УДК 004.054:519.68:681.51

Оценка рефлексивных связей в вероятностной логике

Андрюхин А.И.
Донецкий национальный технический университет
alexandruckin@ramber.ru

Андрюхин А.И. «Оценка рефлексивных связей в вероятностной логике». *В работе представлена система булевых уравнений, которая отражает рефлексивные представления агентов. На основании этой системы булевых уравнений построена системная модель рефлексивных связей. Рефлексивные связи рассматриваются, как обратные связи в нейроморфных сетях. Представлены основные концепты использования вероятностной логики для расчетов обратных связей в логических схемах. Рассмотрено множество базовых моделей для описания состояния последних. Представлены модели известных парадоксов. Анализ парадоксальных высказываний при применении вероятностной логики с ее более мощным алфавитом, нежели булева логика, приводит к исчезновению противоречивых свойств суждений. Приведены результаты расчетов для стационарных вероятностей рефлексивных связей.*

Ключевые слова: вероятностная логика, рефлексия, парадоксы, системная модель

Введение

Модель считают рефлексивной, если в ней отражена способность строить модели себя и других систем и одновременно видеть себя строящими такие модели.

Рефлексия всегда привлекала внимание специалистов различных научных направлений, так как по мнению многих известных специалистов, человека отличает от животного уровень его развития. Поэтому выполнялись исследования рефлексии с различных точек зрения и в различных областях применения [1 – 10].

Известна математико-психологическая модель Лефевра, использующая функцию $X1 = F(x1, F(x2, x3))$. Функция $X2 = F(x2, x3)$ интерпретируется как образ себя, имеющийся у субъекта. Первая переменная этой функции, представляет перцептивный вход, а второй переменной соответствует ментальный образ себя.

Получен известный результат, что функциональное уравнение

$$\Phi(x1, \Phi(x2, x3)) = x1 + (1 - x1)(1 - x2)x3,$$

где $x1, x2, x3$ – числа из [0,1] и все значения Φ(x, y) принадлежат [0,1], имеет единственное решение $\Phi(x, y) = 1 - y + xy = F(x, y)$ [6].

Впервые вероятностная логика (ВЛ) для построения надежных технических систем использовалась в [4]. Активна роль вероятностной логики в исследованиях искусственного интеллекта [11].

В современной компьютерной индустрии известны аппаратные реализации PCMOS «вероятностных комплементарных металлоксидных полупроводников» (Probabilistic Complementary Metal-Oxide Semiconductor) [12]. В ней определение возможности события, ранее требовавшее множества транзисторов, сводится к операции в одном или нескольких вентилях, в которых входные и выходной сигналы – вероятности (эти аспекты рассматриваются ниже).

Целью исследования является оценка возможностей синтетического направления, которое соединяет:

а) известное направление построения надежных систем из ненадежных элементов [1];

б) расширения булевой логики, в частности ВЛ и индуктивной логики;

в) исследования рефлексивных связей как обратных связей в нейроморфных системах.

Задачей данного исследования является построение системной модели отображения рефлексивных связей и их вероятностных равновесных характеристиках в известных логических парадоксах.

Описание булевых связей между агентами

Набор определенных состояний внешней среды и элементов систем в каждый момент времени назовем ситуацией. Ситуация может быть описана множеством отношений R между объектами внешней среды и элементами систем S_i. На базе информации, поступающей из внешней среды и подсистем, система S_i может построить информационную модель ситуации A, которую мы обозначим $J_i(A)$ – информационная модель ситуации A для системы S_i. Само состояние модели $J_i(A)$ описывается путем указания истинности некоторых предикатов $P_{ij}(A)$. Эти предикаты соответствуют отношениям $r_{ij} \in R_i$, где R_i – множество отношений для системы S_i, которые определяются на основании ее данных наблюдений.

Рассмотрим коллектив из N агентов различных типов, каждый из которых может быть описан набором булевых переменных $a_i=(a_{i,1}, a_{i,1},\ldots, a_{i,n}),i=1,N$ [5 – 7]. Будем понимать под состоянием ситуации А в определенные моменты времени $t_1, t_2,..., t_K$ $,k=1,K$ значения компонентов a_i. Каждый агент a_i выполняет оценку ситуации **A** указанием истинности или ложности булевозначной функции $f_i(t_k,A)$, которые принадлежат П. Считаем, что функционирование агентов представимо наиболее применяемой автоматной моделью для описания систем согласно рисунка 1.

