

Витяев Е.Е.¹,

ЕДИНАЯ ФОРМАЛИЗАЦИЯ «ЕСТЕСТВЕННОЙ» КЛАССИФИКАЦИИ, «ЕСТЕСТВЕННЫХ» ПОНЯТИЙ И СОЗНАНИЯ, КАК ИНТЕГРИРОВАННОЙ ИНФОРМАЦИИ по G.TONONI

В работе показывается, что в основе построения «естественных» классификаций, «естественных» понятий и интегрированной информации лежит одно и то же свойство объектов внешнего мира – высокая коррелированность признаков, описывающих объекты «естественных» классов. Выдвигается гипотеза, что информационные процессы работы мозга и сознание настроилась в процессе эволюции на извлечение высоко коррелированной структуры признаков «естественных» объектов путем формирования «естественных» понятий объектов. Эта гипотеза обосновывается ссылками на ряд известных работ. Кроме того, предлагается оригинальная математическая модель, формализующая «естественные» классификации, «естественные» понятия и интегрированную информацию по G.Tononi, основанную на математическом представлении замкнутой на себя системы причинных связей, образующих определённый «резонанс» взаимных предсказаний высоко коррелированной совокупности признаков объектов «естественных» классов. Приводятся результаты компьютерного моделирования построения «естественных» классов и понятий закодированных цифр, иллюстрирующие введенные понятия.

Ключевые слова: Clustering, Categorization, Natural classification, Natural Concepts, Integrated Information, Concepts.

1. Введение

Строение объектов внешнего мира впервые было проанализировано в области «естественной» классификации (см. раздел 2). Было замечено, что «естественные» классы животных или растений обладают и отличаются потенциально бесконечным множеством свойств [22]. Естествоиспытатели, строившие «естественные» классификации, также отмечали, что чем в большем числе существенных признаков сходны сравниваемые предметы, тем вероятнее их одинаковость и в других отношениях и построение «естественной» классификации заключается в «индикации» – от бесконечно большого числа признаков нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки [9-10]. Это означает, что в «естественных» классах признаки сильно коррелированы, например, если классов 128 и признаки бинарные, то независимыми «индикаторными» признаками среди них будут около 7 признаков, т.к. $2^7=128$, а остальные можно определить в зависимости от этих 7 признаков.

Высокая коррелированность признаков для «естественных» классов была подтверждена и в когнитивных исследованиях (см. раздел 3). Eleanor Rosch сформулировала принципы категоризации, один из которых гласит: «воспринимаемый мир не является неструктурированным множеством равновероятно встречающихся свойств, наоборот, объекты воспринимаемого мира имеют высоко коррелированную структуру» [27-31]. Поэтому непосредственно воспринимаемые объекты (basic objects) – информационно богатые связки наблюдаемых и функциональных свойств. В дальнейшем, Bob Rehder выдвинул теорию причинных моделей, в которой отношение объекта к категории основывается уже не на множестве признаков и близости по признакам, а на основании сходства порождающего причинного механизма: «объект классифицируется как член некоторой категории в той степени, в которой его свойства, вероятно, были сгенерированы причинными законами данной категории» [24]. Таким образом, за основу классификации берется уже структура причинных зависимостей между признаками объектов. Для формализации причинных моделей Bob Rehder предложил causal graphical models (CGMs). Однако эти модели основываются на «развертывании» байесовских сетей, которые не допускают циклов.

Основная гипотеза данной работы: информационные процессы работы мозга и сознание настроилась в процессе эволюции на извлечение высоко коррелированной структуры признаков «естественных» объектов путем формирования «естественных» понятий объектов.

¹ Эта работа поддержана Российским Фондом Фундаментальных Исследований № 15-07-03410-а.

Это подтверждается работами G.Tononi (см. раздел 4), который определяет сознание через понятие интегрированной информации – информации, генерируемой системой через причинное взаимодействие между ее частями [21]. Однако он рассматривает интегрированную информацию как внутреннее свойство системы. В отличие от G.Tononi, мы рассматриваем интегрированную информацию как способность системы отражать высокую коррелированность признаков объектов «естественных» классов, а сознание, как способность комплексного иерархического отражения «естественной» классификации объектов внешнего мира.

