

## Новая кибернетика и применение ее принципов в физике

Купервассер Олег

Ключевой идеей в кибернетике была обратная связь [1]. Это замкнутый цикл – объект через посредство других объектов влияет на самого себя. В новой кибернетике [2, 3, 4, 5] эта идея получила дальнейшее развитие – кибернетики (исследователи, наблюдатели) и кибернетика (исследуемая, наблюдаемая система) влияют друг на друга, образуя обратную связь. Новая кибернетика иначе известна как кибернетика кибернетики или кибернетика второго порядка. Heinz von Foerster видит первопричину появления кибернетики второго порядка в попытках классических кибернетиков сконструировать модель мозга [2]: «Мозг требует найти теорию мозга, которая для достижения полноты должна включать и создающего эту теорию. И даже более удивительно, создатель этой теории должен взять в расчет и самого себя при этом. Переходя в рамки кибернетики, это можно сформулировать так: кибернетик, рассматривая свою область должен принимать в расчет и свою собственную активность. Кибернетика в таком случае становится кибернетикой кибернетики или кибернетикой второго порядка».

Можно сформулировать набор основных черт новой кибернетики, связанных с этой идеей:

- **Новый тип обратной связи [2]:** Связь наблюдаемой системы с собой через наблюдателя. Это также и связь наблюдателя с собой через наблюдаемую систему. Часто это связь мала и ее можно пренебречь. Но во многих важных случаях это сделать нельзя. Это связано с **неустойчивостью и хаотичностью** многих реальных систем, обсуждаемой ниже
- **Взаимовлияние и корреляции наблюдателя и наблюдаемой системы [2, 6]:** изучаемую (наблюдаемую) систему нельзя рассматривать в отрыве от исследователя (наблюдателя). Существуют взаимодействия и корреляции между ними, которые часто необходимо учитывать, несмотря на их кажущуюся малость. Это связано с **неустойчивостью и хаотичностью** многих реальных систем, обсуждаемой ниже
- **Относительность наблюдателей [6, 7]:** разные наблюдатели могут видеть наблюдаемую систему по-разному, однако это не ведет к противоречию.
- **Невозможность полного самоописания [6]:** система не может полностью описать и предсказать (понять) саму себя. На самом деле, пытаясь делать полное самописание системы, мы приходим к противоречию. Например, для описания системы мы используем чернила. Но при самоописании они тоже входят в систему. Т.е. их тоже надо описать. Для этого нужны другие чернила. Их тоже надо описать. И т.д. до бесконечности. Тем не менее, не полное, а частичное самописание возможно.
- **Сложность [8]:** реальные системы состоят из большого количества частей, которые взаимодействуют друг другом и окружением
- **Неустойчивость и хаотичность [9]:** многие реальные системы крайне чувствительны к слабым внешним воздействиям. Это приводит к важности учета даже малого взаимодействия наблюдателя с наблюдаемой системой.
- **Эмерджентность [10]:** у целого часто появляются новые свойства, которые нельзя предвидеть на основе даже полного знания частей и взаимодействия между ними

(принципиальная эмерджентность). Действительно, мы знаем, что наблюдаемая система взаимодействует с наблюдателем. Это взаимодействие приводит к тому, что наблюдаемая система становится незамкнутой и подверженной непредсказуемому внешнему шуму. Следовательно, полная система должна включать наблюдателя. Если же мы включаем наблюдателя в систему, она становится непредсказуемой вследствие **невозможности полного самоописания**.

