

***Информационная не редуccionистская теория сознания,
обеспечивающая максимальную точность прогноза реальности***

*Informational non-reductionist theory of consciousness
that providing maximum accuracy of reality prediction*

**意识的信息性非还原论
这确保了现实预测的最大准确性**

Витяев Е.Е.

Vityaev E.E.

Институт математики им. С.Л. Соболева СО РАН,
Новосибирск, Россия,
Sobolev Institute of Mathematics SB RAS, Novosibirsk, Russia,
S.L.Sobolev 数学研究所 SB RAS, 俄罗斯新西伯利亚,
vityaev@math.nsc.ru

Аннотация. В работе рассматривается не редуccionистская теория сознания, которая не сводимая к теориям реальности и к физиологическим или психологическим теориям. Следуя «информационному подходу» Д.И.Дубровского к проблеме «сознание и мозг», будем рассматривать реальность через призму информации о наблюдаемых явлениях, которая, в свою очередь воспринимается субъективной реальностью через ощущения, восприятия, чувства и т.д., которые, в свою очередь, являются информацией относительно соответствующих мозговых процессов. В рамках такого подхода выдвигается следующий принцип построения Информационной Теории Сознания (ИТС): *мозг обнаруживает все возможные причинные связи во внешнем мире и делает все возможные выводы по ним*. В работе показывается, что построенная на нем ИТС: (1) основывается на информационных законах структуры внешнего мира; (2) объясняет строение и функционирование функциональных систем и клеточных ансамблей; (3) обеспечивает максимальную точность прогноза и предвосхищения реальности; (4) разрешает возникающие противоречия и (5) является информационной теорией отражения мозгом реальности.

Ключевые слова. реальность, сознание, мозг, информация, отражение.

Annotation. The paper considers a non-reductionist theory of consciousness, which is not reducible to theories of reality and to physiological or psychological theories. Following D.I.Dubrovsky's "informational approach" to the "Mind-Brain Problem", we consider the reality through the prism of information about observed phenomena, which, in turn, is perceived by subjective reality through sensations, perceptions, feelings, etc., which, in turn, are information about the corresponding brain processes. Within this framework the following principle of the Information Theory of Consciousness (ITS) development is put forward: *the brain discovers all possible causal relations in the external world and makes all possible inferences by them*. The paper shows that ITS built on this principle: (1) also base on the information laws of the structure of external world; (2) explains the structure and functioning of the brain functional systems and cellular ensembles; (3) ensures maximum accuracy of predictions and the anticipation of reality; (4) resolves emerging contradictions and (5) is an information theory of the brain's reflection of reality.

Keywords. reality, consciousness, mind-brain problem, brain, information.

注释, 注释. 著作探討了非還原論的意識理論，不能還原為現實理論以及生理或心理學理論。遵循 D. I. 杜布羅夫斯基對“意識與大腦”問題的“信息方法”，我們將通過觀察到的現象的信息棱鏡來考慮現實，而這些現象又通過感覺、知覺、感受等被主觀現實所感知。……，這又是有關相應大腦過程的資訊。在這個方法的框架內，提出了建構意識資訊理論（ITS）的以下原則：大腦偵測外在世界中所有可能的因果關係，並從中得出所有可能的結論。該工作表明，建立在其之上的智慧交通系統：（1）基於外部世界結構的資訊規律；（2）解釋功能係統和細胞組件的結構和功能；（3）確保對現實的預測和預期的最大準確性；（4）解決了新出現的矛盾，（5）是關於大腦如何反映現實的資訊理論。

關鍵字. 現實、意識、大腦、資訊、反思。

1. Введение

Не редукционистская теория сознания не сводима к какой-либо теории реальности и к какой-либо физиологической или психологической теории.

На Седьмой международной конференции по когнитивной науке К.В.Анохин в своем докладе говорил, что «Проблема не в том, что существующие нейрофизиологические теории несовершенны ... Используемые в них коррелятивные подходы просто не могут дать ответа на вопросы о природе разума и субъективного опыта ... Для этого нужна нередукционная фундаментальная теория» [3].

Макс Тегмарк в своей книге [14] также писал, что между внешней реальностью «External Reality» и «Internal Reality» должна быть промежуточная реальность «Consensus Reality», описывающая наблюдаемую реальность в физических терминах и отражающаяся во внутренней реальности.

Следуя Д.И.Дубровскому [10] будем использовать «информационный подход» к описанию «Internal Reality» как субъективной реальности. При этом субъективная реальность – это реальность осознаваемых состояний индивида - ощущений, восприятий, чувств, мыслей, намерений, желаний и т.д.

В тоже время явления субъективной реальности рассматриваются как информация, связанная с соответствующим мозговым процессом как со своим носителем. Таким образом, в рамках информационного подхода к проблеме «сознания и мозг» мы получаем следующую схему рис. 1. В ней реальность описывается через информацию о наблюдаемых явлениях, которая воспринимается субъективной реальностью через ощущения, восприятия, чувства и т.д., которые в свою очередь являются информацией относительно соответствующих мозговых процессов.

Каково назначение этой информации? Точнее всего он определен в «принципе эволюции живого мира»,



Рис. 1. Сознание и Мозг.

Fig. 1. Consciousness and Brain.

圖 1. 意識 和 大腦。

сформулированного П.К.Анохиным: «Сложилась одна универсальная закономерность в приспособлении организмов к внешним условиям, которая в дальнейшем бурно развивалась на протяжении всей эволюции живого мира: в высшей степени быстрое отражение медленно разворачивающихся событий внешнего мира» [5].

Раскроем этот принцип и сформулируем принципы построения информационной теории сознания.

I. Во-первых, следуя «принципу эволюции живого мира» эта информационная теория должна предвосхищать события внешнего мира. Однако, если бы внешний мир был случаен, то никакое его предвосхищение было бы невозможно.

Но наш мир хорошо структурирован. Если в информационной структуре реальности есть какие-то законы, то естественно предположить, что в процессе эволюции вырабатывались такие нейробиологические механизмы, которые бы использовали эти структуры для обеспечения наиболее точного отражения реальности. Поэтому необходим следующий принцип.

II. Информационная теория отражения должна основываться на законах структуры внешнего мира и описывать одновременно как информационную структуру реальности, так и нейрофизиологические и другие механизмы, которые обеспечивают отражение этой структуры в терминах субъективной реальности.

Далее будут приведены такие законы информационной структуры реальности. Нейробиологические механизмы, соответствующие им и использующие эти структуры, указал в своем докладе «Когнитом – гиперсетевая модель мозга» К.В. Анохин. Это КОГи (Когнитивные Группы Нейронов), обобщающие теории функциональных систем и клеточных ансамблей Д. Хебба [17].

В работе предлагается следующий фундаментальный принцип Информационной Теории Сознания (ИТС), которого достаточно для объяснения основных упомянутых информационных процессов:

ПРИНЦИП неограниченного вывода: *Мозг обнаруживает все возможные причинные связи во внешнем мире и делает все возможные выводы по ним.*

Оказывается, что этого принципа достаточно для построения Информационной Теории Сознания, которая:

1. Объясняет строение и функционирование когов функциональных систем и клеточных ансамблей Д. Хебба.
2. Основывается на приводимых далее информационных законах структуры внешнего мира.
3. Обеспечивает максимальную точность прогноза и реальности.
4. Разрешает возникающие противоречия.

5. Является информационной теорией отражения реальности мозгом.

2. Причинность

Рассмотрим первый закон информационной структуры внешнего мира – его причинность. Причинность является следствием физического детерминизма: «для любой изолированной физической системы некоторое ее состояние определяет все последующие состояния» [12]. Но рассмотрим автомобильную аварию [12]. В чем её причина? Это может быть состояние дорожного покрытия, его влажность, положение солнца относительно водителя, безрассудное вождение, психологическое состояние водителя, неисправность тормозов и т.д. Очевидно, что в этом случае нет определенной причины.

В философии науки причинность сводится к предсказанию и объяснению. «Причинно-следственная связь означает предсказуемость... если известна вся предыдущая ситуация, событие может быть предсказано..., если даны все факты и законы природы, связанные с этим событием» [12]. Понятно, что знать все факты, число которых, как в случае аварии, потенциально бесконечно и все законы невозможно. Кроме того, человек и животные узнают законы внешнего мира путем обучения. Поэтому, причинность сводится к предсказанию с помощью индуктивно-статистического вывода, когда предсказание выводятся из фактов и статистических законов с некоторой вероятностью.

Кроме того, причинно-следственные связи в виде статистических законов, обнаруженных на реальных данных или в результате обучения, сталкиваются с проблемой статистической двусмысленности – из них могут быть выведены противоречивые предсказания [18]. Чтобы избежать этой двусмысленности Гемпель ввел требование максимальной специфичности [18], неформально состоящее в том, что статистические законы должны включать максимум наличной информации.

Нами была решена проблема статистической двусмысленности и определены максимально Максимально Специфические Причинные Связи (МСПС), для которых было доказано, что индуктивно-статистический вывод, использующий их, не приводит к противоречиям [28,32] и тем самым максимально точно реализуют принцип эволюции живого мира П.К.Анохина. Нами был разработан специальный семантический вероятностный вывод [28,32], который обнаруживает МСПС. Он, в частности, удовлетворяет определению Cartwright [15] вероятностной причинной связи относительно некоторого бэкграунда, состоящее в том, что каждое условие посылки причинной связи строго увеличивает условную вероятность заключения. Также нами разработана формальная модель нейрона [30], удовлетворяющая правилу Хебба, которая осуществляет этот вывод и обнаруживает МСПС причинные связи [28,32].

Таким образом, точно анализируя понятие причинной связи мы получаем, что этот информационный закон реальности может быть представлен в ИТС обнаружением причинных связей на уровне нейрона, обеспечивающих его пластичность и проявляющуюся, в частности, в условных реакциях. При этом они могут быть максимально точными посредством использования МСПС условных связей.

Это обеспечивает *первый уровень максимальной точности предсказаний* и предвосхищения реальности, реализующие принцип опережающего отражения действительности.

3. «Естественная» классификация

Перейдем к следующему закону информационного строения объектов внешнего мира - «естественной» классификации. Первый достаточно подробный анализ «естественной» классификации принадлежит Дж. Ст. Миллю [21]. Сначала отделим «искусственные» классификации от «естественных»: «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», ... то мы найдем, что в этом отношении одни классы сильно отличаются от других. ... обладают столь большим количеством признаков, что их нельзя ... перечислить» [21].

