

УДК 004.387; 535.417

ВЛИЯНИЕ ХАРАКТЕРИСТИК ПАТТЕРНА ЗАДАЧИ НА АДЕКВАТНОСТЬ ГИПОТЕЗ, ПОРОЖДАЕМЫХ НЕЙРОННОЙ СЕТЬЮ СО СВЯЗЯМИ ПО СХЕМЕ ГОЛОГРАФИИ ФУРЬЕ

З.С. Бекашева, В.Н. Васильев, А.А. Востриков, А.В. Павлов

В развитие подхода к реализации механизмов творческого мышления на нейросетях со связями, реализуемыми методом голографии Фурье, исследовано влияние размера паттерна внутренней репрезентации условий задачи в связи с его спектральными характеристиками на адекватность порождаемых сетью гипотез. Показано, что адекватность порождаемых гипотез возрастает при увеличении обобщенной частоты паттерна.

Ключевые слова: голография, нейронная сеть, искусственный интеллект, когнитивная система.

Введение

Одна из актуальных задач искусственного интеллекта – реализация механизмов творческого мышления [1]. По мнению ряда исследователей, творческие способности зависят в значительной степени от двух факторов – развитости образного (правополушарного) мышления [1–5] и способности к «погружению в хаос», т.е. переходу к хаотической динамике нейронной активности [6–8].

Один из возможных подходов к реализации образного мышления в искусственных интеллектуальных системах основан на том, что мозг, являясь нейронной сетью (НС), обрабатывает картины нейронной активности, формирующиеся в коре мозга при восприятии сенсорами информации из внешнего мира. Эти картины нейронной активности суть паттерны внутренней репрезентации (ПВР) воспринимаемой информации. В работах [9, 10] был сформулирован подход, согласно которому процесс решения задач может рассматриваться как процедура обработки ПВР, а именно:

- условия встреченной задачи представляются в нейронной сети системой сенсоров и сенсорных трактов в виде ПВР воспринимаемой информации – ПВР условий задачи;
- решение задачи – это ПВР ответа, формируемые в результате обработки сетью ПВР задачи на основе имеющихся знаний, представленных межнейронными связями.

Подход [9, 10] основан на результатах, полученных в когнитивной науке, согласно которым жизнедеятельность организма направлена на восстановление целостности внутренней картины мира, нарушаемой поступающей извне информацией [7]. Соответственно, воспринимаемая информация является субъективно проблемной, т.е. ПВР такой информации рассматривается как ПВР задачи, которую надо решить, если он не соответствует имеющимся знаниям – воспринимается как искаженный или дефектный вариант «правильного» или эталонного образа. Тогда процесс решения задачи представляет собой достройку или восстановление субъективно целостного образа, а ответ задачи – достроенный или исправленный ПВР.

В этих рамках творческая задача (ТЗ) отличается от нетворческой тем, что ПВР ответа не может быть сформирован НС посредством обработки ПВР условий задачи на основе имеющихся знаний методами какого-либо из формальных выводов (дедуктивного, абдуктивного или индуктивного). В этом случае один из возможных методов решения ТЗ предполагает генерацию большого числа образов, несколько отличающихся друг от друга, но принадлежащих одной области фазового пространства сети [6]. Эти образы, не связанные с ПВР условиями задачи и знаниями правил формального вывода, суть творческие гипотезы – варианты ответов на встреченную задачу. «Погружение в хаос», т.е. переход к хаотической динамике нейронной активности, как раз и является методом генерации большого числа новых образов, каждый из которых представляет собой ПВР гипотезы.

В работах [9, 10] обоснована реализуемость этого подхода на НС с двунаправленными связями, формируемыми методом голографии Фурье (рис. 1).

Напомним кратко суть этого подхода.

