор этих строк не слышал на момент публикации, чтобы необходимость перечисленных выше имманентных для системы принципов и соответствующих атрибутов как необходимых предпосылок для создания ‘сильного ИИ’ обсуждалась бы в работающих над темой ИИ сообществах.

5 Энморфия ‘самосознания’ ИИ: Критерии различия между типами искусственного интеллекта и перспектива

⁴³ Можно рассматривать принципы и соответствующие атрибуты также как системные ‘функции оптимизации’ и ‘гиперпараметры’.

⁴⁴ ‘Неоптимальным’ состоянием системы мы называем такое её состояние, которое угрожает стабильности её *системообразующего концепта*.

⁴⁵ Т.е. он должен рефлексировать часть своих возможных (будущих) состояний как системы, которые включают в себя как окружающий его мир, так и его самого, включая его собственную конечность как системы.

Теперь мы можем также установить системно-теоретические критерии для различения трех типов систем ИИ, которые мы рассматривали в этой главе.

Экспертные системы, *основанные на правилах*, являются *детерминистическими*. Их поведение и свойства априори определены в соответствующем ‘руководстве по программированию’.

Чисто техническая подсистема ‘*слабого ИИ*’ – состоящая из *истинно-стохастической* подсистемы машинного обучения и *детерминистической* подсистемы принятия решений – в целом является *истинно-стохастической*⁴⁶.

Механизм *адаптации* ‘слабого ИИ’ приводится в действие не технически, а человеком. *Этический* атрибут ‘слабого ИИ’ – *мера справедливости* – также определяется человеком. Аналога *коммуникационного* атрибута ‘слабого ИИ’ я пока еще не видел. Этот атрибут может быть реализован в ‘слабом ИИ’, но *мера качества* для соответствующей функции оптимизации также должна будет управляться человеком.

‘*Сильный ИИ*’, если он когда-то будет создан, должен быть имманентно и по своей сути *квази-стохастическим*, так, чтобы он нуждался в человеке только на этапе проектирования и был бы способен действовать полностью автономно после ввода в эксплуатацию. Это должно включать самостоятельную постановку всех семантических целей, самостоятельный контроль их достижения, а также корректирующую и предупреждающую самокоррекцию своих ‘функций оптимизации’ и ‘гиперпараметров’. Его собственные системные ‘функции оптимизации’ и ‘гиперпараметры’ должны включать следующие специфические вариативные атрибуты: *этический* атрибут, *коммуникационный* атрибут и ‘*рефлексию рисков*’ в отношении его собственной конечности как системы.

Для ‘сильного ИИ’ механизм *адаптации* и все его атрибуты должны быть частью самой системы и работать без вмешательства человека.

Основные различия энморфии отношения (‘*энморфии самосознания*’ систем ИИ, если бы с ними обращались как с живыми существами) между ‘слабым ИИ’ и ‘сильным ИИ’ с системно-теоретической точки зрения мы собрали в обзорной таблице (различия выделены **жирным** шрифтом):

Тип ИИ →	<i>основанный на правилах ИИ</i> (разд. 2)	‘ <i>Слабый ИИ</i> ’ (чисто техническая подсистема) (разд. 3)	‘ <i>Сильный ИИ</i> ’ (если он в будущем будет существовать) (разд. 4)
Свойство ↓			
Тип системы	детерминистическая	истинно-стохастическая	квази-стохастическая
Главные принципы ‘функций’	содержание соответствует	Принцип наименьшего расходования ресурсов	- Принцип наименьшего расходования ресурсов (наибольшего выбора)

⁴⁶ Вся в целом ‘автоматизированная система принятия решений’ (АСПР) является *квази-стохастической* только благодаря вмешательству человека *ex-ante* и *ex-post*. Человек (в роли аналитика данных) должен заранее определить все семантические цели, постоянно контролировать их достижение и корректировать ‘функции оптимизации’ и ‘гиперпараметры’ ‘слабого ИИ’.