Необходимо упомянуть, что в автоматной модели системы обычно определяют функцию выхода и функцию перехода. Так как мы считаем тип агента неизменным, то функцию перехода по состояниям можно не задавать, т.к. можно считать, что состояние единственно.

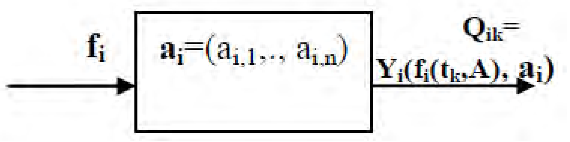

Рисунок 1. – Оценка значения воздействия $f_i(t_k,A)$ агентом a_i логическим значением Q_{ik}

Следовательно, информация, получаемая в моменты времени t_k, $k=1,K$, может быть выражена системой булевых уравнений:

$${}_i(f_i(t_k,A), a_i)\oplus Q_{ik}=0, i=1,N, k=1,K, \qquad (1)$$

где $Q_{ik} \in$ (*True*, *False*), $\oplus$(*Exor*) – операция сумма по модулю 2. Далее считаем, что значения *True* (*False*) кодируются 1(0) соответственно.

Пусть $P1(A)$ и $P2(A)$ – булевозначные функции (предикаты, булевы функции). Определим булеву функцию $T(P1,P2)$ равную 1, если $P1=P2$ и 0 в противном случае. Тогда значение функции $T(P1, P2)$ можно вычислять согласно формуле (2), следуемой из ее таблицы истинности (табл. 1):

$$T(P1,P2)=\neg(P1\oplus P2). \qquad (2)$$

Легко показать, что для булевой переменной x $T(x,0)=\bar{x}$, $T(x,1)=x$.

Важно отметить, что система булевых уравнений (1) может не иметь решений для определенного момента времени t_k или определенного множества *П*.

Примерами являются известные классические парадоксы, которые показывают ограниченность булевой логики для описания человеческого мышления, т.е. нет булевых решений для систем логических уравнений, которыми мы описываем парадоксы.

Таблица 1. – Таблица истинности $T(P1,P2)$

P1	*P2*	*T(P1,P2)*
0	0	1
0	1	0
1	0	0
1	1	1

1. Парадокс лжеца: «Я лгу».

2. Парадокс о Сократе и Платоне: Сократ: «То, что сказал Платон, есть ложь», Платон: «Сократ говорит только правду».

3. Парадокс Журдэна (эквивалент предыдущего парадокса):

«Второе предложение ложно. Первое предложение верно».

4. Парадокс Альберта Саксонского:

*Q*1: «Предложение *Q*2 ложно»;

*Q*2: «Предложение *Q*3 ложно»;

*Q*3: «Предложение *Q*1 ложно».

Для этих парадоксов нет булевых решений.

Рассмотрим парадокс лжеца. Обозначим через *Q* тип высказывающего человека.

Имеем уравнение $Q= T(Q,False)=T(Q,0)$ и получаем противоречие:

$Q= \bar{Q}$.

Рассмотрим парадокс Журдэна. Учитывая кодировку *True*(*False*) соответственно 1(0), получаем уравнения $Q1=T(Q2,0)$, $Q2=T(Q1,1)$. Отсюда $Q1= \bar{Q}2$, $Q1=Q2$. Складывая их по модулю 2, получаем $Q1\oplus Q1= \bar{Q}2\oplus Q2$ или $0=1$.

Рассмотрим парадокс Альберта Саксонского. Имеем уравнения $Q1=T(Q2,0)$, $Q2=T(Q3,0)$, $Q3=T(Q1,0)$. Согласно (2) $Q1= \bar{Q}2$, $Q2= \bar{Q}3$, $Q3= \bar{Q}1$. Складывая их по модулю 2, получаем $Q1\oplus Q2\oplus Q3= \bar{Q}1\oplus \bar{Q}2\oplus \bar{Q}3$. Прибавляя к левым и правым частям $Q1\oplus Q2\oplus Q3$, получим $0\oplus 0\oplus 0=1\oplus 1\oplus 1$ или $0=1$.