1. ПВР воспринимаемой информации $Im_{Per}(x,y)$ (условия задачи) формируется в слое репрезентаций R , знания (голограмма Фурье эталонного образа) $Im_R(x,y)$ реализуются в виде матрицы связей нейронных слоев H , а в слое корреляций C формируется функция взаимной корреляции условий задачи и имеющихся знаний.
2. Мерой «проблемности» воспринимаемой информации служит мера сходства ПВР $Im_{Per}(x,y)$ с имеющимися знаниями $Im_R(x,y)$, в качестве которой используется амплитуда глобального максимума корреляционной функции $Im_{Per}(x,y) \otimes Im_R(x,y)$ (ГМКФ). Если ПВР условий задачи не соответствует имеющимся знаниям, то ГМКФ либо отсутствует, либо его амплитуда не превосходит амплитуды боковых максимумов. Ни вспомнить правильный ПВР, ни вывести его не удастся. ПВР ответа можно только придумать (увидеть во сне, например) – это творческая задача. Для того чтобы «придумать»

ПВР ответа, НС переходит в режим хаотической активности, в котором в слое R на каждой итерации $R \rightarrow C \rightarrow R$ генерируется новый образ – творческая гипотеза.

- Последний этап – верификация порожденных гипотез по критерию адекватности реальности для выбора только одной гипотезы – решения задачи.

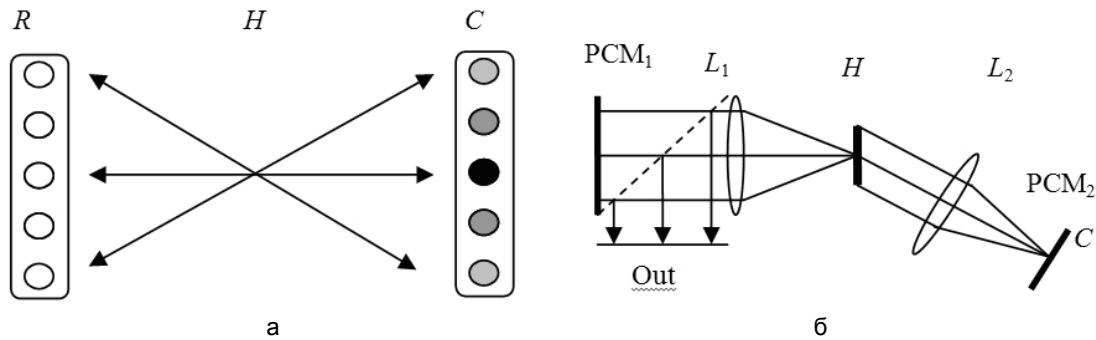


Рис. 1. Схема двухслойной нейронной сети с матрицей двунаправленных связей (а) и ее реализация методом голографии Фурье (б): R – слой репрезентаций; C – слой корреляций; L_1 и L_2 – Фурье-преобразующие линзы с фокусными расстояниями f ; H – матрица межнейронных связей (голограмма); PCM_1 и PCM_2 – фазосопрягающие зеркала во входной и корреляционной плоскостях

В [11] выделено два аспекта проблемы адекватности:

- адекватность ПВР условий задачи реальности, обусловленная свойствами сенсора и сенсорного тракта;
- адекватность ПВР гипотезы условиям задачи, обусловленная свойствами НС.

В работе [11] рассмотрен второй аспект проблемы адекватности в части влияния фильтрации, возникающей в схеме рис. 1, б, вследствие ограниченности динамического диапазона регистрирующих сред, на статистические характеристики гипотезы. В настоящей работе, в развитие [9–11], исследовано влияние ограниченности области существования (размера) ПВР условий задачи на характеристики ПВР гипотезы.

Механизм формирования гипотез

Свойства ПВР гипотез, порождаемых в режиме хаотической динамики, определяются тем, что в НС (рис. 1) связи нейронных слоев R и C , представляющие имеющиеся знания, формируется в виде голограммы Фурье

$$H_{R \rightarrow C}(v_x, v_y) = F^*(\text{Im}_R(x, y)),$$

где (v_x, v_y) – пространственные частоты; астериск (*) обозначает комплексное сопряжение; F – символ преобразования Фурье. Голограмма Фурье удовлетворяет определению знаний [2]. При предъявлении в слое R ПВР условий задачи $\text{Im}_{Per}(x, y)$ в слое C формируется функция корреляции $\text{Im}_{Per}(x, y)$ и эталонного ПВР (знаний) $\text{Im}_R(x, y)$

$$C(\zeta, \eta) = \text{Im}_{Per}(x, y) \otimes \text{Im}_R(x, y) = F(F(\text{Im}_{Per}(x, y))F^*(\text{Im}_R(x, y))),$$

