

# МОДЕЛИ ОРГАНИЗАЦИОННО-ЗАМКНУТЫХ СИСТЕМ И КОНТУРЫ РАЗВИТИЯ НОВЫХ ПОДХОДОВ В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И КОГНИТИВНОЙ НАУКИ

## Часть I

**А.Б. Казанский**



**А.Б. Казанский**, Научно-исследовательское предприятие РАН "Институт эволюционной физиологии и биохимии им. И.М. Сеченова" РАН, с. н. с., канд. биол. наук  
kazansky.a@yandex.ru

*В работе с общих позиций анализируются известные теоретические модели организационно-замкнутых систем в биологии и кибернетике. Такие системы способны к саморазвитию путем «самораскрутки», многоуровневого мета-рекурсивного процесса, называемого часто бутстрапированием ("bootstrapping"). В класс подобных «бутстрап-систем» в частности, вписывается понятие биосферы, биологического организма, экологической и социальной системы, самозагружающегося программного обеспечения компьютера, сети Интернет, эволюционно самообучающегося робота или коллектива роботов.*

*Рост интереса к подобным процессам и системам делает актуальной задачу поиска языка для их адекватного описания, формализации, классификации и построения общего аналитического аппарата.*

*Ключевые слова: бутстрап, автопоэзис, кибернетика второго порядка, голономия.*

### Введение

В биологии и кибернетике широкую известность приобрела системная модель биологической автономии как самопродуцирующейся системы, «автопоэзис» У. Матураны и Ф. Варелы [1, 2], которую авторы провозгласили базовой в понимании живого во всех его аспектах. Не менее популярны близкие ей теоретические модели самовосстанавливающегося метаболизма выдающегося представителя теоретической биологии Р. Розена [3] и «гиперциклов» лауреата Нобелевской премии М. Эйгена. Из биологии и химии эти системные модели перекочевали в социологию благодаря общеизвестным работам Н. Лумана а также легли в основу так называемой «кибернетики второго порядка», или кибернетики наблюдателя [4, 5]. Все эти теоретические построения так или иначе основаны на модели сложной, организационно-замкнутой автономной системы, способной в процессе взаимодействия со средой к спонтанному саморазвитию, «самораскрутке», тому, что в англоязычной литературе часто именуют термином «бутстрапирование» ("bootstrapping").

Универсального и строгого определения понятия *бутстрап* не существует. Бутстрап (англ. "bootstrap") – старое английское слово, означающее кожаные пельтики, «ушки» на задниках походных или рабочих ботинок. В XIX веке в Англии стала популярной поговорка: "to pull himself up from the swamp by his bootstraps" («вытащить себя из болота за ушки на задниках ботинок»). Смысл поговорки означает: «Выбиться в люди благодаря собственным усилиям». Собственно, термин *бутстрапирование* как раз и выражает идею самостоятельного, без участия внешних управляющих воздействий эмерджентного развития, идею «самораскрутки». Эта способность обеспечивается, с одной стороны, особой динамической организацией систем данного типа (назовем их условно «бутстрап-системами») с непрерывным процессом самоконструирования и самоидентификации путем постоянного самообновления элементов и структуры, а с другой стороны, – наличием особой «творческой» внешней среды – метасистемы (см. ниже), с которой каждая такая система взаимодействует особым образом. Бутстрап-системы могут легко *самомодифицировать*

ся, эволюционировать как по дарвиновской схеме (эндогенная изменчивость – наследование – отбор средой наиболее приспособленных вариантов), так и непрямым способом, путем активного влияния на локальную среду, ее селекционные свойства («конструирование ниш», «стигмергия» – см. ниже, раздел 5).

