

# ЧЕТВЕРИЧНАЯ ЛОГИКА КАК ЛОГИКА “УБЕЖДЕННОСТИ” ДЛЯ СИЛЬНОГО ИИ

**Сафронов Алексей Владимирович**

К.т.н., МГУ им. М.В. Ломоносова

alexey.safronov.w@gmail.com

## QUATERNARY LOGIC AS THE LOGIC OF "BELIEF" FOR STRONG AI

**A. Safronov**

*Summary:* The article considers a family of logical systems in which the law of double negation does not apply, but the law of n-dimensional negation does, depending on the n-dimensionality of logic. This approach allows us to describe three- and more-dimensional logics in a new way, and give an interpretation of "uncertainty" in these systems. In particular, it is shown that from this point of view, only logics that are multiples of two are complete: binary, quaternary, octal, etc., since the addition of each new dimension in logic doubles its "n-ness". Tables are given for the main logical operations of the logic of conviction as an extension of classical logic. It is shown that such logic operates in a two-dimensional logical space and is more flexible and intuitive, and theoretically can be used for strong artificial intelligence systems.

*Keywords:* law of double negation, quaternary logic, logic of conviction, uncertainty, strong artificial intelligence.

*Аннотация:* В статье рассматривается семейство логических систем, в которых не действует закон двойного отрицания, но действует закон n-мерного отрицания в зависимости от n-мерности логики. Такой подход позволяет по-новому описать трех- и более мерные логики, и дать интерпретацию «неопределенности» в этих системах. В частности, показано, что с этой точки зрения полными является только логики кратные двум: двоичные, четвертичные, восьмеричные и т.д., так как добавление каждого нового измерения в логике увеличивает его «n-ность» вдвое. Приводятся таблицы для основных логических операций логики убежденности как расширения классической логики. Показано, что такая логика оперирует в двухмерном логическом пространстве и является более гибкой и интуитивной, и теоретически может быть использована для систем сильного искусственного интеллекта.

*Ключевые слова:* закон двойного отрицания, четвертичная логика, логика убежденности, неопределенность, сильный искусственный интеллект.

Понятие “неопределенность”, принятое в троичной логике, весьма неудачно, так как не несёт полезного содержания. Считается, что неопределенность сама по себе ничего не значит, и должна преодолеваться за счёт дополнительных действий, достигая либо значения “истина”, либо “ложь”. В троичной или тернарной логике принято, что суждения или высказывания могут характеризоваться как истинные, ложные или неопределенные. Однако, подход к “неопределенности” выбранный в троичной логике изначально противоречив, так как он приводит к ошибке при операции отрицания. Покажем это.

В логике Клини [1], [2], например, имеет место следующий вид отрицания:

A	¬A
F	T
U	U
T	F

Здесь F – это ложь, U – неопределенность, T – правда. То есть отрицанием истины является ложь, отрицанием лжи – истина, а отрицанием неопределенности – сама же эта неопределенность. При такой форме отрицания получается, что неопределенность как бы закликивается сама на себе, тогда как истина и ложь переходят при

отрицании друг в друга. Фактически, в таком подходе использована “калька” закона двойного отрицания классической двоичной логики, где двойное отрицание тождественно утверждению. Скажем:

Не (Не (Истина)) = Истина

Не (Не (Ложь)) = Ложь

Это совершенно верно для двоичной логики, но уже в троичной логике двойное отрицание не должно быть тождественно утверждению. Таким свойством должно обладать тройное отрицание. А именно:

Не (Не (Не (Истина))) = Истина

Не (Не (Не (Ложь))) = Ложь

Не (Не (Не (Неопределенность))) = Неопределенность

Почему это необходимо? Точнее, почему такой подход к отрицанию имеет основания? Есть основания полагать, что мерность логики определяет ее предельные возможности внутреннего развития. В двоичной логике “истинное” развивается в “ложное”, а “ложное” наоборот в истинное. Цикл развития оказывается очень коротким. Однако, в троичной и более мерных логиках число звеньев цикла необходимо растет, в противном случае, дополнительные звенья теряют свою целесообразность. Покажем, что к такому пониманию отрицания был близок Гегель, но не формализовал ее для своей логики.