«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» [21]. На основании понятия «естественной» классификации Дж. Ст. Милль определяет понятие «образа» класса, как о некотором образце, обладающем всеми признаками данного класса.

Естествоиспытатели писали, что создание «естественной» классификации заключается в «индикации» – от бесконечно большого числа признаков нужно перейти к ограниченному их количеству, которое заменило бы все остальные признаки [13]. Это означает, что в «естественных» классах признаки сильно коррелированы, например, если есть 128 классов и атрибуты двоичные, то независимыми «индикаторными» атрибутами среди них могут быть только 7 атрибутов, т.к. $2^7 = 128$, а другие атрибуты могут быть предсказаны по значениям этих 7 атрибутов. Мы можем выбирать различные 7-10 атрибутов в качестве «индикаторных» и тогда другие атрибуты, которых потенциально бесконечно много, могут быть предсказаны по этим выбранным атрибутам. Поэтому, существует экспоненциальное (относительно числа атрибутов) число причинно-следственных связей, связывающих атрибуты объектов «естественных» классов. Такая избыточность информации при восприятии

объектов внешнего мира подтверждается в когнитивных науках при анализе «естественных» понятий.

4. «Естественные» понятия в когнитивных науках

Высоко коррелированную структуру внешнего мира выявляет также теория «естественных» понятий. При этом, «естественная» классификация выявляет структуру объектов внешнего мира, а «естественные» понятия, исследуемые в когнитивных науках, определяют восприятие этих «естественных» объектов как элементов субъективной реальности.

В работах Eleanor Rosch был сформулирован следующий принцип категоризации «естественных» категорий: «Perceived World Structure ... 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» [25].

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» [24].

В дальнейших исследованиях было обнаружено, что моделей, основанных на признаках, сходстве и прототипах, недостаточно для описания «естественных» классов. 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» [23]. 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.

Для представления причинного знания Bob Rehder использовал Байесовские сети [22]. Однако они не поддерживают циклов и поэтому не могут моделировать циклические причинные связи. Предлагаемая нами далее формализация в виде вероятностных формальных понятий прямо моделирует циклические причинные связи [5-8,28-29,32].

5. Теория интегрированной информации G.Tononi

На высоко коррелированной структуре внешнего мира основана также теория интегрированной информации G.Tononi [20, 26-27]. Если «естественная» классификация описывает объекты внешнего мира, а когнитивные науки – восприятие объектов внешнего мира, то теория интегрированной информации анализирует информационные процессы мозга при восприятии объектов внешнего мира.

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» [27].

Взаимосвязь интегрированной информации с реальностью G.Tononi описывает следующим образом: «Cause-effect matching ... measures how well the integrated conceptual structure ... fits or ‘matches’ the cause-effect structure of its environment», «... 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» [26-27].

G.Tononi определяет сознание как первичное понятие, которое обладает следующими феноменологическими свойствами: composition, information, integration, exclusion [20, 26-27]. 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 «естественную» классификацию объектов внешнего мира. Тогда сознание, в отличие от G.Tononi

определяется не феноменологическими свойствами нейронных структур, а как способность мозга, используя интегрированную информацию нейронных структур, отражать мир, представленный иерархической «естественной» классификацией в системе «естественных понятий» и их причинных моделей.

6. Формализация «естественной» классификации, «естественных» понятий и сознания, как интегрированной информации по G.Toponi

В соответствии с Принципом неограниченного вывода, мозг осуществляет все возможные выводы по причинным связям. Эти причинные связи, которых экспоненциальное число, в процессе восприятия «естественных» объектов заикливаются сами на себя, образуя определенный «резонанс», который является системой с высоко интегрированной информацией в смысле G.Toponi. 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», 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, «прототип» понятия и его «причинную модель». 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 «причинную модель».

Можно показать, что МСПС причинные связи, организованные в клеточные ансамбли, позволяют максимально достоверно идентифицировать объекты внешнего мира и затем максимально точно предсказывать свойства этих объектов, используя эту идентификацию. Такие предсказания получаются более точными, поскольку для предсказания используются только относящиеся к этому классу МСПС причинные связи. Это образует *второй*, ещё более точный, с точки зрения прогноза, *уровень организации информационных процессов*.

Нами предлагается принципиально новый математический аппарат для определения интегрированной информации, «естественной» классификации и «естественных» понятий. Наша формализация основана на вероятностном обобщении анализа формальных понятий [8, 29-32]. Формальные понятия могут быть определены как неподвижные точки детерминированных правил (не имеющих исключений) [19]. Но, как писал Дж. Милль «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». Поэтому нужно уйти от детерминированных правил и заменить их на вероятностные, чтобы определять признаки не точно, а для большинства. Поэтому мы обобщаем формальные понятия на вероятностный случай, заменяя детерминированные правила на МСПС причинные связи и определяя вероятностные формальные понятия как неподвижные точки этих максимально специфических правил [8,29-32]. В силу того, что вывод по максимально специфическим причинным связям непротиворечив, полученная неподвижная точка также будет непротиворечива и не будет содержать одновременно признак и его отрицание, т.е. такое определение вероятностных формальных понятий корректно.

Можно показать [9], что вероятностные формальные понятия адекватно формализуют «естественную» классификацию и, кроме того, полученная «естественная» классификация удовлетворяет всем требованиям, которые естествоиспытатели предъявляли к «естественным» классификациям [9].

Рассмотрим пример компьютерного моделирования обнаружения «естественных» классов, «естественных» понятий и интегрированной информации для закодированных цифр. Пусть $X(a)$ – множество свойств объекта a , заданных некоторым множеством предикатов, а $(P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}) \in MS(X)$ – множество МСПС причинных связей выполненных для свойств X , $\{P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}\} \subset X$.

Тогда оператор предсказания Pr и неподвижная точка могут быть записаны следующим образом [6,9]:

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

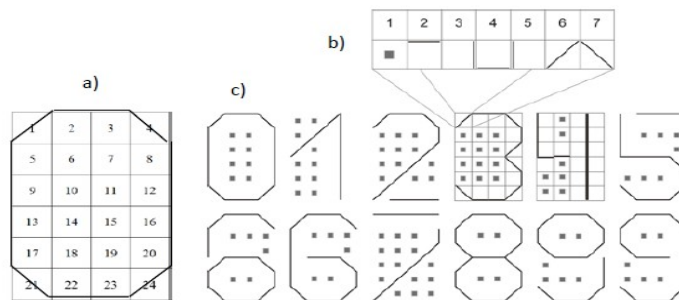


Рис. 2. Кодировка цифр

Fig. 2. Number encoding

圖 2. 數位編碼

где $\Phi_{Krit}(X)$ – оператор, модифицирующий множество признаков X путем добавления или удаления некоторого признака так, чтобы определенный критерий $Krit$ взаимной согласованности причинных связей по взаимному предсказанию признаков из X был максимальным [6,9]. Критерий $Krit$ измеряет

информационную интеграцию признаков по системе причинных связей $MS(X)$ иначе, чем 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 локальный максимум, то неподвижная точка, отражающая некоторый «естественный» объект, обладает максимумом интегрированной информации и свойством «exclusion» по G.Tononi.

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 МСПС причинных связей – общих утверждений об объектах,

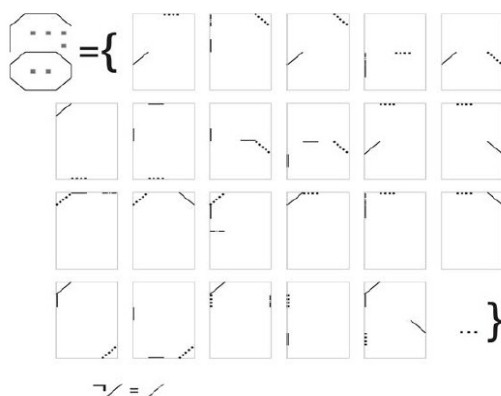


Рис. 3. Неподвижная точка цифры 6.

Fig. 3. Fixed point of number 6.

圖 3. 6 號固定點。

о которых говорил Дж. Ст. Милль. 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 на рис. 2. Первая закономерность цифры 6 рис. 3, представленная в первом прямоугольнике после фигурной скобки означает, что, если в квадрате 13 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 причинную связь как $(13-6 \Rightarrow 3-2)$. Нетрудно проверить, что эта причинную связь выполнена на всех цифрах. Вторая причинная связь 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 на рисунке пунктирной линией, как показано в нижней части рис. 3. Получим причинную связь $(9-5 \& \neg(1-5) \Rightarrow 4-7)$.

Последующие 3 причинных связей в первой строке цифры 6 будут соответственно $(13-6 \Rightarrow 4-7)$, $(17-5 \& \neg(13-5) \Rightarrow 4-7)$, $(13-6 \Rightarrow 16-7)$.

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

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

Вероятностными формальными понятиями, описываются не только «естественные» понятия, но и контексты. Контексты также обладают свойством максимальной точности предсказания – причинные связи, обнаруженные на некотором контексте и составляющие его причинную модель будут точнее предсказывать свойства данного контекста.

Рассмотрим следующий пример рис. 4, содержащий как цифры, так и буквы.

Можно обучиться только на цифрах и построить вероятностные формальные понятия цифр, можно обучиться на буквах и построить вероятностные понятия только букв, а можно обучиться на цифрах и буквах и построить формальные понятия цифр и букв.

1	2	3	4	5	6	7	8	9
А	Б	В	Г	Д	Е	Ё	Ж	З
И	Й	К	Л	М	Н	О	П	Р
С	Т	У	Ф	Х	Ц	Ч	Ш	Щ
Ъ	Ы	Ь	Э	Ю	Я			

Рис. 4. Два контекста - Цифры и Буквы

Fig. 4. Two contexts - Numbers and Letters

圖 4. 兩個上下文 - 數字和字母

В каждом из этих случаев будут обнаружены различные МСПС причинные связи, но МСПС причинные связи, описывающие совместно цифры и буквы будут содержать дополнительные признаки, отделяющие их друг от друга, что получается автоматически по МСПС причинным связям. При рассмотрении (в контексте) и букв и цифр МСПС причинные связи будут иметь более высокую вероятность, чем МСПС отдельно для цифр и букв, поэтому именно они будут срабатывать в этом контексте. Наша формальная модель нейрона,

обнаруживающая максимально специфические причинные связи [30] следует известному физиологическому свойству нейронов – быстрее по времени срабатывают более вероятные условные стимулы.

7. Теория функциональных систем

Формализация когов второго типа – когов функциональных систем, основана на рассмотрении целенаправленного поведения, которое осуществляется путем выработки условных (причинных) связей между действиями и их результатами. Этих условных связей достаточно для моделирования функциональных систем и разработки аниматов.