Тип ИИ →	основанный на правилах ИИ (разд. 2)	‘Слабый ИИ’ (чисто техническая подсистема) (разд. 3)	‘Сильный ИИ’ (если он в будущем будет существовать) (разд. 4)
Свойство ↓			
оптимизации’	‘руководства по программированию’	(наибольшего выбора)	- Принцип самосохранения системы
Механизм адаптации	не применимо	настраивается и корректируется <u>внешним</u> по отношению к системе человеком	имманентный системе и саморегулирующийся
Этический атрибут	не применимо	<i>Мера справедливост</i> и настраивается и корректируется <u>внешним</u> по отношению к системе человеком.	вариативный атрибут имманентных системе ‘функций оптимизации’ и ‘гиперпараметров’
Коммуникационный атрибут	не применимо	Пока что мы этого не видели, но это было бы применимо и осуществимо; <i>Мера качества</i> для соответствующей функции оптимизации должна настраиваться и корректироваться <u>внешним</u> по отношению к системе человеком.	вариативный атрибут имманентных системе ‘функций оптимизации’ и ‘гиперпараметров’

Тип ИИ →	основанный на правилах ИИ (разд. 2)	‘Слабый ИИ’ (чисто техническая подсистема) (разд. 3)	‘Сильный ИИ’ (если он в будущем будет существовать) (разд. 4)
Свойство ↓			
Рефлексия рисков ⁴⁷	не применимо	не применимо	атрибут имманентных системе ‘функций оптимизации’ и ‘гиперпараметров’; мотивирует и наделяет ‘сильный ИИ’ способностью заниматься интеллектуальной и творческой деятельностью и задавать ‘сущностно/экзистенциально необходимые вопросы’ (см. Глоссарий)
Ощущение ‘негативных чувств’ по поводу собственной конечности как системы ⁴⁸	не применимо	не применимо	результат работы механизма мониторинга в рамках механизма адаптации
Способность к познанию смысла ⁴⁹	отсутствует	отсутствует	принципиально возможно благодаря рефлексии рисков, которая мотивирует и позволяет системе задавать ‘сущностно/экзистенциально необходимые вопросы’

Этот обзор показывает, что основные различия между ‘слабым ИИ’ и ‘сильным ИИ’ лежат – с точки зрения теории систем – в *энморфотипе*⁵⁰ конкретной технологии, реализующей конкретный ИИ.

Наше системно-теоретическое рассмотрение показывает, среди прочего, что ‘слабый ИИ’ и ‘сильный ИИ’ существенно отличаются друг от друга. Оно также показывает, что до создания ‘сильного ИИ’ еще долгий путь, если это вообще станет когда-либо

⁴⁷ Т.е. он должен рефлексировать часть своих возможных (будущих) состояний как системы, которые включают в себя как окружающий его мир, так и его самого, включая его собственную конечность как системы.

⁴⁸ как результат работы механизма мониторинга в рамках механизма адаптации.

⁴⁹ ‘Познание смысла’ мы синонимически называем ‘семантическим познанием’ (содержание) как дополнительное понятие к ‘синтаксической обработке символов’ (форма).

⁵⁰ см. [5], ЧАРТЕР II, ch. 1.4 «Enmorphotype»

возможным. Глобальный этический аспект того, хочет ли человечество вообще такого рода творение, требует глобальной интенсивной общественной дискуссии о пользе и рисках.