Основные концепты вероятностной логики

Основными элементами вероятностной логики являются логические связки-операции $(\neg,\wedge,\vee,\leftrightarrow,\rightarrow)$ с индексом *p*, с помощью которого мы указываем вероятностную оценку истинности определенной формулы [12 – 14]. Пример интерпретация вероятностного отрицания $\neg_p$ представлен в таблице 2.

Таблица 2. – Вероятностное отрицание

Вход *X*	Выход $\neg_p X$	
	0	1
0	*1-p*	*p*
1	*p*	*1-p*

Нужно отметить, что ВЛ не является дистрибутивной и ассоциативной.

Общие сведения о свойствах и операциях ВЛ представлены в таблице 3, для которой справедливо $r=pq+(1-p)(1-q)$.

Андрюхин А.И.
Донецкий национальный технический университет

Таблица 3. – Свойства и операции ВЛ

1	Коммутативность	$x \vee_p y \leftrightarrow y \vee_p x$ $x \wedge_p y \leftrightarrow y \wedge_p x$
2	Двойное отрицание	$\neg_q(\neg_p x) \leftrightarrow \neg_p(\neg_q x)$ $\neg_p 0 \leftrightarrow \neg_1(\neg_p 1)$ $\neg_p 1 \leftrightarrow \neg_1(\neg_p 0)$
3	Операции с 1 и 0	$(0 \wedge_p y) \leftrightarrow \neg_p 1$ $(1 \wedge_p y) \leftrightarrow \neg_1(\neg_p y)$ $(0 \vee_p y) \leftrightarrow \neg_1(\neg_p y)$ $(1 \vee_p y) \leftrightarrow (\neg_p 0)$
4	Эквивалентность	$(y \wedge_p y) \leftrightarrow \neg_1(\neg_p y)$ $(y \vee_p y) \leftrightarrow \neg_1(\neg_p y)$
5	Вероятностная тавтология	$(y \wedge_p(\neg_1 y)) \leftrightarrow \neg_p 1$ $(y \vee_p(\neg_1 y)) \leftrightarrow \neg_p 0$
6	Вероятностная формула де Моргана	$\neg_q(x \vee_p y) \leftrightarrow \neg_1 y \wedge_r \neg_1 x$ $\neg_q(x \wedge_p y) \leftrightarrow \neg_1 y \vee_r \neg_1 x$

Применение ВЛ для описания функционирования PCMOS-схем можно проиллюстрировать, анализируя основные примитивы этой технологии, из которых мы рассмотрим инвертор на рисунке 2 с соответствующей таблицей 2.

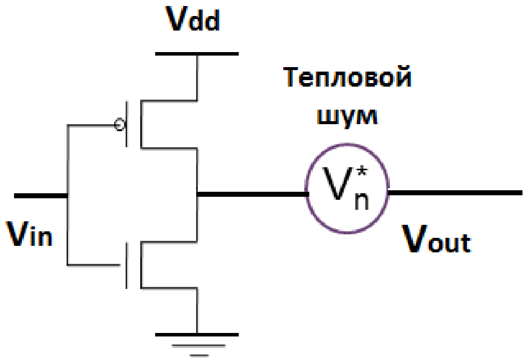

Рисунок 2. – Моп-инвертор

Согласно современным представлениям тепловой шум влияет на выходное значение инвертора согласно зависимости, представленной на рисунке 3.

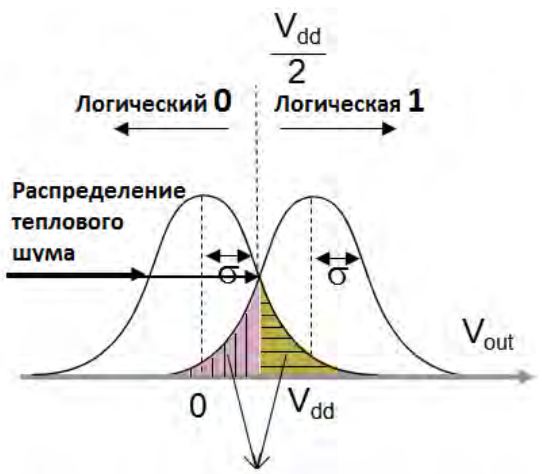

Рисунок 3. – Вероятностная модель инвертора

Вероятность правильной работы инвертора может быть определена выражением (*erf*-функция Лапласа):

$$P=(1+erf(V_{dd}/(2/\sqrt{2}\sigma))/2.$$

Для целей статьи, т.е. для обработки рефлексивных предложений, выполним упрощение рисунка 1 к виду, представленному на рисунке 4. Здесь *Pr*-вероятность правильной оценки воздействия X и соответствующего ответа Y.