Ещё П.К. Анохиным писал, что «Речь идет о коллатеральных ответвлениях пирамидного тракта, отводящих ко многим нейронам “копии” тех эфферентных посылок, которые выходят на пирамидный тракт» [4-5]. Таким образом, когда моторный нейрон отправляет сигнал мышцам о некотором действии, то копии этого возбуждения отправляются в том числе в проекционные зоны, которые могут фиксировать результат выполненного действия. Поэтому мозг обнаруживает все причинные связи между действиями и их результатами.

Покажем на схеме рис 5, что этого достаточно для объяснения основных механизмов формирования функциональных систем работы мозга [7,31]. Предположим, что у нас нет ещё опыта и возникло мотивационное возбуждение, показанное черным треугольником. Тогда для удовлетворения потребности методом проб и ошибок мы можем сделать некоторое действие, которое будет активировано некоторым нейроном, обозначенным белым треугольником.

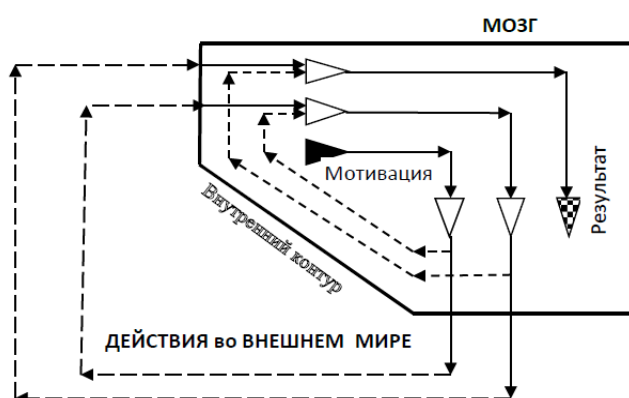


Рис. 5. Формирование МСПС условных связей между действиями и результатами

Fig. 5. Formation of conditional connections between actions by MSPS and results

圖 5. 透過 MSPS 在動作之間形成條件連接 和結果

Одновременно с активацией этого действия «копия» возбуждения этого нейрона будет отправлена в проекционные зоны, где найдется нейрон, который

среагирует на результат действия, поступивший из внешнего мира. Поскольку этот же нейрон предварительно получит возбуждение от активации действия белым нейроном, то он сформирует условную связь между активацией действия белым нейроном и полученным результатом. Если теперь после получения этого результата, изменившего обстановку, осуществить некоторое следующее действие, также обозначенное белым треугольником, то получим следующий результат, для которого также найдется нейрон, который среагирует на результат этого действия. Если в результате потребность будет удовлетворена и цель достигнута, то вся цепочка активных нейронов и условных связей, приведшая к результату будет подкреплена и занесена в память. Таким образом, возникнет *внутренний контур* прогноза достижения результата по причинным связям. Тогда, при следующем возникновении мотивационного возбуждения, эта цепочка действий будет извлечена из памяти и спрогнозирует достижение результата по внутреннему контуру ещё до всяких действий. Так будет сформирован план действий, который по внутреннему контуру, как сказано в цитате П.К.Анохина, активирует нейроны, ожидающие результаты действий, что сформирует *акцептор результатов действий*, подробно исследованный в теории функциональных систем.

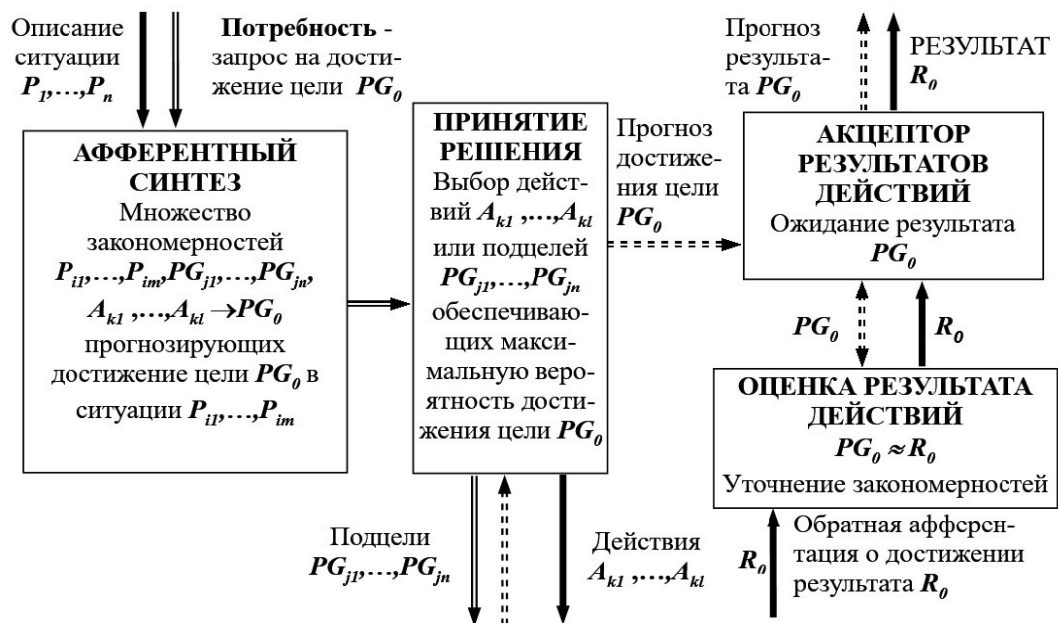


Рис. 6. Схема функциональной системы

Fig. 6. Functional system diagram

圖 6. 功能係統圖

Таким образом, формирование и работу функциональной системы можно объяснить формированием причинных связей между действием и его результатами.

В терминах МСПС причинных связей, схема работы функциональных систем выглядит следующим образом рис 6. [7,31]. Потребность будем рассматривать как запрос к функциональной системе на достижение цели, обозначенной предикатом PG_0 . Этот запрос поступает в блок афферентного синтеза и для функциональных систем, не имеющих функциональных подсистем, извлекает из памяти причинные связи вида $P_{i1}, \dots, P_{im}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$, приводящие к достижению цели PG_0 , где $P_{i1}, \dots, P_{im}$ – свойства обстановки необходимые для достижения цели, а $A_{k1}, \dots, A_{kl}$ – последовательность действий, приводящая к цели. При этом свойства $P_{i1}, \dots, P_{im}$ должны присутствовать в свойствах обстановки $P_1, \dots, P_n$, поступающими в блок афферентного синтеза. Для иерархически организованных функциональных систем этот запрос извлекает из памяти МСПС причинные связи более сложного вида $P_{i1}, \dots, P_{im}, PG_{j1}, \dots, PG_{jn}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$, включающие в себя запросы на достижение подцелей $PG_{j1}, \dots, PG_{jn}$.

Далее извлеченные правила поступают в блок принятия решений, где по каждому из правил осуществляется прогноз достижения цели и вычисляется оценка вероятности достижения цели. Прогноз по правилам, где только действия, осуществляется по вероятности самого правила. Прогноз по правилам с запросами к подцелям осуществляется путем отправки этих запросов в функциональные подсистемы, принятием в них решений и получением от них вероятностей прогноза достижения соответствующих подцелей. Результирующая вероятность прогноза вычисляется произведением вероятности правила на вероятности достижения подцелей. После этого осуществляется принятие решения о способе достижения цели и выбирается правило, имеющее максимальную вероятность прогноза достижения цели.

Затем формируется план действий, включающий все действия, включенные в правило и все действия, которые есть в функциональных подсистемах. Одновременно с планом действий формируется аппарат акцептора результатов действия, включающий ожидание всех спрогнозированных подрезультатов в функциональных подсистемах, а также в самой функциональной системе. После этого план действий начинает выполняться, а ожидаемые результаты сравниваться с получаемыми результатами.

Если все подрезультаты и конечный результат достигнуты и совпали с ожидаемыми результатами, то само правило и все правила функциональных подсистем, которые были выбраны в процессе принятия решений подкрепляются и их вероятность увеличивается. Если в какой-то подсистеме результат не достигнут, то соответствующее правило этой функциональной подсистемы штрафуются. Тогда возникает ориентировочно-исследовательская

реакция, которая пересматривает принятое решение. Данная модель успешно применялась для моделирования аниматов [11,16,31].

8. Сознание как механизм разрешения противоречий

Рассмотрим теперь **третий уровень точности предсказаний**, обеспечивающийся информационной теорией сознания – сознанию как механизму разрешения противоречий.

Мир многогранен как алмаз рис. 7 и не существует его единого непротиворечивого описания и функция сознания – правильно выбрать соответствующий контекст, в рамках которого можно получить максимально точное предсказание. В науке такими контекстами являются парадигмы, формирующими определенный взгляд, точку зрения и соответствующую систему понятий той или иной теории. Эти парадигмы, как правило, не совместимы между собой.

Такую точку зрения на сознания также высказывает В.М.Аллахвердов [2]. В своей работе [1] он пишет: «Сознание, столкнувшись с противоречивой информацией, пытается удалить эту информацию с поверхности сознания или так ее видоизменить, чтобы противоречие исчезло или перестало осознаваться как противоречие». В этой работе он приводит 7 случаев разрешения противоречий сознанием. Все эти случаи объясняются работой или взаимодействием вероятностных формальных понятий, которые определяют рассматриваемые понятия или контексты. Рассмотрим для краткости два из них:

1. Случай 1. Самый простой способ избавиться от противоречия или двусмысленности — выбрать для осознания какую-то одну интерпретацию, а все остальные (несовместимые с ней) не осознавать (негативно выбрать).

Пример. Явление бинокулярной конкуренции, когда испытуемому одновременно предъявляются на разные глаза разные стимулы. Если предъявлять два изображения, одно из которых является более вероятным или более знакомым, испытуемые преимущественно видит только его.

Объяснение. Вероятностное формальное понятие взаимно предсказывает свойства, входящие в понятие, а также отрицание других свойств, которых в нем не должно быть и тем самым вытормаживает альтернативы.

2. Случай 2. При осознании разных сторон противоречия делается попытка найти способ объяснения — соединение разных сторон в непротиворечивое целое.



Рис. 7. Многогранная реальность

Fig. 7. Multifaceted reality

圖 7. 多層面的現實

Пример. В условиях бинокулярной конкуренции, если предъявить на один глаз красный круг, а на другой — черный треугольник, то испытуемый увидит черный треугольник на красном фоне.

Объяснение. Если воспринимаемые признаки не противоречат друг другу и не вытормаживают друг друга, то могут образовать объединенное вероятностное формальное понятие и соответственно восприниматься.