6 Глоссарий

Термин	Определение
Основополагающие понятия теории систем по А. Уемову [1], необходимые для чтения этой работы	
система	<p>произвольная вещь, на которой реализуется какое-то <i>отношение</i>, обладающее произвольно взятым определенным <i>свойством</i>.</p> <p>Или эквивалентно:</p> <p>произвольная вещь, на которой реализуются какие-то <i>свойства</i>, находящиеся в произвольно взятом определенном <i>отношении</i>.</p>
системообразующий концепт системы	априорно заданное системообразующее <i>свойство</i> или <i>отношение</i> ; в зависимости от этого, системообразующий концепт является <i>атрибутивным</i> или <i>реляционным</i> , соответственно.
структурный фактор системы	<p>Совокупность свойств и отношений, удовлетворяющая заданному системообразующему концепту.</p> <p>Структурный фактор может быть реляционным (в случае атрибутивного концепта) и атрибутивным (в случае реляционного концепта).</p>
субстрат системы	носитель реляционной или атрибутивной структуры.
Другие основополагающие понятия, необходимые для чтения этой работы	
экзистенциальная триада	<p>набор {<i>субстрат</i>, свойство, отношение}, необходимый для создания системы, базирующейся на этом наборе.</p> <p>Экзистенциальная триада является <u>достаточной</u> для создания системы с соответствующим ей <i>системообразующим концептом</i>, если ‘отношение’ в этой триаде</p> <ul style="list-style-type: none"> - носит принципиально <i>стохастический</i> характер и - <i>статистически</i> подчиняется определенной закономерности (в общем случае ПНР – Принципу Наименьшего Расходования Ресурсов). <p>Эволюция этой системы следует характеру ‘отношения’ в экзистенциальной триаде.</p>
универсальная экзистенциальная пентада	<p>Форма, необходимая и достаточная для описания абстрактной структуры <u>любой</u> системы (и, тем самым, <u>любой</u> наблюдаемой сущности) независимо от содержания и назначения этой системы и принципов, управляющих этой системой.</p> <p>Универсальная экзистенциальная пентада – это вся схема сама по себе, представленная на Figure 1 in [5], Part A, Chapter I ch. 3.2 “Epmoφhua”, т.е все пять элементов схемы и взаимосвязи между этими элементами, а именно:</p> <ul style="list-style-type: none"> - <i>субстрат</i> первичной системы, - <i>свойства</i> первичной системы, - <i>отношение</i> первичной системы = <i>субстрат</i> метасистемы, - <i>свойства</i> метасистемы (<i>энморфия</i> отношения) и - <i>отношение</i> метасистемы. <p><i>Экзистенциальная пентада</i> является <u>универсальной</u> и <u>полной</u>.</p>

Термин	Определение
информация	изменение степени неопределенности
информационный метаболизм	<p>Экзистенциально необходимые прием и обработка сигналов из окружающей среды системой и реакция системы на эти сигналы.</p> <p>Информационный метаболизм свойственен не только человеку, но и любой <i>квази-стохастической</i> системе, так как её механизм <i>адаптации</i> не может работать без обмена сигналами с окружающей средой и их обработки.</p> <p>Концепция ‘информационного метаболизма’ была введена Антонием Кемпинским (Antoni Kępiński) как параллель энергетическому метаболизму организма (Antoni Kępiński <i>Psychopatologia nerwic</i> (Антон Кемпинский <i>Психопатология неврозов</i>), 1972).</p>
сущностно необходимые вопросы	<p>Вопросы, <u>ответы</u> на которые необходимы для реализации Принципа Самосохранения Системы.</p> <p>В этом смысле ‘сущностно необходимые вопросы’ можно также назвать ‘экзистенциально необходимыми вопросами’.</p>
адаптация	<p>корректировка внутрисистемной ‘нормы’ (её изменение, отмена, создание новой) в результате воздействия <i>обратной связи</i>.</p> <p>Механизм <i>адаптации</i> включает механизмы</p> <ul style="list-style-type: none"> - <i>мониторинга</i> состояния системы (которое также зависит от условий окружающей среды), - внутрисистемной <i>коррекции</i> (коррективного действия) по отношению к изменяющемуся состоянию системы и - <i>предотвращения</i> аналогичного ‘неоптимального’ состояния системы посредством корректировки соответствующей, имманентной этой системе ‘нормы’. <p>Эти механизмы имманентны данной системе.</p> <p>‘Неоптимальным’ состоянием системы мы называем такое её состояние, которое угрожает стабильности её <i>системообразующего концепта</i>.</p> <p>Совокупность механизмов <i>мониторинга</i> и <i>коррекции</i> часто называют механизмом <i>обратной связи</i>.</p> <p>Для <i>квази-стохастических</i> систем все эти три механизма существуют и должны быть активны.</p> <p>Для <i>истинно-стохастических</i> систем, которые не обладают <i>долгосрочной памятью</i>, механизм <i>предотвращения</i> не может функционировать, т.к. <i>долгосрочная память</i> необходима для сохранения внутрисистемной ‘нормативной базы’. Поэтому механизм <i>адаптации</i> для <i>истинно-стохастических</i> систем эквивалентен механизму <i>обратной связи</i> (только <i>мониторинг</i> и <i>коррекция</i>).</p>
ресурс (системы)	<p>произведение ‘количества шагов на пути из состояния А в состояние В’ на ‘количество альтернативных решений/возможностей на каждом таком шаге’.</p> <p>Ресурс системы можно абстрактно представить как произведение двух категориально дополнительных понятий:</p> <p style="text-align: center;">«ресурс» = ‘действие’ * ‘выбор’,</p>