Анализируя Спинозу, Гегель писал: «Определенность есть отрицание» – таков абсолютный принцип спинозовской философии; этим истинным и простым взглядом обосновывается абсолютное единство субстанции. Но Спиноза не идет дальше отрицания как определенности или качества; он не переходит к познанию отрицания как абсолютного, т.е. себя отрицающего, отрицания; тем самым спинозовская субстанция сама не содержит абсолютной формы; и познание этой субстанции не есть имманентное познание» [3].

Гегель критикует Спинозу за то, что онтология последнего не подразумевает отрицания как развития познания. Его онтология как бы статична для Гегеля. В то же время, закон диалектики немецкого философа подразумевает, то двойное отрицание приводит к тождеству, но не к простому возвращению в исходную точку, а к возвращению с новым качеством. «Отрицания отрицания закон, один из основных законов диалектики, характеризующий направление процесса развития, единство поступательности и преемственности в развитии, возникновения нового и относительной повторяемости некоторых моментов старого» [4].

Это означает, что, следуя посылке Гегеля, мы могли бы получить несколько иной вывод, чем он сам. А именно мы могли бы дополнить его соображение идеей, что само «отрицание» зависит от мерности логики. Далее из изложения и рассмотрения этой идеи, однако, будет понятно, что для этого требуется, чтобы логика была полной (кратной двум). рассмотрим это.

Таблица при операции отрицания в тернарной логике может выглядеть так:

<b>A</b>	<b>¬A</b>
<b>F</b>	<b>U</b>
<b>U</b>	<b>T</b>
<b>T</b>	<b>F</b>

или так:

<b>A</b>	<b>¬A</b>
<b>F</b>	<b>T</b>
<b>U</b>	<b>U</b>
<b>T</b>	<b>F</b>

Возможны две симметричные тернарные логики. В одной из них (в последнем примере) отрицание «лжи» суть «истина», а отрицание «истины» суть неопределённость, тогда как отрицание «неопределённости» суть «ложь». Во второй происходит симметричное смещение. Впрочем, уже здесь видно, что тернарная логика является неполной, так как неопределённость не имеет своей противоположности. Так как у чего-то логически бесмысленного ее просто не может быть. Однако, покажем

это более четко, вводя неопределённость в модель естественным образом.

Рассмотрим это так. Что на самом деле означает характеристика «неопределённость»? Чтобы ответить на этот вопрос, заметим, что троичная логика является расширением двоичной, и в какой-то степени оказывается менее строгой системой суждений. В этом смысле ни одна из перечисленных характеристик суждений не должна использоваться в тернарной логике. Действительно, под «истиной» здесь следует понимать «истинно, что A», а под «ложью» - «Истинно, что не A». Соответственно, «неопределённость» тогда становится здесь «Не истинно, что A». Но здесь обнаруживает себя неполнота такой системы, так как возможно ещё «Не истинно, что не A». Именно поэтому возникает неопределённость, которая в троичной логике непреодолима.

Распишем это более подробно, и покажем, как происходит расширение от двоичной логики к троичной. Скажем, если мы в классической логике считаем, что выражение «A» истинно, то мы, на самом деле утверждаем «Истинно, что A». (Заимствуя такую подмену в модальной логике [5]). Теперь оказывается, что отрицание как будто может иметь два вида: 1) «Не истинно, что A» и 2) «Истинно, что не A». Но какой из вариантов будет соответствовать классической характеристике «Ложно, что A»? При должной осмотрительности, мы признаем, что искомым отрицанием исходной посылки могут являться оба выражения, в зависимости от выбранного нами определения, что есть отрицание. Это создаёт неопределённость или вариативность. Собственно говоря, именно это обстоятельство и дает возможность для расширения логики до троичной, и далее до четверичной.

Фактически троичная и четверичная логики могут содержать следующие характеристики как расширения классической логики:

Классическая (двоичная логика)	Троичная логика (неполная)	Четверичная логика (полная)
Истинно, что A	Истинно, что A	Истинно, что A
	Не истинно, что A	Не истинно, что A
		Не истинно, что не A
Истинно, что не A	Истинно, что не A	Истинно, что не A

Как видно из таблицы троичная логика является неполной, и в ней невозможно построить непротиворечивую интерпретацию характеристик истинности. В ней присутствуют «Истинное утверждение», «Не истинное утверждение» и «Истинное отрицание». Более полной, очевидно, является четверичная логика, которую мы назвали бы» логикой убежденности. В ней «отрицание» в общем виде имеет несколько иной смысл, чем в двоичной логике. Отрицание здесь есть переход одной границы логического пространства. Проще всего это

можно изобразить как переход границы на следующей диаграмме, но также как в примере с троичной логикой здесь необходимо выбрать направление (по часовой или против часовой стрелки).