Есть одна характерная отличительная черта в организации и поведении бутстрап-систем: все они в той или иной мере демонстрируют идею нетрадиционной, «коллективной» формы элементарности, связанности, голографичности. В определенном смысле их можно назвать *рефлексивными*, по крайней мере, с формально-логической точки зрения [6]. Устройство таких систем выражает парадоксальную буддийскую идею «все во всем», «часть, содержащая целое». Это означает полную самосогласованность, взаимообусловленность частей, неклассическую (коллективную) концепцию элементарности, когда каждая часть составлена из всех остальных («взаимопроникновение»). Реализуется такая система за счет своего рода рекурсивности в ее организации. При этом отвергается классическая механистическая форма элементарности в виде линейной иерархии системных уровней, кончающейся внизу системой неделимых фундаментальных сущностей. В физической бутстрап-модели мира фактически не существует выделенных элементарных неделимых абсолютных сущностей. Эту идею ярко выразил Фритьоф Капра [7], давно пропагандирующий идею адронного бутстрапа физика Дж. Чу: «Вселенная рассматривается в качестве сети взаимосвязанных событий. Ни одно из свойств того или иного участка этой сети не имеет фундаментального характера; все они обусловлены свойствами остальных участков сети, общая структура которой определяется универсальной согласованностью всех взаимодействий» [7]. Выдающийся физик Д. Бом [8] называл данное свойство «*голономия*» (не путать с известным математическим понятием), когда в каждом фрагменте реальности в явном или неявном виде содержится информация обо всех других фрагментах. Даосский миф о чудесном ожерелье бога Индры, в каждой бусинке которого отражаются все остальные бусинки, хорошо выражает идею физического бутстрапа. Комната с зеркальными стенами – еще один часто используемый образ. Таким образом, физическую бутстрап-модель Дж. Чу правильнее было бы вслед за Д. Бомом назвать моделью физической голономии.

Другая, динамическая, процессуальная трактовка понятия *голономия* означает буквально "коллективное воспроизводство", когда все элементы непрерывно участвуют в производстве друг друга, коллективным воспроизводстве системы связей, ответственных

за это производство и часто даже границы, отделяющие систему от окружающей среды. Такой логически замкнутый круг причинности демонстрирует упомянутый выше *автопоззис*, или минималистская модель живого организма, как биологической автономии, обобщенного метаболизма (см. следующий раздел).

Несмотря на грандиозные замыслы отдельных ученых и мыслителей, холистическая модель мира в трактовках Дж. Чу, Д. Бом, Н. Бора, В. фон Гейзенберга и др. пока не стала ведущей парадигмой в физике и не привела к третьей революции. Хотя кварки, как физические элементы мироздания, и ведут себя необычно, так называемая стандартная модель в целом находится в рамках классической парадигмы элементарности.

Совсем иная картина наблюдается в биологии, экологии, науках о Земле, социологии, теории когнитивных систем, социальной и эволюционной психологии, в области искусственного интеллекта. В данных областях происходит «анти-картезианская» революция, поэтому они нуждаются в принципиально новом формализме. Релятивизм индивидуальной и социальной (биосоциальной) картины мира, зависимость их от положения наблюдателя, совместное творение («сотворение») социального (биосоциального) мира делают позицию абстрактного отстраненного наблюдателя в ряде случаев нереальной.

По мнению автора, *бутстрап*, это довольно расплывчатое понятие, весьма широко и некритически используемое в мировой научной и технической литературе, требует развития, уточнения и формализации. Это понятие выражает суть поведения многих сложных систем в биологии, физике, социологии, кибернетике, информатике и может дать имя целой объединяющей системной концепции. Развитая концепция бутстрапа поможет объединить такие подходы к анализу сложных явлений, как синергетика, автопоззис, киберсемиотика, теория когнитивных систем и переосмыслить понятие автономии.

Скромная цель настоящей работы заключается в том, чтобы обратить внимание на феномен бутстрап-систем, в первую очередь, на примере биологии, кибернетики и информатики. Рост интереса к подобным процессам и системам делает актуальной задачу поиска языка для их адекватного описания, формализации, классификации и построения общего аналитического аппарата их исследования. После изложения и критического анализа моделей саморазвивающейся биологической автономии и связанных с ними концепций в различных областях, делается попытка выявить некоторые характерные свойства подобных процессов, наметить некоторые пути к их формализации и моделированию.

