УДК 510.164

ФОРМАЛИЗАЦИЯ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ ЗАМКНУТЫХ СИСТЕМ

Протько М.А., студент гр.050502

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Борисенко О.Ф. – канд. физ.-мат. наук

Аннотация. В данной статье будет представлена выборка, позволяющая доказать ограничения нейронных сетей на пути к сильному интеллекту. Также изложенных в этой работе сведений будет достаточно для преобразования системы естественного языка в язык формальной логики с возможной корреляцией с семантическим смыслом. Данные свойства формальных систем позволят выбирать функции активации, целевые функции и архитектуру нейронных сетей для систем машинного обучения

Ключевые слова. Формальная система, теоремы Гёделя о неполноте, машина Гёделя.

Цель данной работы не в том, чтобы просто сформулировать правило рассуждений и проиллюстрировать технику их использования, а в том, чтобы обосновать эти правила и выяснить границу их применимости (аналогично цели [1]). Таким образом мы очертим круг задач, пока не доступный чисто формальному подходу.

Многие из последующих заключений сделаны на основании источников [1-5]. Большая часть определений взята из [1].

Под формализацией здесь и далее будем понимать процесс представления какой-либо содержательной области в виде формальной системы или исчисления.

Формальной системой назовем тройку S = (L, P, R), где L ⊆ N – разрешимое множество всех предложений системы, а P, R ⊆ L – перечислимые множества доказуемых и опровержимых предложений соответственно. Считаем, что все три множества заданы фиксированными алгоритмами (машинами Тьюринга), обозначаемыми M_L, M_P и M_R [2].

Под замкнутой системой будем понимать такую S, что любое предположение, выражаемое через предположение из L, будет также принадлежать L.

Под исследованием операций в системе S будем рассматривать определение различий между $p| = q$ (1) и $p| - q$ (2). Т.е. изучение противоречий между семантическим следованием (1) (q истинно в любой интерпретации [3], в которой истинно p) и синтаксическим отношением (2).

Основное свойство синтаксических отношений в том, что они вычислимы на машине Тьюринга [2].

Если возможно доказать, что

$$p| = q \ \leftrightarrow \ p| - q \qquad (3)$$

то можно утверждать, что формальные теории T, полученные из множества S алгоритмически разрешимы, если мы сможем с достаточной степенью точности воспроизвести свойства (1) в (2), когда p и q принадлежат S.

Под алгоритмической разрешимостью будем понимать определение, основанное на тезисе Чёрча-Тьюринга, что всякая вычислимая функция вычислима машиной Тьюринга и выразима некой частично рекурсивной функцией [4].

В данной работе частично рекурсивной функцией будем называть:

Под формальной теорией T будем понимать некое множество предложений или формул, принадлежащих языку L. В данном случае теория T состоит из всех формул, которые могут быть выведены в соответствии с некоторым фиксированным отношением следования из множества аксиом принадлежащих L.

Под языком L будем понимать множество всех предложений системы S (в данном случае, не гарантировано разрешимой), где элементами L являются некие определенные функции, аксиомы и численные (символьные) элементы, а также теории.

Под теорией T будем понимать язык L с множеством T предложений или формул, принадлежащих L. Семантическое определение T состоит из всех формул L, которые предполагаются истинными в любой интерпретации, принадлежащей некому множеству интерпретаций. Синтаксическое определение T состоит из всех формул, выводимых посредством аксиом (ранее определенная формальная теория T – синтаксическая).

Формула языка L – это выражение, являющееся утверждением. Предложение – это формула языка L, интерпретация которой не зависит от входящих в нее переменных.

Язык L является пропозициональным, когда из множества L исключаются символы ∀ и ∃, переменные, функциональные символы и все символы отношений рангов, больших 0.