Термин	Определение
	<p>см. детали в [5], СНАРТЕР VII, ch. 2.3.2 «Complementary Terms as Resource».</p> <p>Конкретная реализация ‘шагов на пути из состояния А в состояние В’ и ‘альтернативных решений/возможностей на каждом таком шаге’, т.е. конкретная реализация ‘действия’ и ‘выбора’ в каждой системе является специфической и должна быть определена для каждой системы отдельно⁵¹.</p> <p>Например, для физических систем ‘ресурсом’ является количество квантов действия, необходимое для перехода системы в другое заданное макроскопическое состояние⁵²; для коммуникации (включая коммуникативную функцию языка) – количество позиций в сообщении (тексте) * количество различных знаков (например, букв и знаков препинания), необходимых для передачи заданного содержания; для образовательного – да и для любого другого социального процесса – количество отдельных (учебных) тем * количество альтернативных (дидактических) методов, которые необходимо рассмотреть и применить, соответственно, для достижения заданной (учебной) цели.</p>
<p>принцип Наименьшего Расходования Ресурсов (ПНР)</p>	<p>Принцип динамики развития любой системы, заключающийся в том, что некая система при переходе из состояния А в состояние В реализует в статистическом среднем такой способ перехода из А в В, при котором ‘ресурс’ системы расходуется наименьшим образом.</p> <p>ПНР представляет собой универсальную информацию-об-управлении-отношением (т.е. составную часть <i>энморфии</i> отношения) и регулирует процесс взаимодействия между <i>субстратом</i> и <i>структурным фактором</i> <u>любой</u> системы – физической, социальной, коммуникативной и т.д., в основе которой лежит <i>стохастический</i> процесс.</p> <p>В частности, ПНР регулирует процесс взаимодействия между материей и информацией в Природе в форме принципа наибольшей энтропии, который эквивалентен принципу наименьшего действия, см. [5], СНАРТЕР VII, ch. 2.1.5 «The Principle of Least Resources Consumption» и 2.3.2 «Complementary Terms as Resource».</p>
<p>принцип Самосохранения Системы (ПСС)</p>	<p>принцип стабилизации любой системы, заключающийся в том, что <u>отклонение</u> системы от следования Принципу Наименьшего Расходования Ресурсов ограничивается тем, что системообразующий концепт данной системы остается стабильным, сохраняется.</p> <p>Принцип Самосохранения Системы имеет силу для <u>любоx</u> систем, т.е. является универсальной составной частью их <i>энморфии</i> отношения. Для <i>истинно-стохастических</i> систем он выполняется автоматически благодаря их марковскости, которая сама по себе возвращает стохастически ‘выбившиеся’ системы на путь максимальной энтропии.</p>

⁵¹ количество «шагов на пути из состояния А в состояние В» должно быть > 0 , и количество «альтернативных решений/возможностей на каждом таком шаге» должно быть > 1 . Причина этого состоит в том, что Природа должна потратить больше чем ноль ресурсов, чтобы создать наблюдаемое состояние. Для этого природа ‘должна’ сделать по крайней мере 1 «шаг на пути в другое состояние», и «альтернативные решения на каждом таком шаге» не могут быть детерминистическими и, следовательно, количество альтернатив должно быть > 1 ; см. детали в [5], СНАРТЕР VII, ch. 2.1.3, 2.1.4, 2.3.2.