Убедительно / А	Не убедительно / А
Убедительно / не А	Не убедительно / не А

Допустим, что было выбрано направление по часовой стрелке для отрицания (в обратную сторону - анти-отрицание). В этом случае, отрицание в такой логической модели будет означать пересечение границы по часовой стрелке. Легко заметить, что все пространство здесь разделено на четыре зоны двумя линиями (осьми) - осью "степени убежденности" и осью "степени утвердительности". Вертикальная ось, в данном случае, делит пространство на две зоны - «Убедительно» слева и «Неубедительно» справа. Горизонтальная ось делит пространство на «А» (утверждение) сверху и «Не А» (отрицание или анти-утверждение) снизу. Фактически это эквивалентно следующим характеристикам по часовой стрелке: «Убедительное утверждение», «Неубедительное утверждение», «Неубедительное отрицание» и «Убедительное отрицание».

Например, есть исходное выражение: «Если наступает утро, то встает солнце». Допустим, что это Надежное утверждение. Тогда его отрицание «Если не наступает утро, то встает солнце» будет Ненадежным утверждением. В свою очередь его отрицание «Если наступает утро, то не встает солнце» окажется Ненадежным отрицанием. А его отрицание «Если не наступает утро, то не восходит солнце» будет Надежным отрицанием. Следующее отрицание приведет к исходному выражению. Очевидно, что в такой логике само «отрицание» имеет специальное значение, отличное от классического. К счастью, в полной логике ему можно подобрать математическую интерпретацию.

В троичной логике (в силу ее неполноты) нам не удалось найти математическую интерпретацию смысла указанных характеристик, но для четверичной логики это возможно. В частности, можно было бы придать этим характеристикам следующие математические значения: "Убеждены, что А" = "1", "Не убеждены, что А" = "i", "Не убеждены, что не А" = "-i" и "Убеждены, что не А" = "-1", где "i" - мнимая единица. В этом случае логическое отрицание можно расценивать в математическом смысле как умножение на мнимую единицу. Тогда получится следующая таблица, показывающая, что четверное отрицание в данной логике равносильно одному утверждению. То есть  $\text{He}(\text{He}(\text{He}(\text{He}(A)))) = A$

A	не (A)	не (не (A))	не (не(не(A)))	не(не(не(не(A))))
1	i	-1	-i	1
i	-1	-i	1	i
-i	1	i	-1	-i
-1	-i	1	i	-1

Рассмотрим основные логические операции для такой математической интерпретации.

Конъюнкция

И	1	i	-i	-1
1	1	i	-i	-1
i	i	i	-i	-1
-i	-i	-i	-i	-1
-1	-1	-1	-1	-1

Дизъюнкция

ИЛИ	1	i	-i	-1
1	1	1	1	1
i	1	i	i	i
-i	1	i	-i	-i
-1	1	i	-i	-1

Конъюнкция и дизъюнкция в такой логике имеет классический вид минимума (Min) и максимума (Max) соответственно, однако импликация заметно отличается от классической.

Импликация

ИМП	1	i	-i	-1
1	1	i	i	i
i	1	i	-i	-1
-i	1	1	1	1
-1	1	i	-i	-i

Рассмотрим импликацию сужений «я не получил зарплату» и «я купил машину». Допустим, что «я не получил зарплату» есть Неубедительное отрицание. А «я купил машину» есть Неубедительное утверждение. Тогда выражение «если я не получил зарплату, то я купил машину» окажется Убедительным утверждением (См. таблицу). Аналогично и любое другое Неубедительное отрицание в начале импликации приводит к Убедительному утверждению в результате. Заметим, однако, что если мы выберем другое направление для отрицания (против часовой стрелки), то окажется, что Неубедительное утверждение как первый член импликации всегда приводит к Убедительному утверждению в его результате. В этом случае отрицание было бы эквивалентно умножению на мнимую единицу со знаком минус.

Подобная четверичная логика, несомненно, не может быть использована для выведения строгих логических суждений об истинности или ложности чего-либо, но она может быть базой для моделирования поведения в ситуации оценки некоей характеристики убежденности. Например, субъективной убежденности в чем-то. В этом смысле очень важно учитывать, что каждый чело-

век имеет свои представления о критериях убедительности тех или иных получаемых сведений, поэтому «логика убежденности» – это инструмент математического описания более сложных интуитивных суждений, нежели классические бинарные.

