

# Начальное порождение понятий при обработке образов на алгебре фурье-дуальных операций<sup>1</sup>

**Аннотация.** Рассмотрено индуктивное порождение понятий при обработке образов на нейросетях, реализующих алгебру фурье-дуальных определяющих модель операций. Данна модель первоначального, т.е. при отсутствии ранее наличествовавших, порождения понятий. Показана биологическая мотивированность модели. Предложен феномен когнитивного насыщения, выражаящийся в остановке роста оценки эффективности формирования индуктивной гипотезы с ростом числа примеров. Дан анализ его механизма и связь с информационными характеристиками паттернов внутренних репрезентаций воспринимаемых образов, обусловленных индивидуальными особенностями когнитивного агента. Даны результаты моделирования на примере обращения силлогизма «Darii» при регистрации весов связей в плоскости и объеме.

**Ключевые слова:** понятийное мышление, образное мышление, понятие, индуктивное порождение понятий, фурье-дуальность, преобразование Фурье, нейронные сети, зрительная кора, согласованная фильтрация.

## Введение

В искусственном интеллекте и когнитивных науках традиционно выделяют две формы мышления: понятийное и образное. Первое рассматривается как более высокая форма, формирующаяся на основе образного [1-3]. Согласно классификации «образ – понятие», изложенной в [2], образ и понятие суть крайние случаи внутренних представлений информации. Образ в абстрактном пространстве признаков представлен унимодальным распределением с хорошо локализованным максимумом, а понятие – пологим холмом или плато, охватывающим широкий диапазон значений только тех признаков, что существенны для понятия. Образ конкретен и имеет преимущественно сенсорное происхождение, а понятие формируется как результат обработки набора образов, направленной на абстрагирование от частных признаков, важных для конкретизации образа, но несущественных для понятия. Соответственно, размерность пространства признаков

понятия может быть меньше размерности пространства признаков образа.

Переход от образного к понятийному мышлению предполагает два условных, но обязательных этапа: порождение нового понятия, установление связи нового понятия с уже наличествующими.

Один из методов порождения абстрактных понятий при обработке конкретных образов – индуктивное обобщение [4]. Правомочность применения здесь индукции определяется тем, что только индуктивный вывод дает расширение теории [5] и преобразование данных в знания [6].

В статье [7] предложен метод индуктивного порождения понятий в рамках бионического подхода, предполагающего внутреннее представление воспринимаемой информации картинами активности нейронных ансамблей [8] и их обработку на нейронных сетях (НС), реализующих алгебру фурье-дуальных операций. Эти картины нейронной активности в статье будем называть паттернами внутренней репрезентации информации (ПВР) или просто паттернами.

<sup>1</sup>Работа выполнена при финансовой поддержке РФФИ, проект № 15-01-04111-а.

Обратим внимание, что использование кратко изложенной в работах [9, 10] алгебры фурье-дуальных операций и, соответственно, НС со связями, веса которых формируются и запоминаются в пространстве Фурье, относит данный подход к категории биологически мотивированных. Это утверждение основано на результатах многолетних экспериментальных исследований зрительной системы, проводимых под руководством В.Д. Глазера и Ю.Е. Шелепина в Институте нейрофизиологии РАН им. И.П. Павлова, согласно которым в зрительной системе высших животных и человека реализуется преобразование Фурье [11-15]. Более того, имеет место не однократное, но двукратное преобразование Фурье, дающее в итоге реализацию закона Де-Моргана и модели согласованной фильтрации. Гипотеза согласованной фильтрации в зрительной системе была высказана Н.Н. Красильниковым еще в 1957 году [16] и затем экспериментально подтверждена в нейрофизиологических исследованиях [17-22].

Согласно [17], биологические структуры зрительной системы реализуют два типа согласованных фильтров:

1) нелинейные, согласованные с деталями воспринимаемого изображения – информация с нелинейных фильтров в норме поступает в левое полушарие головного мозга;

2) линейные, согласующие число периодов в весовой функции с размером изображения и создающие абстрактное, т.е. без деталей, описание изображения – информация с этих фильтров поступает в правое полушарие.

Взаимодействие обоих полушарий дает в итоге фильтр, согласованный с полным детальным описанием изображения, что подтверждено психофизиологическими экспериментами с переносом кодов из одного полушария в другое [17].

Описанный в [7] метод индуктивного порождения понятий предполагает существование шкалы общности признаков, в качестве которой принята реальная шкала пространственных частот в пространстве Фурье и, соответственно, априорного критерия отнесения признаков к частным или общим. Наличие в модели такого критерия не противоречит результатам нейрофизиологических исследований [11-22], но ограничивает ее применимость. Поэтому в работе [23] дана модель индуктивного перехода к понятийному мышлению, включающая оба этапа и свободная от ограничения метода [7] –

единственным критерием отнесения признака к частным или общим в ней служит частота его появления в наборе образов.

Вместе с тем, модель [23] предполагает наличие понятий, уже сформированных к моменту порождения нового. Таким образом, строго говоря, это модель не формирования понятийного мышления «с нуля», а развития уже существующего. Представляет интерес модель именно порождения понятийного мышления при обработке образов как модель перехода от образного мышления к понятийному, т.е. порождения понятий при отсутствии уже наличествующих.

В статье предложено решение этой задачи развитием модели [23]. Вопросы изменения объема понятия и его связи с характеристиками обучения оставлены за рамками статьи.

## 1. Подход и модель

### 1.1. Подход к задаче

Рассмотрим задачу индуктивного порождения понятий на примере обращения классического силлогизма Darii:

Частные высказывания  
(результаты):

*Сократ – смертен,  
Платон – смертен,  
etc.;*

Частные высказывания  
(случаи):

*Сократ – человек,  
Платон – человек,  
etc.;*

Общий вывод (правило):  
**Человек вообще  
смертен.**

Здесь можно выделить два конкретных образа: персоны, например, «Сократ» и ее (и только ее) смертности, и два понятия: «Человек вообще» и смертности вообще. Для удобства различения образов от понятий, последние даны полужирным шрифтом. Обозначим образы персоны и ее смертности *S* и *P*, а понятия человека и смертности – *M* и *C*. По условиям нашей задачи есть только представленная высказываниями типа «Сократ – смертен» связь двух образов – «результат» в терминах Ч.С. Пирса. Ни случая, ни правила нет, так как нет понятий, они пока еще скрыты в образах, их надлежит выявить – найти и выделить из образов, а связи «случай» и «правило», вербализованные высказываниями «Сократ – человек» и «Человек

**«вообще смертен»** (или в другой формулировке **«Все люди смертны»**) – сформировать.

В рамках декларированного во Введении биологически мотивированного подхода и образы, и понятия представлены картинами нейронной активности или паттернами. Далее термин паттерн в соответствующем контексте без дополнительных пояснений употребляется как синоним образа и понятия. Критерии их различия обсудим ниже. Паттерны являются функциями двух координат, но во избежание излишнего загромождения статьи, по возможности представим их функциями одной координаты, имея в виду две координаты тогда, когда это принципиально – при определении отношения дисперсий.

Представим образы как сумму двух субпаттернов:

$$\begin{cases} S_i(x) = S_i^C(x) + S_i^{Ind}(x) \\ P_i(x) = P_i^C(x) + P_i^{Ind}(x), \end{cases} \quad \begin{cases} S_i^C(x) = m^S \cdot S_i(x) \\ S_i^{Ind}(x) = (1 - m^S) \cdot S_i(x), \end{cases} \quad (1)$$

$$\begin{cases} P_i^C(x) = m^P \cdot P_i(x) \\ P_i^{Ind}(x) = (1 - m^P) \cdot P_i(x) \end{cases}$$

где субпаттерн с индексом  $C$  – включает признаки, общие для ряда образов, а с индексом  $Ind$  – частные, присущие только одному конкретному образу,  $m$  с соответствующими верхними индексами – удельные веса общих, т.е.