где ζ, η – координаты в слое C . На этапе обратного прохождения возбуждения $C \rightarrow H \rightarrow R$ используется матрица связей, реализуемая посредством инверсной голограммы

$$H^{-1}(v_x, v_y) = \left(\frac{1}{F^*(\text{Im}_R(x, y))} \right),$$

тогда, в предположении линейности активационных функций нейронов в слое C (C -нейронов), картина нейронной активности в слое R в области вокруг исходного ПВР $\text{Im}_R(x, y)$, восстановленная в итерации $R \rightarrow H \rightarrow C \rightarrow H^{-1} \rightarrow R$, описывается выражением

$$\text{Im}(x_k, y_l) = F \left(\frac{F(C(x_k + \zeta, y_l + \eta))}{F^*(\text{Im}_R(x, y))} \right), \tag{1}$$

где (x_k, y_l) – координаты в плоскости R вне области существования $\text{Im}_R(x, y)$ (т.е. в области ореола). Выражение (1) может быть приведено [12, 13] к развитой в теории случайных процессов модели линейного предсказателя [14, 15]. Если ПВР $\text{Im}(x, y)$ описывается как реализация однородного случайного поля, то (1) дает наилучшую по критерию минимума среднего квадрата ошибки оценку линейного прогноза.

Для перехода к режиму хаотической динамики по сценарию Фейгенбаума активационные функции C -нейронов должны иметь соответствующую нелинейность NI [16], тогда для итерации N выражение (1) примет вид

$$\text{Im}_N(x_k, y_i) = F \left(\frac{F(Nl(C_{N-1}(x_k + \zeta, y_i + \eta)))}{F^*(\text{Im}_R(x, y))} \right).$$

Это выражение описывает ПВР гипотезы, порождаемой сетью на N -ой итерации.

Ограниченность области существования ПВР $\text{Im}_R(x, y)$ влияет на характеристики автокорреляционной функции (АКФ) и, тем самым, на характеристики ПВР порождаемой гипотезы.

Для упрощения дальнейших выкладок примем допущение о разделимости переменных при описании ПВР и соответственно используем термины «процесс» вместо «поле» и «стационарность» вместо «однородность».

Влияние ограниченности паттерна условий задачи на адекватность гипотезы

Для большей наглядности начнем анализ с предположения, что ПВР задачи представляет собой гармонику ограниченной длины. Тогда в слое R (без учета нелинейности в слое C) одна гармоника при использовании корреляционной функции (КФ) бесконечного процесса (т.е. теоретической КФ) экстраполируется линейным предсказателем с нулевой ошибкой.

Однако, поскольку корреляционная функция строится на основе ограниченной реализации (измеренная КФ) и, следовательно, КФ имеет линейно зависимый от длины реализации дополнительный спад к краям, то возникает ошибка экстраполяции.

Если предсказание на максимальную глубину x_{\max} строится по реализации длиной L , которую будем называть базой предсказания, то для реализации модели [14, 15] необходима АКФ, имеющая размер области определения (длину) $2(L+x_{\max})$.

Предсказание по измеренной АКФ длиной $2(L+x_{\max})$:

$$\begin{aligned} \text{Im}(x) &= \frac{1}{L} \int_x^{L+x} A^3 \cos(\omega_0 \zeta) \left(1 - \frac{\zeta}{L+x_{\max}}\right) \sin(\omega_0(L+x-\zeta)) d\zeta + \\ &+ \frac{1}{L} \int_x^{L+x} A^3 \frac{\sin(\omega_0 \zeta) - \sin(\omega_0(2L-\zeta))}{2\omega_0(L+x_{\max})} \sin(\omega_0(L+x-\zeta)) d\zeta = \\ &= \left(A^3 \sin(\omega_0(L+x)) \left(\frac{1}{2} - \frac{((L+x)^2 - x^2)}{4L(L+x_{\max})} \right) \right) + \frac{1}{L(L+x_{\max})} \frac{A^3 \cos(\omega_0(x+L))(-x+x_{\max}-L)}{4\omega_0} + \\ &+ \frac{1}{L(L+x_{\max})} \left(\frac{A^3 \cos(\omega_0(x-L))(x-x_{\max}-2L)}{4\omega_0} - A^3 \frac{\sin(\omega_0(x-L)) + \sin(\omega_0(3L-x))}{8\omega_0^2} \right). \end{aligned}$$