Нами предлагается принципиально новый математический аппарат для определения интегрированной информации, «естественной» классификации и «естественных» понятий, основанный на прямой формализации «резонанса» причинных связей. Мы предполагаем, что мозг осуществляет все возможные выводы по причинным связям, отражая высоко коррелированную структуру объектов внешнего мира. Эти причинные связи при восприятии «естественных» объектов замыкаются на себя образуя определенный «резонанс», что является системой с высоко интегрированной информацией в смысле G.Tononi. При этом, «резонанс» возникает тогда и только тогда, когда эти причинные связи отражают некоторый целостный «естественный» объект, в котором потенциально бесконечное множество признаков взаимно предполагают друг-друга. Возникающие при этом циклы выводов по причинным связям математически описываются «неподвижными точками» (см раздел 5), которые характеризуются тем, что дальнейшее применение выводов к рассматриваемым свойствам не предсказывает новых свойств. Полученное в неподвижной точке множество взаимно предсказывающихся свойств дает «образ» класса и «прототип» объекта. Поэтому мозг воспринимает «естественный» объект не набором признаков, а как «резонирующую» систему причинных связей замыкающихся на себя через одновременный вывод всей совокупности признаков «образа» или «прототипа» образующих целостный паттерн. В разделе 6 приведен пример компьютерного моделирования обнаружения «естественных» классов, «естественных» понятий и интегрированной информации на примере закодированных цифр.

2. «Естественная» классификация

Первый достаточно подробный философский анализ «естественной» классификации принадлежит Дж. Ст. Миллю [22]. В этом анализе выявлены все основные свойства «естественных» классификаций, которые в дальнейшем были подтверждены естествоиспытателями строившими «естественные» классификации и представителями когнитивных наук, исследовавшими «естественные» понятия.

По Дж. Ст. Миллю [22] «искусственные» классификации отличаются от «естественных» тем, что они могут быть основаны на любом одном или несколькими признаках, так что разные классы различаются только тем, что включают объекты, обладающие различными значениями этих признаков. Но если рассмотреть классы «животных» или «растений», то они отличаются столь большим (потенциально бесконечным) количеством свойств, что их нельзя перечислить. И все эти свойства будут основаны на утверждениях, подтверждающих это различие. При этом, среди свойств некоторого «естественного» класса будут как наблюдаемые, так и не наблюдаемые свойства. Чтобы учитывать скрытые признаки, надо находить их причинные проявления в наблюдаемых признаках. «Так, например, естественная классификация животных должна основываться преимущественно на их внутреннем строении; однако было бы странно, как замечает О. Конт, если бы мы были в состоянии определить род и вид того или другого животного, только предварительно убив его» [22].

Дж. Ст. Милль дает следующее определение «естественной» классификации: это такая классификация, которая объединяет объекты в группы относительно которых можно высказать наибольшее число общих предложений и она основывается на таких свойствах, которые служат причинами многих других или по крайней мере составляют их верные признаки. Он определяет также понятие «образа» класса, которое является предтечей «естественных» понятий: «наше понятие о классе – тот образ, которым этот класс представлен в нашем уме, – есть понятие о некотором образце, обладающем всеми признаками данного класса ... в самой высокой степени».

Рассуждения Дж. Ст. Милля были подтверждены естествоиспытателями. Например, Уэвель В.: «Чем больше общих утверждений об объектах дает возможность сделать классификация, тем она естественней» [7]. О неисчерпаемом количестве общих свойств у «естественных» классов пишет Рутковский Л. [9]: «Чем в большем числе существенных признаков сходны сравниваемые предметы, тем вероятнее их одинаковость и в других отношениях». Аналогичное высказывание

делает Смирнов Е.С. [10]: «Таксономическая проблема заключается в “индикации”: от бесконечно большого числа признаков нам нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки».

В результате была сформулирована проблема определения «естественных» классификаций, которая до сих пор обсуждается в литературе [5,34]. Однако, с нашей точки зрения, до настоящего времени нет достаточно адекватной формализации «естественной» классификации.

3. Принципы категоризации в когнитивных науках.