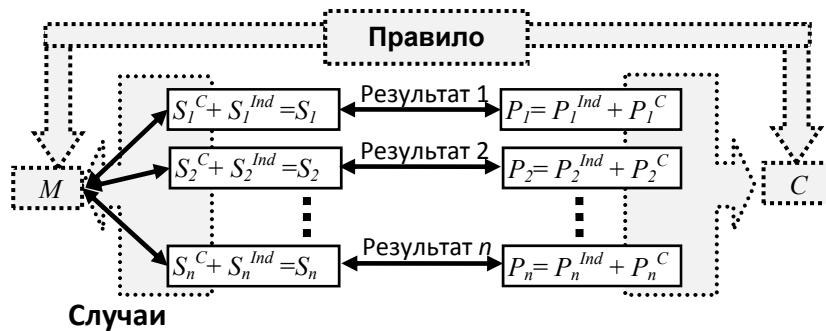


Рис. 1. Индуктивное формирование понятий при обработке образов

$S_i$  – индексные образы (*Сократ, Платон, etc.*),  $P_i$  – индуцируемые образы (*смертность Сократа, Платона, etc.*), все образы имеют индивидуальные  $S_i^{Ind}, P_i^{Ind}$  и общие  $S_i^C, P_i^C$  признаки,  $M$  и  $C$  подлежащие формированию понятия «человек вообще» и «смертность вообще»; стрелки – связи (в терминах Ч.С. Пирса): Результаты – наличествующие при обучении связи пар образов «Сократ – смертен», Случаи – подлежащие формированию связи порождаемого понятия с образами «Сократ – человек», Правило – подлежащая формированию связь двух понятий «все люди смертны»

коррелированных фрагментов паттернов для соответствующих наборов  $\{S_i\}$  и  $\{P_i\}$ .

Общие субпаттерны подлежат обнаружению и выделению как паттерны понятий. Принципиальная схема этой задачи графически представлена на Рис. 1.

На Рис. 1 виден критерий отличия паттернов образа и понятия – это характеристики не самих паттернов, а структуры их связей. В результате обучения НС каждый образ (его паттерн) связан только с одним, парным ему, образом – это структура связей типа «результат» (в терминах Ч.С. Пирса). Понятие же должно быть связано сразу со всеми образами соответствующего набора, эту структуру можно определить как «звезду случаев». Таким образом, в отличие от подхода [2, 3], здесь использован нейросетевой критерий, так как сами по себе паттерны не могут быть атрибутированы как образы или понятия – критичны не их свойства как паттернов, а структура их связей. Например, понятием может стать один из образов обучающего набора, если он содержит только общие для всего набора признаки. И, наоборот, при разрушении структуры связей «звезда случаев» понятие превратится в образ, без изменения самого паттерна. Это полностью соответствует важнейшему атрибуту НС парадигмы – информация хранится не в виде самих паттернов, а в структуре их связей.

Напомним, что работе [23] предполагалось, что понятие  $M$  и связи типа «случай» уже существуют, т.е. образы  $S_i$  включали только фрагмен-

ты, содержащие частные признаки  $S_i^{Ind}$ . В текущей постановке схема задачи (Рис. 1) симметрична – разделение образов на индексные и индуцируемые условно, так как и из паттернов образов, обозначенных как индексные, надлежит выделить субпаттерны общих признаков  $S_i$ , сформировать понятие  $M$  и связи типа «случай».

## 1.2. Модель

Рассмотрим возможность решения задачи на двухслойной НС с двунаправленными связями, веса которых формируются и регистрируются в пространстве Фурье. Принципиальная схема НС приведена на Рис. 2.

Образы представляются паттернами в слоях НС, а их связи (результаты в рамках примера) формируются в пространстве Фурье:

$$H(v_x) = \sum_{i=1}^n \left[ F(P_i(x)) F^*(S_i(x)) \exp(-j2\pi v_x (x_{P_i} - x_{S_i})) + F(S_i(x)) F^*(P_i(x)) \exp(-j2\pi v_x (x_{S_i} - x_{P_i})) \right] \quad (2)$$

где  $n$  – число обучающих примеров (пар ассоциированных образов  $S_i$  и  $P_i$ );  $F$  – символ преобразования Фурье; астериск – символ комплексного сопряжения; экспоненциальные члены описывают пространственное положение паттернов в соответствующих слоях; слагаемые в квадратных скобках – односторонние связи слоев S→P и P→S, соответственно.

Можно видеть, что поскольку (2) представляет собой частный вариант формирования связей по модели внешнего произведения векторов, то данная НС соответствует модели двунаправленной ассоциативной памяти Б. Коско [24]. Отличие от модели Коско в том, что концепция памяти предполагает восстановление эталона, наиболее близкого к входному (или ассоциированному с ним) образу. В обсуждаемой задаче этот вариант ответа неприемлем – необходимо найти и выделить субпаттерны, общие для всех запомненных образов в наборах  $\{S_i\}$  и  $\{P_i\}$ .

Рассмотрим первое слагаемое в (2), описывающее связи S→P. Для связей P→S, описываемых вторым слагаемым в (2), анализ совершенно аналогичен. Для того чтобы не загромождать статью чрезвычайно длинными выкладками, детальный вывод основных выражений опустим. Заинтересованный читатель найдет вывод всех выражений

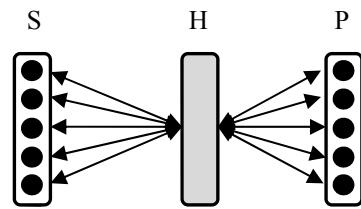


Рис. 2. Принципиальная схема двухслойной нейросети с двунаправленными связями

S и P – слои репрезентации образов  $S_i$  и  $P_i$ ;  
H – веса связей в пространстве Фурье;  
фурье-преобразующие элементы между S и H, H и P не показаны

применительно к голограммической реализации схемы Рис. 2 в работе [25].

При предъявлении в слое S паттерна  $S_i$ , в слое P формируется отклик НС – паттерн:

$$\begin{aligned} P_R(x) = P_R^C(x) + P_R^N(x) = & \left\{ 1 + (m^S)^2 (n-1) \right\} \cdot \\ & \cdot \left\{ P_i^C(x + x_p) * [S_i(x) \otimes S_i(x)] \right\} + \\ & + P_i^{Ind}(x + x_p) * [S_i(x) \otimes S_i(x)] + \\ & + (m^P)^2 \sum_{i \neq j} P_j^{Ind}(x + x_p) * [S_i(x) \otimes S_i(x)] + \\ & + \sum_{i \neq j} P_j(x + x_p) * [S_i^{Ind}(x) \otimes S_j^{Ind}(x)] + \\ & + \sum_{i \neq j} P_j(x + x_p) * [S_i^{Ind}(x) \otimes S_j^C(x)] + \\ & + \sum_{i \neq j} P_j(x + x_p) * [S_i^C(x) \otimes S_j^{Ind}(x)] \end{aligned} \quad (3)$$

где \* и  $\otimes$  – символы операций свертки и корреляции, соответственно, нижний индекс  $R$  означает восстановленный паттерн, т.е. с учетом влияния импульсного отклика системы.

Все слагаемые в (3) пространственно налагаются друг на друга. Первое описывает искомый субпаттерн общих признаков  $P_R^C(x)$  – паттерн понятия как индуктивной гипотезы. Остальные в сумме дают паттерн, который в рамках задачи представляет собой помеху  $P_R^N(x)$ .

Второе слагаемое не зависит от числа примеров  $n$  и дает постоянный вклад в помеху. Остальные зависят от  $n$  и, если с ростом  $n$  их вклад уменьшается, то возможно решение задачи – выделение общих, т.е. коррелированных субпаттернов  $P_R^C(x)$ .