## 1. Теоретические модели самопродуцирующейся автономии в биологии и химии

Автономия – это греческое слово, означающее буквально «само-законие», само-основание (“self-grounding”), самодостаточность, замкнутость на себя. В мифологическом сознании подобные свойства демонстрировали не только боги, такие, например, как бог Солнца Атум в Египте, который сам себя оплодотворил, проглотив свое семя. Так, птица Феникс, сжигает себя в гнезде и возрождается вновь из пепла, а змея Уроборус, кусает свой хвост.

В современной философии идея автономии ярче всего представлена в онтологическом монизме М. Хайдеггера, его экзистенциальной феноменологии самообосновывающегося и самораскрывающегося человеческого бытия – *Dasein*. Правда, что важно в контексте настоящей статьи, этот главный философ двадцатого века изначально определил человеческое бытие через экзистенциалы совместного приобщения к бытию (*Inder-Welt-sein*) и бытие-с-другими, совместного бытия, *Mit-Sein*, что в чем-то созвучно идее совместного эмерджентного развития, коллективного бутстрапа (“mutual bootstrapping”), отстаиваемую автором (см. Заключение). Довольно парадоксальные идеи фундаментальной онтологии М. Хайдеггера в настоящее время привлекают внимание теоретиков искусственного интеллекта и когнитологов. Более того, некоторые социобиологи пытаются адаптировать эту онтологию культуры применительно к биологии и искусственному интеллекту (см. часть II, раздел 7).

В постмодернистских «текстах» емкое понятие «ризомы», введенное в философию Ж. Делезом и Ф. Гваттари во второй части их фундаментального труда «Капитализм и шизофрения» в системном плане перекликается с моделью автопоэзиса, речь о которой пойдет ниже.

Натурфилософия и биология давно беременны идеей организма как продукционного бутстрапа, или обобщенного метаболизма, когда компоненты целостной системы, организма участвуют в устойчивом непрерывном процессе взаимного производства и таким образом одновременно отстаивают свою автономию, постоянно самоидентифицируются и порождают внутренний контекст и смысл. Фактически, автономия, в данном случае, есть конкретизация слишком абстрактного и недостаточно конструктивного понятия «целостность». По крайней мере, идея самодостаточной и самоподдерживающейся автономии интересовала Платона, встречается в последнем трактате Ч. Дарвина о роли дождевых червей в поддержании плодородия почвы, на словесном уровне формализована И. Кантом и основателем биосемиотики фон Иксюлем. Весьма яркий философ и мыслитель

Ганс Йонас, один из слушателей курса М. Хайдеггера во Фрайбурге, посвятил целый трактат философии жизни, делая акцент на подобной модели метаболизма. Русский ученый А.А. Богданов в своей «Тектологии», пожалуй, впервые попытался формализовать понятие самопродуцирующейся организации с позиции теории систем. В начале семидесятых эти идеи получили новое формальное выражение в развернутой концепции автопоэзиса Умберто Матураны и Франциско Варелы.

В узком смысле слова, *автопоэзис* – это минималистская базовая теоретическая модель жизни, живого организма, представленная формально в виде самопродуцирующейся автономной системы, то есть как самоорганизации особого типа. Развивая и формализуя подход У. Матураны, авторы пришли к определению *автопоэтической системы*, формальной теоретической модели живого организма [1].

*Автопоэзис (autopoiesis)* – неологизм, термин, составленный авторами из двух латинских слов древне-греческого происхождения: “auto” – «сам» и “poiesis” – «строить, творить, производить». Он означает динамическое «самопроизводство» в смысле самопроизводство, строительство себя самого внутри себя (не следует путать с самовоспроизведением копии себя вовне при саморепликации, размножении), причем процесс продуцирования реализуется самими элементами, воспроизводящими коллективным образом себя, продукционные связи и связи связей – *автопоэтическую организацию*. Последняя является инвариантом, в отличие от *структуры* системы, которая может подвергаться серьезным изменениям в процессе взаимодействия со средой.