Мы интерпретируем *X,Y* как вероятности. Изображенный узел является простейшим вероятностным преобразователем, который при *Pr*, блтзкому к 1, выполняет операцию *T(X, Pr)* при *X=0* или *1*. Это означает, что если *Pr* булева переменная, а *X* – действительное число (вероятность**)**, мы продолжаем функции *T(X,0)*, *T(X,1)* на вещественный отрезок (0,1) функциями *1-X* и X соответственно**.**

Подчеркнем, что *T(p,q)= r=pq+(1-p)(1-q).*

Заметим, что предлагаемый преобразователь явно ассоциируется с «предложение есть речение, в котором что-нибудь утверждается или отрицается относительно чего-нибудь другого» [15, с.73].

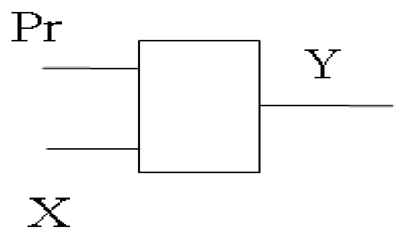

Рисунок 4. – Вероятностный преобразователь

Отметим, что в конкретной рассматриваемой ситуации (момент времени) преобразователь выполняет только одну из этих функций (отрицания или подтверждения).

Расчет для схем без обратных связей

При расчете выходных частот для логических схем мы используем методику из [16]. Для этого выполняем замену булевых выражений алгебраическими выражениями, согласно $X \vee Y=X+Y-XY$, $X \wedge Y=XY$, $\bar{X}=1-X$, $X \oplus Y= X+Y-2XY$, где *X*, *Y* в правых частях равенств являются положительными действительными переменными не больше 1. Мы интерпретируем их как частоту появления 1 в двоичной случайной последовательности. Таким образом, мы просто находим распределение конкретной функции от случайных величин с известными распределениями.

На рисунках 5, 6, 7 приведены функции распределения для действительных выражений (соответствующих булевым) и *X,Y* имеют равномерное распределение на (0,1).

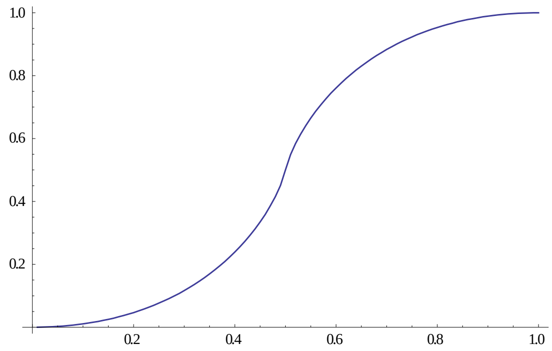

Рисунок 5. – Функция распределения X+*Y-2XY*

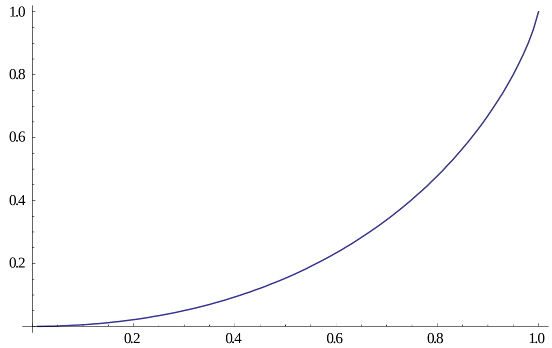

Рисунок 6. – Функция распределения *X+Y-XY*

Необходимо отметить, что для базовых элементов мы можем иногда построить аналитические выражения для функции распределения для конкретного распределения *X,Y* (равномерного, треугольного и т.п.), но для наших задач мы всегда можем построить эмпирическую функцию распределения, используя методы статистического моделирования.