Обратим внимание на члены в квадратных скобках в выражении (3) – авто- и кросскорреляционные функции, играющие роль импульсных откликов системы. В силу свойства сдвиговой инвариантности преобразования Фурье эти члены – мультимодальные функции, размер области определения которых равен сумме областей определения соответствующих паттернов. В результате, формируемый в слое Р паттерн, как результат свертки с мультимодальными функциями, будет характеризоваться чрезвычайно существенной потерей разрешения, т.е. потерей информации о мелких деталях паттернов. В живых системах межнейронные связи фиксированы в пространстве, поэтому свойство инвариантности преобразования Фурье к сдвигу не сохраняется и, следовательно, члены в квадратных скобках в (3) суть уни-модальные функции, что и используем при дальнейшем анализе.

В качестве оценок общей модулированности (контрастности) восстановленных субпаттернов  $P_R^C(x)$  и  $P_R^N(x)$  примем их дисперсии  $D_R^C$  и  $D_R^N$ , соответственно[23]. Тогда отношение дисперсий будет оценкой эффективности формирования паттерна индуктивной гипотезы в зависимости от  $n$ . Рост этого отношения будет означать увеличение контрастности субпаттерна общих признаков на фоне субпаттерна частных  $P_R^N(x)$ , превращающегося в равномерный серый фон. Детальный вывод, подробно приведенный в статье [25], здесь опустим ниже и дадим итоговые выражения.

Зависимость дисперсии субпаттерна общих признаков  $P_R^C(x)$ , как показано в [25], зависит от числа обучающих примеров следующим образом:

$$D_R^C = \left(1 + (n-1)K(m^s)\right)^2 D_{iR}^C, \quad (4)$$

где  $D_{iR}^C$  – дисперсия члена во вторых фигурных скобках в (3), т.е.  $i$ -го субпаттерна, восстановленного в слое Р с учетом ее изменения относительно дисперсии эталонного субпаттерна  $P_i^C$  в результате свертки с импульсным откликом схемы  $[S_i(x) \otimes S_i(x)]$ , а

$$K(m^s) = \left( (m^s)^2 + \sqrt{(1-m^s) \frac{2\kappa}{\Omega^s}} + 2\sqrt{t \frac{2\kappa}{\Omega^s}} \right),$$

где  $\kappa$  – коэффициент, зависящий от информационных характеристик индексных образов:

вида корреляционной функции и корреляционной оценки информационной емкости паттерна

$$\Omega^s = \frac{L_x^s L_y^s}{\pi r_R^2},$$

где  $r_R$  – радиус корреляции с учетом влияния импульсного отклика;  $t$  – коэффициент, учитывающий отношение размеров субпаттернов (1) следующим образом:

$$\begin{cases} t = m^s & \text{если } m^s \leq 0.5 \\ t = 1 - m^s & \text{если } m^s \geq 0.5 \end{cases}.$$

Согласно результатам [26], для экспоненциальной функции корреляции при не очень низких оценках информационной емкости, ориентировочно  $\Omega > 20$ , можно принять  $\kappa \approx 0.25$ .

Дисперсия  $P_R^N(x)$  может быть представлена выражением [25]:

$$D_R^N = D_{iR}^{Plnd} \cdot \left[ 1 + \left( K(m^s) \right)^2 (n-1) \left( 1 + (n-2) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \right) + \left[ + 2K(m^s) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \sqrt{(n-1) \left( 1 + (n-2) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \right)} \right] \right], \quad (5)$$

где  $D_{iR}^{Plnd}$  – дисперсия восстановленного, т.е. с учетом влияния импульсного отклика, субпаттерна частных признаков;  $\Omega^{Plnd}$  – корреляционная оценка его информационной емкости.

Отсюда получим интересующую нас в качестве оценки эффективности выделения субпаттернов общих признаков на фоне субпаттернов частных, зависимость отношения дисперсий субпаттернов от размера обучающей выборки (числа примеров)  $n$ :

$$V(n) = \frac{D_R^C}{D_R^N} \approx \frac{D_{iR}^C}{D_{iR}^{Plnd}} \cdot \frac{\left(1 + (n-1)K(m^s)\right)^2}{\left[ 1 + \left( K(m^s) \right)^2 (n-1) \left( 1 + (n-2) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \right) + \left[ + 2K(m^s) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \sqrt{(n-1) \left( 1 + (n-2) \sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \right)} \right] \right]}. \quad (6)$$

### 1.3. Анализ теоретической оценки и феномен когнитивного насыщения

Из анализа выражения (6) можно сделать ряд выводов.

1. Из (6) следует, что условие решения первого этапа задачи – формирования индуктивной гипотезы – частичная коррелированность не только индуцируемых  $\{P_i\}_{i=1}^n$ , но также и индексных образов  $\{S_i\}_{i=1}^n$  в обучающей выборке. Это означает наличие во всех паттернах  $\{S_i\}_{i=1}^n$  общих субпаттернов – скрытых прототипов индуктивно формируемого понятия  $M$ . В рамках используемого примера для всех индексных образов общее то, что все они – люди. Таким образом, поскольку НС запоминает не образы сами по себе, но связи пар образов, то и порождается не одно понятие, но сразу пара – и  $C$  и  $M$ .

Обратим внимание также на то, что обученной НС (2) может быть предъявлен любой образ  $S_i$ , даже при  $i > n$ , т.е. не обязательно принадлежащий обучающей выборке – главное, чтобы он содержал субпаттерн  $S_i^C$ , представляющий свойство быть человеком.

2. С ростом числа обучающих примеров  $n$  оценка эффективности выявления паттерна общих признаков на фоне субпаттернов частных (6) выходит в насыщение. Вновь опуская громоздкие выкладки, дадим достаточную для анализа механизма приближенную оценку:

$$\lim_{n \rightarrow \infty} V(n) \approx \sqrt{\frac{\Omega^{Plnd}}{2\kappa}}. \quad (7)$$

Зависимость уровня насыщения (7) только от оценки информационной емкости субпаттернов частных признаков отражает механизм выявления субпаттернов общих признаков – сглаживание (снижение контраста) с ростом числа обучающих примеров  $n$  субпаттерна частных признаков как фона, на котором выявляется субпаттерн общих признаков: нормированная дисперсия последнего от  $n$  не зависит.

Этот эффект выхода оценки (6) в насыщение представляется правомочным трактовать как феномен когнитивного насыщения, так как (6) есть оценка зависимости эффективности обучения от объема обучающей выборки. Начиная с некоторого числа обучающих примеров, дальнейшее обучение оказывается неэффективным.

Из (7) следует, что уровень насыщения субъективен, так как определяется только индивидуальными особенностями когнитивного агента, а именно:

– оценка информационной емкости субпаттерна

$$\Omega^{Plnd} = \frac{L_x^{Plnd} L_y^{Plnd}}{\pi r_R^2}$$

зависит от его (*sic!* паттерна внутренней представлении, а не объекта внешнего мира) размера  $L_x^{Plnd} L_y^{Plnd}$ , т.е. от числа нейронов, активированных частными признаками, и радиуса корреляции  $r_R$ , определяемого фильтрующими свойствами сенсоров, сенсорных трактов и НС – полосой пропускания сенсорных сигналов [27];

– коэффициент  $\kappa$  также определяется свойствами сенсоров, сенсорных трактов и НС [26].

3. Если оценка информационной емкости паттерна, подлежащего выявлению понятия, много меньше аналогичной оценки субпаттерна частных признаков и последняя оценка достаточно велика, т.е. частные признаки явно доминируют в образах, то имеет силу

$$\sqrt{\frac{2\kappa}{\Omega^{Plnd}}} \ll 1 \text{ и } K(m^s) \approx (m^s)^2.$$

Отсюда можно получить более простое, сравнительно с (6), аппроксимативное выражение:

$$V(n) \approx \frac{D_{kout}^{(s)c}}{D_{kout}^{(s)u}} \frac{\left[ 1 + (n-1)(m^{(R)})^2 \right]^2}{\left[ 1 + (m^{(R)})^4 (n-1) \left( 1 + (n-2) \sqrt{\frac{2\kappa}{\Omega^{(s)u}}} \right) \right]}. \quad (8)$$

Выражение (8) дает заметную ошибку в диапазоне  $n < n_{\text{эфф}}$ , но при приближении  $n$  к  $n_{\text{эфф}}$  оценки (6) и (8) сравнимы.