Автопоэзис находится в диалектическом единстве со своей противоположностью – *аллопоэзисом*. Аллопоэтическая система производит нечто «на экспорт», не имеющее отношение к собственным компонентам. Она не способна самовоспроизводиться.

Фактически, автопоэзис – это самоорганизующийся продукционный процесс особого рода, позволяющий поддерживать автономию системы, выделять себя из среды (самоидентифицироваться) в процессе особого взаимодействия с ней (так называемого «структурного сопряжения»). Он реализуется рекурсивным образом – в процессе реализации самих продукционных процессов. Этот процесс является базовым инвариантом, гарантирующим сохранение автономности системы. Он определяет характер взаимодействия системы с окружающей средой, ее эволюцию и адаптивное поведение, которые (в случае сохранения автопоэтической организации) считаются когнитивными.

Таким образом, в автопоэтической системе *организация* (комплекс взаимоотношений между компонентами и производящими компоненты процессами)

и структура (конкретное пространственно-временное размещение конкретных материальных компонент в существующем физическом пространстве) выступают как особая комплиментарная пара. Организация должна учитывать специфическую (динамическую) конфигурацию компонент, которая определяет систему как класс, а структура должна учитывать, как ее конкретные компоненты вступают в данные взаимоотношения, которые конституируют ее.

Ф. Варела дает следующее упрощенное определение автопоэтической системы [2, р. 13].

*«Автопоэтическая система организована (определена как единство) в виде сети процессов (продукции, синтеза, трансформации и деструкции) компонент, которые производят компоненты, которые: (1) – путем взаимодействия друг с другом и трансформации непрерывно регенерируют и реализуют сеть процессов (отношений), которые производят их; и (2) – конституируют ее (автопоэтическую машину) как конкретное единство в пространстве, в котором они (элементы) существуют путем определения замкнутой топологической области ее реализации как такая сеть».*

Из этого следует, что автопоэтическая машина постоянно генерирует и определяет свою собственную организацию через свою работу как системы по производству своих собственных компонент, и делает это в бесконечном круговороте (обороте) компонент при условиях постоянных возмущений, вызванных воздействием среды и компенсации этих возмущений. Поэтому автопоэтическая машина – это гомеостатическая машина (скорее, статическая в плане системы отношений, а не состояния (relation-static)), система, которая имеет свою собственную организацию, определенную как сеть связей отношений в качестве фундаментального инварианта. Любое единство, объект имеет собственную организацию, определяемую как сеть отношений между элементами, процессами или теми, и другими. Среди всех этих возможных случаев автопоэтические машины – единства, чья организация определяется специфической автопоэтической сетью процессов (отношений) продукций компонент, а не самими компонентами или их статическими отношениями.

Таким образом, реализация автопоэтической организации – результат ее собственной работы. В такой системе компоненты и организация одновременно являются продуцентами и продуктами. Иначе говоря, автопоэтическая организация рекурсивна и автореферентна (в процессе работы обращается сама к себе). Она реализует вариант бутстрапа или эмерджентное саморазвитие путем постоянного самоконструирования, которое является особой формой его функционирования.

Следует сразу оговориться, что это так называемое «неконструктивное» определение, оно не опера-

ционально, слишком схематично, что характерно для определений в духе Общей Теории Систем. В отличие от строгих математических определений, оно неполно и содержит неявные допущения (тип топологии и др.). Как пишет в своем обзоре Мак Малин, за тридцать лет так и не удалось смоделировать конкретную модель типа клеточного автомата, полностью согласующуюся с этим определением. Собственно говоря, настоящего формализма автопоэзиса не построено. То, что сделали У. Матурана и Ф. Варела – это лишь гипотетические наброски к созданию теории и настоящего формализма биологической автономии.