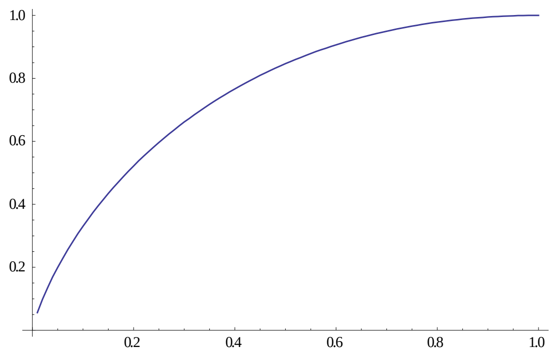

Рисунок 7. – Функция распределения *XY*

Так, кривую на рисунке 4 для *X+Y-XY (X,Y* имеют равномерное распределение) мы можем описать выражением $t-(-1+t)\,Ln(1-t)$, *где* $t \in (0,1)$.

Расчет для схем с обратными связями

При расчете значений выходных величин *Y* мы предполагаем их стационарность в установившемся режиме и, следовательно, мы решаем уравнения

$$Y=F(X,Y).$$

Решение существует, так как F есть сжимающее отображение (является суперпозицией функций соответствующих логическим операциям-связкам).

Так согласно этому предположению на рисунке 8 стационарная выходная частота *Y* равна $1/(1+X)$, *0.5 и 0.5* для примитивных операций И-НЕ, НЕ и отрицанию суммы по модулю 2 (исключающее ИЛИ) соответственно.

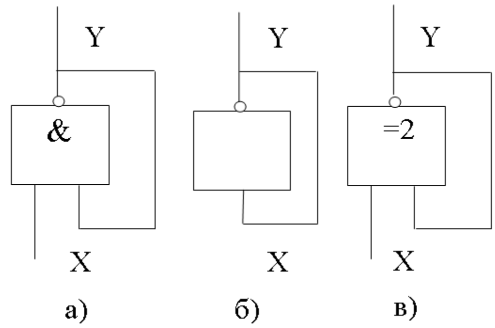

Рисунок 8. – Расчет параметров ОС для примитивов

Рассмотрим менее тривиальную схему, представленную на рисунке 9 и график функции ее выхода *Y* при различных *P,Q* на рисунке 10.

На основании вида этой функции распределения можно сделать вывод о достоверности логического утверждения *Y*.

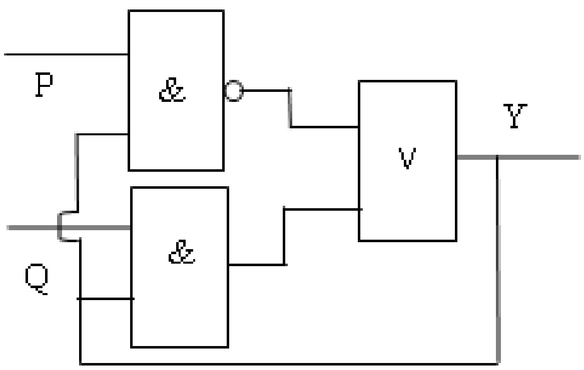

Рисунок 9. – Пример схемы с ОС

В данном примере мы можем аналитически определить равновесную частоту для выхода *Y*

$$Y=(1+P-(1+2P+P^2-4PQ)^{0.5})/2PQ.$$

Андрюхин А.И.
Донецкий национальный технический университет

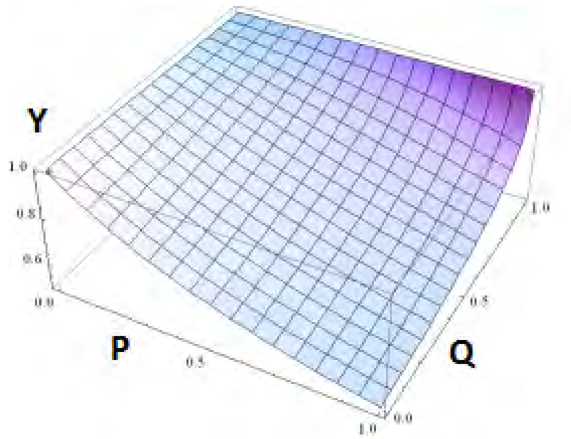

Рисунок 10. – График функция Y

Расчет параметров рефлексивных связей для языковых парадоксов

Для языковых парадоксов необходим системный анализ всей информации при рассмотрении конкретной вербальной ситуации и восстановление умалчиваемой или подразумеваемой информации.