В работах Eleanor Rosch [27-31] на основании проведенных экспериментов сформулированы принципы категоризации «естественных» категорий [27,31], подтверждающие высказывания Дж. Ст. Милля и естествоиспытателей:

“Cognitive Economy. The first principle contains the almost common-sense notion that, as an organism, what one wishes to gain from one's categories is a great deal of information about the environment while conserving finite resources as much as possible. To categorize a stimulus means to consider it, for purposes of that categorization, not only equivalent to other stimuli in the same category but also different from stimuli not in that category. On the one hand, it would appear to the organism's advantage to have as many properties as possible predictable from knowing any one property, a principle that would lead to formation of large numbers of categories with as fine discriminations between categories as possible».

«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. ... In short, 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 sometimes with one, sometimes another attribute; others are rare; others logically cannot or empirically do not occur».

Понятно, что первый принцип невозможен без второго – когнитивная экономия не возможна без структурированности мира. Непосредственно воспринимаемые объекты (basic objects) – информационно богатые связи наблюдаемых и функциональных свойств, которые образуют естественную разрывность, создающую категоризацию. Эти связи формируют «прототипы» объектов классов (образец у 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» [30-31]. «Rosch and Mervis (1975) have shown that the more prototypical of a category a member is rated, the more attributes it has in common with other members of the category and the fewer attributes in common with members of the contrasting categories» [28-29].

В дальнейшем теория «естественных» понятий Eleanor Rosch получила название прототипической теории понятий (prototype theory). Основные ее черты описываются следующим образом: «The prototype view (or probabilistic view) keeps the attractive assumption that there is some underlying common set of features for category members but relaxes the requirement that every member have all the features. Instead, it assumes there is a probabilistic matching process: Members of the category have more features, perhaps weighted by importance, in common with the prototype of this category than with prototypes of other categories» [32].

К сожалению, в такой формулировке исчезают принципы категоризации, сформулированные Eleanor Rosch, говорящие об отражении в «естественных» понятиях высоко коррелированной структуры внешнего мира.

В дальнейших исследованиях было обнаружено, что моделей, основанных на признаках, сходстве и прототипах, недостаточно для описания 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.

Многие исследователи считают, что наиболее важным теоретическим знанием является причинное знание, ввиду его функциональности. Оно позволяет организму вмешиваться во внешние события и получать контроль над окружением. Исследования показали, что знания людей о категориях не сводится к перечню свойств, а включает богатое множество причинных взаимосвязей между этими свойствами, которые можно ранжировать. Важность свойств категории также зависит от их причинных взаимосвязей. В некоторых экспериментах [15,33] было показано,

что свойство важнее, если оно сильнее включено в причинную сеть взаимосвязей признаков. В других [23] экспериментах было показано, что свойство важнее, если у него больше причин.

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. In other words, theories might make certain combinations of features either sensible and coherent ... in light of the relations linking them, and the degree of coherence of a set of features might be an important factor determining membership in a category» [24].

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, как это понималось в генетическом подходе к «естественной» классификации: «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. As a result, causal-model theory will assign a low degree of category membership to objects that have many broken feature correlations (i.e., cases where causes are present but effects absent or vice versa). Objects that have many preserved correlations (i.e., causes and effects both present or both absent) will receive a higher degree of category membership because it is just such objects that are likely to be generated by causal laws» [24].

Для представления причинного знания некоторыми исследователями были использованы Байесовские сети или причинные модели [17,19-20]. Однако эти модели не могут моделировать циклические причинные связи, потому что Байесовские сети не поддерживают циклов.

В работе [26] Bob Rehder предложил модель причинных циклов, основанную на «раскрытии» причинных графических моделей (causal graphical models, CGMs). «Раскрытие» осуществляется путем создания байесовской сети, развертывающейся во времени. Такая сеть уже достаточно хорошо подходит для моделирования «естественных» понятий, отражающих высоко коррелированную структуру внешнего мира. Однако, она, все-таки, не в явном виде включает в себя формализацию циклов причинных связей.

Предлагаемая нами далее модель прямо основана на циклических причинных связях, представленных принципиально новыми математическими моделями – неподвижными точками предсказаний по причинным связям.

4. Сознание как интегрированная информация.

Если «естественная» классификация описывает объекты внешнего мира, а когнитивные науки – восприятие объектов внешнего мира, то теория интегрированной информации сознания G.Tononi анализирует информационные процессы мозга по восприятию объектов внешнего мира.