4. Информационные характеристики паттернов индексных образов – коэффициент их корреляции  $\rho_{ij}^{(s)}$  и оценка информационной емкости  $\Omega^{(s)}$ , входящие в выражение для члена  $K(m^{(s)})$ , играют заметную роль по преимуществу в диапазоне небольших  $n$  при высоких оценках информационной емкости индексных образов  $\Omega^{(s)}$ , влияя лишь на скорость формирования понятия на начальном этапе, но не на величину насыщения (7).

5. Если субпаттерны частных признаков индуцируемых образов ортогональны, то  $\sqrt{\frac{2\kappa}{\Omega^{(S)u}}} = 0$  и (6) принимает вид:

$$V(n) = \frac{D_R^C}{D_R^N} \approx \frac{D_{iR}^C}{D_{iR}^{Plnd}} \frac{\left[1 + (n-1)K(m^{(S)})\right]^2}{\left[1 + (n-1)\left(K(m^{(S)})\right)^2\right]}. \quad (9)$$

В первом приближении для не малых  $n$  выражение (8) может быть аппроксимировано линейной зависимостью, что совпадает с результатами, представленными в [23]. Если паттерны индексных образов полностью коррелированы, т.е. не содержат субпаттернов частных признаков, то  $K(m^{(S)}) = 1$  и получим линейную зависимость:

$$V(n) = \frac{D_R^C}{D_R^N} = n \frac{D_{iR}^C}{D_{iR}^{Plnd}}. \quad (10)$$

## 2. Моделирование

Моделировалась работа НС применительно к используемому в статье примеру на основе обращенного силлогизма Darii (Рис. 1). Использован биологически мотивированный метод организации весов связей в объеме, т.е. с потерей свойства инвариантности преобразования Фурье к сдвигу. Паттерны наборов  $\{S_i\}$  и  $\{P_i\}$  моделировались реализациями размером  $256 \times 256$  пикселей случайных полей, генерированных по моде-

лям фрактального броуновского движения (ФБД–процесса со стационарными приращениями) и однородных полей с разными радиусами корреляции, включая дельта-коррелированные. Для изменения радиуса корреляции от  $r < 1$  до  $r = 10$  пкс. и, соответственно, оценок информационной емкости  $\Omega$ , использована фильтрация Гаусса. Статистические и спектральные характеристики образов в рамках каждого набора  $\{P_i\}_{i=1}^n$  и  $\{S_i\}_{i=1}^n$  были одинаковы.

Исходя из соображений совмещения биологической мотивированности и наглядности иллюстрации, каждый паттерн содержал:

1) коррелированный, т.е. общий для всех паттернов последовательности субпаттерн как надписи «Mortal» (для индуцируемых образов) и «Men» (для индексных образов), заполненные одной для всех паттернов соответствующего набора реализацией случайного поля; их удельные веса (1)  $(1 - m^S) = 0.369$  и  $(1 - m^P) = 0.375$ , соответственно;

2) некоррелированные, т.е. представляющие индивидуальные признаки субпаттерны: они включали личные имена «Socrates», «Plato» и т.п., заполненные различными реализациями полей, а также и всю остальную часть поля, исключая коррелированные части паттернов.

Информационные характеристики обучающих образов даны в Табл. 1, радиусы корреляции соответствуют радиусам размывающего дельта-коррелированное поле фильтра Гаусса.

Табл. 1. Информационные характеристики обучающих образов

Номер	Радиус корреляции $r$	Оценка информационной емкости субпаттерна частных признаков $\Omega^{Plnd}$	Оценка информационной емкости субпаттерна общих признаков $\Omega^{Pc}$	Уровень насыщения теоретический (7)
1	$r < 1$ (дельта-коррелированное поле)	$\infty$	$\infty$	$\infty$
2	1	13050	7815	161
3	2	3260	1954	80
4	3	1450	868	54
5	4	815	488	40
6	5	521	312	32
7	7.5	232	139	21.5
8	10	130	78	16

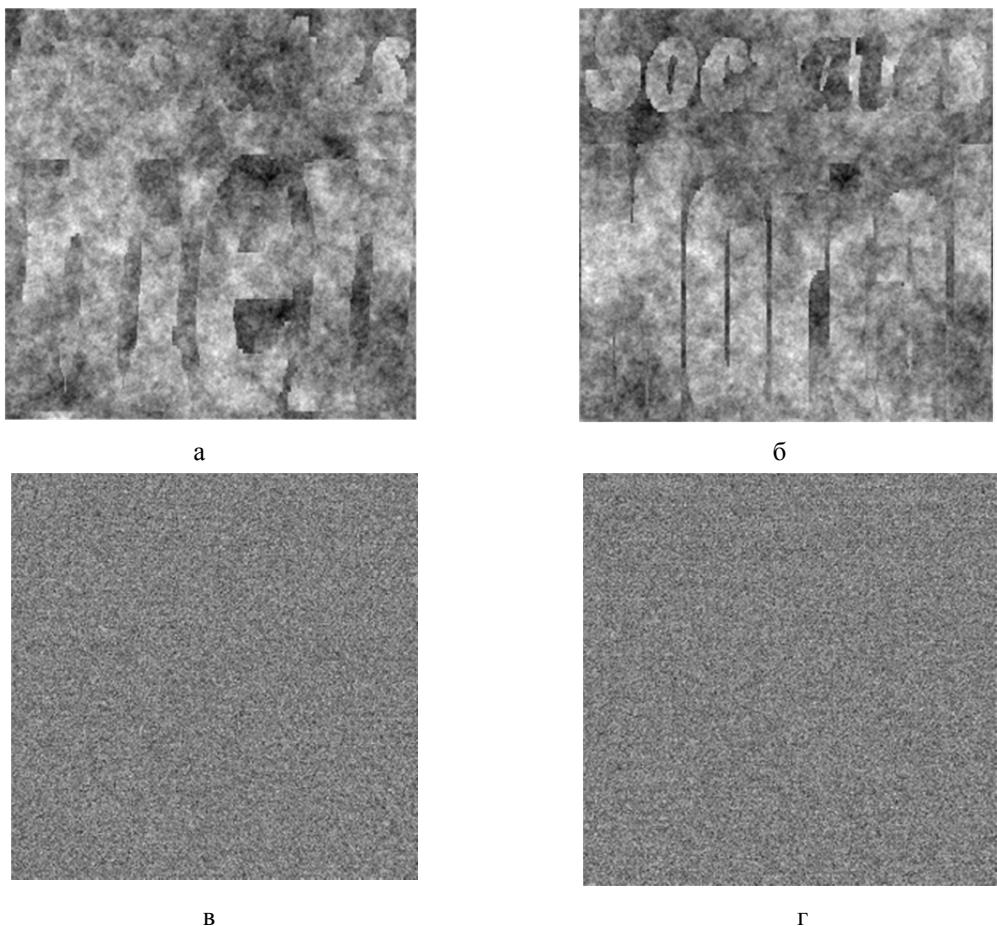


Рис. 3. Примеры обрабатываемых пар паттернов

- а – индексный образ «Сократ человек» – реализация двумерного ФБД
- б – индуцируемый образ «Сократ смертен» – реализация двумерного ФБД
- в – индексный образ «Сократ человек» – реализация дельта-коррелированного поля
- г – индуцируемый образ «Сократ смертен» – реализация дельта-коррелированного поля

На Рис. 3 даны примеры обрабатываемых пар паттернов. На Рис. 3, а и Рис. 3, б представлено поле – реализация двумерного ФБД с параметром Хёрста  $H=0.1$ , на Рис. 3, в и Рис. 3, г – дельта-коррелированное поле. Видно, что надписи в структуре паттерна визуально не выделяются; это иллюстрирует факт разрушения информации и ее превращения в белый шум при переходе к дельта-коррелированности.