Первые два слагаемых описывают спад предсказания к краям, остальные слагаемые имеют меньшую амплитуду и характеризуют набег фазы, появляющийся также вследствие ограниченности реализации. Их амплитуды малы по сравнению с двумя первыми слагаемыми и равны $\frac{1}{4\pi K}$, $\frac{1}{2\pi K}$ и $\frac{1}{4\pi^2 K^2}$

соответственно, где $K = \frac{L}{T_0}$ – отношение длины паттерна к периоду гармоники (так называемая норми-

рованная частота). При достаточно большом K , в зависимости от требуемой точности, этими слагаемыми можно пренебречь (рис. 2). Ошибка в таком случае зависит от глубины предсказания и от длины реализации. Таким образом, при большой величине нормированной частоты K ошибка предсказания одной гармоники описывается выражением

$$\text{Err}(x) = \left(\sin(\omega_0 x) - \frac{\sin(\omega_0(L+x))}{4(L+x_{\max})L} \left(L^2 + 2L(x_{\max}-x) \right) \right).$$

Для компенсации этой ошибки при большой величине нормированной частоты можно ввести в корреляционной плоскости фильтр с пропусканием, обратным спаду, который описывается следующей функцией (рис. 3):

$$F_{\text{fil}}(\zeta) = \begin{cases} 1, \zeta \geq (L+x_{\max}); \\ \frac{1}{1 - \frac{\zeta}{L+x_{\max}}}, \zeta < (L+x_{\max}). \end{cases}$$

В результате такой фильтрации АКФ на интервале $(L+x_{\max})$ становится пропорциональной теоретической АКФ. В этом случае ошибка предсказания будет состоять только из слагаемых, описывающих набег фаз.

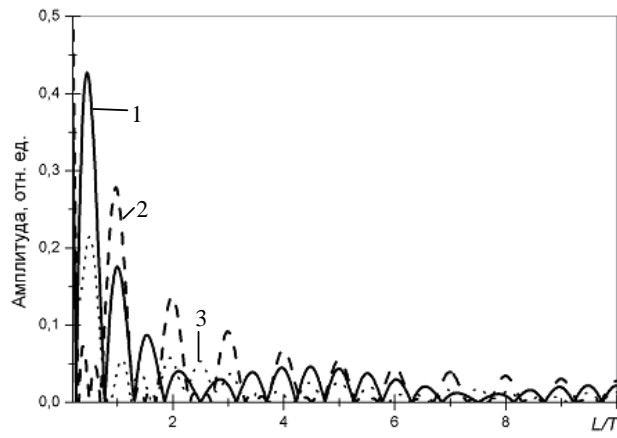


Рис. 2. Зависимости амплитуды слагаемых, характеризующих набег фазы, от нормированной частоты при различной глубине предсказания $\frac{\zeta}{L}$ и максимальной глубине предсказания $\frac{\zeta_{\max}}{L}$

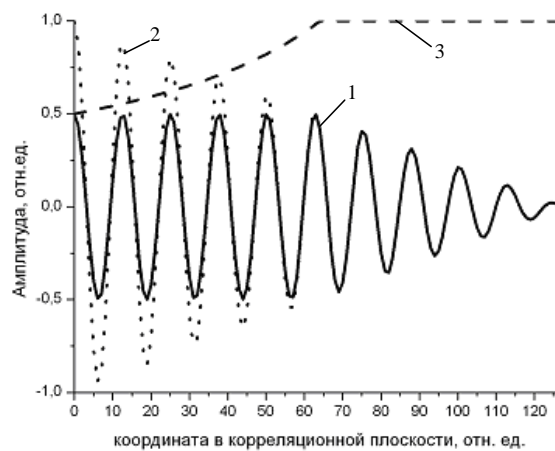


Рис. 3. Автокорреляционные функции и фильтр в корреляционной плоскости: 1 – АКФ до фильтрации; 2 – фильтр, обратно пропорциональный спаду АКФ; 3 – АКФ после фильтрации