Во всех случаях можно правильно выбрать наиболее подходящее вероятностное формальное понятие или контекст для разрешения противоречия и получения наиболее точного предсказания в соответствии с причинными связями выбранного понятия или контекста.

1. Introduction

The non-reductionist theory of consciousness is not reducible to any theory of reality and to any physiological or psychological theory.

At the Seventh International Conference on Cognitive Science, K.V.Anokhin said "The problem is not that the existing neurophysiological theories are imperfect ... The correlative approaches used in them simply cannot answer questions about the nature of mind and subjective experience ... This requires a non-reductionist fundamental theory" [3].

Max Tegmark in his book [14] also write that

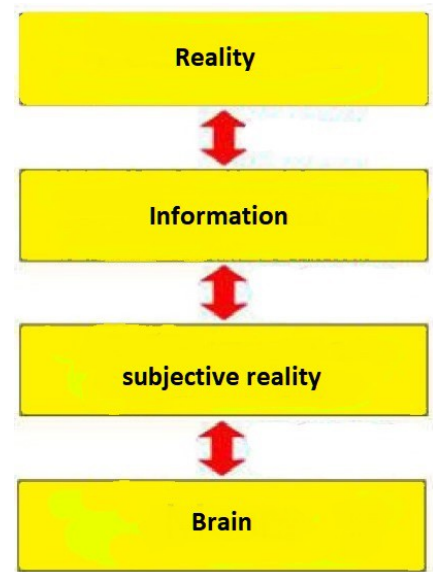


Fig. 1. Consciousness and Brain

between the external reality "External Reality" and "Internal Reality" there should be an intermediate "Consensus Reality", describing the external reality in physical terms and at the same time is reflected in the internal reality.

Following D.I. Dubrovsky [10], we will use an "informational approach" to the description of "Internal Reality" as a *subjective reality*. Herewith, the subjective reality is the reality of an individual's conscious states – sensations, perceptions, feelings, thoughts, intentions, desires, etc. At the same time, the phenomena of subjective reality are considered as information related to the corresponding brain process as its carrier.

Thus, within this framework of the information approach to the "Mind-Brain Problem", we get the following scheme fig. 1., where reality is described through information about observed phenomena, which are perceived by subjective reality through sensations, perceptions, feelings, etc., which in turn are information about the corresponding brain processes.

What is the purpose of this information? Most precisely, it is defined in the "principle of the evolution of the living world", formulated by P.K. Anokhin: "There was one universal pattern in the adaptation of organisms to external conditions, which later developed rapidly throughout the evolution of the living world: a highly rapid reflection of the slowly unfolding events of the external world" [5].

Let us reveal this principle and formulate the principles of creating an information theory of consciousness.

I. First, following the "principle of the evolution of the living world", this information theory should anticipate the events of the outside world.

However, if the external world were accidental, then anticipation of it would be impossible. But our world is well structured. If there are any laws in the information structure of reality, then it is natural to assume that in the process of evolution there were developed such neurobiological mechanisms that would use these structures to produce the most accurate reflection of reality. Therefore, the following principle is necessary.

II. The information theory of reflection should be based on the laws of the structure of the external world and describe simultaneously both the information structure of reality and neurophysiological and other mechanisms that ensure the reflection of this structure in terms of subjective reality.

The following laws of the information structure of reality will be given. The corresponding neurobiological mechanisms that use these structures were indicated by K.V. Anokhin in his report "Cognitome – hypernetwork model of the brain". These are KOGi (Cognitive Groups of Neurons), generalizing the ideas of the functional systems theory and D. Hebb's cellular ensembles [17].

We propose the following fundamental principle of the Information Theory of Consciousness (ITS), which is sufficient to explain the basic information processes:

THE PRINCIPLE OF unlimited inference: *The brain detects all possible*

causal connections in the external world and makes all possible conclusions on them.

It turns out that this principle is sufficient to build an ITS, which:

1. Explains the structure and functioning of KOGs of functional systems and D. Hebb's cellular ensembles.
2. Based on the following information laws of the structure of the external world.
3. Provides maximum accuracy of predictions and anticipations of reality.
4. Resolves emerging contradictions.
5. It is an information theory of reflection of reality by the brain.

2. Causality

Let's consider the first law of the information structure of the external world – its causality. Causality is a consequence of physical determinism: "for any isolated physical system, some of its state determines all subsequent states" [12]. But consider a car accident [12]. What is the reason for it? This may be the condition of the road surface, its humidity, the position of the sun relative to the driver, reckless driving, the psychological state of the driver, brake malfunction, etc. Obviously, there is no definite reason in this case.

In the philosophy of science, causality is reduced to prediction and explanation. "Causality means predictability ... if the entire previous situation is known, the event can be predicted ... if all the facts and laws of nature related to this event are given" [12]. It is clear that to know all the facts, the number of which, as in the case of an accident, is potentially infinite and all the laws are impossible. In addition, humans and animals learn the laws of the outside world through training. Therefore, causality is reduced to prediction by inductive statistical inference, when the prediction is derived from facts and statistical laws with some probability.

In addition, causal relationships in the form of statistical laws found on real data or as a result of training face to the problem of statistical ambiguity – contradictory predictions can be derived from them [18]. To avoid this ambiguity, Hempel introduced the requirement of maximum specificity [18], informally consisting in the fact that statistical laws should include the maximum available information.

We solved the problem of statistical ambiguity and determined the Maximum Specific Causal Relationships (MSCR), for which it was proved that inductive statistical inference using them does not lead to contradictions [28,32] and thereby most accurately implement the principle of evolution of the living world by P.K.Anokhin. We have developed a special semantic probabilistic inference [28,32] that detects MSCR. In particular, it satisfies Cartwright's definition of a probabilistic causal relationship with respect to some background, which consists in the fact that each condition of the premise of causal relationship strictly increases the conditional

probability of the conclusion. In addition, we have developed a formal neuron model [30] satisfying the Hebb's rule, which implements this inference and detects MSCR causal relationships [28,32].

Thus, accurately analyzing the concept of causality, we get the informational law of reality that can be represented in ITS by detecting causal connections at the level of a neuron, ensuring its plasticity and manifesting itself, in particular, by conditional reactions. At the same time, they can be as accurate as possible through the use of MSCR conditional links.

This provides the *first level of the maximum accuracy of predictions* and anticipations of reality, implementing the principle of anticipatory reflection of reality.

3. "Natural" classification

Let's move on to the next law of the informational structure of the external world objects – the "natural" classification. The first rather detailed analysis of the "natural" classification belongs to J.S. Mill [21]. First, we will separate the "artificial" classifications from the "natural" ones: "Let's take any attribute, and if some things have it and others do not, then we can base the division of all things into two classes on it.", "But if we turn to ... the class of "animal" or "plant", ... then we will find that in this respect some classes are very different from others. ... have so many features that they cannot be ... enumerated" [21].

J.S. Mill defines the "natural" classification as follows: "Most of all, it corresponds to the goals of scientific (natural) classification when objects are combined into such groups regarding which the greatest number of general proposals can be made" [21]. Based on the concept of "natural" classification, J. S. Mill defines the concept of an "image" of a class as a certain pattern that has all the characteristics of this class.

Naturalists wrote that the creation of a "natural" classification consists in "indication" – from an infinitely large number of features it is necessary to move to a limited number of them, which would replace all other features [13]. This means that in "natural" classes, the attributes are strongly correlated, for example, if there are 128 classes and the attributes are binary, then only 7 attributes can be independent "indicator" attributes among them, since $2^7 = 128$, and other attributes can be predicted by the values of these 7 attributes. We can choose various 7-10 attributes as "indicator" and then other attributes, of which there are potentially infinitely many, can be predicted from these selected attributes. Therefore, there is an exponential (relative to the number of attributes) number of causal relationships linking the attributes of objects of "natural" classes.

Such redundancy of information, but already when perceiving objects of the external world, is confirmed in cognitive sciences when considering "natural"

concepts.

4. "Natural" concepts in cognitive sciences

The highly correlated structure of the objects of the external world is also revealed by the theory of "natural" concepts. "Natural" classification reveals the structure of the objects of the external world, and "natural" concepts, studied in cognitive sciences, determine the perception of these "natural" objects as elements of subjective reality.

In the works of Eleanor Rosch, the following principle of categorization of "natural" categories was formulated: «Perceived World Structure ... 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» [25].

Directly perceived objects (basic objects) are information-rich bundles of observable properties that create categorization (an image in the J.S. Mill definition): «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» [24].

In further research, it was found that models based on features, similarities and prototypes are not enough to describe "natural" classes. Considering these studies, Bob Rehder put forward a theory of causal models, according to which: "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" [23]. In the theory of causal models, the relation of an object to a category is no longer based on a set of signs and proximity by signs, but on the basis of the similarity of the generative causal mechanism.

Bob Rehder used Bayesian networks to represent causal knowledge [22]. However, they do not support cycles and therefore cannot model cyclic causal relationships. The formalization we propose further in the form of probabilistic formal concepts directly models cyclic causal relationships [5-8,28-29,32].

5. The integrated information theory by G.Tononi

The theory of integrated information by G.Tononi is also based on the highly correlated structure of the external world [20,26-27]. If the "natural" classification describes objects of the external world, and cognitive sciences describe the perception of objects of the external world, then the theory of integrated information

analyzes the information processes of the brain on the perception of objects of the external world.

Integrated information is considered by G.Tononi as a property of a system of cyclic causal relationships: «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» [27].

The relationship of integrated information with reality G.Tononi describes as follows: «Cause-effect matching ... measures how well the integrated conceptual structure ... fits or ‘matches’ the cause-effect structure of its environment», «... 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» [26-27].

G.Tone defines consciousness as a primary concept that has the following phenomenological properties: composition, information, integration, exclusion [20,26-27]. We present the formulations of these properties together with our interpretation of these properties (given in parentheses) from the point of view of "natural" classification of the external world objects.

1. composition – elementary mechanisms (causal interactions) can be combined into higher-order ones ("natural" classes form a hierarchy);

2. information – only mechanisms that specify 'differences that make a difference' within a system count (only the system of "resonating" causal relationships forming the class is significant);

3. integration – only information irreducible to non-independent components counts (only the system of "resonating" causal relationships is significant, which not reducible to the information of individual components, indicating an excess of information and the perception of a highly correlated structure of a "natural" object);

4. exclusion – only maxima of integrated information count (only values of features that are maximally interconnected by causal relationships form an "image" or "prototype").