На Рис. 4 даны примеры паттернов, восстановленных в слое Р при предъявлении в слое S первого паттерна «Socrates Men»: для дельта-коррелированного поля (Рис. 4, а), радиусов корреляции 2 пкс. (Рис. 4, б) и 10 пкс. (Рис. 4, в) (строки 1, 3 и 8 Табл. 1, соответственно). На Рис. 5. даны измеренные зависимости  $V(n)$  для разных оценок информационной емкости различающихся субпаттернов индуцируемых об-

разов, а на Рис. 6 – относительные ошибки экспериментальных и теоретических оценок эффективности выявления.

Из анализа Рис. 5 и Рис. 6 можно заключить:

1. С ростом числа обучающих примеров  $n$  точность теоретической оценки (6) возрастает. Имея в виду аналоговый характер обработки, точность оценки (6) при  $n \geq 15$  для характеристик образов 3 – 7 (Табл. 1) следует признать удовлетворительной.

2. Оценка (6) неудовлетворительна для паттернов – реализаций дельта-коррелированных полей (кривые 1 и 2). Дельта-коррелированность как полное разрушение внутренней связности образов ведет к превращению информации в белый шум, что сопровождается ростом дисперсии. Последний фактор не учитывался при выводе (6). В то же время, метод эффективно выявил из шума

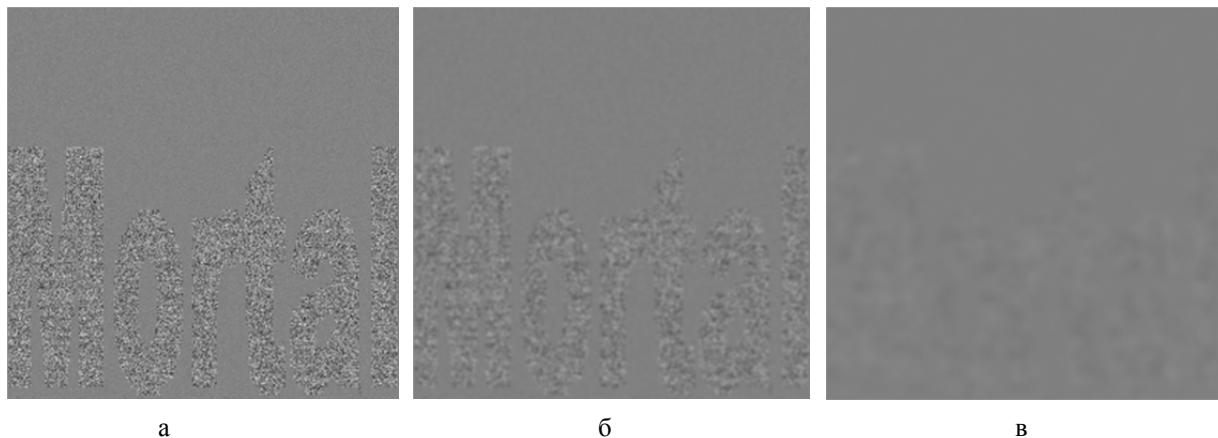


Рис. 4. Примеры паттернов, восстановленных в слое Р

а – для дельта-коррелированного поля  
 б – при радиусе корреляции  $r = 2$   
 в – при радиусе корреляции  $r = 10$

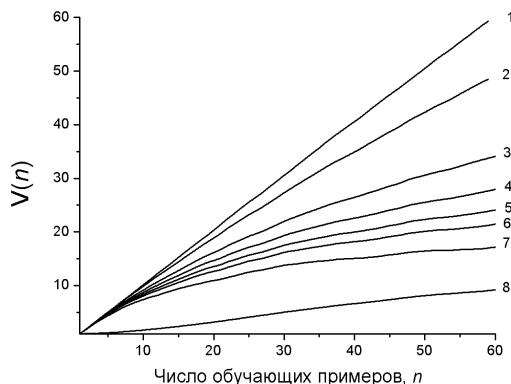


Рис. 5. Зависимости от числа обучающих примеров  $n$  отношения дисперсий в восстановленном в слое Р паттерне субпаттернов общих и частных признаков для характеристик обучающих паттернов, данных в Табл. 1

скрытую в нем информацию – надпись, не читавшаяся на эталонном образе, (Рис. 3, б), восстановлена НС, обученной 59-ю примерами (Рис.4, а).

3. Выражение (6) дает также неудовлетворительную оценку при предельно низкой оценке информационной емкости субпаттерна общих признаков (кривая 8). На Рис. 4, в наглядно видна причина этого – механизм сглаживания фона работает все также эффективно, но сам подлежащий выявлению субпаттерн столь малоинформативен, что даже на ровном фоне практически не читается.

Для иллюстрации важности несохранения свойства инвариантности преобразования Фурье

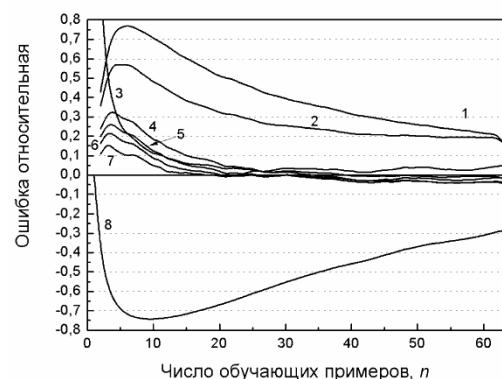


Рис. 6. Относительные ошибки экспериментальных (Рис.5) и теоретических (6) зависимостей от числа обучающих примеров  $n$  отношения дисперсий субпаттернов общих и частных признаков в восстановленном паттерне

к сдвигу численные эксперименты были повторены при сохранении свойства сдвиговой инвариантности. На Рис. 7 приведены паттерны, восстановленные НС (Рис. 2) в слое Р для различного числа обучающих примеров  $n$ , а на Рис. 8 – измеренные зависимости дисперсий субпаттернов и их отношения от числа обучающих примеров  $n$ .

На Рис. 7 и Рис. 8 можно видеть, что сохранение свойства инвариантности преобразования Фурье к сдвигу, важного в ряде практических применений, ведет к чрезвычайному, сравнительно с Рис. 5, снижению оценки эффективности формирования индуктивной гипотезы.