Так, при анализе парадокса о Сократе и Платоне: Сократ: «То, что сказал Платон, есть ложь», Платон: «Сократ говорит только правду», необходимо строить модель на основании всего текста, а не рассматривать модель каждого предложения (известный системный принцип целостности).

Каждый из субъектов (Платон и Сократ) может быть представлен преобразователем моделью согласно рисунка 4, а общей вербальной ситуации соответствует рисунку 11, где визуально имеем неявную в тексте обратную связь.

Согласно модели на рисунке 12 имеем $X_1=Y_2$, $X_1=Y_2$, $X_3=Y_1$ $X_2=Y_3$ и $Y_1=P(1-X_1)$, $Y_2=R(1-X_2)$, $Y_3=Q(1-X_3)$.

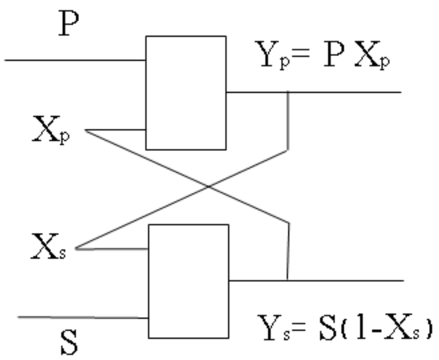

Рисунок 11. – Парадокс Платон-Сократ

Согласно рисунка 11 имеем $X_s=PY_p=PX_p$ и $X_p=SY_s=S(1-PX_p)$. Следовательно, стационарные частоты $X_p=S/(1+PS)$ и $X_s=PS/(1+PS)$.

Выполним аналогичный расчет для парадокса Альберта Саксонского (рис. 12):

*Q*1: «Предложение *Q*2 ложно»;

*Q*2: «Предложение *Q*3 ложно»;

*Q*3: «Предложение *Q*1 ложно».

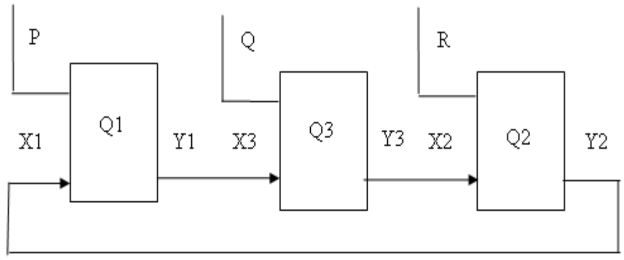

Рисунок 12. – Модель парадокса Альберта Саксонского

Отсюда определяем стационарные (равновесные) частоты:

$$Y_1=P(1-R+RQ)/(1+PRQ),$$
$$Y_2=R(1-Q+PQ)/(1+PRQ),$$
$$Y_3=Q(1-P+RP)/(1+PRQ).$$

Выводы

Научная новизна работы заключается в том, что впервые была предложена методика построения системы булевых уравнений, которые отражают рефлексивные представления агентов относительно идентификации логического состояния других агентов и себя в частности. Получены вероятностные характеристики рефлексивных связей в равновесных состояниях.

Практическая значимость результатов состоит в получении оценок состояния агентов коллектива, что позволяет более точно и эффективно прогнозировать поведение последних. Это приводит к уменьшению временных затрат при взаимодействии и является основой более эффективного сотрудничества.

Результаты работы, помимо теории мультиагентных систем и создания нейроморфных систем, могут быть использованы в технической диагностике, логической идентификации систем.

Литература

1. Тейяр де Шарден П. Феномен человека. М., 2001.
2. Д. Хофштадтер. Гедель, Эшер, Бах: эта бесконечная гирлянда. Самара, 2001. – 752 с.
3. Налимов В.В. Вероятностная модель языка. М.: Наука, 1979. – 303 с.
4. Нейман Дж. Вероятностная логика и синтез надежных организмов из ненадежных компонент // Шеннон К.Э., Маккарти Дж. Автоматы Сборник статей. – Пер. с англ. – М.: Издательство иностранной литературы, 1956. – 403 с.

