

## Е.Е. ВИТЯЕВ

Институт математики им. С.Л.Соболева  
[vityaev@math.nsc.ru](mailto:vityaev@math.nsc.ru)

### ФОРМАЛИЗАЦИЯ КОГНИТОМА<sup>1\*</sup>

Формализация когнитома осуществлена путем формализации когов функциональных систем и когов феноменального опыта путем построения единой формализации «естественной» классификации, «естественных» понятий и интегрированной информации G.Toponi. Данная формализация основана на принципе: мозг, в процессе эволюции, настроился на выделение высоко коррелированной структуры причинных связей «естественных» объектов путем обнаружения «естественной» классификации и «естественных» понятий.

**Ключевые слова:** когнитом, понятие, теория функциональных систем

#### Введение

Понятие когнитома – когнитивной гиперсети головного мозга – является в настоящее время одним из наиболее развитых подходов к изучению фундаментальных механизмов работы мозга и когнитивных функций [1]. Когнитом состоит из взаимосвязанных когов («когнитивных групп нейронов, представляющих элементы субъективного опыта») двух типов – функциональных систем и элементов феноменального опыта. Вторые возникают из первых: «за счет активности первых и перекрытия на некоторых популяциях клеток – клеточных ансамблей Д.Хебба» [1]. Эти перекрытия возникают тогда, когда в разных функциональных системах фигурируют одни и те же объекты, ситуации, способы действия, возбуждающие эти ансамбли клеток.

Гиперсеть когнитома – это сети, сетей, сетей ..., состоящие из гиперсимплексов. К.В.Анохин видит «задачу фундаментальной теории мозга и разума в описании этих структур, их происхождения, функций и процессов в них» [1]. В такой постановке к гиперсетям предъявляются достаточно сильные требования – они должны объяснить весь комплекс когнитивных процессов: сознание, формирование «образа мира», целенаправлен-

---

<sup>1</sup> Автор выражает благодарность К.В. Анохину за множество полезных советов и обсуждения.

\* Данная работа выполнена при поддержке РФФИ, проект № 15-07-03410-а.



ное поведение, формирование понятий и концептов, восприятие, обучение и т.д. Однако, гиперсети – это, прежде всего, математический и информационный объект. Очевидно, что одним только определением гиперсетей как сеть, сетей, сетей ... гиперсимплексов все эти процессы объяснить невозможно. Необходимо включить в рассмотрение формальную модель внешнего мира, а также информационных процессов нейронов и связей между ними. Только такая комплексная формализация позволит понять смысл происходящих в когнитоме информационных процессов.

Такой подход ставит следующую задачу: промоделировать фундаментальные механизмы работы мозга, основываясь на построении математической и компьютерной модели когнитома, включающей формализацию внешнего мира, нейронов и сети нейронов.

Для решения данной задачи, прежде всего, нужна формальная модель внешнего мира. В качестве такой модели нами предлагается оригинальная формализация «естественной» классификации объектов внешнего мира [4-5], основанная на вероятностном обобщении формальных понятий [3,23]. Понятие «естественной» классификации формализует представления естествоиспытателей о строении «естественных» объектов внешнего мира [4]. Мы покажем, что данная формализация является также формализацией «естественных» понятий, как это определено в работах [14-16]. Определение «естественных» понятий основано на принципе категоризации Eleanor Rosch, постулирующей структурность воспринимаемого мира: «воспринимаемый мир не является неструктурированным множеством равновероятно встречающихся свойств, наоборот, объекты воспринимаемого мира имеют высоко коррелированную структуру» [16].

На высоко коррелированной структуре внешнего мира основана также интегрированная информация, определенная G.Tononi [12,17-20]. Но у G.Tononi нет модели внешнего мира, и интегрированная информация рассматривается как внутреннее свойство системы причинных связей, проявляющаяся в сознании. Он определяет понятие концепта как системы причинных связей с максимально интегрированной информацией. Поскольку у него нет внешнего мира, то он не может сказать, что его понятие концепта отражает «естественные» понятия и «естественные» объекты.