В качестве доказательства (3) используется следующее:

Рассматривается характеристика синтаксическими средствами множества T формул некой теории, описанной в семантических терминах. Иными словами, необходимо найти такое множество аксиом, или же теорем, строящихся на некоторых аксиомах, которые позволяют охарактеризовать рассматриваемую формулу.

Также рассматриваемая характеристика будет удовлетворять условию, если по предположению Лоренца, аксиоматзировать данную теорию T в некоторой конструктивной метатеории T’ (T – предметная теория, строящаяся средствами L предметного языка, T’ теория, изучающая предметную теорию. L’ – метаязык, язык теории T’.), причем данная метатеория будет доказывать (3).

В результате перехода из семантического определения в синтаксическое, возможно получить следующие множества S:

(i) L – пропозициональный язык, T охватывает все общезначимые формулы L. Имеется способ проверки принадлежности любой формулы к T. T возможно дедуктивно аксиоматизировать. (используется пропозициональная логика)

(ii) T включает в себя общезначимые формулы L в полном его объеме. В L найдется функциональный символ ранга, не низшего чем 1, или же один символ отношения ранга, не низшего чем 2 (используется логика первого порядка, или же, логика предикатов)

(iii) T включает в себя все формулы L. L содержит конкретные символы =, 0,1, + и * и множество натуральных чисел. Причем каждая формула из T общезначима при каждой интерпретации некой семантической структуры. (используется арифметика)

Теория T называется полной, если для каждого предложения p из языка L либо само p либо $\overline{\mathrm{p}}$ принадлежит T.

Под рангом некого символа понимаем число операций из L, с помощью которых он образован.

Общезначимостью q определяется такое свойство q, что $\varphi 1 = 1$ для любой интрепретации φ, иными словами $| = q$, иначе, q – истина. (φ – интерпретация языка L, подробнее об ее свойствах в главе «интерпретации и структуры» в [1])

По первой теореме Гёделя (семантическое определение) (*):

Если T – перечислимая арифметическая теория. Если T непротиворечива, то T Π_1-неполна. Если к тому же T Σ_1-корректна, то в T существует недоказуемое и неопровержимое Π_1-предложение.

Где сигнатура Σ – набор символов языка L, такой, что Σ = (R, F, C, r), где R – множество символов отношений, F - множество функциональных символов, C - множество констант r – функция, сопоставляющая элементы R и F.

Π_n – множество всех истинных предложений сигнатуры Σ_n.

Теория Σ_1-корректна тогда, когда для любого символа a, принадлежащего множеству арифметических формул Γ справедливо:

$$\mathbb{N}| = \mathrm{a} \ \rightarrow \ T \ | - a \tag{4}$$

Теория T полна тогда, когда для любого предложения a из Γ справедливо:

$$T \ | - a \rightarrow \mathbb{N}| = \mathrm{a} \tag{5}$$

Формулы (4) и (5) справедливы для арифметики Пеано.

Противоречием называем отношение двух взаимоисключающих высказываний.

Из теоремы Гёделя следует:

i – разрешима и аксиоматизируема, из чего следует, что если рассматриваемая теория T описывается пропозициональным языком, то она доказывает (3), из чего следует, что семантические термины разрешимы, или, что эквивалентно, выполнимы машиной Тьюринга.

ii – аксиоматизируема и неразрешима, из чего следует опровержение (3). Такую задачу нельзя решить на машине Тьюринга.

iii – разрешима и неаксиоматизируема, что дает аналогичный ii вывод относительно (3).

Сказанное справедливо, поскольку доказательство теоремы * требует опровержения (6):

$$\mathbb{N} \ | = \varphi \rightarrow T \ | - \ \varphi \tag{6}$$

Что противоречит следствию 1 из [2] (подробнее на с.18).