- **Целостность и взаимосвязь (корреляции и взаимодействия) [6, 10, 11, 12]:** наш мир это не набор случайных тел и событий, а довольно целостный набор взаимодействующих (пусть иногда и очень слабо) и коррелированных друг с другом объектов. Т.е. наш мир скорее напоминает взаимосвязанный организм, чем комплекс случайных событий (голографическая модель Вселенной или голографическая модель мозга, например)
- **Научная непредсказуемость [6, 10]:** многие системы нельзя во всех деталях описать чисто логическими, научными методами, например математическим моделированием. Многие системы, хоть и непредсказуемы, но описываются вероятностными законами. Здесь идет речь о более сильной непредсказуемости, которая не описывается даже теорией вероятностей. Часто эта невозможность принципиальна, и не связана со сложностью системы. На самом деле, взаимодействие наблюдателя и наблюдаемой системы приводит к тому, что наблюдаемая система становится незамкнутой и подверженной непредсказуемому внешнему шуму. Если же мы включаем наблюдателя в систему, она становится непредсказуемой вследствие **невозможности полного самоописания**
- **Главные параметры системы [6, 13]:** часто сложная система включает много параметров, описывающих ее. Однако во многих случаях динамику системы и ее взаимосвязи можно достаточно точно описать, используя лишь небольшое количество параметров. Они называются главными параметрами системы. Главные параметры – один из методов борьбы со **сложностью и непредсказуемостью** системы. Примерами главных параметров служат термодинамические переменные в физике, или «фичеры» - характерные черты многопиксельных объектов, используемые в теории распознавания образов.
- **Интуиция (озарение) [6, 14]:** В психологии интуиция проявляется, когда решение проблемы приходит само быстро и внезапно. То, что невозможно достичь научными методами, можно достичь интуицией. Интуиция по определению не вытекает из науки. Интуиция – метод преодоления научной **непредсказуемости** системы *на практике*. На самом деле, наблюдатель взаимодействует с окружающим миром, коррелирован с ним и является его частью. Это позволяет ему, просто действуя по интуиции (из неких необъяснимых внутренних побуждений), достигать тех целей, которых невозможно в принципе достичь чисто научными методами.
- **Междисциплинарность [15]:** очень похожие свойства и явления происходят в системах, относящихся к совершенно разным областям знаний.

Проиллюстрируем основные черты новой кибернетики с помощью примеров из различных областей знаний.

## Математика и логика:

Обратная связь в математике часто встречается и приводит к доказательству самых фундаментальных теорем о невозможности. Приведем ряд примеров:

- теорема Геделя о неполноте [16-17]. Гедель ввел функцию однозначного соответствия между набором символов и целыми числами. Это позволило ему утверждения о свойствах аксиом формальной математической системы формулировать как теоремы об арифметике целых чисел. Он смог таким образом сформулировать утверждение «Я лгу», которое ссылается на самого себя (**обратная связь**). Такое утверждение взаимно противоречиво, и не может быть ни истинным, ни ложным. Таким образом, его нельзя ни доказать, ни опровергнуть, если система аксиом непротиворечива. Это ведет к теореме Геделя о неполноте: всякая непротиворечивая формальная система аксиом, включающая арифметику, принципиально неполна.
- теорема Геделя-Тьюринга: Можно доказать невозможность алгоритма, который может за конечное время определить, будет ли другой алгоритм работать бесконечно долго [18], для данных входных параметров, или остановится. Попытка подать на вход этого алгоритма его самого (**обратная связь**) приводит к неизбежному противоречию, что доказывает его невозможность.
- В теории множеств также возникают такие противоречия [19]. Определим множество всех множеств, которые не включают себя в качестве подмножеств. Будет ли такое множество включать само себя (**обратная связь**)? Попытка ответить на этот вопрос неизбежно приводит к противоречию. Это означает, что для непротиворечивости теории множеств необходимо ввести ограничения на возможность некоторых совокупностей быть множествами.

## Нейронная теория мозга:

Для иллюстрации принципов новой кибернетики и доказательства теорем о невозможности приведем несколько примеров из **нейронной теории мозга**:

- **Возможно ли создание электронного мозга, полностью копирующего работу нейронов человеческого мозга и равный по силе человеческому?** [6, 20, 21, 22] Ответ на этот вопрос следующий – в полной мере нет. Действительно, работа нейронов определяется внутренним состоянием нейронов, которое во всей своей полноте эквивалентно полному набору составляющих его атомов. Далее, работа нейронов мозга зависит не только от самих нейронов, но и окружающих их клеток (глиальные клетки). Работа этих клеток зависит от работы всего организма. А работа всего организма зависит от его окружающей среды. Т.е., чтобы смоделировать работу мозга во всей полноте необходимо смоделировать весь окружающий мир во всей полноте с точностью до атомов. Что в принципе невозможно из-за принципиальной **невозможности полного самоописания** [6]. Конечно, это не означает, что работу мозга нельзя смоделировать. В каком-то приближении, введя **главные параметры**, описывающие нейроны, это возможно. Но сделать это во всей полноте, включая наиболее тонкую интуицию, вполне возможно не удастся. Приведем немного шуточный пример. Яблоко упало на голову Ньютона, что привело к гениальному открытию закона гравитации. Следовательно, чтобы сделать электронный мозг,