G.Tononi определяет сознание как первичное понятие, которое обладает следующими феноменологическими свойствами: composition, information, integration, exclusion [21]. Для более точного определения этих свойств G.Tononi вводит понятие интегрированной информации: «интегрированная информация, характеризующая редукцию неопределенности, это информация, генерируемая системой, приходящей в некоторое состояние через причинное взаимодействие между ее частями, которая превосходит информацию, генерируемую независимо ее частями самими по себе» [21]. В терминах интегрированной информации феноменологические свойства формулируются следующим образом. В скобках мы приводим интерпретацию этих свойств с точки зрения «естественной» классификации.

1. composition – elementary mechanisms (causal interactions) can be combined into higher-order

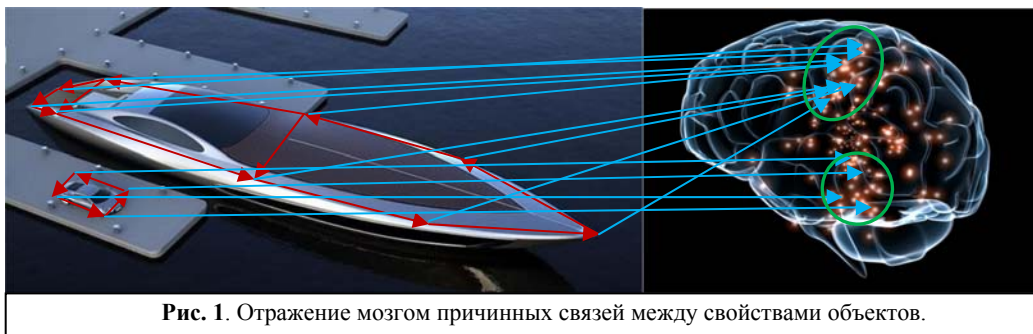


Рис. 1. Отражение мозгом причинных связей между свойствами объектов.

ones (образовывать «естественные» классы в виде причинных циклов и иерархию «естественных» классов);

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

3. integration – only information irreducible to non-interdependent components counts (значима только система «резонирующих» причинных связей, свидетельствующая об избытке информации и восприятию 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»).

Поскольку G.Tononi не рассматривает «естественную» классификацию объектов внешнего мира, то эти свойства определяются как внутренние свойства системы.

Мы рассмотрим эти свойства не как внутренние свойства системы, а как способность системы отражать комплексы причинных связей внешних объектов, а сознание – как способность комплексного иерархического отражения «естественной» классификации внешнего мира.

Рассмотрим процесс отражения причинных связей (рис. 1). Он включает:

1. объекты внешнего мира (машина, лодка, причал), относящиеся к некоторым «естественным» классам;
2. процесс отражения мозгом свойств объектов и связывающих их причинных связей, обозначенных синими линиями;
3. объединение возбужденных структур мозга в системы, обозначенные зелеными овалами.

В теории G.Tononi рассмотрен только третий пункт процесса отражения. Всю совокупность возбужденных групп нейронов в максимально интегрированную концептуальную структуру (maximally integrated conceptual structure) G.Tononi определяет как квалиа (quale).

Интегрированная информация у G.Tononi рассматривается как система циклических причинных связей. Однако остается непонятным, что же отражает интегрированная информация.

Нами выдвигается гипотеза о том, что «естественная» классификация, «естественные» понятия и интегрированная информация G.Tononi описываются одним и тем же формализмом и в определенном смысле тождественны друг другу. И мозг с помощью интегрированной информации настраивается на восприятие «естественных» объектов внешнего мира. Но для этого требуется принципиально новый формализм, описываемый ниже.

5. Единая формализация «естественной» классификации, «естественных» понятий и сознания, как интегрированной информации по G.Tononi

Мы предполагаем, что мозг осуществляет все возможные выводы, которые можно сделать по причинным связям, представляющим высоко коррелированную структуру внешнего мира. Тогда возникающие при этом циклы выводов математически описываются неподвижными точками, которые характеризуются тем, что дальнейшее применение выводов к рассматриваемым свойствам не предсказывает наличие дополнительных свойств. Полученное в неподвижной точке множество взаимно предсказывающихся свойств дает образ класса или прототип. Если искать неподвижные точки по всем возможным причинным связям, наблюдаемым на некотором множестве объектов, то получим «естественную» классификацию этих объектов.