Важно отметить, что четверичная логика по сути является двухмерной. А это означает, что она допускает более сложные модели, чем нечеткая вероятностная логика, оперирующая в рамках одной оси, например, оси «истинности» (вероятности) от «0» до «1». В четверичной логике, как было отмечено выше, есть ось «убежденности» и ось «утвердительности». Суждения здесь могут быть не только «убедительными» или «неубедительными», но и «утверждающими» и «отрицающими» или «анти-утверждающими». (Поэтому здесь, в частности, открывается путь для разработки двухмерной нечеткой вероятностной логики).

Выше мы привели пример суждения «я не получил зарплату» как суждения анти-утверждающего типа, однако, отрицательная форма самого этого суждения была выбрана только для наглядности. По сути «анти-утверждением» в смысле характеристики «отрицательности» может обладать любое суждение, в том числе, и «я получил зарплату». «Отрицательность» суждения здесь означает то, что оно, например, внутренне противоречиво или обладает «интонацией» смысла. В целом введение второй логической оси связано с расширением «интуитивных» описательных возможностей для формальной логики. Например, говоря «да» или «конечно», с помощью интонации голоса человек может обогащать смысл этих суждений вплоть до «нет», «никогда», «ни в коем случае».

Интересно заметить, что четверичная логика является полной, в том смысле, что допускает естественную математическую интерпретацию, именно потому, что

она двухмерная. При этом каждое измерение задается парой полюсов «Убедительно»-«Неубедительно», «Утвердительно»-«Анти-утвердительно» (или, скажем, «Противоречиво»-«Непротиворечиво»). Классическая двоичная также логика полна и определена на оси «Истинно»-«Ложно», но, например, троичная логика не обладает полнотой. Следующей полной логикой, которая бы ещё больше приблизила математику машин к математике интуитивных логических суждений человека была бы, видимо, восьмеричная (или трехмерная) логика. Мы бы назвали ее, например, логикой «надежды» в соответствии с идеями развития форм человеческого мышления, изложенных в статье [6].

Не вызывает сомнений тот факт, что человек в отличие от машины использует более «гибкие» математические модели. Причем не только для решения тех или иных общих задач (artificial general intelligence), которые сегодня успешно решаются компьютером с помощью нейронных сетей и систем машинного обучения, но и для формирования собственной мотивации. В конечном счете, машина пока еще не может «чувствовать», так как мы не знаем о математической природе «чувствительности». В этом смысле, развитие полных (в терминах данной статьи) логик - четверичной (двухмерной), восьмеричной (трехмерной), шестнадцатеричной (четырёхмерной) и т.д., выглядит как возможный путь повышения гибкости искусственных логических систем и приближения их к реальным.

Идея двухмерной и более-мерной логики, обсуждаемая в статье, пересекается с двухаспектной теорией Девида Чалмерса или с двумерной семантикой в модальной логике [7]. Однако, обе концепции обладают слабостью, а точнее необходимостью выводить связь двух аспектов или измерений. Логика, обсуждаемая в данной статье, вводит оба «измерения» органически, поэтому их связь является здесь исходной позицией.

#### ЛИТЕРАТУРА

1. Priest, G., 'Logic of Paradox', Journal of Philosophical Logic 8 (1979), 219–241.
2. Клини С.К. Введение в метаматематику. — М., 1957.
3. Гегель Г.В.Ф. Наука логики, т. 2 Москва, 1971
4. Воробьев М., Закон отрицания отрицания, М., 1958.
5. Мочь, быть и знать. Система модальностей // Эпштейн М.Н. Философия возможного: модальности в мышлении и культуре. — СПб.: Алетейя, 2001. — 336 с.
6. Сафронов А.В. К проблеме эволюции сознания. Мульти-субстанциональная эволюция наблюдателя и иерархия модальных логик в каузальном дуализме // Медицина. Социология. Философия. Прикладные исследования. — 2021. — № 5. — С. 73–83.
7. Ball, D., 2014, "Two-Dimensionalism and the Social Character of Meaning", Erkenntnis, 79: 567–595.

© Сафронов Алексей Владимирович (alexey.safronov.w@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»