Близкий по смыслу подход к формализации понятия биологической автономии и биологического организма предложил Роберт Розен [3]. Это так называемая модель (M-R System) – системы, демонстрирующей свойства самовосстановления благодаря замкнутому на себя метаболизму (Metabolism-Repair System). По определению Розена, материальная система является организмом, если она замкнута по отношению к так называемой «эффективной причине», названной так еще Аристотелем. В этом определении эффективная причина аналогична производственным правилам, использованным Матураной и Варелой, то есть мы имеем здесь организационное замыкание по продукции. Базовая идея Розена состоит в том, что эффективная причина, например, автомобиля или дома – это его строитель или производитель, а вот в случае с живыми организмами сама система является собственным строителем. Подобная система, как и в случае с автопоэзисом, организационно полностью замкнута. Правда, понятие эффективной причины более операционально и больше подходит для моделирования.

Автономия требует определенной способности к самомоделированию (автомодельность). Из схемы Розена следует, что моделирование – тоже замкнутый на себя процесс в живой системе. Но даже если система предрасположена к моделированию себя, как замкнутой достаточной причине, из этого отнюдь не следует, что это моделирование отражает реальность или что оно не ведет к порочному логическому кругу. Самовосстанавливающийся метаболизм укореняет замыкание эффективной причинности в чисто материальной причинности, за что его и критикуют.

Слабость всех концепций биологической автономии как раз заключается в полной организационной замкнутости. Знаковая, семиотическая составляющая связи со средой возникает эмерджентно из чисто невербального способа взаимодействия и общения с внешним миром как у Ж. Пиаже в его генетической психологии. Согласно этим представлениям, коэволюцию системы и среды можно свести к последовательности взаимных возмущений и изменений гра-

ницы, то есть формы («структурное сопряжение» в автопоэзисе). Неудивительно, что реально автопоэзис стимулирует развитие теории морфодинамики. На самом деле, взаимоотношение организма и среды нельзя полностью свести к подобной схеме эволюции взаимных возмущений. Организм познает себя через мир и через активные действия в мире. Кроме того, внутри достаточно сложного организма знаковые процессы просто необходимы для его выживания и поддержания автономии. Так, Г. Патти обосновывает неизбежность исходного дуализма знаковой и материальной составляющих в организме и сложных кибернетических системах. Это дало основание говорить о появлении новой области – киберсемиотики.

Разделение в автопоэтической системе продукционной «организации», как метаструктуры, самосохраняющегося инварианта внутри структуры и вариабельной части структуры (названной просто «структура»), весьма проблематично. Опять-таки, это разделение декларировано, но возможность реализации такого разделения, как впрочем, и реализуемость автопоэтической системы в полной мере не доказаны, несмотря на более чем тридцатилетние усилия. В конце концов, автопоэтическая организация как устойчивая метаструктура и изменчивая во времени транзитная структура относятся к разным логическим типам, и их диалектическое сопоставление некорректно. При реализации бутстрапа явно происходит порождение новой информации, нового смысла. Механизм порождения нового не раскрывается в первоначальном варианте чисто механической модели автопоэзиса. Позднее была предложена концепция *инактивизма, или воплощенного действия* (см. часть II, раздел 7).

Как уже отмечалось, основное отличие автопоэзиса от различного рода автокаталитических моделей заключается в дополнительном требовании создавать и поддерживать топологически замкнутую границу, отделяющую систему от среды. Возможно, это требование также делает проблематичной задачу модельной реализации автопоэзиса и вообще ставит под вопрос возможность такой реализации. По-видимому, стоит различать границу метаболизма и когнитивную границу, связанную с поведенческой активностью организма в среде (см. часть II, раздел 7 о концепции расширенного разума).

Наконец, У. Матурана и Ф. Варела попытались подобно М. Хайдеггеру построить своего рода фундаментальную онтологию жизни, где роль *Dasein* играет то, что по аналогии можно было бы назвать *Dalieben* («здесь-жизнь») или *автопоэтический наблюдатель*. Некоторые критики отмечают, что эта попытка вырваться из оков картезианского дуализма на деле приводит даже не к удвоению, а учетверению реальности.

Все сказанное не принижает большую методологическую ценность автопоэзиса и других чисто материальных механистических концепций биологической автономии в кибернетике, теории систем, когнитологии, психологии, социологии и в биологии, как критической теории, давшей импульс для появления новой парадигмы, которая в настоящее время находится в стадии формирования.