Наша формализация является одновременной формализацией «естественной» классификации, «естественных» понятий и интегрированной информации, поскольку в ее основе лежит одно и то же свойство внешнего мира – высоко коррелированная структура свойств «естественных» объектов [22]. Это свойство максимально точно отражено в нашей оригинальной формализации в виде неподвижных точек взаимных предска-



ний свойств «естественных» объектов. Эта формализация дает также формализацию восприятия [6-7] и когов феноменального опыта. Наша формализация принципиально отличается от других формализаций систем причинных связей [9], основанных на байесовских сетях, поскольку байесовские сети не поддерживают циклы.

В когнитоме гиперсимплексы сети, сетей ... тогда есть отражение иерархии «естественных» классов внешнего мира, где вершина гиперсимплекса – имя класса, а основание гиперсимплекса – максимально интегрированная информация по G.Toponі или неподвижная точка в нашей формализации.

Приводимая нами единая формализация приводит к следующему принципу работы мозга: ***мозг, в процессе эволюции, настроился на выделение высоко коррелированной структуры причинных связей «естественных» объектов за счет обнаружения «естественной» классификации внешнего мира гиперсимплексами.***

Формализация когов второго типа – когов функциональных систем, нами была ранее осуществлена в большом количестве работ [2,6-8,21-22]. Формализация этих когов была апробирована путем разработки виртуальных аниматов, на которых показала свою эффективность [2,11,21]. Поэтому, в данной публикации она приведена не будет, но будет указано как два вида когов – функциональных систем и элементов феноменального опыта – взаимодействуют между собой.

### **Что такое «естественная» классификация**

Высоко коррелированную структуру внешнего мира выявляют две теории – «естественной» классификации и «естественных» понятий. При этом, «естественная» классификация говорит о структуре объектов внешнего мира, а «естественные» понятия, исследуемые в когнитивных науках, о восприятии этих «естественных» объектов.

Первый достаточно подробный философский анализ «естественной» классификации принадлежит Дж. Ст. Миллю [13]. Приведем его точку зрения, чтобы стал понятен смысл «естественной» классификации. Сначала отделим «искусственные» классификации от «естественных». Приведем определение Дж. Ст. Милля: *«Возьмем любой признак, и если одни вещи обладают им, а другие не обладают, то на нем можно основать деление всех вещей на два класса». «Но если мы обратимся к ... классу «животное» или «растение», ... и посмотрим, какими особенностями индивидуумы, обнимаемые данным классом, отличаются от индивидуумов, не входящих в него, то мы найдем, что в этом отношении одни*



*классы сильно отличаются от других. ... обладают столь большим количеством признаков, что их нельзя ... перечислить» [13].*

«Естественную» классификацию Дж. Ст. Милль определяет следующим образом: *«Всего более соответствует целям научной (естественной) классификации, когда предметы соединяются в такие группы, относительно которых можно высказать наибольшее число общих предложений» [13].* На основании понятия «естественной» классификации Дж. Ст. Миллем определяет понятие «образа» класса, которое является предтечей «естественных» понятий, исследуемых в когнитивной психологии: *«Естественные группы ... определяются признаками, ... при этом принимаются во внимание не только признаки, безусловно общие всем включаемым в группу предметам, но вся совокупность тех признаков, из которых все встречаются в большинстве этих предметов, а большинство – во всех. Вследствие этого наше понятие о классе – тот образ, которым этот класс представлен в нашем уме, – есть понятие о некотором образце, обладающем всеми признаками данного класса».*

Проблема определения «естественных» классификаций до сих пор обсуждается в литературе [17]. Однако, с нашей точки зрения нет достаточно адекватной формализации «естественной» классификации.