Вывод по причинным связям является вероятностным (индуктивно-статистическим) предсказанием и неподвижные точки будут неподвижными точками предсказаний. Однако, вывод предсказаний индуктивно-статистическим выводом сталкивается с проблемой статистической двусмысленности, когда в процессе индуктивного вывода (обучения) получаются правила, из которых выводится противоречие – какое-то свойство одновременно должно быть и не должно быть. Отсюда возникает нетривиальная проблема – определить обнаружение причинных связей, предсказание и неподвижные точки так, чтобы противоречий не возникало.

Эта проблема была нами решена путем определения специального семантического вероятностного вывода причинных связей, который позволил получить нужные математические результаты:

1. доказано, что семантический вероятностный вывод обнаруживает причинные связи в виде максимально специфических правил, которые учитывают всю доступную информацию и предсказывают без противоречий, что решает проблему статистической двусмысленности [2,35];

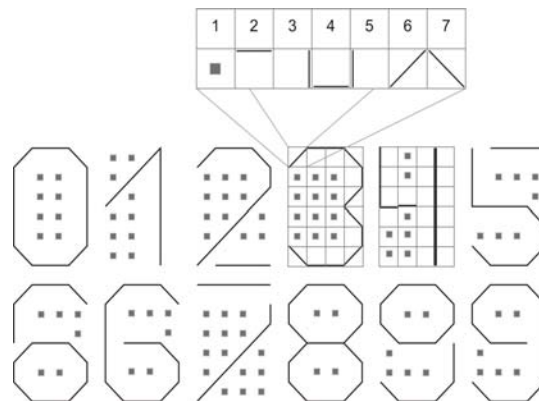


Рис. 2. Кодировка цифр.

2. доказано, что неподвижные точки по максимально специфическим правилам непротиворечивы [2,35];
3. доказано, что неподвижные точки по максимально специфическим правилам являются вероятностным обобщением формальных понятий [3-4, 38-39], исследуемых в анализе формальных понятий [17];
4. семантический вероятностный вывод может быть рассмотрен как формальная модель нейрона, а неподвижные точки как клеточные ансамбли [37].

Пусть $X(a)$ – множество свойств объекта a , заданных некоторым множеством предикатов, $a(P_{i_1} \& \dots \& P_{i_k} \Rightarrow P_{i_0}) \in MS(X)$ – множество максимально специфических условных связей, выполненных для свойств X , $\{P_{i_1}, \dots, P_{i_k}\} \subset X$. Тогда оператор предсказания Pr и неподвижная точка могут быть записаны следующим образом [39]:

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

где $\Phi_{Krit}(X)$ – оператор, модифицирующий множество признаков X путем добавления или удаления некоторого из признаков так, чтобы определенный критерий $Krit$ взаимной согласованности причинных связей по взаимному предсказанию признаков из X был максимальным. Критерий $Krit$ по-своему измеряет информационную интеграцию признаков по системе причинных связей $MS(X)$, как это делается в теории G.Tononi. Неподвижная точка достигается тогда, когда $Pr^{n+1}(X(a)) = Pr^n(X(a))$, для некоторого n , где Pr^n – n кратное применение оператора Pr . Поскольку при каждом применении оператора Pr значение критерия $Krit$ увеличивается и в неподвижной точке достигает локального максимума, то неподвижная точка, отражающая некоторый «естественный» объект, обладает максимумом интегрированной информации и свойством «exclusion» по G.Tononi. Однако принципиальным отличием данного подхода от теории G.Tononi является то, что система причинных связей в неподвижной точке сама, без внешних по отношению к ней оценок, формирует набор признаков «образа» или «прототипа» при восприятии «естественного» класса.

Проиллюстрируем формирование неподвижных точек компьютерным экспериментом по обнаружению «естественных» классов/понятий закодированных цифр.

Закодируем цифры как показано на рис. 2. Сформируем обучающее множество, состоящее из 360 перетасованных цифр (12 цифр рис. 2 продублированных в 30-ти экземплярах без указания, где какая цифра). На этом множестве семантическим вероятностным выводом было обнаружено 55089 максимально специфических закономерностей – общих утверждений об объектах, о которых говорит Дж. Ст. Миллем и Уэвель В.