способный повторить это открытие, нам нужно смоделировать не только мозг Ньютона, но и весь окружающий мир, приведший к падению яблока на его голову в нужный момент! (**новый тип обратной связи** между наблюдателем и наблюдаемой системой)

- **Возможно ли создание человека и его мозга, полностью копирующего другого человека?** [20] Нет. Ответ и его причины те же, что и в предыдущем пункте. Принципиальной разницы между точным физическим копированием и копированием на электронный носитель нет.
- Современная наука считает, что сознание (или, по крайней мере, его наблюдаемые проявления) полностью определяется физическими процессами. **Возможно ли создание «кода мозга», устанавливающего однозначное соответствие между мыслями и работой нейронов? А также возможно ли создание прибора, способного таким образом «читать мысли»?** [6, 20, 21] Ответ тот же – во всей полноте - нет. Это легко проиллюстрировать с помощью **обратной связи** [20]. Предположим, что такой прибор создан, и сам человек может читать свои мысли, например, определять свой гнев. Более того, проникнувшись полным доверием к прибору, человек может с помощью него судить о своем состоянии. Но, что показывает прибор, если именно его человек использует, чтобы понимать себя. Что же происходит? Появляются новые механизмы обратной связи мозга, и образуется новый «код мозга», который включает уже не только нейроны, но и сам прибор для чтения мыслей. Более того, такие обратные связи могут возникать естественно. Например, человек о своем гневе может судить по поту на лбу или по испуганным лицам окружающих. Они играют роль прибора для чтения мыслей. Таким образом, код мозга во всей своей полноте зависит не только от нейронов, но и от других клеток тела и даже окружения. Кроме того, он может динамически меняться со временем.  
Это не означает, что «код мозга» совсем невозможен. **Главные параметры** нейронов позволяют сделать это. Но существуют естественные ограничения для его полноты, приведенные выше.

#### **Компьютерные науки:**

- **Можно ли создать «китайскую комнату» хотя бы в принципе в реальности?** [6, 20, 21] Во всей полноте - нет. Просто, как мы уже показали выше, полный свод правил для человека в «китайской комнате» равен размеру Вселенной и должен включать его самого. Чего не может быть из-за принципиальной **невозможности полного самоописания**.
- **Может ли компьютер вести себя непредсказуемо?** [6] Да. Примером может служить квантовый компьютер (или его классический аналог – неустойчивый аналоговый компьютер [23-25]). Действительно, о работе такого компьютера может судить лишь человек запустивший его. Любой сторонний наблюдатель, не присутствующий при его запуске, при попытке понять его работу нарушит эту работу. Такие системы являются непредсказуемыми.
- **Может ли интеллект, сравнимый с человеческим, работать лишь на основе научной логики и не обладать интуицией?** [6, 26] Нет. Мы выше говорили, что

наличие **непредсказуемых** систем ограничивает возможности чисто научного познания окружающего мира.