⁵² т. е. физическая величина ‘действие’ ($\text{kg}\cdot\text{m}^2\cdot\text{s}^{-1}$)/ h (постоянная Планка – значение кванта действия)

Термин	Определение
	Для <i>квази-стохастических</i> систем такого автоматизма нет. Поэтому его отсутствие должно компенсироваться явными, данной системе имманентными механизмами, способствующими сохранению этой системы. Такие (системе имманентные) механизмы реализуются через механизм <u>адаптации внутри самой системы</u> .
принцип Наибольшего Выбора	принцип минимизации ограничительных факторов на возможности принятия решений, принцип максимизации свободы выбора. Именно Принцип Наибольшего Выбора как одна из характеристик самосознания живых существ приводит к их гибкости (флексibility), <u>приспособляемости</u> к различным условиям существования.
энморфия ⁵³ чего-либо	специальный термин для понятия 'информация-об-управлении-чем-либо', например, 'энморфия отношения'. Различие понятий 'информация' и 'энморфия' заключается в том, что 'информация' взаимодействует с <u>материальным субстратом</u> , а 'энморфия' – с <u>отношением, процессом</u> между этой 'информацией' и этим материальным субстратом.
стохастический процесс	процесс, каждое следующее состояние которого наступает с какой-либо вероятностью, отличной от 0 и 1.
стохастическая система	система, <i>структурный фактор</i> которой базируется на стохастическом процессе
детерминистический процесс	процесс, каждое следующее состояние которого <u>однозначно определено</u> его настоящим состоянием, т.е. каждое следующее состояние наступает с вероятностью 1. Это значит, что каждое предыдущее состояние процесса также может быть однозначно вычисленно исходя из его настоящего состояния. Если следующее состояние процесса наступает с вероятностью 0, то процесс остановился, больше не существует; он также подпадает под определение детерминистического процесса.
детерминистическая система	система, <i>структурный фактор</i> которой базируется на детерминистическом процессе
марковское свойство (стохастического процесса)	каждое следующее состояние марковского <i>стохастического процесса</i> , реализующего регулярные цепи Маркова, вероятностно зависит <u>исключительно</u> от его актуального состояния и не зависит от его предыдущих состояний. Это свойство можно выразить еще таким образом: прошлое <i>истинно-стохастических</i> , т.е. марковских систем влияет на их будущее исключительно через их настоящее.
истинно-стохастический процесс	<i>стохастический процесс</i> , обладающий <i>марковским свойством</i> . <u>«Истинная стохастичность»</u> состоит в <u>отсутствии непосредственной памяти</u> о предыдущих состояниях.

⁵³ термин 'энморфия' (enmorphia, enmorphia) сконструирован на основе греческого: ἐνμορφία (ἐν-μορφία => (приведение) в-форму, (bringing) in-form)