Unlike G.Tononi, we consider these properties not as internal properties of the system, but as the ability of the system to reflect the "natural" classification of the objects of the external world. Then consciousness, unlike G.Tononi, is defined not by the phenomenological properties of neural structures, but as the ability of the brain, using the integrated information of neural structures, to reflect the world represented by a hierarchical "natural" classification and the system of "natural concepts" and their causal models.

6. Formalization of "natural" classification, "natural" concepts and consciousness as integrated information by G.Tononi

In accordance with the Principle of unlimited inference, the brain carries out all possible conclusions on causal relationships. These causal relationships, of which there is an exponential number, in the process of perceiving "natural" objects, loop on themselves, forming a certain "resonance", which is a system with highly integrated information in the sense of G.Tononi. At the same time, "resonance" occurs if and only if these causal relationships reflect some "natural" object in which a potentially infinite set of features mutually assume each other. The resulting cycles of conclusions on causal relationships are mathematically described by "fixed points", which are characterized by the fact that further application of conclusions to the properties under consideration does not predict new properties. The set of mutually related properties obtained at a fixed point gives the "image" of the class or "prototype" of the concept and its "causal model". Therefore, the brain perceives a "natural" object not as a set of features, but as a "resonating" system of causal connections that close on themselves through the simultaneous inference of the entire set of features of the "image" or "prototype" forming a "causal model".

It can be shown that the MSCR causal relationships organized into cellular ensembles make it possible to identify objects of the external world as reliably as possible and then predict the properties of these objects as accurately as possible using this identification, since only MSPS causal relationships related to this class are used for predictions. This forms a *second*, even more accurate, from the point of view of forecasting, *level of organization of information processes*.

We propose a fundamentally new mathematical apparatus for determining integrated information, "natural" classification and "natural" concepts. Our formalization is based on a probabilistic generalization of the formal concepts analysis [8, 29-32]. Formal concepts can be defined as fixed points of deterministic rules (with no exceptions) [19]. But, as J. Mill wrote: "Natural groups ... are determined by features, ... while taking into account not only the features that are certainly common to all the objects included in the group, but the whole set of those features, of which all occur in most of these objects, and the majority in all." Therefore, it is necessary to get away from deterministic rules and replace them with probabilistic ones in order to determine the features not exactly, but for the majority. Therefore, we generalize formal concepts to the probabilistic case, replacing deterministic rules with MSPS causal relationships and defining probabilistic formal concepts as fixed points of these maximally specific rules [8, 29-32]. Due to the fact that the conclusions, based on the most specific causal relationships are consistent, the resulting fixed point will also be consistent and will not contain both a feature and its negation, i.e. such a definition of probabilistic formal concepts is correct.

It can be shown [9] that probabilistic formal concepts adequately formalize

"natural" classification and, in moreover, the resulting "natural" classification satisfies all the requirements that naturalists imposed on "natural" classifications [9].

Let's consider an example of computer simulation of the "natural" classes, "natural" concepts and integrated information discovery for the encoded digits. Let $X(a)$ – be the set of properties of object a given by some set of predicates, and let $(P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}) \in MS(X)$ – be the set of MSPS of causal relationships performed for properties X , $\{P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}\} \subset X$ then the prediction operator Pr and the fixed point can be written as follows [6, 9]:

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

where $\Phi_{Krit}(X)$ – is an operator that modifies the set of features X by adding or removing some feature, so that a certain criterion $Krit$ of mutual consistency of causal relationships by mutual prediction of features from X is maximal [6,9]. The *Krit* criterion measures the informational integration of features according to the system of causal relationships differently than it is done in the G.Tononi's theory. A fixed point is reached when $Pr^{n+1}(X(a)) = Pr^n(X(a))$, for some n , where Pr^n – n is a multiple application of the operator Pr . Since with each application of the operator Pr , the value of the *Krit* criterion increases and reaches a local maximum at a fixed point, then a fixed point, which reflect some "natural" object, has a maximum of integrated information and the "exclusion" property according to G.Tononi.

Let us encode the digits as shown in fig. 2. and form a training set, consisting of 360 shuffled digits (12 digits of fig. 2 duplicated in 30 copies without specifying where which digit is). On this set, a semantic probabilistic inference revealed 55089 MSCR causal relationships – general statements about objects that J.S. Mill spoke about. According to these causal relationships, exactly 12 fixed points were found that correspond to numbers.

An example of a fixed point for the digit 6 is shown in fig. 3. Consider what this fixed point is. Let's number the signs of the digits as shown in Fig. 2. The first pattern of figure 6 in fig. 3, represented in the first rectangle after the curly bracket, means that if there is a sign 6 in square 13 (let's denote it as 13-6), then there should be a sign 2 in square 3 (let's denote it as (3-2)). The predicted sign is indicated by a dotted line. Let's write this causal relationship as $(13-6 \Rightarrow 3-2)$. It is not difficult to verify that this causal relationship is carried out on all figures. The second causal relationship means that from the sign (9-5) and the negation of the value 5 of the first sign $\neg(1-5)$ (the first sign should not be equal to 5), the sign (4-7) follows. Negation is indicated in the figure by a dotted line, as shown at the bottom of fig. 3. We get a causal relationship $(9-5 \& \neg(1-5) \Rightarrow 4-7)$. The next 3 causal relationships in the first row of the digits 6 will be respectively $(13-6 \Rightarrow 4-7)$, $(17-5 \& \neg(13-5) \Rightarrow 4-7)$, $(13-6 \Rightarrow 16-7)$.

Fig. 3 shows that the causal relationships and the signs of the number 6 form a fixed point – mutually predict each other. Note that the causal connections used in the

fixed point are fulfilled on all digits, and the fixed point itself identifies only the digit 6. This illustrates the phenomenological property 2 'differences that make a difference', in which the system of causal connections perceives "realizes" an integral object. Therefore, the figures are distinguished not by causal relationships in themselves, but by their systemic relationship.

A fixed point forms a "prototype" according to Eleanor Rosch or an "image" according to J. S. Mill. The program does not know in advance which combinations of features are maximally correlated with each other.

Probabilistic formal concepts describe not only "natural" concepts, but also contexts. Contexts also have the property of maximum prediction accuracy – causal relationships found on a certain context and its causal model will more accurately predict the properties of this context.

Consider the following example on fig. 4, containing both numbers and letters. You can learn only on numbers and build probabilistic formal concepts of numbers, you can learn on letters and build probabilistic concepts of letters only, and you can learn both on numbers and letters and build formal concepts of numbers and letters. In each of these cases, various MSCR causal relationships will be found, but MSCR causal relationships describing numbers and letters together will contain additional signs separating them from each other, which is obtained automatically by MSCR causal relationships. When considering (in context) both letters and numbers of MSCR causal relationships will have a higher probability, then MSCR on numbers or letters and therefore they will be triggered in the formal model of the neuron. Our formal neuron model, which detects the most specific causal connections [30], follows the well-known physiological property of neurons - more probable conditional stimuli are triggered faster in time.

7. Theory of functional systems

The formalization of the second type KOGs – the KOGs of functional systems, is based on the consideration of purposeful behavior, which is carried out by developing conditional (causal) links between the actions and its results. These conditional connections are sufficient for modeling functional systems and developing animats.

P.K. Anokhin wrote that "We are talking about the collateral branches of the pyramidal tract, diverting to many neurons "copies" of impulsations that go to the pyramidal tract" [4-5]. Thus, when a motor neuron sends a signal to the muscles about some action, copies of this excitation are sent, including to the projection zones, which can record the result of the action performed. Therefore, the brain detects all causal connections between actions and their results.

We show in the diagram fig. 5 that this is sufficient to explain the basic mechanisms of the functional systems of the brain formation [7, 31]. Let's assume that we have no experience yet and a motivational excitement has arisen, shown by

the black triangle. Then, to meet the need by trial and error, we can do some action that will be activated by some neuron, indicated by a white triangle. Simultaneously with the activation of this action, a "copy" of the excitation of this neuron will be sent to the projection zones, where there will be a neuron that will react to the result of the action received from the outside world. Since this neuron will first receive excitation from the activation of an action by a white neuron, it will form a conditional relation between the activation of an action by a white neuron and the result obtained. If now, after receiving this result that has changed the situation, we carry out some next action, also indicated by a white triangle, then we will get the following result, for which there will also be a neuron that will react to the result of this action. If, as a result, the need was satisfied and the goal is achieved, then the entire chain of active neurons and conditional connections that led to the result will be reinforced and stored in the memory. Thus, there will be an internal contour of forecasting the results achievement by the causal relationships. Then, at the next occurrence of motivational excitement, this chain of actions will be extracted from memory and will predict the achievement of the result along the inner contour even before any actions. So an action plan will be formed, which, according to the inner contour, as stated in the quote by P.K. Anokhin, activates neurons waiting for the results of actions, which will form an *acceptor of the results of actions*, studied in detail in the theory of functional systems. Thus, the formation and operation of the functional system can be explained by the formation of causal relationships between the action and its results.

In terms of MSCR causal relationships, the scheme of functional systems is as follows fig. 6. [7, 31]. We consider the need as a request to the functional system to achieve the goal indicated by the predicate PG_0 . This request enters the afferent synthesis block and, for functional systems that do not have functional subsystems, extracts causal relationships of the form $P_{i1}, \dots, P_{im}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$ from memory, leading to the goal PG_0 achievement, where $P_{i1}, \dots, P_{im}$ are the properties of the environment necessary to achieve the goal achievement, and $A_{k1}, \dots, A_{kl}$ – a sequence of actions leading to the goal. At the same time, the properties of $P_{i1}, \dots, P_{im}$ must be present in the properties of the environment $P_1, \dots, P_n$ entering the afferent synthesis block. For hierarchically organized functional systems, this query extracts causal relationships of a more complex type $P_{i1}, \dots, P_{im}, PG_{j1}, \dots, PG_{jn}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$ from the MSCR memory, including requests to achieve the sub-goals $PG_{j1}, \dots, PG_{jn}$.

Then the extracted rules are sent to the decision-making block, where forecast of the goal achievement is made for each rule and a probability estimation of the goal achievement is calculated. Prediction according to the rules, where only actions are performed, is carried out according to the probability of the rule itself. The forecast according to the rules which requests sub-goals is carried out by sending these requests to functional subsystems, decision-making in them and receiving from them the probability estimations of the corresponding sub-goals achievement. The resulting probability of the forecast is calculated by the product of the probability of the rule on

the probability of achieving of its sub-goals. After that, a decision is made by the rule selecting that has the maximum probability estimation of the goal achievement.