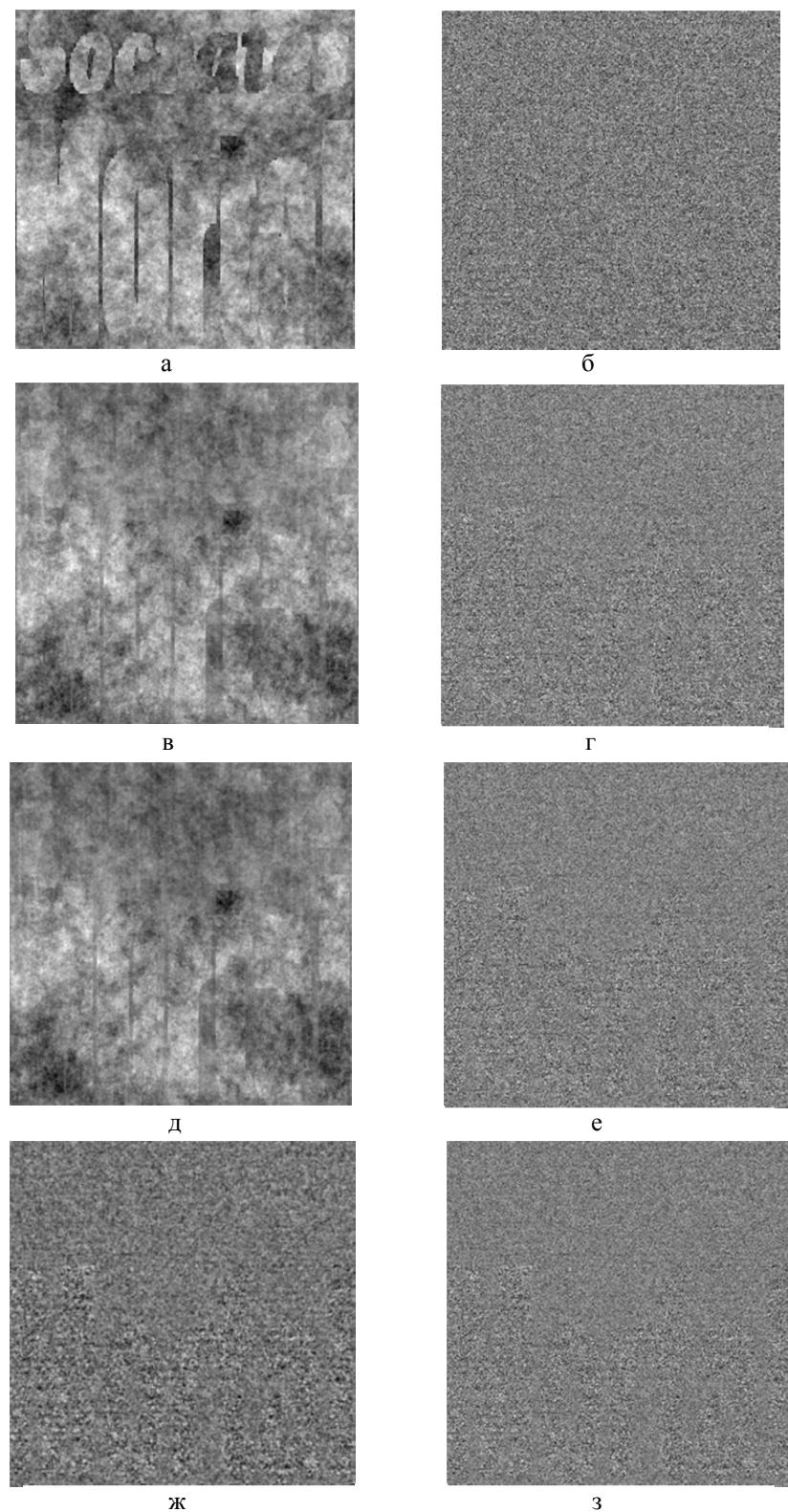


Рис. 7. Восстановленные паттерны для разного числа обучающих примеров  $n$

а – поле по модели ФБД, один обучающий пример; б – дельта-коррелированное поле, один обучающий пример; в – поле по модели ФБД, 10 обучающих примеров; г – дельта-коррелированное поле, 10 примеров; д – поле по модели ФБД, 20 обучающих примеров; е – дельта-коррелированное поле, 20 примеров; ж – однородное поле, 59 обучающих примеров, радиус корреляции  $r=2$ ; з – дельта-коррелированное поле, 59 обучающих примеров.

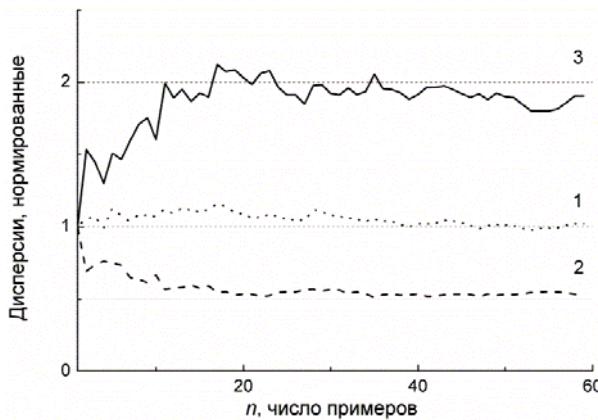


Рис. 8. Зависимости дисперсий коррелированных (1) и некоррелированных (2) субпаттернов и их отношения (3) от числа обучающих примеров

Для иллюстрации зависимости оценки эффективности (6) от коэффициента корреляции паттернов внутри обучающего набора, т.е. удельного веса субпаттерна общих признаков, работа НС моделировалась применительно к представлению паттернов реализациями стационарного случайного процесса с экспоненциальным спектром амплитуд и случайным спектром фаз с нулевым мат. ожиданием и дисперсией  $2\pi$ , различными радиусами корреляции. Длины всех реализаций составляли 1024 отсчета. Размеры субпаттернов  $P^C$  и  $P^{Ind}$  были равны и составляли 512 отсчетов. Коэффициент корреляции  $\rho^S = (m^S)^2$  индуцируемых паттер-

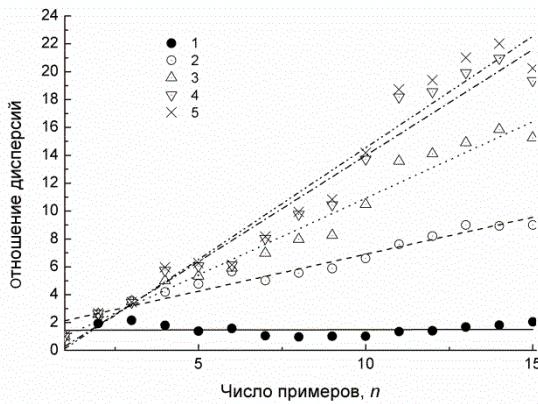


Рис. 9. Зависимости отношения дисперсий (10) от числа примеров для коэффициентов корреляции  $\rho^S$ : 1 - 0.01, 2 - 0.25, 3 - 0.49, 4 - 0.81, 5 - 1; оценки информационной емкости  $\Omega^C = \Omega^U = 22$

нов менялся от 0 до 1. Импульсный отклик системы приводился к дельта функции «разбеливанием» фурье-образов:

$$|F(S_i(x))| = \text{Const}(\nu_x).$$

На Рис. 9 даны экспериментальные зависимости отношения дисперсий  $V(n) = \frac{D_R^C}{D_R^N}$  от числа примеров  $n$  для ряда значений  $\rho^S$ , а также их линейные аппроксимации для одной реализации. Зависимость оценки  $V(n) = \frac{D_R^C}{D_R^N}$  от числа примеров  $n$  практически линейна, т.е. соответствует теоретическим выводам.

На Рис. 10 для ансамбля из 20 реализаций даны зависимости мат. ожидания (Рис. 10, слева) тангенса угла наклона аппроксимирующих прямых (Рис. 9), от коэффициента корреляции  $\rho^S$  для ряда оценок информационной емкости паттернов  $\Omega$ , а на Рис. 10, справа – их среднеквадратические отклонения. Тангенс угла наклона отношения  $V(n) = \frac{D_R^C}{D_R^N}$  дает оценку чувствительности к коэффициенту корреляции образов в обучающем наборе  $\rho^S$ .