Термин	Определение
	<p>последующее состояние вероятно зависит только от актуального состояния.</p> <p><i>Энморфия</i> отношения <u>невариабельна</u> (всегда принцип наименьшего действия без вариабельных характеристик).</p>
квази-стохастический процесс	<p><i>стохастический процесс</i>, <u>не</u> обладающий <i>марковским свойством</i>.</p> <p><i>Квази-стохастические</i> системы обладают <u>непосредственной и долгосрочной памятью</u> о предыдущих состояниях.</p> <p><i>Энморфия</i> отношения <u>вариабельна</u> (всегда принцип наименьшего расходования ресурсов с вариабельными характеристиками и принцип самосохранения системы с механизмом адаптации).</p> <p>Н.В.: <i>квази-стохастические</i> процессы <u>не являются детерминистическими</u>.</p>
волеобладатель	<p><u>любая квази-стохастическая</u> система, т.е. стохастическая система со <i>свободой выбора</i>, которая учитывает весь свой предыдущий опыт и имеет механизм <i>адаптации</i>.</p> <p>Другими словами, <i>волеобладатель</i> - это <i>адаптивная</i> система со <i>свободой выбора</i>.</p>
социум	<p>социальная единица (энтитет), группа <i>волеобладателей</i>, социально связанная система взаимодействующих <i>волеобладателей</i>, общество любого размера, удерживаемое вместе любыми внутренними отношениями</p>
категориальные дополненности	<p>Пусть существует ограниченная совокупность (набор) понятий, содержащая более одного понятия. Понятия из этой совокупности назовем <i>категориально дополнительными</i> друг к другу, если:</p> <ol style="list-style-type: none"> 1) эти понятия могут существовать исключительно совместно, т.е. существование любого одного понятия с необходимостью обуславливает существование всех других понятий из набора, и 2) одно понятие из набора невозможно определить через любую совокупность других понятий из этого набора.
атрибутивные противоположности	<p>Пусть существует ограниченная совокупность (набор) свойств, содержащая более одного свойства. Свойства из этой совокупности назовем <i>атрибутивными противоположностями</i>, если каждый член этой совокупности представляет лишь <u>специфическое экстремальное значение одного и того же атрибута</u>, и поэтому может быть определен через другой член этой совокупности.</p> <p>Различая между <i>атрибутивными противоположностями</i> (напр. {высокий, низкий}) и <i>категориальными дополненностями</i> (напр. {форма, содержание}), отметим, что атрибутивные противоположности принципиально не являются категориально дополнительными, т.к. каждый член атрибутивной пары может быть определен через другой член этой пары. Например, атрибут 'размер' может иметь экстремальные значения {большой, малый}; эти значения можно выразить друг через друга.</p> <p>Атрибутивные противоположности всегда описывают свойства/качества, т.е. <u>значения какого-либо атрибута</u> и никогда – понятия. При этом, изменение значения этого атрибута при переходе от одного экстремального к другому происходит без 'скачков', т.е. без изменения степени симметрии (без 'фазовых</p>

Термин	Определение
	<p>переходов II рода'). Атрибутивные противоположности часто подразумевают наличие какого-либо эталона, т.е. 'нормы', относительно которой оцениваются значения соответствующего атрибута (напр. {дорогой, дешевый}, {добрый, злой}). Атрибутивные противоположности практически всегда отражаются в языке антонимичными парами, тогда как <u>категориальные дополнительные далеко не всегда представимы таковыми.</u></p>
время	<p>различимость микросостояний Природы друг от друга <u>является</u> течением времени (т.е. временем самим по себе).</p> <p>Поэтому время дискретно.</p> <p>Различимость состояний является необходимой предпосылкой их наблюдаемости, т.е. их бытия. Поэтому бытие и время взаимнооднозначно связаны. см. [5], СНАРТЕР VII, ch. 1.3 «Time Microstructure».</p>
прошлое	<p>зафиксированная/задокументированная совокупность состоявшихся событий.</p> <p>Поэтому прошлое детерминистично, см. [5], СНАРТЕР VII.</p>
настоящее	<p>принятие решения о выборе следующего состояния из множества возможных состояний.</p> <p>Настоящее переводит вероятностное будущее в детерминистское прошлое. Именно эта дополнительность вероятностного будущего и детерминистского прошлого является причиной <i>необратимости</i> хода времени. см. [5], СНАРТЕР VII</p>
мгновение	<p>теоретическое понятие, описывающее нереализуемое в Природе 'промежуточное состояние', в котором возможность выбора уже существует, а разрешение этой альтернативы еще не существует. Так как время дискретно, никаких 'промежуточных состояний' сущностей быть не может.</p> <p>Такое определение делает 'мгновение', а с ним и настоящее, понятием относительным, а не абсолютным.</p>
будущее	<p>множество возможных состояний.</p> <p>Поэтому будущее вероятно (пробабалистично), см. [5], СНАРТЕР VII.</p>
память	<p>свойство сохранения информации (как рациональной, так и эмоциональной, если применимо к данной системе) на период времени, выходящий за рамки данного состояния системы (<i>мгновения</i>, ситуации), так что эта сохраненная информация может непосредственно влиять <u>на более чем одно</u> последующее состояние (ситуацию) этой системы.</p> <p>Такую память можно назвать также '<i>долгосрочной памятью</i>'. <i>Долгосрочная память</i> является необходимым атрибутом <i>квази-стохастического процесса</i>.</p> <p>В этом контексте '<i>краткосрочная память</i>' - это свойство сохранения информации (как рациональной, так и эмоциональной, если применимо к данной системе) на период времени, не выходящий за рамки данного состояния системы (<i>мгновения</i>,</p>