Поскольку формулу (3) возможно привести к виду (6) с помощью рекурсивных функции, которые преобразуют сигнатуру Σ к сигнатуре арифметики Пеано (как это сделать подробно описано в [5], также как и синтаксическое определение теорем о неполноте, дающее, в данном случае, аналогичные выводы), можно заявить, что для систем S типа ii и iii не являются формальными, из чего следует невыразимость семантического определения формальными теоремами, из чего следует, что задача неразрешима и невыполнима на машине Тьюринга.

Заключение: формулы (1) и (2) не равносильны друг другу. Пример этому – семантическое определение арифметики, по которому ее теоремы T полны и разрешимы, а значит, аксиоматизированы. Но при переходе к синтаксическому определению, по теореме Гёделя, получаем обратный результат. (получаем парадокс лжеца, «Это предложение ложно» - нельзя доказать истинность или ложность этого утверждения. Аналогично арифметическому «2+2=4».)

Т.е., в случае, когда необходимо доказать нечто используя машину Тьюринга, делается искусственная надстройка над сигнатурой, позволяющая определить понятие истины. Чаще всего, это некоторые перечисления, свойства и теоремы, которые задаются в качестве аксиом.

С изученными областями использование машины Тьюринга не составляет проблем, но в случае добавления новой аксиомы, или же расширения сигнатуры, мы можем прийти к внутреннему противоречию в системе. Из чего следует, что большая часть вычислений окажется неверна. Предотвратить автоматический возникновение такой ситуации не представляется пока выполнимым.

Из вышеописанного можно убедиться в необходимости построения некой машины, которая способна оперировать понятиями истинности в отрыве от доказуемости и адаптировать свои алгоритмы исходя из этого понятия.

Согласно [6], такая машина – машина Гёделя. Данная машина оперирует некими изначальными аксиомами, где вход программы – некие теоремы, а выход – доказательство. Если в данной программе находится недоказуемая теорема, совершается поиск возможного доказательства, до тех пор, пока не будет доказано его отсутствие.

С использованием данной машины возможно реализовать алгоритмы шифрования (поскольку, по определению, единственное отличие шифра от кода в «понимании», или наличии семантического смысла), причем с помощью доказуемых теорем, а не заранее выбранных аксиом (которые, в некоторых случаях, могут привести к противоречию). Таким образом возможно создавать новые системы шифрования и автоматически проверять их принадлежность к множеству шифров.

Список использованных источников:

1. *Линдон Р. Заметки по логике / перевод с англ. Гостева Ю.А., ред. Яглома Н.М. – Издательство «Мир», Москва, 1968. – 126 с.*
2. *Беклемишев Л.Д. Теоремы Геделя о неполноте и границы их применимости // «Успехи математических наук» т. 65, вып. 5, 2010 – с. 62-103.*
3. *Gruber Monika Alfred Tarski and the "Concept of Truth in Formalized Languages"- a running commentary with consideration of the Polish original and the German translation.2016 – P.225.*
4. *Шень А. и др. Лекция 16. Машины Тьюринга. Тезис Чёрча-Тьюринга. // Лекции по дискретной математике, 2016, с.21*
5. *The undecidable. Basic Papers On Undecidable Propositions, Unsolvable Problems And Computable Functions / edited by Martin Davis / Raven Press, Hewlett, New York, 1965, - P. 434.*
6. *Jürgen Schmidhuber's GÖDEL MACHINE [эл. Источник]; URL: https://people.idsia.ch/~juergen/goedelmachine.html*

UDC 510.164

FORMALIZATION AND RESEARCH OF CLOSED SYSTEMS OPERATIONS

Protsko M.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Borisenko O.F. – PhD in Physics and Mathematics

Annotation. This article will present a sample that allows you to prove the limitations of neural networks on the way to strong intelligence. Also, the information presented in this work will be sufficient to transform the natural language system into the language of formal logic with a possible correlation with semantic meaning. These properties of formal systems will allow you to choose activation functions, target functions and the architecture of neural networks in systems of machine learning.

Keywords. Formal system, Godel's incompleteness theorems, Godel's machine.