- **Может ли компьютер обладать интуицией и угадывать лучше генетического алгоритма?** [6, 26] Да. Для моделирования интуиции в компьютерах часто используют генератор случайных чисел. Но такая модель примитивна [27]. Для более точного моделирования интуиции нужно позволить окружающему миру влиять на этот генератор. Возникают корреляции между окружающим миром и генератором. Это может позволить ему лучше «угадывать» свойства окружающего мира, чем это делает простой случайный генератор в генетическом алгоритме. Следует отметить, что за эту интуицию мы должны заплатить высокую цену – сам компьютер становится **непредсказуемой** системой. Что может привести, например, к бунту роботов.
- **Может ли интуиция искусственного интеллекта быть полностью эквивалентной по силе человеческой?** [6] Если бы человек мог четко формализовать свои цели и желания, то да. Однако многие желания человека не формализуемы и не понятны даже ему самому. Как мы писали выше, смоделировать человека во всей полноте невозможно в принципе. Следовательно, невозможно и создание идеальной машинной интуиции. Что нужно человеку никто не может знать лучше него самого.
- **Можно ли создание искусственного интеллекта по силе эквивалентного человеческому?** [3, 6] Нет. Поскольку человеческая интуиция не может быть скопирована во всей полноте. А значит, и некоторые задачи человек будет всегда решать лучше машины, поскольку лучше понимает свои желания. И возможный бунт роботов – лучший пример такой неадекватности искусственного интеллекта. Поскольку он оказывается неспособным решать задачи, нужные человеку, как человек или лучше человека.
- **Каким образом можно создать интеллект, сравнимый по силе с человеческим, для решения формализуемых задач (с четкой постановкой задачи и цели)?**  
Думается, что разумная система должна создаваться по образцу эволюции и воспитания человека, но в более ускоренном темпе.  
Это должна быть обучающаяся и самообучающаяся система.  
Скорее всего, наиболее подходящей моделью будет подобие нейронной сети. Причем даже существующие эвристические алгоритмы стоит переработать в форму нейронных сетей.  
Разумная система должна создаваться от простого к сложному: начать надо с создания роботов, копирующего поведение простейших насекомых, имеющих самую простую нервную систему. А затем постепенно усложнять задачу, создавая все более сложные роботы, используя на последующих этапах результаты предыдущих.  
Необходимо нечто вроде естественного отбора, регулируемого генетическим алгоритмом. Если мы захотим подключить и интуицию, для создания таких систем или для решения уже неформализуемых задач, необходимо подключить генератор случайных чисел, коррелированный с окружением.

**Экономические, политические, психологические и социологические науки:**

Тут важность указанных выше принципов очевидна [28]. Влияние наблюдателя на наблюдаемую систему очень велико. Любая «теория» захватив многие умы, становится сама частью изучаемой системы. Например, невозможна теория выигрыша на бирже – если она верна, то все должны выиграть. Но это невозможно. Следовательно, в полной мере эти науки науками считать нельзя. Тут огромную роль всегда играет интуиция, много больше, чем в других науках.

### **Физика:**

Казалось бы, самая точная наука и влияние наблюдателя несущественно, в противоположность к экономическим, политическим, психологическим и социологическим наукам [28]. Как может наблюдатель повлиять на результаты опытов в ускорителе? Но для многих реальных физических систем это не так – они обладают свойством неустойчивости и хаотичности. Это приводит к разрешению многих парадоксов в физике. Этому посвящена вся следующая часть данной работы. Здесь приведем лишь один пример.

**Почему начало Вселенной столь низкоэнтропийное? Почему наша Вселенная не может быть результатом гигантской флуктуации из термодинамического равновесия с максимальной энтропией?** Ответ прост – на самом деле это может быть. Более того – это более вероятно согласно законам статистической физики. Но проверить экспериментально, какая из двух вышеописанных теорий происхождения мира верна, невозможно в принципе из-за **невозможность полного самоописания** и самонаблюдения. Фейнман [29] приводит аргумент, что в случае гигантской флуктуации вокруг нас должно было бы быть «океан» термодинамического равновесия. И видит тут противоречие. Но это не так. Наблюдатель может быть просто неспособен наблюдать этот «океан». Ведь он сам часть этой флуктуации и взаимодействует с этим миром. Он не может вырваться за его пределы, поскольку будет просто разрушен некомфортной окружающей равновесной средой. Таким образом, низкоэнтропийная теория Вселенной – лишь удобная точка зрения, а не установленный научный факт. Мы видим лишь малую наблюдаемую часть Вселенной, а пытаемся судить обо всей Вселенной лишь по малой ее части. Что очень чревато ошибкой...