5. Андрюхин А.И., Кузнецов А.В. Компьютерное исследование физических аспектов рефлексивности мышления человека // Научные труды ДГТУ. Серия: «Проблемы моделирования и автоматизации проектирования динамических систем». Вып. 29.2002 г. С. 218 – 226.
6. Lefebvre, V.A. A Psychological Theory of Bipolarity and Reflexivity. Lewiston, N.Y.: The Edwin Mellen Press. 1992.
7. Андрюхин А.И., Артеменко В.А. Рефлексивная компьютерная модель и логическая идентификация состояния коллектива агентов // Наукові праці ДонНТУ. Вип. 15 (203) Серія “Інформатика, кібернетика та обчислювальна техніка”. С. 101 – 106.
8. Лепский В.Е., Зорина Г.И. Рефлексивное предприятие XXI века // Рефлексивные процессы и управление, № 2, 2005, том 5. С. 21 – 40.
9. Wenpin Jiao. Multi-agent cooperation via reasoning about the behavior of others // Computational Intelligence, Vol. 26, Num. 1, 2010. – pp. 57 – 83.
10. J.M. Vidal.Fundamentals of Multiagent Systems with NetLogo Examples.March 1, 2010.
11. Витяев Е.Е., Перловский Л.И, Ковалерчук Б.Я., Сперанский С.О. Вероятностная динамическая логика мышления // Нейроинформатика, 2011, том 5, № 1.
12. Pinar Korkmaz, “Probabilistic CMOS (PCMOS) in the Nanoelectronics Regime”, PhD Thesis, Georgia Institute of Technology, December 2007.
13. A. Darwiche. Modeling and Reasoning with Bayesian Networks. – New York,Cambridge University Press, 2009.
14. Pearl D. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. – Morgan-Kaufmann, 1988.
15. Милль Д.С. Система логики силлогистической и индуктивной: Изложение принципов доказательства в связи с методами научного исследования. – М.: ЛЕНАНД, 2011. – 832 с.
16. Малюгин В.Д. Параллельные логические вычисления посредством арифметических полиномов. – М.: Наука, Физматлит, 1997.

Андрюхін О.І. “Оцінка рефлексивних зв'язків в ймовірнісної логіці”. *В роботі представлена система булевих рівнянь, яка відображає рефлексивні уявлення агентів. На підставі цієї системи булевих рівнянь побудована системна модель рефлексивних зв'язків. Рефлексивні зв'язку розглядаються, як зворотні зв'язки в нейроморфних мережах. Представлено основні концепти використання ймовірнісної логіки для розрахунків зворотних зв'язків у логічних схемах. Розглянуто безліч базових моделей для опису стану останніх. Представлено моделі відомих парадоксів. Аналіз парадоксальних висловлювань при застосуванні ймовірнісної логіки з її більш потужним алфавітом, ніж булева логіка, призводить до зникнення суперечливих властивостей суджень. Наведено результати розрахунків для стаціонарних ймовірностей рефлексивних зв'язків.*

Ключові слова: імовірнісна логіка, рефлексія, парадокси, системна модель

Andruckin A.I. “Evaluation of reflexive relations in probabilistic logic”. *The aim of the study is to assess the capabilities of the synthetic direction, which connects a known direction of building reliable systems from unreliable elements, expansion of Boolean logic, in particular, of the PBL and inductive logic and study of reflexive relations as feedbacks in neuromorphic systems. The objective of this study is to construct a system model of reflexive relations and their probabilistic characteristics of equilibrium in certain logical paradoxes. The article presents a system of Boolean equations, which reflects the resubmission of reflexive agents. On the basis of this system of Boolean equations built system model of reflexive relations. Reflexive communication are considered as feedback in neuromorphic networks. The basic concepts of using probabilistic logic is presented for calculation of feedbacks in logic circuits. Set of base models are considered to describe the state. Models of famous paradoxes are presented. Analysis of paradoxical statements in the application of probabilistic logic with its more powerful alphabet than Boolean logic leads to the disappearance of conflicting judgments properties. he work proposed a method for constructing a system of Boolean equations that reflect reflexive representation of agents with respect to the identification of the logical state of other agents and himself in particular. Probabilistic characteristics of reflexive relations were obtained in equilibrium. The results, in addition to the theory of the creation of multi-agent systems, and neuromorphic systems may be used in the technical logical identification and diagnosis systems. The results of calculation for the stationary probabilities of reflexive relations are presented for logical paradoxes of Plato and Socrates, Albert of Saxony.*

Keywords: probabilistic logic, reflection, paradoxes, the system model.

Статья поступила в редакцию 29.01.2014
Рекомендована к публикации канд. техн. наук А.В. Григорьевым

Андрюхин А.И.
Донецкий национальный технический университет