Таким образом, сравнение зависимостей отношения  $V(n) = \frac{D_R^C}{D_R^N}$ , принятого в качестве оценки эффективности метода выделения паттерна гипотезы из фона-помехи, представленных на Рис. 5 и Рис. 8, наглядно демонстрирует

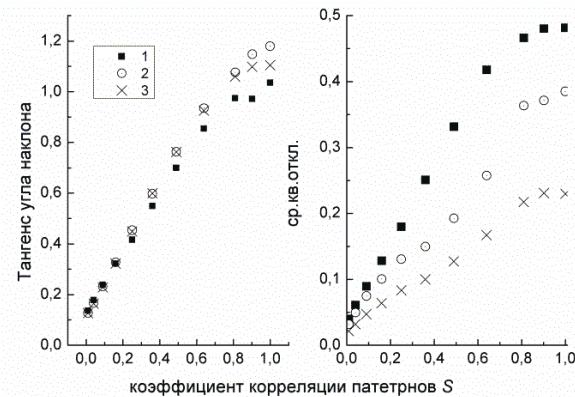


Рис. 10. Зависимости мат. ожидания (слева) и их среднего квадрата отклонения (справа) тангенсов угла наклона линейной аппроксимации (Рис.9) от  $\rho^S$ , усредненного по 20 реализациям: 1 -  $\Omega^C = \Omega^U = 22$ ; 2 -  $\Omega^C = \Omega^U = 46$ ; 3 -  $\Omega^C = \Omega^U = 93$

эффективность отказа от сохранения свойства инвариантности преобразования Фурье к сдвигу за счет объемной организации весов связей, что и имеет место в живых системах. При регистрации весов связей в объеме, как следует из теоретического анализа выражения (6) и экспериментальных данных (Рис. 9), модель индуктивного порождения первоначальных понятий работает даже при крайне слабой коррелированности обрабатываемых образов (низком удельном весе субпаттерна общих признаков).

## Заключение

Таким образом, рассмотрена и экспериментально подтверждена модель первоначального, т.е. при отсутствии ранее существовавших, индуктивного порождения понятий при обработке набора пар образов на алгебре фурье-дуальных операций. Модель реализуется двухслойной нейронной сетью и свободна от априорных критериев отнесения признаков к частным или общим, кроме частоты их появления в обучающем наборе.

Как показывают результаты нейрофизиологических исследований [11-22], модель, основанная на использовании преобразования Фурье в качестве операции, задающей дуальность, отвечает критерию биологической мотивированности. Схема двухслойной НС со связями в пространстве Фурье, в свою очередь, соответствует универсальной структуре корковых колонок зрительной коры [21].

Эффективность формирования паттерна понятия, как выявления субпаттерна общих признаков на основе обработки набора образов, определяется оценкой информационной емкости субпаттернов частных признаков образов, исключая случай крайне низкой оценки информационной емкости подлежащих выявлению субпаттернов общих признаков. Это отражает механизм работы модели – сглаживание по мере роста числа обучающих примеров субпаттернов индивидуальных признаков. Полученная теоретическая оценка дает удовлетворительную точность сравнительно с экспериментальными данными при «разумных» значениях информационных характеристик образов – в первую очередь, необходима достаточная информационная емкость самих общих фрагментов.

Эффективность индуктивного формирования паттерна понятия имеет нелинейный характер и

с ростом числа обучающих примеров выходит в насыщение, уровень которого определяется оценкой информационной емкости субпаттернов частных признаков. Последняя, в свою очередь, определяется индивидуальными характеристиками когнитивного агента: фильтрацией сенсорных сигналов при их передаче на кору головного мозга и числом нейронов, активированных воспринимаемыми образами – богатством внутренних представлений. Таким образом, отмеченный эффект насыщений отражает индивидуальную способность к формированию абстрактных понятий на основе обработки набора образов и может быть определен как эффект когнитивного насыщения.

Индуктивное порождение понятия возможно при коррелированности индексных образов. Это означает, что порождается не одно понятие, но обязательно их пара в результате выявления понятий как в наборе индуцируемых образов, так и в индексных.

Автор считает приятным долгом выразить благодарность проф. И.Б. Фоминых и проф. О.П. Кузнецовой за обсуждения и замечания, способствовавшие формированию развивающегося подхода.

## Литература

1. Кузнецов О.П. Быстрые процессы мозга и обработка образов // Новости искусственного интеллекта. 1998. №2.
2. Голицын Г.А., Фоминых И.Б. Нейронные сети и экспериментальные системы: перспективы интеграции // Новости искусственного интеллекта. 1996. №4. С.121-145.
3. Фоминых И.Б. Интеграция логических и образных методов отражения информации в системах искусственного интеллекта // Новости искусственного интеллекта. 1998. № 3. С. 76 – 85.
4. Hayes B.K., Heit E., Swendsen H. Inductive reasoning // Wiley Interdisciplinary Reviews: Cognitive Science.2010. V. 1, Issue 2. P. 278-292.
5. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. Изд. 2-е. М.: ФИЗМАТЛИТ. 2008.
6. Барский А.Б. Логические нейронные сети. М.:Бином. 2007. 351 С.
7. Павлов А.В. Реализация правдоподобных выводов на нейросетях со связями по схеме голограмии Фурье // Искусственный интеллект и принятие решений. 2010. №1. С.3-14.
8. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р. Модели динамики нейронной активности при обработке информации мозгом – итоги десятилетия // Успехи Физических наук. 2002. Т.172. №10. С.1189-1214.
9. Павлов А.В. Математические модели оптических методов обработки информации // Известия Академии

- Наук. Серия: Теория и системы управления. 2000. №3. С.111 – 118.
10. Павлов А.В. Об алгебраических основаниях голограммической парадигмы в искусственном интеллекте: алгебра Фурье-дуальных операторов // V-я Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте», 28-30 мая 2009 г., Коломна. Труды конференции. М.:Физматлит.2009. Т.1. С.140-148.
  11. Глазер В.Д., Иванов В.А., Щербач Т.А. Ответ рецептивных полей зрительной коры кошки на сложные стимулы // Физиологический журнал СССР. 1972. Т. 58. № 3.
  12. Глазер В.Д., Иванов В.А., Щербач Т.А. Исследование рецептивных полей нейронов зрительной коры кошки как фильтров пространственных // Физиологический журнал СССР. 1973. Т. 59. № 2.
  13. Глазер В.Д., Дудник К.Н., Куперман А. М., Лушина Л. И., Невская А.А., Подвигин Н.Ф., Праздникова И.В. Зрительное опознание и его нейрофизиологические механизмы. Л.: Наука. 1975.
  14. Глазер В.Д. Зрение и мышление. Спб.: Наука. 1993. 341 С.
  15. Глазер В.Д. О роли пространственно-частотного анализа, примитивов и межполушарной асимметрии в опознании зрительных образов // Физиология человека. 2000. Т. 26. №5. С. 145-150.
  16. Н.Н.Красильников Влияние шумов на контрастную чувствительность и разрешающую способность приемной телевизионной трубы // Техника телевидения. 1958. Вып. 25. С.26 – 43.
  17. Глазер В.Д. Согласованная фильтрация в зрительной системе // Оптический журнал. 1999. Т.66. № 10. С. 10 – 14.
  18. Красильников Н.Н., Красильникова О.И., Шелепин Ю.Е. Экспериментальное исследование согласованной пространственной фильтрации в зрительной системе человека при наблюдении чисто хроматических изображений // Оптический журнал. 1999. Т.66. №10. С.22 – 25.
  19. Красильников Н.Н., Шелепин Ю.Е. Функциональная модель зрения // Оптический журнал. 1997. Т.64. №2. С.72 – 82.
  20. Шуваев В.Т. Согласованная фильтрация – одна из возможных функций базальных ганглиев? // Оптический журнал. 1999. Т.66. №10. С.41 – 45.
  21. Алексеенко С.В. Архитектура связей в зрительной коре и зрительное опознавание // Оптический журнал. 1999. Т.66. №10. С.52 – 54.
  22. Кропотов Ю.Д. Система селекции действий мозга как фильтр высокого уровня // Оптический журнал. 1999. Т.66. №10. С.55 – 57.
  23. Павлов А.В., Кочетков П.В. Формирование понятийного мышления на алгебре фурье-дуальных операций // Искусственный интеллект и принятие решений. 2017. № 1. С. 27-42.
  24. Kosko B. Bi-directional associative memories. IEEE Transactions on Systems, Man and Cybernetics. 1987. V.18. №1. Р. 49-60.
  25. Павлов А.В. Выявление общих фрагментов в последовательности образов наложенным объемными голограммами Фурье // Квантовая электроника. 2017. Т. 47. № 4. С. 335–342.
  26. Шубников Е.И. Отношение сигнал/помеха при корреляционном сравнении изображений // Оптика и спектроскопия. 1987. Т.62. №.2. С. 450 – 456.
  27. Кулешов А.М., Шубников Е.И. Влияние нелинейности среды и пространственных ограничений фильтра на параметры сигнала в голограммическом корреляторе // Оптика и спектроскопия. 1986. Т. 60. № 3. С. 606 – 609.