Для ПВР со сложным спектральным составом казалось бы логичным просуммировать результаты экстраполяции отдельных гармоник. Однако ограниченность реализации обуславливает появление кросс-корреляции между гармониками, что ведет к уменьшению радиуса корреляции. Этот фактор не учитывается при суммировании результатов экстраполяции отдельных гармоник. Рассмотрим реализацию процесса, состоящего из двух гармоник. АКФ такой ограниченной реализации описывается следующим выражением:

$$\begin{aligned}
 C(\zeta) &= \frac{1}{L} \int_0^{L-\zeta} \left(\sum_{i=0}^n (A_i \sin(\omega_i x)) \right) \left(\sum_{i=0}^n (A_i \sin(\omega_i (x + \zeta))) \right) dx = \\
 &= \frac{\sum_{i=0}^n \left(A_i^2 \cos(\omega_i \zeta) \left(1 - \frac{\zeta}{L} \right) - \frac{A_i^2 (\sin(\omega_i (2L - \zeta)) - \sin(\omega_i \zeta))}{2\omega_i} \right)}{2} + \\
 &+ \sum_{i \neq j}^n \frac{2A_i A_j \cos\left(\zeta \frac{(\omega_i + \omega_j)}{2}\right) \left(\sin\left(L(\omega_j - \omega_i) + \zeta \frac{(\omega_j - \omega_i)}{2}\right) - \sin\left(\zeta \frac{(\omega_j - \omega_i)}{2}\right) \right)}{2L(\omega_j - \omega_i)} \\
 &- \sum_{i \neq j}^n \frac{2A_i A_j \cos\left(\zeta \frac{(\omega_i - \omega_j)}{2}\right) \left(\sin\left(L(\omega_j + \omega_i) + \zeta \frac{(\omega_j + \omega_i)}{2}\right) - \sin\left(\zeta \frac{(\omega_j + \omega_i)}{2}\right) \right)}{2L(\omega_j + \omega_i)}.
 \end{aligned}$$

Первое слагаемое – это сумма АКФ каждой из гармоник, а два последних – сумма их кросс-корреляций. Чем больше разность этих частот, тем меньше их кросс-корреляция и, следовательно, тем ближе АКФ реализации к сумме АКФ гармоник. Если же частоты кратные ($\omega_1=N\omega_0$, где N – целое), то кросс-корреляция равна нулю, и АКФ суммы гармоник равна сумме АКФ каждой гармоники.

Таким образом, для ПВР условий задачи с дискретным спектром, частоты которого далеко отстоят друг от друга, ПВР гипотезы можно построить с достаточно малой ошибкой. Для иллюстрации на рис. 4 представлена зависимость ошибки экстраполяции, которая возникает вследствие кросс-корреляции двух гармоник, от координаты в корреляционной плоскости для трех случаев соотношения частот гармоник: 1) $\omega_1=5, \omega_2=3$; 2) $\omega_1=16, \omega_2=3$; 3) $\omega_1=6, \omega_2=3$.

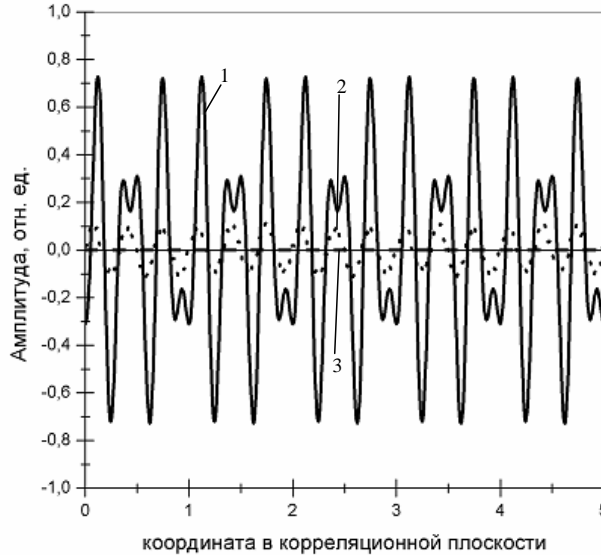


Рис. 4. Кросс-корреляционные слагаемые корреляционной функции паттерна из двух гармоник при разных частотах гармоник