### **Что такое «естественные» понятия**

В работах Eleanor Rosch был сформулирован следующий принцип категоризации «естественных» категорий: *«Perceived World Structure. The second principle of categorization asserts that ... perceived world – is not an unstructured total set of equiprobable co-occurring attributes. Rather, the material objects of the world are perceived to possess ... high correlational structure ... combinations of what we perceive as the attributes of real objects do not occur uniformly. Some pairs, triples, etc., are quite probable, appearing in combination ... with one, sometimes another attribute; others are rare; others logically cannot or empirically do not occur» [16].*

Непосредственно воспринимаемые объекты (basic objects) – информационно богатые связки наблюдаемых свойств, создающие категоризацию (образ в определении Дж. Ст. Милля): *«Categories can be viewed in terms of their clear cases if the perceiver places emphasis on the correlational structure of perceived attributes ... By prototypes of categories we have generally meant the clearest cases of category membership» [16].* В дальнейшем теория «естественных» понятий Eleanor Rosch получила название прототипической теории понятий (prototype theory).

В дальнейших исследованиях было обнаружено, что моделей, осно-



ванных на признаках, сходстве и прототипах, недостаточно для описания «естественных» классов. Необходимо учитывать теоретические, причинные и онтологические знания, относящиеся к объектам классов. Например, люди не только знают, что птицы имеют крылья, могут летать и вить гнезда на деревьях, но также и то, что птицы выют гнезда на деревьях, потому что могут летать, и летать, потому что они имеют крылья.

Учитывая эти исследования, Bob Rehder выдвинул теорию причинных моделей (causal-model theory), в соответствии с которой: «people's intuitive theories about categories of objects consist of a model of the category in which both a category's features and the causal mechanisms among those features are explicitly represented» [14]. В теории причинных моделей отношение объекта к категории основывается уже не на множестве признаков и близости по признакам, а на основании сходства порождающего причинного механизма: «Specifically, a to-be-classified object is considered a category member to the extent that its features were likely to have been generated by the category's causal laws, such that combinations of features that are likely to be produced by a category's causal mechanisms are viewed as good category members and those unlikely to be produced by those mechanisms are viewed as poor category members» [14].

Для представления причинного знания были использованы Байесовские сети. Однако они не поддерживают циклов и поэтому не могут моделировать циклические причинные связи. Предлагаемая нами формализация прямо моделирует циклические причинные связи, представленные неподвижными точками предсказаний по причинным связям [5,23].

### **Теория интегрированной информации G.Tononi**

На высоко коррелированной структуре внешнего мира основана также теория интегрированной информации G.Tononi [12,19-20]. Интегрированная информация у G.Tononi рассматривается как свойство системы циклических причинных связей: «Indeed, a “snapshot” of the environment conveys little information unless it is interpreted in the context of a system whose complex causal structure, over a long history, has captured some of the causal structure of the world, i.e. long-range correlations in space and time» [19].

В соответствии с нашей гипотезой мозг с помощью интегрированной информации настраивается на восприятие «естественных» объектов внешнего мира путем формирования «естественных» понятий, определяемые в теории G.Tononi через локальные максимумы интегрированной информации: «Cause-effect matching ... measures how well the integrated conceptual structure ... fits or ‘matches’ the cause-effect structure of its envi-



ronment», «... matching should increase when a system adapts to an environment having a rich, integrated causal structure. Moreover, an increase in matching will tend to be associated with an increase in information integration and thus with an increase in consciousness» [20].

G.Tononi определяет сознание как первичное понятие, которое обладает следующими феноменологическими свойствами: composition, information, integration, exclusion [12]. Приведем формулировки этих свойств вместе с нашей интерпретацией этих свойств (приведенной в скобках) с точки зрения «естественной» классификации объектов внешнего мира.

1. composition – elementary mechanisms (causal interactions) can be combined into higher-order ones («естественные» классы объектов образуют иерархию);

2. information – only mechanisms that specify ‘differences that make a difference’ within a system count (только система «резонирующих» причинных связей, формирующая класс, является значимой);

3. integration – only information irreducible to non-interdependent components counts (значима только система «резонирующих» причинных связей, не сводимая к информации отдельных компонент, свидетельствующая об избытке информации и восприятии высоко коррелированной структуры «естественного» объекта);

4. exclusion – only maxima of integrated information count (только значения признаков, которые максимально взаимосвязаны причинными связями формируют «образ» или «прототип»).