По этим закономерностям было обнаружено ровно 12 неподвижных точек, точно соответствующих цифрам. Пример неподвижной точки для цифры 6 приведен на рис. 3. Рассмотрим, что представляет собой эта неподвижная точка. Занумеруем признаки цифр, как указано в таблице 1.

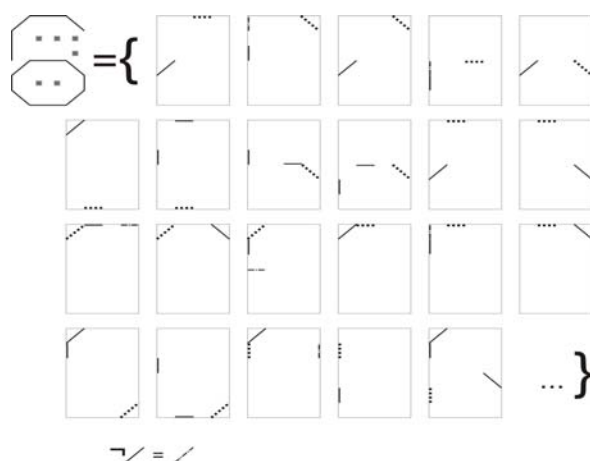


Рис. 3. Неподвижная точка цифры 6.

Первая закономерность цифры 6 рис. 3, представленная в первом прямоугольнике после фигурной скобки означает, что, если в квадрате 13 (табл. 1) стоит признак 6 (обозначим это как 13-6), то в квадрате 3 должен стоять признак 2 (обозначим как (3-2)). Предсказываемый признак обозначается точечной линией. Запишем эту закономерность как $(13-6 \Rightarrow 3-2)$. Нетрудно проверить, что эта закономерность действительно выполнена на всех цифрах. Вторая закономерность означает, что из признака (9-5) и отрицания значения 5 первого признака $\neg(1-5)$ (первый признак не должен быть равен 5) следует признак (4-7). Отрицание обозначается пунктирной линией, как показано в нижней части рис. 3. Получим закономерность $(9-5 \& \neg(1-5) \Rightarrow 4-7)$. Последующие 3 закономерности в первой строке цифры 6 будут соответственно $(13-6 \Rightarrow 4-7)$, $(17-5 \& \neg(13-5) \Rightarrow 4-7)$, $(13-6 \Rightarrow 16-7)$.

Таблица 1. Кодировка полей цифр.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24

На рис. 3 видно, что закономерности и признаки цифры 6 образуют неподвижную точку – взаимно предсказывают друг друга. Заметим, что при этом закономерности, используемые в неподвижной точке, выполнены на всех цифрах, а сама неподвижная точка выделяет только одну цифру. Это иллюстрирует феноменологическое свойство 2 ‘differences that make a difference’. Таким образом, система причинных связей воспринимает «осознает» целостный объект. Поэтому цифры выделяются не закономерностями самими по себе, а их системной взаимосвязью.

Неподвижная точка формирует «прототип» по Eleanor Rosch или «образом» у Дж. Ст. Миллю. Программа не знает заранее, какие сочетания признаков максимально коррелируют между собой.

ЛИТЕРАТУРА

1. Витяев Е.Е. Классификация как выделение групп объектов, удовлетворяющих разным множествам согласованных закономерностей // Анализ разнотипных данных (Выч. сист. 99), Новосибирск, 1983, С. 44-50.
2. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. НГУ, Новосибирск, 2006. С. 293.
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., Мартынович В.В. Вероятностные формальные понятия на контекстах с отрицаниями // Информационные технологии в гуманитарных исследованиях, Вып.19, ГПНТБ СО РАН, Новосибирск, 2014. С. 5-20.
5. Забродин В.Ю. О критериях естественной классификации // НТИ, сер.2, 1981, №8.
6. Кожара В.Л. Функции классификации // Теория классификаций и анализ данных, Новосибирск, 1982.
7. Мейен С.В., Шрейдер С.А. Методологические аспекты теории классификаций // Вопросы философии, 1976, №12.
8. Розова С.С. Классификационная проблема в современной науке. Новосибирск: Наука, 1986. 224 с.