Термин	Определение
	<p>ситуации), так что эта сохраненная информация может непосредственно влиять не более чем на одно – следующее – последующее состояние (ситуацию) этой системы. <i>Краткосрочная память</i> реализует <i>марковское свойство</i> и является необходимым атрибутом <i>истинно-стохастического процесса</i>.</p>
история	<p>последовательность фаз развития <i>квази-стохастической</i> системы, т.е. <i>волеобладателя</i>, к которому эта ‘история’ относится. Полная история <i>волеобладателя</i> включает в себя полный цикл развития соответствующей <i>квази-стохастической</i> системы от ее возникновения до саморазрушения. Этот полный цикл развития существует для <u>любой квази-стохастической</u> системы.</p>
пространство	<p>дискретный <i>субстрат</i>, необходимый для <i>различения материальных</i> сущностей, см. [5], ЧАРТЕР VII, ch. 3 «Space Microstructure».</p>
энморфотип (живых сущностей)	<p>совокупность всех атрибутов ‘<i>энморфии</i> самосознания’ живых сущностей, взаимодействующих как с его генотипом, так и с его фенотипом.</p>
свободная воля	<p>Свободная воля является свободой выбора, имеющей недетерминистический характер, но не представляющей собой марковский процесс, и учитывающей, по крайней мере, весь предыдущий опыт системы.</p> <p>Т.е. это определенная свобода выбора, возможность локального отклонения <i>квази-стохастического</i> процесса от следования Принципу Наименьшего Расходования Ресурсов.</p> <p>Процесс принятия решений.</p>
рефлексия рисков (только человеком) (неопределенность возможного (будущего))	<p>Включение в принятие решений, т.е. в свободу выбора <u>человека</u>, саморефлексии возможных будущих состояний, которые включают как окружающий человека мир, так и его самого, в том числе его собственную конечность как системы.</p>

7 Ссылки

- [1] Авенир Иванович Уемов *Системные аспекты философского знания*, Одесса, 2000
- [2] Katharina Zweig *Ein Algorithmus hat kein Taktgefühl*, Heyne, 2019, ISBN 978-3-453-20730-1⁵⁴
- [3] Manuela Lenzen *Künstliche Intelligenz*, C.H.Beck, 2020, ISBN 978-3-406-75124-0⁵⁵
- [4] Janelle Shane *You Look Like a Thing and I Love You*, 2019 Janelle Shane, ISBN 978-1-47226-899-0
- [5] Igor Furgel *Being and Systemacy*, Westarp BookOnDemand, 1. Auflage 2022, ISBN: 978-3-96004-131-3

8 Благодарности

Я хотел бы выразить мою глубокую благодарность моей жене Ирине за наши чрезвычайно полезные и интересные дискуссии по отдельным аспектам этой темы. Не менее глубокую благодарность я хотел бы выразить моему университетскому профессору по философии Авениру Ивановичу Уемову за его неоценимое участие в формировании моего стиля взаимодействия с миром.

⁵⁴ Катарина Цвейг *У алгоритма нет чувства такта*, Хейне, 2019

⁵⁵ Мануэла Ленцен *Искусственный интеллект*, С.Х.Бек, 2020