В отличие от G.Tononi, мы рассматриваем эти свойства не как внутренние свойства системы, а как способность системы отражать комплексы причинных связей внешних «естественных» объектов, а сознание – как способность комплексного иерархического отражения «естественной» классификации внешнего мира гиперсетью когнитомата.

### **Вероятностные формальные понятия**

В анализе формальных понятий [10] объекты классифицируются следующим образом: в группы выделяются те объекты, которые имеют общий набор признаков, и никакой другой объект ровно этим же набором признаков не обладает. Однако, если нам известны результаты не только одного эксперимента, но целой серии экспериментов, и для построения классификации объектов нам надо использовать всю совокупность данных, то объекты в одних экспериментах будут иметь определенные признаки, а в оставшихся нет, как в высказывании Дж. Ст. Милля: *«при этом принимаются во внимание не только признаки, безусловно общие всем*



включаемым в группу предметов, но вся совокупность тех признаков, из которых все встречаются в большинстве этих предметов, а большинство – во всех». Что бы учесть этот вероятностный характер классификации, нами было разработано очень не тривиальное вероятностное обобщение формальных понятий [3,5,23]. Формальные понятия были представлены неподвижными точками импликаций, далее, импликации были обобщены на вероятностный случай, как максимально специфические вероятностные правила и, затем, вероятностные формальные понятия были определены как неподвижные точки этих вероятностных импликаций. Было доказано, что в неподвижной точке максимально специфические вероятностные правила логически непротиворечивы [5,23].

Формально неподвижные точки определяются следующим образом. Пусть  $X(a)$  – множество свойств объекта  $a$ , заданных некоторым множеством предикатов,  $(P_{i_1} \& \dots \& P_{i_k} \Rightarrow P_{i_0}) \in MS(X)$  – множество максимально специфических [3] условных связей, выполненных на  $X$ ,  $\{P_{i_1}, \dots, P_{i_k}\} \subset X$ . Оператор предсказания  $Pr$  запишем следующим образом [6]:

$$Pr(X) = \Phi_{Krit}(X \cup \{P_{i_0} \mid (P_{i_1} \& \dots \& P_{i_k} \Rightarrow P_{i_0}) \in MS(X)\} \cup \{\neg P_{i_0} \mid (P_{i_1} \& \dots \& P_{i_k} \Rightarrow \neg P_{i_0}) \in MS(X)\}),$$

где  $\Phi_{Krit}(X)$  – оператор, модифицирующий множество признаков  $X$  путем добавления или удаления некоторого из признаков так, чтобы определенный критерий  $Krit$  согласованности причинных связей по взаимному предсказанию признаков  $X$  был максимальным. Критерий  $Krit$ , по-своему,