9. Рутковский Л. Элементарный учебник логики. Санкт-Петербург, 1884.
10. Смирнов Е.С. Конструкция вида таксономической точки зрения // Зоол. Журн. Т. 17, №3, 1938, С. 387-418.
11. Субботин А.Л. Классификация. М.: ИФ РАН, 2001.
12. Шрейдер С.А. Систематика, типологии, классификация // Теория и методология биологических классификаций, М.: Наука, 1983.
13. Ahn, W. (1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, 69, 135–178.
14. Ahn, W., & Kim, N. S. (Eds.). (2001). *The causal status effect in categorization: An overview*. San Diego, CA: Academic Press.
15. Ahn, W., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, 41, 361–416.
16. Alexander Demin, Denis Ponomarev, Evgenii Vityaev. Probabilistic Concepts in Formal Contexts // Preliminary Proceedings of the Ershov Informatics Conference PSI Series, 8-th Edition (June 27 – July 1, 2011, Novosibirsk), Novosibirsk, 2011. P. 29-38.
17. Cheng, P. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104, 367-405.
18. Ganter, B. Formal Concept Analysis: Methods, and Applications in Computer Science. TU Dresden, Germany, 2003.
19. Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., & Kushnir, T. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3-23.
20. Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, 116, 56.
21. 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.
22. Mill, J.S. System of Logic, Ratiocinative and Inductive. L., 1983.
23. Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, 130, 323–360.
24. Bob Rehder. Categorization as causal reasoning // *Cognitive Science* 27 (2003) 709–748.
25. Rehder, B. (2003b). A causal-model theory of conceptual representation and categorization. *J. of Exper. Psych.: Learning, Memory, and Cognition*, 29, 1141-1159.
26. Bob Rehder, Jay B. Martin. Towards A Generative Model of Causal Cycles // 33rd Annual Meeting of the Cognitive Science Society 2011, (CogSci 2011), Boston, Massachusetts, USA, 20-23 July 2011, V.1 pp. 2944-2949.
27. Rosch, E.H. Natural categories // *Cognitive Psychology* 4. P. 328-350.
28. Rosch, E., Mervis, C.B. Family resemblances. Studies in the internal structure of categories // *Cognitive Psychology*, 7. P. 573–605.
29. Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., Boyes-Braem, P. Basic objects in natural categories // *Cognitive Psychology*, 8. P. 382–439.
30. Rosch, Eleanor and Lloyd, Barbara B. (eds), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum, 1978, P. 27-48.
31. Rosch, E., Principles of Categorization // Rosch, E. & Lloyd, B.B. (eds), *Cognition and Categorization*, Lawrence Erlbaum Associates, Publishers, (Hillsdale), 1978. P. 27–48
32. B. H. Ross, E. G. Taylor, E. L. Middleton, and T. J. Nokes. Concept and Category Learning in Humans // H. L. Roediger, III (Ed.), *Cognitive Psychology of Memory*. Vol. [2] of *Learning and Memory: A Comprehensive Reference*, 4 vols. (J. Byrne Editor), Oxford: Elsevier, 2008, P. 535-556.
33. Sloman, S., Love, B. C., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, 22, 189–228.
34. The Nature of Classification. Relationships and Kinds in the Natural Sciences. Palgrave Macmillan. 2013. 208p.
35. Evgenii Vityaev. The logic of prediction // *Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference* (August 16-19, 2005, Novosibirsk, Russia), edited by S.S. Goncharov, R. Downey, H. Ono, World Scientific, Singapore, 2006, P. 263-276.
36. Vityaev E.E. A formal model of neuron that provides consistent predictions // *Biologically Inspired Cognitive Architectures* 2012. Proceedings of the Third Annual Meeting of the BICA Society (A. Chella, R. Pirrone, R. Sorbello, K.R. Johannsdottir, Eds). In *Advances in Intelligent Systems and Computing*, v.196, Springer: Heidelberg, New York, Dordrecht, London. 2013, P. 339-344.
37. E.E. Vityaev, A.V. Demin, D. K. Ponomarev. Probabilistic Generalization of Formal Concepts // *Programming and Computer Software*, 2012, Vol. 38, No. 5. P. 219–230.
38. E.E. Vityaev, V.V. Martinovich. Probabilistic Formal Concepts with Negation // A. Voronkov, I. Virbitskaite (Eds.): *PCI 2014, LNCS 8974*, P. 385-399.
39. 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.