Then an action plan is formed, including all the actions included in the rule and all the actions that are in the functional subsystems. Simultaneously with the action plan, the acceptor of the actions result is formed, including the expectation of all predicted sub-results in functional subsystems, as well as in the functional system itself. After that, the action plan begins to be implemented, and the expected results are compared with the results obtained.

If all the sub-results and the final result are achieved and coincide with the expected results, then the rule itself and all the rules of the functional subsystems that were selected in the decision-making process are reinforced and their probability increases. If the result is not achieved in some subsystem, then the corresponding rule, selected by decision-making block, of this functional subsystem is penalized. Then there is a tentative research reaction that revises the decision. This model has been successfully used to model animates [11, 16, 31].

8. Consciousness as the tool for resolving contradictions

Let us consider the *third level of prediction accuracy* provided by the information theory of consciousness – consciousness as a mechanism for resolving contradictions.

The world is multifaceted like a diamond fig. 7 and there is no single consistent description of it, and the function of consciousness is to choose the appropriate context correctly, within which you can get the most accurate prediction. In science, such contexts are paradigms that form a certain view, point of view and the corresponding system of concepts of a particular theory. These paradigms, as a rule, are not compatible with each other.

This point of view on consciousness is also expressed by V.M. Allakhverdov [2]. In his work [1] he writes: "Consciousness, faced with contradictory information, tries to remove this information from the surface of consciousness or modify it so that the contradiction disappears or ceases to be perceived as a contradiction." In this work, he cites 7 cases of resolving contradictions by consciousness. All these cases are explained by the properties or interaction of probabilistic formal concepts that define the concepts or contexts in question. Let's consider two of them for brevity:

1. Case 1. The easiest way to get rid of contradiction or ambiguity is to choose one interpretation for awareness, and not to realize all the others (incompatible with it) (negatively choose).

Example. The phenomenon of binocular competition, when different stimuli are simultaneously presented to the different eyes to the subject. If two images are presented, one of which is more likely or familiar, the subjects mostly see only it.

Explanation. A probabilistic formal concept mutually predicts the properties included in the concept, as well as the negation of other properties that should not be in it, and thereby inhibits alternatives.

2. Case 2. When realizing the different sides of the contradiction, an attempt is made to find a way of explanation — the connection of different sides into a consistent whole.

Example. In the conditions of binocular competition, if you present a red circle on one eye and a black triangle on the other, the subject will see a black triangle on a red background.

Explanation. If the perceived features do not contradict each other and do not inhibit each other, then they can form a combined probabilistic formal concept and be perceived accordingly.

In all cases, it is possible to correctly choose the most appropriate probabilistic formal concept or context to resolve the contradiction and obtain the most accurate prediction in accordance with the causal relationships of the chosen concept or context.

8. 介紹

非還原論的意識理論不能還原為任何現實理論以及任何生理或心理學理論。

在第七屆國際認知科學會議上，K.V. Anokhin 在報告中表示，「問題不在於現有的神經生理學理論不完善…其中使用的相關方法根本無法回答有關心靈本質和主觀體驗的問題。．．因為這需要一個非還原性的基本理論」[3]。

Max Tegmark 在他的書[14]中也寫道，在外部現實“外部現實”和“內部現實”之間必須有一個中間現實“共識現實”，它以物理術語描述觀察到的現實，並反映在內在現實中。

繼 D. I. 杜布羅夫斯基[10]之後，我們將使用「資訊方法」將「內在現實」描述為主觀現實。同時，主觀現實是個體意識狀態的現實——感覺、知覺、感覺、思想、意圖、慾望等。同時，主觀現實的現像被認為是與相應的大腦過程相關的訊息作為其載體。

因此，在「意識與大腦」問題的資訊方法框架內，我們得到如下圖所示：

1. 其中，現實是透過有關觀察到的現象的資訊來描述的，這些現像是由主觀現實透過感覺、知覺、感覺等感知的，而這些訊息又是有關相應大腦過程的訊息。

這些資訊的目的是什麼？ P.K. Anokhin 提出的「生命世界進化原理」對此有著最精確的定義：「在生物體對外條件的適應過程中出現了一種普遍模式並隨後在生命世界的整個進化過程中迅速發展：外部世界緩慢展開的事件的極

其快速的反映” [5]。

讓我們揭示這一原理並制定建構意識資訊理論的原則。

一、首先，遵循“生命世界進化原理”，資訊理論必須預測外在世界的事件。

然而，如果外在世界是隨機的，那麼就不可能預測它。但我們的世界結構良好。如果現實的資訊結構存在一些規律，那麼很自然地可以假設，在演化過程中，這種神經生物學機制被發展出來，可以利用這些結構來提供對現實最準確的反映。因此，以下原則是必要的。

二、反映資訊理論應以外在世界的結構法則為基礎，同時描述現實的資訊結構以及保證這一結構在主觀現實方面的反映的神經生理學等機制。

下面我們將介紹現實訊息結構的這樣的規律。K.V. 的報告「Cognitome – 大腦的超網路模型」中指出了與它們相對應的神經生物學機制以及使用這些結構。阿諾金。這些是 COG（神經元認知組），概括了 D. Hebb 的功能係統和細胞整體理論 [17]。

這項工作提出了意識資訊理論（ITC）的以下基本原理，足以解釋所提到的主要資訊過程：

無限制推理原理：大腦偵測外在世界所有可能的因果關係，並從中得出所有可能的結論。

事實證明，這項原則足以建構意識資訊理論，該理論：

1. 解釋 D. Hebb 的功能係統和細胞群的結構和功能。
2. 基於下面給出的外在世界結構的資訊規律。
3. 提供最準確的預測和現實。
4. 化解新出現的矛盾。
5. 是關於大腦如何反映現實的資訊理論。

9. 因果關係

讓我們考慮外在世界資訊結構的第一定律——因果關係。因果關係是物理決定論的結果：「對於任何孤立的物理系統，其某些狀態決定所有後續狀態」[12]。但考慮一下一場車禍[12]。其原因何在？這可能是路面狀況、濕度、太陽相對於駕駛者的位置、魯莽駕駛、駕駛者的心理狀態、煞車故障等顯然，在這種情況下沒有具體的原因。

在科學哲學中，因果關係被簡化為預測和解釋。「因果意味著可預測性…如果已知整個先前的情況，則可以預測事件…如果給出與該事件相關的所有事實和自然法則」[12]。顯然，不可能知道所有的事實，就像事故的情況一樣，事實的數量可能是無限的，並且了解所有的法律。此外，人類和動物透過訓練來學習外在世界的規律。因此，因果關係歸結為使用歸納統計推理的

預測，即預測是從具有一定機率的事實和統計規律得出的。

此外，在實際數據中發現的統計規律形式的因果關係或作為訓練結果的因果關係面臨著統計模糊性的問題——可能從中得出不一致的預測[18]。為了避免這種歧義，Hempel 引入了最大特異性的要求 [18]，非正式地包括這樣的事實：統計法應包括最大程度的可用資訊。

我們解決了統計模糊性問題，並確定了最大特定因果關係（MSPL），已證明使用它們進行歸納統計推理不會導致矛盾[28, 32]，從而最準確地實現了該原理生命世界的進化 P.K. Anokhin。我們開發了一種特殊的語意機率推理 [28, 32] 來檢測 MSPS。特別是，它滿足卡特賴特相對於某些背景的機率因果關係的定義[15]，即因果關係的每個前提條件嚴格地增加了結論的條件機率。我們還開發了一個滿足 Hebb 規則的正式神經元模型 [30]，該模型執行這一結論並檢測 MCPS 因果關係 [28,32]。

因此，透過準確分析因果關係的概念，我們發現這種現實的資訊規律可以透過檢測神經元層面的因果關係在 ITS 中表示，確保其可塑性並在特別是條件反應中表現出來。同時，透過使用 IMS 條件鏈接，它們可以盡可能準確。

這提供了對現實的預測和預期的第一級最大準確性，實現了對現實的高級反映的原則。

10. “自然”分類

讓我們繼續討論外在世界物件資訊結構的下一個法則——「自然」分類。第一個對「自然」分類進行足夠詳細的分析屬於 J. St. 磨坊[21]。首先，讓我們把「人為」的分類與「自然」的分類分開：「讓我們取任何一個符號，如果某些事物擁有它，而另一些事物則沒有，那麼就可以根據它來將所有事物分為兩類。」「但是，如果我們轉向…『動物』或『植物』類別，…我們會發現，在這方面，某些類別與其他類別有很大不同。……具有如此多的特徵，以至於它們無法... 被列出」[21]。

J. St. 的「自然」分類 穆勒將其定義如下：「當物件被組合成可以提出最多數量的一般命題的群體時，它最符合科學（自然）分類的目的」[21]。基於 J. St. 的「自然」分類概念 穆勒將類別「圖像」的概念定義為具有給定類別的所有特徵的特定樣本。

自然主義者寫道，「自然」分類的創建包括「指示」——必須從無限的特徵轉移到有限數量的特徵，這將取代所有其他特徵[13]。這意味著在「自然」類中，屬性是高度相關的，例如，如果有 128 個類，並且屬性是二元的，那麼其中只有 7 個屬性可以是獨立的「指標」屬性，因為 $2^7 = 128$ 可以根據這 7 個屬性的值來預測其他屬性。我們可以選擇 7-10 個不同的屬性作為“指

標”，然後可以從這些選定的屬性中預測可能有無限多個的其他屬性。因此，連接「自然」類別物件屬性的因果關係數量呈指數級（相對於屬性數量）。認知科學在分析「自然」概念時證實了感知外在世界物件時的這種資訊冗餘。

11. 認知科學中的「自然」概念

「自然」概念理論也揭示了外在世界的高度相關結構。同時，「自然」分類揭示了外在世界物體的結構，認知科學中研究的「自然」概念決定了對這些「自然」物體作為主觀現實要素的感知。

在 Eleanor Rosch 的著作中，制定了以下「自然」類別的分類原則：「感知世界結構…不是等機率共現屬性的非結構化總集。相反，世界上的物質對象被認為擁有……高度相關的結構……我們所認為的真實對象的屬性的組合並不均勻地出現。有些對、三元組等很可能與一個屬性、有時是另一個屬性組合出現；其他的則很少見；其他的在邏輯上不能或在經驗上不會發生」[25]。直接感知的物件（基本物件）是可觀察屬性的資訊豐富的集合，可以創建分類（J. St. Mill 定義中的圖像）：「如果感知者將重點放在物件上，則可以根據其清晰的情況來查看類別」感知屬性的相關結構…透過類別原型，我們通常指的是類別成員資格的最清晰案例」[24]。