Кроме самовосстанавливающегося метаболизма Роберта Розена автопоэзису близка модель так называемой самомодифицирующейся компонентной системы Камписа [9]. Правда, Розен и Кампис не предлагают, как в случае автопоэзиса, новой экзистенциальной феноменологии наблюдателя и не приравнивают познание к самоконструированию, как это делает У. Матурана. Кроме того, только в автопоэзисе придается важное значение границе системы (т.е. форме), как элементу самовосстанавливающейся структуры. Все три системно-теоретических подхода по-настоящему еще не формализованы и даже не полностью реализованы в моделях. В то же время, менее теоретически нагруженные модели коллективного молекулярного автокатализа вполне реализуемы [10].

## 2. Кибернетика второго порядка и «собственное поведение»

В начале своей известной монографии по теории биологической автономии Ф. Варела [2] четко определяет область применения этой теории. Теория автономии дополняет классическую теорию управления на случай взаимоотношения адаптивных коэволюционирующих самообучающихся и развивающихся систем, демонстрирующих автономные и когнитивные свойства. Невозможно однозначно количественно прогнозировать реакцию таких систем, своего рода «вещей в себе», на наши действия в условиях существенной неопределенности в знании их структуры и поведения. В соответствии с принципом *структурной детерминации*, реакция каждой автономной системы на внешние воздействия определяется ее внутренней структурой, при этом универсальный смысл, единый контекст отсутствуют, но он и необязателен для сосуществования по типу взаимной дополненности. В данном случае, управление в классическом смысле, как в физической механике, когда структура системы известна и существует единый смысл для всех взаимодействующих систем, невозможно. Тем не менее, как ни странно, в повторяющемся процессе рекурсивно организованных сенсомоторных актов взаимных возмущений и реакций на возмущения возникают инварианты во взаимоотношениях таких автономных систем. Происходит взаимное обучение (коэволюция), приводящее часто к притирке, согласованию индивидуальных моделей поведения, коа-

даптации, конгруэнции, сопряжения систем без необходимости введения универсальных единых смыслов, понятий и моделей. Так, «вещи в себе» с очень разными внутренними мирами, разной логикой и картиной мира вполне могут сосуществовать без особых конфликтов. Взаимоотношение людей – наглядный тому пример. Основатель кибернетики второго порядка (кибернетики наблюдателя) Гейнц фон Фёрстер назвал эти инварианты «собственное поведение» ('eigenbehavior') по аналогии с давно известным в математике понятием «собственный вектор» ("eigenvector") [11]. Фактически, собственное поведение – это неподвижная точка оператора, остающаяся неизменной при его действии. Фон Фёрстер постулировал существование подобных решений для уравнения, описывающего рекурсивную структуру механизма порождения смысла наблюдателем при оценке взаимоотношения субъекта и объекта в теории генетической психологии Жана Пиаже [12]. Согласно этому уравнению, то, что наблюдается в какой-то определенный момент времени  $t$ , то есть  $obs(t)$ , есть результат действия когнитивной сенсомоторной операции (**COORD**) на результат предыдущего наблюдения,  $obs(t-1)$ :

$$obs(t) = COORD(obs(t-1)).$$

Особенностью этого уравнения является то, что оно рекурсивно и не содержит начальной точки или начальных условий. Таким образом, любой конкретный акт наблюдения является результатом неопределенной последовательности когнитивных операций. Бесконечность этого процесса есть математическое выражение той конструктивистской позиции, согласно которой наблюдаемые образы (феномены) не относятся непосредственно к реальным объектам мира (ноуменам), а лишь являются результатом бесконечного каскада когнитивных и сенсомоторных операций, представленных оператором **COORD** в определенном процессе *структурного сопряжения* субъекта и среды. Как уже упоминалось, понятие «структурное сопряжение» ("structural coupling") детально анализируется в концепции автопоэзиса и описывает взаимодействие автопоэтических систем по принципу комплиментарности, взаимного дополнения. Как предполагается, «решения»  $O(i)$  этого уравнения, которые не существуют в строго математическом смысле, поскольку отсутствуют начальные условия, выражают устойчивость к действию цепи последовательных операций **COORD**. То есть эти решения дают те значения, которые сохраняют свою структуру (действие, или функцию), несмотря на воздействие на них когнитивных/сенсомоторных операций, воздействующих на них снова и снова в неопределенном рекурсивном процессе:

$O(i) = obs(t) = COORD(COORD(...COORD(obs(t-n))))...$ .  
Другими словами, для длинной последовательности

когнитивных операций структура  $obs(t)$  не изменяется (устойчивость к воздействию операций). Когда это достигается,  $obs(t)$  называется «собственным поведением» и представлено  $O(i)$ . Можно сказать, что собственное поведение обладает самоопределением, или автореферентностью (самоссылочностью) в границах своей устойчивости к операции **COORD**, подразумеваемая комплиментарные отношения (кругообразность, замкнутость) между собственными поведением и когнитивными/сенсомоторными операторами. Одни подразумевают или определяют другие и наоборот.

### 3. Эпистемология конструктивизма и бутстрап – модель представления знаний

Традиционный искусственный интеллект (GOFAI – "Good Old-Fashion Artificial Intelligence") придерживается эпистемологии соответствия или корреспонденции, которая видит знания как простое отображение или «отражение» внешнего мира. Предполагается, что каждый концептуальный объект (символ) в познавательной модели субъекта соответствует физическому объекту в окружающей среде. Структуру модели можно рассматривать как гомоморфный образ, карту или кодирование структуры внешней реальности.

Такая эпистемология «отражения – корреспонденции» приводит к практическим и концептуальным проблемам, связанным с отображением внешней реальности в символическое представление. Так как когнитивная система не имеет прямого доступа к реальности, кроме как через восприятия посредством органов чувств, которые есть уже внутренние модели, – как мы убедимся, что используем верные отображения? Или, в другой формулировке, это проблема, которая известна как "symbol grounding" (укорененность символа): как символы, элементы модели «укоренены» во внешней реальности, которую они предполагают отражать?

Проблема не может быть решена с помощью одной только модели. Это следует из принципа «лингвистической комплиментарности», который обобщает классические логические и эпистемологические ограничения, такие как теорема Геделя или принцип неопределенности Гейзенберга. Согласно данному принципу, ни один язык не может полностью описать сам себя, дать свое собственное описание или описание процессов интерпретации. Другими словами, модели не могут включать в себя представление об отображении, которое связывает их символы с их интерпретацией.

Модели традиционного ИИ всегда неполны и содержат произвольные допущения. Проблема самоопределения обобщает знаменитую проблему границы ("frame problem") в ИИ. Если бы модель непрерывно адаптивно изменялась, то проблему еще можно было бы как-то разрешить, но философия корреспонденции

истины не позволяет этого. Новые символы нельзя вывести на основе уже существующих, приходится вводить их произвольно. Модели ИИ статичны и произвольны. Границы их применимости трудно обозначить. Они не универсальны и эффективны только для случаев хорошо формализованных и структурированных задач.

Согласно альтернативной, конструктивистской эпистемологии, знания - суть не пассивные отражения внешней реальности, а активные конструкции самого субъекта. Предполагается, что конструкция не отражает внешнюю реальность, а помогает объекту адаптироваться или подстроиться к реальному миру.

Это означает, что субъект попытается построить модели, которые будут когерентны с моделями, которыми он либо уже обладает, либо получает с помощью своих органов чувств, или, используя средства коммуникации, получает от других. Поскольку модели только сравниваются с другими моделями, отсутствие доступа к внешней реальности более не является, как в случае эпистемологии корреспонденции, препятствием для их развития. В такой эпистемологии знания обосновываются или объявляются «верными» не на основе соответствия внешней реальности, а на основе соответствия, когерентности с другими частями знания. Проблемой остается определение того, что понимать под когерентностью. Простое соответствие здесь явно недостаточно, как, например, в случае несвязанных данных.