**Павлов Александр Владимирович.** Доцент университета ИТМО. Окончил Ленинградский институт точной механики и оптики в 1980 году. Доктор физико-математических наук. Количество печатных работ: более 100. Область научных интересов: искусственный интеллект, обработка образов, голограммия. E-mail: pavlov@phoi.ifmo.ru

## Conceptual thinking generation by patterns processing by algebra of fourier-dual operations

A.V. Pavlov

Inductive concepts generation by set of examples processing by the algebra of Fourier-dual operations is considered. A model of initial, that is under yet existed concepts absence, concepts generation is proposed. Biological inspiration of the model is discussed. A phenomenon of cognitive saturation, manifesting by the estimation of the inductive hypothesis pattern determining efficiency stopping by the number of learning examples increasing is demonstrated. The phenomenon mechanism is analyzed with respect of information capacity of the patterns of inner representation of perceived information, that is determined by the individual peculiarities of cognitive agent. The phenomenon manifestation is decreased by using volume recording media for the interconnections registration. Theoretical conclusions are confirmed by numerical simulations for “Darii” syllogism inversion.

**Keywords:** figurative thinking, Conceptual thinking, Concepts, Induction, Inductive concepts generation, Fourier-duality, Fourier transform, Neural networks, Visual Cortex, Matching Filtering

### References

1. Kuznetsov O.P. Bystrye processy mozga i obrabotka obrazov // Novosti iskusstvennogo intellekta. 1998. №2. <http://raai.org/library/ainews/1998/2/DISTR.ZIP>.
2. Golitsyt G.A., Fominyh I.B. Neironnye seti I expertnye sistemy: perspektivy integratcii // Novosti iskusstvennogo intellekta.

- 1996, №4. P.121-145.
3. Fominyh I.B. Integratcija logicheskikh i obraznyh metodov otrazeniya informacii v sistemah iskusstvennogo intellekta // Novosti iskusstvennogo intellekta, 1998. № 3. P. 76 – 85.
  4. Hayes B.K., Heit E., Swendsen H. Inductive reasoning // Wiley Interdisciplinary Reviews: Cognitive Science.2010. V. 1, Issue 2. P. 278-292.
  5. Vagin, V.N., Golovina, E.Yu. and Zagoryanskaya, A.A., Dostoverniy i pravdopodobnyi vydov v intellektual'nykh sistemakh (Reliable and Probable Conclusion in Intelligent Systems), Vagin, V.N. and Pospelov, D.A., Eds., Moscow: FIZMALIT, 2008.
  6. Barskii A.B. Logicheskie neironnye seti.- M.: Binom, 2007. 351 P.
  7. Pavlov A.V. Realizatsia pravdopodobnyh vydov na neirosetiyah so sviaziami po scheme holographii Fourier // Iskusstvennyy intellekt i priinyatie resheniy. 2010, №1. P.3-14.
  8. Borisuk G. N., Borisuk R. M., Kazanovich Ya. B., Ivanitskii G. R. Models of neural dynamics in brain information processing — the developments of 'the decade' // Phys. Usp. 451073–1095 (2002); DOI: 10.1070/PU2002v045n10ABEH001143.
  9. Pavlov A.V. Mathematical Models of Optical Methods in Data Processing// Journal of Computer and Systems Sciences International. 2000, №3. P.441 – 448.
  10. Pavlov A.V. Ob algebraicheskikh osnovaniyah holograficheskoi paradigm v iskusstvennom intellekte: algebra Fourier-dual'nyh operatorov // V mezdunarodnaya nauchno-prakticheskaya konferentsiya: Integrirovannye modeli i myagkie vychisleniya v iskusstvennom intellekte", 28-30.05. 2009 ., Kolomna. Trudy konferentsii.- M.Fizmatlit,2009. T.1, P.140-148.
  11. Glezer V.D., Ivanov V.A., Sherbach T.A. Otvet retseptivnykh polei zritel'noi kory koshki na slognye stimuly // Phys. Journal USSR. 1972. V.58. №3.
  12. Glezer V.D., Ivanov V.A., Sherbach T.A. Issledovanie retseptivnykh polei neironov zritel'noi kory koshki kak filtrov prostranstvennyh chastot // Phys. Journal USSR. 1973. V.59. №2.
  13. Glezer V.D.,Dudnik K.N. et al. Zritel'noe opoznaniye i ego neirofiziologicheskie mechanizmy. L.: Nauka, 1975.
  14. Glezer V.D. Zrenie i myshlenie. SPb.: Nauka, 1993. 341 P.
  15. Glezer V.D. The Role of Spatial-Frequency Analysis, Primitives, and Interhemispheric Asymmetry in the Identification of Visual Images // Human Physiology. 2000. V.26. №.5. P.636 – 640.
  16. Krasilnikov N.N. Vliyanie shumov na kontrastnuyu chuvstvitelnost' i razreshayushuyu sposobnost' priemnoii televizionnoii trubki // Technika televideniya. 1958. 25. p.26 – 43.
  17. Glezer V.D. Matched filtering in the visual system // Journal of Optical Technology. Vol. 66, Issue 10, October 1999, Pages 853-856.
  18. Krasil'nikov, N.N., Krasil'nikova, O.I., Shelepin, Yu.E. Experimental study of matched spatial filtering in the human visual system when purely chromatic images are observed // Journal of Optical Technology Volume 66, Issue 10, October 1999, Pages 862-864.
  19. Krasil'nikov, N.N., Shelepin, Yu.E. A functional model of vision // Начало формы.
  20. Journal of Optical Technology. Volume 64, Issue 2, February 1997, Pages 136-144.
  21. Shubaev, V.T. Matched filtering - a possible function of the basal ganglia? // Journal of Optical Technology Volume 66, Issue 10, October 1999, Pages 877-880.
  22. Alekseenko S.V. Architecture of connections in the visual cortex and visual recognition// Journal of Optical Technology Volume 66, Issue 10, October 1999, Pages 886-887.
  23. Kropotov Yu.D. System for selecting the actions of the brain as a high-level filter // Journal of Optical Technology Volume 66, Issue 10, October 1999, Pages 888-893.
  24. Pavlov A.V. Inductive concepts formation by algebra of fourier-dual operations // Int. conf. "Integrated models and soft computing in artificial intelligence", Proceedings, 2015, V.1, p.98 – 115.
  25. Kosko B. Bi-directional associative memories. IEEE Transactions on Systems, Man and Cybernetics. 1987. V.18. №1. P. 49-60.
  26. Pavlov A.V. Detection of common fragments in a series of images by superimposed volume Fourier holograms // Quantum Electronics? 2017. V.47. #7. P.335 – 342.
  27. Subnikov E.I. Signal to noise ratio under images correlation comparison // Optics and Spectroscopy. 1987. V.62. i.2. P .450 – 456.
  28. Kuleshov A.M., Subnikov E.I. Dependence of holographic correlator signal on the medium nonlinearity and filter limitation // Optics and Spectroscopy. 1986. V.60. i.3. P .606 – 609.

**Pavlov Alexander V.** Docent of the ITMO University. Graduated from the Leningrad Institute of Fine Mechanics and Optics in 1980. Dr. Sc in Physics and Mathematics. Has more than 100 publications in indexed journals. Area of interests: Artificial Intelligence, Cognitive Systems, Information Processing, Pattern Recognition, Holography. E-mail: Pavlov@phoi.ifmo.ru