在進一步的研究中，人們發現基於特徵、相似性和原型的模型不足以描述「自然」類別。考慮到這些研究，Bob Rehder 提出了因果模型理論（因果模型理論），根據該理論：「人們關於物件類別的直觀理論由類別模型組成，在該模型中，類別的特徵和因果關係都包含在類別模型中。」這些特徵之間的機制得到了明確的體現」[23]。在因果模型理論中，物件與類別的關係不再基於一組特徵和特徵的相似性，而是基於生成因果機制的相似性。

為了表示因果知識，Bob Rehder 使用貝葉斯網路 [22]。但是，它們不支援循環，因此無法對循環因果關係進行建模。我們進一步以機率形式概念的形式提出的形式化直接模擬循環因果關係[5-8, 28-29, 32]。

12. 綜合資訊理論 G. Tononi

G. Tononi 的整合資訊理論也是基於外在世界的高度相關結構 [20, 26-27]。如果「自然」分類描述了外在世界的物體，認知科學描述了對外在世界物體的感知，那麼整合資訊理論則分析了大腦感知外在世界物體時的資訊過程。

G. Tononi 中的綜合資訊被認為是循環因果關係系統的屬性：「事實上，環境的『快照』傳達的訊息很少，除非它在一個系統的背景下來進行解釋，該系統的因果結構很複雜，長期存在。歷史，已經捕捉了世界的一些因果結構，即空間和時間的長程相關性」[27]。

G. Tononi 將綜合資訊與現實的關係描述如下：“因果匹配……衡量綜合概念結構……與其環境的因果結構的契合或‘匹配’程度”，“……當系統適應具有豐富、綜合因果結構的環境時，匹配度應該會增加。此外，匹配的增加往往與資訊整合的增加有關，從而與意識的增強相關” [26-27]。

G.托諾尼將意識定義為具有以下現象學屬性的主要概念：組成、資訊、整合、排除 [20, 26-27]。讓我們從外部世界物件的「自然」分類的角度來展示這些屬性的表述以及我們對這些屬性的解釋（在括號中給出）。

1. 組合—基本機制（因果交互作用）可以組合成更高階的機制（物件的「自然」類別形成層次結構）；
2. 資訊—只有在系統中指定「產生差異的差異」的機制才算數（只有形成類別的「共振」因果關係系統才是重要的）；
3. 整合—只有不可簡化為非相互依賴的組成部分的信息才算數（只有“共振”因果聯繫的系統才是重要的，不能簡化為單個組成部分的信息，表明信息過多以及對高度相關的結構的感知）“自然”物體）；
4. 排除—僅綜合資訊計數的最大值（僅透過因果關係最大程度地互連的特徵的值形成「影像」或「原型」）。

與 G. Tononi 不同，我們認為這些屬性不是系統的內部屬性，而是系統反映外部世界物件「自然」分類的能力。那麼，與托諾尼不同，意識不是由神經結構的現象學特性決定的，而是由大腦利用神經結構的綜合資訊來反映由系統中的分層「自然」分類所代表的世界的能力決定的。“自然概念”及其因果模型。

13. 根據 G. Tononi, 「自然」分類、「自然」概念和意識作為綜合資訊的形式化

根據無限推理原理，大腦會根據因果關係做出所有可能的推理。這些呈指數級數量的因果聯繫，在感知“自然”物體的過程中不斷循環，形成一定的“共振”，這是托諾尼意義上的信息高度集成的系統。同時，當且僅當這些因果關係反映了某種完整的「自然」物件時，「共振」才會發生，其中潛在的無限數量的特徵相互預設。基於因果關係得出的結論循環在數學上用「不動點」來描述，其特徵在於，將結論進一步應用於所考慮的屬性並不能預測新的屬性。在固定點獲得的一組相互可預測的屬性給出了類別的「圖像」、概念的「原型」及其「因果模型」。因此，大腦將「自然」物體感知為不是一組特徵，而是一個因果連結的「共振」系統，透過「影像」或「原型」的整套特徵的同時輸出來閉合自身形成「因果模型」。

可以證明，組織成細胞係綜的 MSPS 因果關係使得能夠最可靠地識別外部

世界中的對象，然後使用這種識別盡可能準確地預測這些對象的屬性。這樣的預測更準確，因為只有屬於此類 MSPS 的因果關係才會用於預測。從預測的角度來看，這形成了第二個、甚至更準確的資訊過程組織層級。

我們提出了一種全新的數學工具來定義綜合資訊、「自然」分類和「自然」概念。我們的形式化是基於形式概念分析的機率概括[8, 29-32]。形式概念可以定義為確定性規則的固定點（沒有例外）[19]。但是，正如 J. Mill 所寫，「自然群體……是由特徵決定的，……在這種情況下，不僅要考慮該群體中所有對象所共有的特徵，還要考慮整個集合這些特徵，所有這些特徵都存在於這些物體的大多數中，並且大多數- 在所有物體中」。因此，有必要放棄確定性規則，代之以機率性規則，以便不精確地確定特徵，而是為大多數人確定特徵。因此，我們將形式概念推廣到機率情況，以 MSPS 因果關係取代確定性規則，並將機率形式概念定義為這些最大具體規則的不動點[8,29-32]。由於基於最具體的因果關係得出的結論是一致的，因此得到的不動點也將是一致的，並且不會同時包含符號及其否定，即 機率形式概念的這個定義是正確的。

可以證明[9]，機率形式概念充分形式化了「自然」分類，此外，所得的「自然」分類滿足自然科學家對「自然」分類的所有要求[9]。

考慮一個電腦模擬檢測「自然」類別、「自然」概念和編碼數字的整合資訊的範例。讓 $X(a)$ - 物件屬性集 a ，由一組謂詞給出，並且 $(P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}) \in MS(X)$ - 為屬性建立的一組 MSPS 因果連接 X ， $\{P_{i_1} \wedge \dots \wedge P_{i_k} \Rightarrow P_{i_0}\} \subset X$ 那麼預測算子 Pr 和不動點可以寫成如下[6, 9]

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

在哪裡 $\Phi_{Krit}(X)$ - 透過加入或刪除某個特徵來修改一組特徵 X 的算子，使得從 X 中相互預測特徵的因果關係相互一致性的某個標準 $Krit$ 最大[6, 9]。 $Krit$ 標準根據因果關係系統來衡量特徵的資訊整合度 $MS(X)$ 與 G. Tononi 的理論不同。當達到固定點時 $Pr^{n+1}(X(a)) = Pr^n(X(a))$ ，對於某些 n ，其中 Pr^n - n Pr 運算子的多次使用。由於隨著算子 Pr 的每次應用， $Krit$ 準則的值都會增加並在固定點達到局部最大值，因此反映某個「自然」物件的固定點根據 G 具有最大的綜合資訊和「排除」屬性. 托諾尼。

讓我們對數字進行編碼，如圖所示。2. 讓我們建立一個由 360 個打亂的數字組成的訓練集（圖 2 中的 12 個數字，重複 30 個副本，但不標明哪個數字是哪個數字）。在這個集合上，語意機率推理揭示了 55089 MSPS 的因果關係——關於 J. St. 談到的對象的一般陳述。磨。根據這些模式，準確地發現了 12 個與數字完全對應的固定點。

數字 6 的固定點示例如圖 1 所示。3. 讓我們來考慮一下這個不動點是什麼。讓我們對數字的符號進行編號，如圖所示。2、圖 6 的第一個圖案 3，

出現在大括號後的第一個矩形中，表示如果方格 13 包含屬性 6（表示為 13-6），則方格 3 必須包含屬性 2（表示為（3-2））。預測的特徵以虛線表示。我們把這個因果關係寫成 $(13-6 \rightarrow 3-2)$ 。很容易驗證這種因果關係適用於所有數字。第二個因果關係意味著從符號（9-5）和第一個符號 $\neg(1-5)$ 的值 5 的否定（第一個符號不應等於 5）得出符號（4-7）。否定在圖中以虛線表示，如圖底部所示。3. 我們得到因果關係 $(9-5 \& \neg(1-5) \Rightarrow 4-7)$ 。數字 6 第一行接下來的 3 個因果關係將是 $(13-6 \Rightarrow 4-7)$ ， $(17-5 \& \neg(13-5) \Rightarrow 4-7)$ ， $(13-6 \Rightarrow 16-7)$ 、分別。

在圖中。圖 3 顯示數字 6 的因果關係和符號形成一個固定點——它們互相預測。請注意，固定點中使用的因果連接是在所有數字上進行的，而固定點本身僅突出顯示數字 6。這說明了現象學屬性 2 “差異產生差異”，其中因果連接系統感知“實現”整個對象。因此，這些數字的差異不是在於其本身的因果關係，而是在於它們的系統互連。

根據 Eleanor Rosch 的說法，固定點形成“原型”，或根據 J. St. 米盧。程式事先並不知道哪些特徵組合彼此最相關。

機率形式概念不僅描述「自然」概念，也描述上下文。上下文還具有最大預測準確性的屬性——在某個上下文中發現的因果關係並構成其因果模型將更準確地預測該上下文的屬性。

考慮下面的範例圖。4. 同時包含數字和字母。

您可以僅從數字學習並建立數字的機率形式概念，您可以從字母學習並僅建立字母的機率概念，或者您可以從數字和字母學習並建立數字和字母的形式概念。在每種情況下，都會發現不同的 MSPS 因果關係，但是一起描述數字和字母的 MSPS 因果關係將包含將它們彼此分開的附加特徵，這是由 ASMS 因果關係自動獲得的。當（在上下文中）同時考慮 ASHI 的字母和數字時，因果關係將比單獨考慮數字和字母的 ASHI 具有更高的機率，因此它們將在這種上下文中被觸發。我們的正式神經元模型可以檢測最大程度的特定因果關係[30]，遵循神經元的已知生理特性——更可能的條件刺激會更快地及時觸發。

14. 功能係統理論

第二種類型的齒輪——功能係統的齒輪——的形式化是基於對目標導向行為的考慮，這是透過在行動與其結果之間建立條件（因果）聯繫來實現的。這些條件連接足以對功能係統進行建模和開發動畫。

甚至 P.K Anokhin 也寫道，「我們正在談論錐體束的側支分支，導致許多神經元「複製」那些傳出訊號到錐體束」[4-5]。因此，當運動神經元向肌肉發送有關某些動作的信號時，該興奮的副本就會被發送，包括發送到投影區域。投影區域可以記錄所執行動作的結果。因此，大腦會檢測到行為與其結果之