Построение модели – это процесс проб и ошибок с выработкой вариантов и селекцией только тех из них, которые согласуются с остальным эмпирическим материалом. Вместо того, чтобы инструктировать субъект, как ему строить модель, (внутренняя или внешняя) среда помогает отобрать наиболее подходящие варианты модели среди тех, что сгенерировал сам субъект.

Существуют различные оттенки конструктивизма. Так, социальный конструктивизм подчеркивает роль социальных взаимодействий в селекции моделей (консенсус в обществе).

Психологи Глазерсфельд и Пиаже подчеркивают роль индивида, который пытается найти когерентность между своими собственными различными моделями и восприятиями (интернализм).

Селекционисты, вдохновленные эволюционной эпистемологией и концепцией фальсифиционизма К. Поппера, подчеркивают роль внешней реальности в отбраковке неадекватных моделей (экстернализм). Последний подход, естественно, популярен среди ученых – естественников.

Выдающийся кибернетик Гордон Паск в «Теории Разговора» [13] детально описал механизм процесса селекции когерентной истины. Преимущество эпистемологии когерентности состоит в том, что она не

нуждается во внешнем обосновании понятия для построения моделей: когерентность – это двустороннее отношение. Другими словами, когерентные понятия поддерживают друг друга. Динамику, эквивалентную этой общей поддержке отношений, часто также называют «бутстрапированием». Модель **A** может быть использована при конструировании модели **B**, в то время как **B** может быть использована при конструировании **A**. Постепенно происходит своего рода рекурсивный синтез. Получается, что более сложное значение возникает – «бутстрапируется» – из менее сложного. Это признак самоорганизации, порождения смысла без внешней интервенции. Хейлигхен использовал эту философию бутстрапа для открытого представления знаний в сети ИНТЕРНЕТ, (концептоорганизатор на базе модели Г. Паска) (см. часть II, раздел 7).

### Библиография

1. *Varela, F.J., H.R. Maturana and R. Uribe.* The organization of living systems, its' characterization and model. – *Byosystems*, 1974, №5, p. 187–196.
2. *Varela F.* Principles of Biological Autonomy. Elsevier/Noth-Holland, New York, 1979. 306 pp.
3. *Rosen, R.* Life Itself: A Comprehensive Inquiry onto the Nature Origin and Fabrication of Life. – Columbia University Press, New York, 1991.
4. *von Foerster, H.* On self-organizing systems and their environments, in: *Self-Organizing Systems*, (M. Yovis and S Cameron, eds.), Pergamon Press, London, 1966.
5. *von Glasersfeld, E.* Steps Piaget and the Radical Constructivist epistemology, in C.D. Smock and E. von Glasersfeld (eds), *Epistemology and Education*. – Athens, GA: Follow Through Publications, 1974.
6. *Левевр В.А.* Кибернетика второго порядка в Советском Союзе и на Западе. – *Рефлексивные процессы и управление*, №1, 2002, т. 2, с. 96–103.
7. *Капра, Ф.* Дао физики. СПб., "ОРИС", "ЯНА-ПРИНТ", 1994. 302 с.
8. *Bohm, D.* Wholeness and the Implicate Order. Routledge&Kegan Paul, London, Boston, 1980.
9. *Kampis, G.* Self-Modifying Systems in Biology and Cognitive Science. Pergamon Press, 1991. 565 pp.
10. *Эйген М.* Самоорганизация материи и эволюция биологических макромолекул. М., изд. "Мир", 1973. 216 с.
11. *von Foerster, H.* Notes for an epistemology of living things, in: *L'Unite de l'Homme* (E. Morin and M. Piatelli, eds.). Seuil, Paris, 1974.
12. *Piaget, J.* Biologie et Connaissance. Gallimard, Paris, 1969.
13. *Pask, G.* The Foundations of Conversation Theory and Lp. In.: Heylighen F., Rosseel E.& Demeyere F.(eds.). *Self-Steering and Cognition in Complex Systems.Toward a New Cybernetics*, (Gordon and Breach Science Publishers,New York, 1990, pp. 240–247.