Беликов А. А. Замечания об импликации Фаррелла // Философия. Журнал Высшей школы экономики. — 2024. — Т. 8, № 2. — С. 123–141.

Александр Беликов*

Замечания об импликации Фаррелла**

Получено: 29.08.2023. **Рецензировано:** 31.10.2023. **Принято:** 21.04.2024.

Аннотация: В статье предлагается анализ трехзначной импликации, предложенной Робертом Фарреллом в работе «Material Implication, Confirmation, and Counterfactuals» (1979). Посредством предлагаемого анализа установлено, что, во-первых, импликация Фаррелла обладает всеми свойствами, достаточными, чтобы отнести ее к категории коннексивных импликаций. Во-вторых, импликация Фаррелла, будучи сходной по некоторым своим свойствам с конъюнкцией, может быть использована для формального моделирования некоторых особенностей понимания условных высказываний, присущих человеку на ранних этапах развития его когнитивных способностей. Основным техническим результатом работы является построение и мета-теоретическое исследование трехзначной логики $\mathbf{F}_3^*$, содержащей импликацию Фаррелла.

Ключевые слова: условные высказывания, соединительные высказывания, импликация, конъюнкция, коннексивная логика, трехзначная логика, ментальные модели.

DOI: 10.17323/2587-8719-2024-2-123-141.

1. ВВЕДЕНИЕ: ИМПЛИКАЦИЯ ФАРРЕЛЛА

Отправной точкой нашего исследования являются трехзначные матрицы Роберта Фаррелла, предложенные им в Farrell, 1979. Фаррелл обращается к давней проблеме адекватной экспликации семантического содержания условных высказываний, то есть таких высказываний, которые образованы при помощи союза «если..., то...» и его аналогов (например, «если воду нагреть до ста градусов по Цельсию, то она закипит»). Основной трудностью здесь является вопрос о том, какое истинностное значение мы должны приписывать условному высказыванию в тех случаях, когда его антецедент ложен. Фаррелл предлагает

*Беликов Александр Александрович, к. филос. н., старший преподаватель, Московский государственный университет имени М. В. Ломоносова (Москва); старший научный сотрудник, Институт философии Санкт-Петербургского государственного университета (Санкт-Петербург), belikov@philos.msu.ru, ORCID: 0000–0003–1395–8878.

Благодарности: Результаты, изложенные в параграфах 1, 2, 3, 4 и 6, являются частью исследования выполненного в рамках проекта № 20–18–00158–П, финансируемого Российским научным фондом и осуществляемого в Санкт-Петербургском государственном университете. Результаты, изложенные в параграфе 5, являются частью исследования выполненного в рамках проекта № 23–78–01034, финансируемого Российским научным фондом и осуществляемого в Московском государственном университете имени М. В. Ломоносова.

в этих случаях объявлять употребление условного высказывания «неподходящим» (*inappropriate*). Тем самым Фаррелл предлагает вдобавок к классическими истинностным значениям **t** (истина) и **f** (ложь) ввести третье истинностное значение **i** (неподходящий).

Пусть у нас имеется множество истинностных значений $\{\mathbf{t}, \mathbf{i}, \mathbf{f}\}$, на котором определены следующие «функции истинности»

	$f_\neg$
t	**f**
i	**i**
f	**t**

$f_\&$	**t**	**i**	**f**
t	**t**	**i**	**f**
i	**i**	**i**	**f**
f	**f**	**f**	**f**

$f_\vee$	**t**	**i**	**f**
t	**t**	**t**	**t**
i	**t**	**i**	**i**
f	**t**	**i**	**f**

$f_\to$	**t**	**i**	**f**
t	**t**	**i**	**f**
i	**i**	**i**	**f**
f	**i**	**i**	**i**

Каждая из них используется в качестве семантической модели для соответствующей логической связки стандартно определяемого пропозиционального языка $\mathbb{L}$, содержащего конъюнкцию $\&$, дизъюнкцию $\vee$, импликацию $\to$ и отрицание $\neg$[1].

Нетрудно убедиться в том, что табличные определения конъюнкции, дизъюнкции и отрицания соответствуют стандартным определениям этих связок во многих трехзначных логиках. Специфичным, как мы уже отмечали ранее, является лишь табличное определение для импликации.

В этой статье мы проведем анализ трехзначной импликации Фаррелла, оставляя за рамками его оригинальную мотивацию для этой импликации. Основная цель нашей статьи заключается в обосновании следующих двух положений:

1. импликация Фаррелла может быть отнесена к классу коннексивных импликаций;
2. импликация Фаррелла может быть использована для формального моделирования особенностей понимания условных высказываний на ранних этапах развития когнитивных способностей человека.

Будучи обоснованными, эти положения указывают на то, что трехзначная импликация Фаррелла — это довольно простая и интересная логическая связка, изучение которой объединяет сразу две актуальных тенденции в современной логике. Во-первых, использование дедуктивных теорий, содержащих импликацию Фаррелла, относятся к особой категории неклассических логик — к контр-классическим логикам, поскольку именно за счет импликации Фаррелла в таких теориях возникают принципиально новые логические законы, не имеющие места

[1] Для обозначения множества всех пропозициональных переменных этого языка будем использовать $\mathbb{P}$, а для множества всех формул будем использовать $\mathbb{F}$.