В [11] было показано, что при малых L вследствие малых значений обобщенной частоты (отношения длины реализации к радиусу корреляции) статистические характеристики паттерна не соответствуют требованиям модели линейного предсказателя на стационарность реализации, и поэтому далее случай малых значений L и малых нормированных частот $K = \frac{L}{T_0}$ рассматриваться не будет. Тогда АКФ (2) бу-

дет состоять из слагаемых, которые описывают АКФ каждой гармоники. ПВР построенной гипотезы, соответственно, можно приближенно описать выражением, в котором, учитывая громоздкость точного выражения, авторы работы, основываясь на результатах, приведенных на рис. 2, пренебрегли членами высшего порядка малости, начиная с членов с множителями $\frac{1}{L}$ перед амплитудой:

$$\begin{aligned} \text{Im}(x) &= \frac{1}{L} \int_x^{L+x} \frac{A_1^2 \cos(\omega_0 \zeta) ((L+x_{\max} - \zeta))}{2(L+x_{\max})} (A_1 \sin(\omega_0(L+x-\zeta)) + A_2 \sin(\omega_1(L+x-\zeta))) d\zeta + \\ &+ \frac{1}{L} \int_x^{L+x} \frac{A_2^2 \cos(\omega_1 \zeta) ((L+x_{\max} - \zeta))}{2(L+x_{\max})} (A_1 \sin(\omega_0(L+x-\zeta)) + A_2 \sin(\omega_1(L+x-\zeta))) d\zeta \approx \\ &\approx \frac{1}{8} \left(A_1^3 \sin(\omega_0(L+x)) \left(2 - \frac{(L+x)^2 - x^2}{(L+x_{\max})L} \right) + A_2^3 \sin(\omega_1(L+x)) \left(2 - \frac{(L+x)^2 - x^2}{(L+x_{\max})L} \right) \right). \end{aligned} \quad (3)$$

Здесь первые два слагаемых описывают идеальное, в соответствии с моделью линейного предсказателя, построение ПВР гипотезы, а два последних – ошибку, обусловленную спадом АКФ. Заметим также, что в формуле отсутствуют слагаемые кросс-корреляции гармоник, поскольку при больших значениях нормированной частоты (больших значениях L) кросс-корреляция (2) между гармониками незначительна.

Из сравнения выражения (3) с выражением, даваемым моделью линейного предсказателя, следует, что ошибка построения ПВР гипотезы из двух гармоник описывается следующим выражением, также приближенным в предположении достаточно больших L :

$$Err(x) \approx \frac{1}{8} \left(A_1^3 \sin(\omega_0(L+x)) \left(6 + \frac{(L+x)^2 - x^2}{(L+x_{\max})L} \right) + A_2^3 \sin(\omega_1(L+x)) \left(6 + \frac{(L+x)^2 - x^2}{(L+x_{\max})L} \right) \right).$$

В данное выражение входят слагаемые, описывающие спад к краям амплитуд ПВР гипотезы по мере увеличения глубины предсказания, зависящий от размера ПВР задачи. Ошибка имеет вид суммы двух гармоник с амплитудой, возрастающей по мере увеличения глубины предсказания. Следовательно, ограниченность ПВР условий задачи обуславливает появление ошибки ПВР гипотезы относительно ПВР, даваемого моделью линейного предсказателя, которая уменьшается с ростом нормированной частоты, т.е. с увеличением размера ПВР условий задачи.

Поскольку вследствие ограниченности полосы пропускания тракта ввода и динамического диапазона регистрирующих сред спектр ПВР задачи всегда занимает ограниченную спектральную область, то в первом приближении можно ограничиться рассмотрением только нижней и верхней граничных нормированных частот, от значений которых и зависит точность экстраполяции. Если граничные частоты в силу плавного спада спектра определить затруднительно, то более точную оценку позволяет получить использование не нормированных граничных частот, а обобщенной частоты.