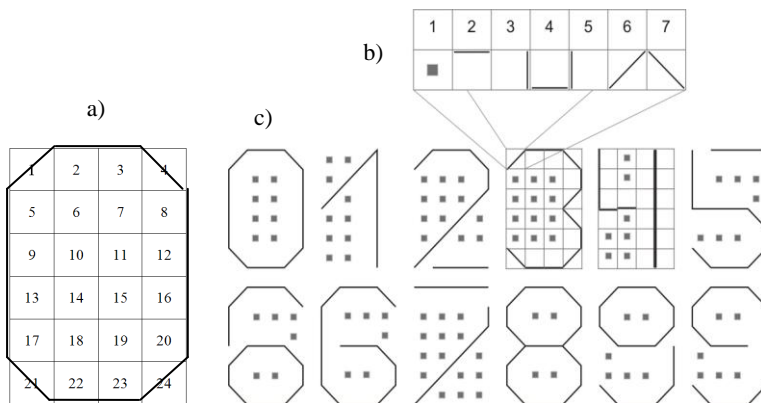


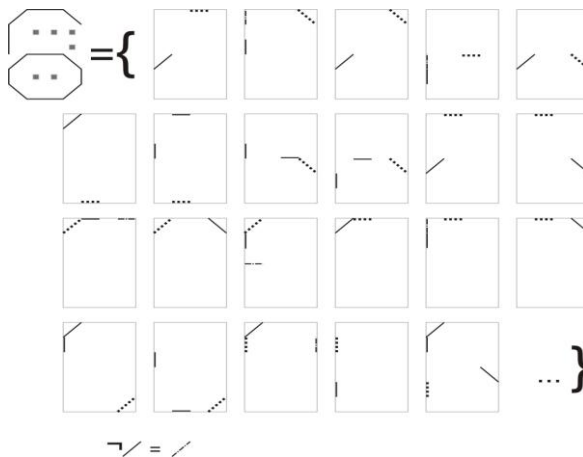
Рис.1. Кодировка цифр



измеряет интегрированную информацию системы причинных связей  $MS(X)$ , по сравнению с теорией G.Тонопи [12]. Неподвижная точка достигается, когда  $Pr^{n+1}(X(a)) = Pr^n(X(a))$  для некоторого  $n$ .

Приведем пример формирования «естественных» классов и «естественных» понятий. Закодируем цифры как показано на рис. 1b. Для этого было сформировано обучающее множество, состоящее из 360 перетасованных цифр (12 цифр рис. 1 продублированных в 30-ти экземплярах без указания, где какая цифра). На нем было обнаружено 55089 закономерности (общих утверждений об объектах, о которых говорит Дж. Ст. Миллем и Уэвель В). По этим закономерностям было обнаружено ровно 12 неподвижных точек, совпадающих с нашими цифрами, хотя алгоритму ничего не было известно о содержании объектов, кроме значений признаков.

Пример неподвижной точки для цифры 6 приведен на рис. 2, где в фигурных скобках приведено описание вероятностных правил, образующих неподвижную точку (посылки правил обозначены сплошной линией, предсказываемый признак – точечной линией, отрицание признака - пунктирной линией). Заметим, что закономерности, используемые в неподвижной точке, выполнены на всех цифрах, а сама неподвижная точка выделяет только одну цифру. Поэтому цифры выделяются не закономерностями, а их взаимосвязью или, в терминах G.Тонопи, системой причинных связей, обладающих максимальной интегрированной информацией или максимальным значением критерия Krit. Неподвижная точка сама формирует паттерн – набор признаков, составляющий неподвижную точку, ко-



**Рис. 1.** Неподвижная точка цифры 6.



торый является прототипом у Eleanor Rosch и образом у Дж. Ст. Милля.

Специальным экспериментом показано, что данная задача не решается сетями Хопфилда, как оригинальной, так и взятой из системы MATLAB [7]. Поэтому наша формализация, отчасти, превосходит сети Хопфилда.

### **Заключение**

Объединение когов функциональных систем и когов феноменального опыта осуществляется за счет замены вероятностных правил, используемых в функциональных системах на правила, включенные в неподвижные точки феноменального опыта, определяющие «контекст» их применения. Такая модель когнитивного опыта способна моделировать следующие когнитивные функции: восприятие [6-7], целенаправленное поведение и функциональные системы [2,21], организацию движений по А.Н. Бернштейну [2,21,24], эмоции и принятие решений [25]. На нейронном уровне гиперсеть когнитивного опыта может быть представлена нейронами, формальная модель которых описана в [26] и удовлетворяет правилу Хебба.