в классической логике. Таким образом, предлагаемые ниже результаты вносят вклад в исследование контр-классических логик в целом. Во-вторых, изучение импликации Фаррелла открывает дорогу к построению более комплексной формально-логической теории человеческого познания, поскольку использование импликации Фаррелла дает нам возможность привычными формальными средствами моделировать рассуждения с учетом динамики развития когнитивных способностей человека. Это, в свою очередь, имеет важные приложения к таким научным областям, как теория аргументации и логический анализ естественных рассуждений.

Обсуждение импликации Фаррелла будет происходить в контексте особой дедуктивной теории $\mathbf{F}_{\mathbf{3}}^{*}$, построение которой также является одним из важных результатов нашей работы. Именно поэтому в статье присутствует отдельный параграф, где мы предлагаем исследование мета-теоретических свойств $\mathbf{F}_{\mathbf{3}}^{*}$.

2. Импликация Фаррелла в свете семантики обобщенных истинностных значений

Для наших целей удобным будет переформулировать приведенные выше матрицы с использованием так называемых обобщенных истинностных значений. Идея обобщенных истинностных значений заключается в том, чтобы рассматривать истинностные значения в рамках той или иной логической теории не как специально созданные для конкретных приложений этой теории абстрактные объекты, а как такие объекты, которые получены в результате естественного обобщения традиционных истинностных значений «Истина» и «Ложь». Договоримся обозначать классическое значение «Истина» через 1, а «Ложь» через 0.

Напомним, как определяется семантика для классической пропозициональной логики. Понятно, что множество истинностных значений классической логики равно $\{1,0\}$. Пусть V есть функция, которая каждой пропозициональной переменной языка $\mathbb{L}$ приписывает элемент из множества истинностных значений $\{1,0\}$. Для того чтобы устанавливать истинностное значение формул, образованных с помощью логических связок, необходимо расширить функцию V в соответствии со следующими семантическими постулатами:

1. $V(\neg A)=1 \Leftrightarrow V(A)\neq 1$,
2. $V(A \mathbin{\&} B)=1 \Leftrightarrow V(A)=1$ и $V(B)=1$,
3. $V(A\vee B)=1 \Leftrightarrow V(A)=1$ или $V(B)=1$,
4. $V(A\to B)=1 \Leftrightarrow$ если $1=V(A)$, то $1=V(B)$.

Отметим, что условия ложности для сложных формул получаются по остаточному принципу, то есть как отрицание условий истинности. Например, условие ложности для импликации — это просто отрицание условия под номером 4. Давайте его выпишем:

$$V(A \to B) = 0 \Leftrightarrow V(A) = 1 \text{ и } V(B) = 0.$$

Аналогично можно сделать и для всех остальных логических связок. Их условия ложности хорошо известны, поэтому мы не будем останавливаться на этом подробно. Для того чтобы закончить построение классической логики высказываний, нужно как минимум определить отношение логического следования. Это определение тоже хорошо известно, мы его опускаем.

Итак, вернемся к вопросу о том, как можно сформулировать матрицы Фаррелла с помощью техники обобщенных истинностных значений. Рассмотрим множество всех непустых подмножеств множества $\{1, 0\}$. В результате такой операции мы получим трехэлементное множество $\{\{1\}, \{1, 0\}, \{0\}\}$. Именно элементы этого нового множества будут истинностными значениями той логики, которую мы определим ниже.

Определение 1. $\mathbf{F}_3^*$-матрица для языка $\mathbb{L}$ есть структура $\langle \mathcal{V}, \mathcal{D}, \mathcal{O} \rangle$, где

- $\mathcal{V} = \{\{1\}, \{1, 0\}, \{0\}\}$,
- $\mathcal{D} = \{\{1\}, \{1, 0\}\}$,
- для всякой n-местной логической связки $*$ в языке $\mathbb{L}$, множество $\mathcal{O}$ содержит n-местную функцию f_* с областью определения $\mathcal{V}^n$ и областью значений $\mathcal{V}$; соответствующие функции определены ниже.

x	$f_\neg$
$\{1\}$	$\{0\}$
$\{1,0\}$	$\{1,0\}$
$\{0\}$	$\{1\}$

$f_\wedge$	$\{1\}$	$\{1,0\}$	$\{0\}$
$\{1\}$	$\{1\}$	$\{1,0\}$	$\{0\}$
$\{1,0\}$	$\{1,0\}$	$\{1,0\}$	$\{0\}$
$\{0\}$	$\{0\}$	$\{0\}$	$\{0\}$

$f_\vee$	$\{1\}$	$\{1,0\}$	$\{0\}$
$\{1\}$	$\{1\}$	$\{1\}$	$\{1\}$
$\{1,0\}$	$\{1\}$	$\{1,0\}$	$\{1,0\}$
$\{0\}$	$\{1\}$	$\{1,0\}$	$\{0\}$