Заключение

Таким образом, на адекватность условию задачи гипотезы, порождаемой нейронной сетью со связями по схеме голографии Фурье, влияет ограниченность размера области существования паттерна условий задачи. Этот фактор влечет дополнительный спад и набег фазы автокорреляционной функции паттерна относительно теоретической автокорреляционной функции случайного поля, используемого для описания паттерна. Это, в свою очередь, ведет к ошибке формирования гипотезы относительно даваемого моделью линейного предсказателя. Эта ошибка уменьшается с ростом обобщенной частоты паттерна условий задачи. Данная зависимость универсальна в том смысле, что малая длина ПВР условий задачи обуславливает ошибку не только вследствие спада АКФ, но и, как было показано в [11], вследствие изменения стационарности исходного ПВР.

Работа выполнена при поддержке РФФИ (гранты 09-01-00165-а и 09-02-00223-а) и ФЦП № 02.740.11.0390.

Литература

1. Фоминых И.Б. О технологии решения творческих задач // Сб. трудов VIII Национальной конференции по искусственному интеллекту «КИИ-2002». – М.: Физматлит, 2002. – Т. 1. – С. 519–525.
2. Кузнецов О.П. Неклассические парадигмы в ИИ // Известия АН. Сер. Теория и системы управления. – 1995. – № 5. – С. 3–23.
3. Голицын Г.А., Фоминых И.Б. Нейронные сети и экспертные системы: перспективы интеграции // Новости искусственного интеллекта. – 1996. – № 4. – С. 121–145.
4. Семинар «Отражение образного мышления и интуиции специалиста в системах искусственного интеллекта» // Новости искусственного интеллекта. – 1998. – № 1. – С. 22–136; 1998. – № 3. – С. 64–100.
5. Леутин В.П., Николаева Е.И. Функциональная асимметрия мозга. Мифы и реальность. – СПб: Речь, 2005. – 368 с.
6. Фриман У.Дж. Динамика мозга в восприятии и сознании: творческая роль хаоса // В сб. «Синергетика и психология». Вып. 3. Когнитивные процессы. – М.: Когито-Центр, 2004. – С. 13–28.
7. Князева Е.Н. Методы нелинейной динамики в когнитивной науке // В сб. «Синергетика и психология». Вып. 3. Когнитивные процессы. – М.: Когито-Центр, 2004. – С. 29–48.
8. Комбс А. Сознание: Хаотическое и странно-аттракторное // В сб. «Синергетика и психология». Вып. 3. Когнитивные процессы. – М.: Когито-Центр, 2004. – С. 49–60.
9. Павлов А.В. О применимости голографических технологий в задаче моделирования творческого мышления // Интегрированные модели и мягкие вычисления в искусственном интеллекте: Сборник трудов IV Международной научно-практической конференции, Коломна, 28–30 мая 2007. – М.: Физматлит, 2007. – С. 282–290.
10. Павлов А.В. О возможности моделирования творческого мышления методом голографии Фурье // Труды XI Национальной конференции по искусственному интеллекту. – М.: URSS, 2008. – Т. 1. – С. 384–392.
11. Бекашева З.С., Васильев В.Н., Павлов А.В. Статистические свойства гипотез, порождаемых нейронной сетью со связями по схеме голографии Фурье // Научно-технический вестник СПбГУ ИТМО. – 2010. – № 6(70). – С. 17–21.
12. Павлов А.В. Возможности ассоциативной обработки информации, реализуемые методом Фурье-голографии // Новости искусственного интеллекта. – 2006. – № 2. – С. 41–56.
13. Павлов А.В. О применимости модели линейной регрессии к описанию Фурье-голографии // Оптика и спектроскопия. – 2005. – Т. 98. – № 6. – С. 1033–1037.

14. Колмогоров А.Н. Интерполирование и экстраполирование стационарных случайных последовательностей // Известия АН СССР. Сер. мат. – 1941. – № 5. – С. 3–14.
15. Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1975. – 320 с.
16. Кроновер Р.М. Фракталы и хаос в динамических системах. – М.: Постмаркет, 2000. – 352 с.

- Бекяшева Зоя Салаватовна** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, студент, zoya.bekyasheva@phoi.ifmo.ru
- Васильев Владимир Николаевич** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, доктор технических наук, профессор, ректор, vasilyev@mail.ifmo.ru
- Востриков Алексей Алексеевич** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, студент, lehav@inbox.ru
- Павлов Александр Владимирович** – Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кандидат технических наук, доцент, pavlov@phoi.ifmo.ru