間的所有因果關係。

讓我們用圖 5 的圖表表明，這足以解釋大腦功能系統形成的基本機制 [7,31]。假設我們還沒有經驗，動機喚醒已經出現，如黑色三角形所示。然後，為了滿足需求，透過反覆試驗，我們可以執行一些操作，該操作將由某個神經元激活，如白色三角形所示。在激活該動作的同時，該神經元的興奮的「副本」將被發送到投影區域，其中將有一個神經元將對來自外界的動作結果做出回應。由於同一個神經元首先會從白色神經元的動作活化中接收到激勵，因此它將在白色神經元的動作活化與所獲得的結果之間形成條件連接。如果現在，在收到改變情況的結果後，我們執行一些下一個動作，也由白色三角形表示，那麼我們將得到下一個結果，對於該結果，還將有一個神經元對結果做出反應這個動作。如果結果滿足了需求並實現了目標，那麼導致結果的整個活躍神經元鍊和條件連接將得到加強並儲存在記憶中。因此，將出現根據因果關係實現結果的預測的內部輪廓。然後，下一次發生動機喚醒時，將從記憶中檢索這一系列行動，並在採取任何行動之前預測內部迴路中結果的實現。這樣，就會形成一個行動計劃，沿著內部迴路，正如 P.K. Anokhin 所引用的那樣，激活等待行動結果的神經元，這將形成行動結果的接受者，詳細研究在功能系統理論。因此，功能系統的形成和運作可以透過行為與其結果之間因果關係的形成來解釋。

就 MSPS 因果關係而言，功能系統的運作方案如下所示：圖 6。[7,31]。我們將把需求視為對功能系統的請求，以實現謂詞 PG_0 指示的目標。此請求進入傳入綜合模組，對於沒有功能子系統的功能系統，從記憶體中檢索 $P_{i1}, \dots, P_{im}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$ 形式的因果關係，從而實現結果目標 PG_0 ，其中 $P_{i1}, \dots, P_{im}$ 是實現目標所需的環境屬性， $A_{k1}, \dots, A_{kl}$ 是實現目標的動作序列。在這種情況下，屬性 $P_{i1}, \dots, P_{im}$ 必須存在於環境屬性 $P_1, \dots, P_n$ 中，進入傳入合成區塊。對於分層組織的功能系統，此請求從 MSPS 的記憶體中檢索更複雜類型 $P_{i1}, \dots, P_{im}, PG_{j1}, \dots, PG_{jn}, A_{k1}, \dots, A_{kl} \Rightarrow PG_0$ 的因果關係，包括要求實現子目標 $PG_{j1}, \dots, PG_{jn}$ 。

接下來，提取的規則進入決策區塊，其中針對每個規則對實現目標進行預測，並計算實現目標的機率估計。基於規則的預測，在只有動作的情況下，是根據規則本身的機率進行的。根據對子目標的請求的規則進行預測是透過將這些請求發送到功能子系統，在它們中做出決策並從它們接收預測實現相應子目標的機率來進行的。透過將規則的機率乘以實現子目標的機率來計算得到的預測機率。之後，決定如何實現目標，並選擇具有最大機率預測目標實現的規則。

然後形成一個行動計劃，其中包括規則中包含的所有行動以及功能子系統中的所有行動。在製定行動計畫的同時，形成了用於接受行動結果的裝置，其中包括對功能子系統以及功能系統本身中所有預測子結果的期望。此後，

行動計劃開始實施，並將預期結果與所獲得的結果進行比較。

如果所有子結果和最終結果都達到並與預期結果一致，則規則本身以及決策過程中選擇的功能子系統的所有規則都會得到強化，其機率也會增加。如果某個子系統沒有達到結果，則對該功能子系統的相應規則進行處罰。然後出現指示性探索性反應，從而修改所做的決定。此模型已成功用於模擬動畫[11, 16, 31]。

15. 八、意識作為解決矛盾的機制

現在讓我們考慮意識資訊理論所提供的預測準確性的第三個層次——意識作為解決矛盾的機制。

世界就像鑽石一樣是多面向的。7 且對此沒有單一一致的描述，意識的功能是正確選擇合適的上下文，從而獲得最準確的預測。在科學中，這樣的脈絡是形成特定理論的某種觀點、觀點和相應概念體系的典範。通常，這些範式彼此不相容。

V.M. Allakhverdov 也表達了這種關於意識的觀點[2]。在他的著作[1]中，他寫道：“意識在面對矛盾的信息時，試圖將這些信息從意識表面移除或修改它，以使矛盾消失或不再被認為是矛盾。”在這部作品中，他給了7個用意識解決矛盾的案例。所有這些情況都是透過定義相關概念或上下文的機率形式概念的操作或交互作用來解釋的。讓我們簡要地考慮其中兩個：

1. 案例 1. 擺脫矛盾或歧義的最簡單方法是選擇一種解釋進行意識，而不是意識到所有其他解釋（與其不相容）（消極選擇）。

例子。當受試者的不同眼睛同時看到不同的刺激時，出現雙眼競爭現象。

如果呈現兩張影像，其中一張更有可能或更熟悉，受試者主要只會看到這張。

解釋。機率形式概念相互預測概念中包含的屬性，以及否定不應該包含在其中的其他屬性，從而抑制替代方案。

2、案例 2。當認識到矛盾的不同面向時，試圖找到一種解釋方式—將不同面向連結成一個一致的整體。

例子。在雙眼競爭的條件下，如果向一隻眼睛呈現紅色圓圈，向另一隻眼睛呈現黑色三角形，則受試者將在紅色背景上看到一個黑色三角形。

解釋。如果感知到的符號不互相矛盾、不互相抑制，那麼它們就可以形成一個統一的機率形式概念並被相應地感知。

在所有情況下，人們都可以根據所選概念或上下文的因果關係，正確選擇最合適的機率形式概念或上下文來解決矛盾，並獲得最準確的預測。

1. V.M. Allakhverdov, O.V. Naumenko, M.G. Filippova, O.B. Shcherbakova, M.O. Avanesyan, E.Y. Voskresenskaya, A.S. Starodubtsev. How consciousness gets rid of contradictions // SHAGI/STEPS 2015(1) The journal of the School of advanced studies in Humanities (in Russian)
2. Allakhverdov V.M. Consciousness as a paradox. Exp. Psy. St. Pet.: DNA, 2000. (in Russian)
3. Anokhin K.V. Cognitom: in search of a general theory of cognitive science // The Sixth International Conference on Cognitive Science, Kaliningrad, 2014, pp. 26-28. (in Russian)
4. Anokhin P.K. The problem of decision-making in psychology and physiology // Problems of decision-making. Moscow: Nauka, 1976. pp. 7-16. (in Russian)
5. Anokhin P.K. Biology and neurophysiology of the conditioned reflex and its role in adaptive behavior. Oxford a.o.: Pergamon press, 1974. 574 p.
6. Vityaev E.E., Neupokoev N.V. Formal model of perception and image as a fixed point of anticipation. Neuroinformatics. 2012, volume 6, No. 1, pp. 28-41. (in Russian)
7. Vityaev E.E. Logic of brain work // Approaches to modeling thinking. ed. V.G. Redko. URSS Editorial, Moscow, 2014, pp. 120-153. (in Russian)
8. Vityaev E.E., Demin A.V., Ponomarev D.K. Probabilistic generalization of formal concepts // Programming. Vol.38, No.5, 2012, pp. 219-230. (in Russian)
9. Vityaev E.E., Martynovich V.V. Formalization of "natural" classification and systematics through fixed points of predictions // SEMR. News. V. 12, IM SB RAS, 2015, pp. 1006-1031. (in Russian)
10. Dubrovsky D.I. The problem of "consciousness and the brain": An information approach. Knowledge, Understanding, Skills. 2013, №4. (in Russian)
11. Mukhortov V.V., Khlebnikov S.V., Vityaev E.E. Improved algorithm of semantic probabilistic inference in the 2-dimensional animate problem // Neuroinformatics, 6(1). 50-62. (in Russian)
12. Rudolf Carnap. Philosophical foundations of Physics. M., "Progress", 1971, p. 388. (in Russian)
13. Smirnov E.S. The construction of a species from a taxonomic point of view. Zool. Journal. (1938). 17:3, pp. 387-418. (in Russian)
14. Max Tegmark. Our mathematical universe. ACT, 2016, p. 592. (in Russian)
15. Cartwright, N. Causal Laws and Effective Strategies. Noûs. (1979) 13(4): 419-437.
16. 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.
17. Hebb D.O. The Organization of Behavior. Wiley: New York; 1949.
18. Hempel C.G. Maximal Specificity and Lawlikeness in Probabilistic Explanation. Philosophy of Science. (1968) 35, pp. 116-33.
19. Ganter B. Formal Concept Analysis: Methods, and Applications in Computer Science. TU Dresden (2003).
20. 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.
21. Mill J.S. System of Logic. Ratiocinative and Inductive. L., 1983.
22. 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.

23. Rehder B. Categorization as causal reasoning // *Cognitive Science*, 27. 2003, pp. 709–748.
24. Rosch E.H. Natural categories // *Cognitive Psychology* 4. 1975, P. 328-350.
25. Rosch E. Principles of Categorization // Rosch E.&Lloyd B.B. (eds), *Cognition and Categorization*. Lawrence Erlbaum Associates, Publishers, 1978. pp. 27–48.
26. Tononi G. Information integration: its relevance to brain function and consciousness. *Arch. Ital. Biol.*, 148: 299-322, 2010.
27. Tononi G. Integrated information theory of consciousness: an updated account. *Arch Ital Biol* 150, 2012, 56–90.
28. Vityaev E.E. The logic of prediction. // *Proceedings of the 9th Asian Logic Conference Mathematical Logic in Asia*. World Scientific, Singapore. 2005, pp. 263–276.
29. Vityaev E.E., Martinovich V.V. Probabilistic Formal Concepts with Negation // A. Voronkov, I. Virbitskaite (Eds.). *PCI 2014, LNCS 8974*, 2015, pp. 385-399.
30. Vityaev E.E. A formal model of neuron that provides consistent predictions // *Biologically Inspired Cognitive Architectures 2012 // Advances in Intelligent Systems and Computing*, v.196, Springer, 2013, pp. 339-344.
31. Vityaev E. Purposefulness as a Principle of Brain Activity // *Anticipation: Learning from the Past. Cognitive Systems Monographs*, V.25, Springer, 2015, pp. 231-254.
32. Vityaev, E., Odintsov, S. How to predict consistently? *Trends in Mathematics and Computational Intelligence // Studies in Computational Intelligence* 796. Mara Eug. Cornejo (ed), 2019, 35-41.