### **Bibliography**

1. Norbert Wiener, “Cybernetics Or Control and Communication in the Animal and the Machine”, USA, MIT Press, 1965, p. 212
2. Heinz von Foerster (2003), Understanding Understanding: Essays on Cybernetics and Cognition, New York : Springer-Verlag, 362pp
3. Thomas Marlowe, “Systemics and Requirements: A Missing Dimension?” , WMSCI 2013, 9-12 July, Orlando, Florida, USA
4. Louis H. Kauffman, “Circularity, Topology and Cybernetics: Second Order Science”, WMSCI 2013, 9-12 July, Orlando, Florida, USA
5. Karl H. Muller, “Unfolding and Expanding Science with the Help of the New Science of Cybernetics (NSC)”, WMSCI 2013, 9-12 July, Orlando, Florida, USA

6. Oleg Kupervasser, "The Universal Arrow of Time is a Key for the Solution of the Basic Physical Paradoxes", EJTP, Vol 10, No 29, July 2013
7. Einstein A. (1916 (translation 1920)), Relativity: The Special and General Theory, New York: H. Holt and Company
8. Waldrop, M. Mitchell (1992). Complexity: The Emerging Science at the Edge of Order and Chaos. New York: Simon & Schuster.
9. Paul Glendinning, "Stability, Instability and Chaos. An Introduction to the Theory of Nonlinear Differential Equations", Cambridge University Press, 1994
10. Licata, I., "Emergence and Computation at the Edge of Classical and Quantum Systems". In: Licata, I., Sakaji, A. Eds. Physics of Emergence and Organization, World Scientific, 2008:1-25
11. Susskind, Leonard (1995). "The World as a Hologram". Journal of Mathematical Physics 36 (11): 6377–6396. arXiv:hep-th/9409089
12. Karl Pribram, "Hologomic Brain Theory," Scholarpedia, 2(5):2735, 2007, [http://www.scholarpedia.org/article/Hologomic\\_brain\\_theory](http://www.scholarpedia.org/article/Hologomic_brain_theory)
13. Abdi. H., & Williams, L.J. (2010). "Principal component analysis.". Wiley Interdisciplinary Reviews: Computational Statistics, 2: 433–459
14. Robinson-Riegler, Bridget Robinson-Riegler, Gregory. "Cognitive psychology : applying the science of the mind", (3rd ed. ed.). Boston: Pearson Allyn & Bacon
15. T. Grandon Gill, "Interdisciplinary Research, Education, and Communication through Case Studies and Methodologies", WMSCI 2013, 9-12 July, Orlando, Florida, USA
16. Kurt Gödel , 1931, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatshefte für Mathematik und Physik 38: 173-98.
17. Kurt Gödel , 1931, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. and On formally undecidable propositions of Principia Mathematica and related systems I in Solomon Feferman, ed., 1986. Kurt Gödel Collected works, Vol. I. Oxford University Press: 144-195
18. Perlovsky, Leonid I., "Neural networks and intellect : using model-based concepts", 2001 by Oxford University Press, New York.
19. H. Poincaré, Science and Hypothesis (1905) (New York: Dover Publications, 1952)
20. Douglas R. Hofstadter, Daniel C. Dennett, The Mind's I, New York, Basic Books, 2000
21. Stuart J. Russel, Peter Norvig "Artificial Intelligence", Prentice Hall, New Jersey, 2010
22. Nick Bostrom "HOW LONG BEFORE SUPERINTELLIGENCE? " (Originally published in Int. Jour. of Future Studies, 1998, vol. 2, Reprinted in Linguistic and Philosophical Investigations, 2006, Vol. 5, No. 1, pp. 11-30.), <http://www.nickbostrom.com/superintelligence.html>
23. Siegelmann, H.T. "Neural Network and Analog Computation: Beyond the Turing Limit", Birkhauser, 1998
24. Calude, C.S., Paun, G. "Bio-steps beyond Turing", BioSystems, 2004, v 77, 175-194
25. Nicolas H. Voelcker; Kevin M. Guckian; Alan Saghatelian; M. Reza Ghadiri "Sequence-addressable DNA Logic", Small Weinheim an der Bergstrasse Germany, 2008, Volume 4, Issue 4, Pages 427 – 431

26. Michael B. Mensky, "Consciousness and Quantum Mechanics: Life in Parallel Worlds", *Miracles of Consciousness from Quantum Reality*, Imperial college press, P. 250 (2010)
27. D.R. Hofstadter, *Godel, Escher, Bach: An Eternal Golden Braid*, Bask Books, USA, 1979.
28. Stuart A. Umpleby, "Expansion of Science", WMSCI 2013, 9-12 July, Orlando, Florida, USA
29. Feynman, Richard, "The Character of Physical Law", Published by Modern Library, New York, 1994