### *Список литературы*

1. Анохин К.В. Когнитом: в поисках общей теории когнитивной науки // Шестая международная конференция по когнитивной науке, Калининград, 2014, 26-28.
2. Витяев Е.Е. Логика работы мозга // Подходы к моделированию мышления. ред. В.Г. Редько. УРСС Эдиториал, Москва, 2014г., стр. 120-153.
3. Витяев Е.Е., Демин А.В., Пономарёв Д.К. Вероятностное обобщение формальных понятий // Программирование, Т.38, №5, 2012, С. 219-230.
4. Витяев Е.Е., Костин В.С. Естественная классификация, систематика, онтология // Информационные технологии в гуманитарных исследованиях, Вып. 13, ИАЭТ СО РАН, Новосибирск, 2009, стр. 65-75
5. Витяев Е.Е., Мартынович В.В. Формализация «естественной» классификации и систематики через неподвижные точки предсказаний // Сибирские электронные математические известия, Том 12, ИМ СО РАН, 2015, стр. 1006-1031
6. Витяев Е.Е., Неупокоев Н.В. Формальная модель восприятия и образа как неподвижной точки предвосхищений // Подходы к моделированию мышления. УРСС Эдиториал, Москва, 2014г., стр. 155-172.
7. Витяев Е.Е., Неупокоев Н.В. Математическая модель восприятия и образа. Информационные технологии в гуманитарных исследованиях, Вып.17, ИАЭТ СО РАН, Новосибирск, 2012, 63-72.
8. Мухортов В.В., Хлебников С.В., Витяев Е.Е. Улучшенный алгоритм семантического вероятностного вывода в задаче 2-мерного анимата // Нейроинформатика, 2012, том 6, № 1, стр. 50-62
9. Karl Friston. The free-energy principle: a unified brain theory? Nature Reviews Neuroscience 11, 127-138



10. B. Ganter, R. Wille. Formal Concept Analysis. Mathematical Foundations. Springer Verlag, 1999.
11. Demin, A.V., Vityaev E.E. Learning in a virtual model of the *C. elegans* nematode for locomotion and chemotaxis // *Biologically Inspired Cognitive Architectures*. 2014, v.7, pp.9–14.
12. Masafumi Oizumi, Larissa Albantakis, Giulio Tononi. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0 // *PLOS Computational Biology*, May 2014, V.10. Issue 5.
13. Mill, J.S. System of Logic, Ratiocinative and Inductive. L., 1983
14. Rehder, B. Categorization as causal reasoning//*Cognitive Science*,27,2003,709–748.
15. Rosch, E.H. Natural categories // *Cognitive Psychology* 4. P. 328-350.
16. Rosch, E., Principles of Categorization // Rosch, E. & Lloyd, B.B. (eds), *Cognition and Categorization*, Lawrence Erlbaum Associates, Publishers, 1978. P. 27–48
17. [http://www.math.nsc.ru/AP/ScientificDiscovery/index\\_rus.html](http://www.math.nsc.ru/AP/ScientificDiscovery/index_rus.html)
18. The Nature of Classification. Relationships and Kinds in the Natural Sciences. Palgrave Macmillan. 2013. 208p.
19. Tononi G. Information integration: its relevance to brain function and consciousness. *Arch. Ital. Biol.*, 148: 299-322, 2010
20. Tononi, G. Integrated information theory of consciousness: an updated account. *Arch Ital Biol* 150, 2012, 56–90.
21. Vityaev, E. Purposefulness as a Principle of Brain Activity // *Anticipation: Learning from the Past. Cognitive Systems Monographs*, V.25, Springer, 2015, pp. 231-254.
22. Vityaev, E. Unified formalization of «natural» classification, «natural» concepts, and consciousness as integrated information by Giulio Tononi // *The Sixth international conference on Biologically Inspired Cognitive Architectures, Procedia Computer Science*, v.71, Elsevier, 2015. pp 169-177
23. Vityaev, E.E., Martinovich, V.V. Probabilistic Formal Concepts with Negation // A. Voronkov, I. Virbitskaite (Eds.): *PCI 2014, LNCS 8974*, P. 385-399.
24. Витяев Е.Е. Объяснение Теории Движений Н.А.Бернштейна. VII Всероссийская научно-техническая конференция «Нейроинформатика-2005», Сборник научных трудов, часть 1, Москва, 2005, стр.234-240.
25. Витяев Е.Е. Принятие решений. Переключающая и подкрепляющая функции эмоций // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006», Сборник научных трудов, Москва, 2006, с.24-30
26. Vityaev E.E. A formal model of neuron that provides consistent predictions // *Biologically Inspired Cognitive Architectures 2012. In Advances in Intelligent Systems and Computing*, v.196, Springer, 2013, pp. 339-344