$f_\to$	$\{1\}$	$\{1,0\}$	$\{0\}$
$\{1\}$	$\{1\}$	$\{1,0\}$	$\{0\}$
$\{1,0\}$	$\{1,0\}$	$\{1,0\}$	$\{0\}$
$\{0\}$	$\{1,0\}$	$\{1,0\}$	$\{1,0\}$

Оценка в $\mathbf{F}_3^*$-матрице есть функция $v : \mathbb{F} \mapsto \mathcal{V}$, которая удовлетворяет следующему условию для всякой логической связки $*$ в $\mathbb{L}$ и $\phi_1, \ldots, \phi_n \in \mathbb{F}$:

$$v(*(\phi_1, \ldots, \phi_n)) = f_*(v(\phi_1), \ldots, v(\phi_n)).$$

Теперь определим отношение логического следования и понятие общезначимости в логике $\mathbf{F}_3^*$.

Определение 2. Для всяких $\Gamma \cup \phi \subseteq \mathbb{F}$, из Γ *логически следует* ϕ в $\mathbb{F}_3^*$ (символически, $\Gamma \vDash_{\mathbf{F}_3^*} \phi$), если и только если для всякой оценки v в $\mathbf{F}_3^*$-матрице верно, что если $v(\gamma) \in \mathcal{D}$, для всякой $\gamma \in \Gamma$, то $v(\phi) \in \mathcal{D}$. Формула ϕ называется *общезначимой*, если и только если $v(\phi) \in \mathcal{D}$ при любой оценке v в $\mathbf{F}_3^*$-матрице.

Замечание 1. Под логикой $\mathbf{F}_3^*$ в контексте вышеприведенных определений понимается пара $\langle \mathbb{L}, \vDash_{\mathbf{F}_3^*} \rangle$, то есть это множество всех формул языка $\mathbb{L}$, замкнутых относительно отношения логического следования из определения 2.

Табличные определения логических связок, которые изображены выше, структурно совпадают с тем, как эти же связки определяются в терминах истинностных значений Фаррелла. Единственное отличие состоит в том, что нынешние истинностные значения в прямом смысле слова состоят из классических значений. Теперь мы можем неформально прочитать значение $\{1\}$ как «истинно и не ложно», значение $\{0\}$ как «ложно и не истинно», и значение $\{1, 0\}$ как «истинно и ложно».

В такой семантике возникает два уровня интерпретации категорий «истина» и «ложь» применительно к семантическим оценкам высказываний. С одной стороны, теперь, когда мы говорим, что то или иное высказывание «истинно», мы можем иметь в виду, что оно «только истинно», то есть принимает значение $\{1\}$. Когда же мы говорим, что высказывание «ложно», то мы можем иметь в виду, что оно «только ложно», то есть принимает значение $\{0\}$. При этом остается и третья возможность — оценить высказывание как «истинное» и «ложное» одновременно. Это соответствует той ситуации, когда выказывание принимает значение $\{1, 0\}$. С другой стороны, мы можем использовать оценки «истинно» и «ложно» в более нейтральном или обобщенном смысле. Когда по отношению к какому-то высказыванию мы используем оценку «истинно», мы можем иметь в виду, что оно принимает одно из двух истинностных значений, содержащих 1, то есть какое-то значение из множества $\{\{1\}, \{1, 0\}\}$. Аналогично мы можем использовать и оценку «ложно», имея в виду, что высказывание принимает одно из двух истинностных значений, содержащих 0, то есть какое-то значение из множества $\{\{0\}, \{1, 0\}\}$. При таком обобщенном подходе мы как бы остаемся в рамках мнимой двузначности и признаем, что любое высказывание может быть истинным или ложным. Однако мы подразумеваем при этом, что истинность и ложность высказывания не являются взаимоисключающими свойствами.

Такой подход очень удобен, потому что он, во-первых, позволяет представить многозначную логику, сохранив внешнюю элегантность и простоту формальной семантики для классической логики, а во-вторых, он позволяет представить табличные определения логических связок в более привычном виде, формулируя их как условия истинности и ложности.

В частности, табличные определения функций $f_{\neg}$, $f_{\wedge}$, $f_{\vee}$ и $f_{\to}$ могут быть просто переписаны в виде следующих семантический условий.

1. $1 \in v(\neg A) \Leftrightarrow 0 \in v(A)$,
2. $0 \in v(\neg A) \Leftrightarrow 1 \in v(A)$,
3. $1 \in v(A \,\&\, B) \Leftrightarrow 1 \in v(A)$ и $1 \in v(B)$,
4. $0 \in v(A \,\&\, B) \Leftrightarrow 0 \in v(A)$ или $0 \in v(B)$,
5. $1 \in v(A \vee B) \Leftrightarrow 1 \in v(A)$ или $1 \in v(B)$,
6. $0 \in v(A \vee B) \Leftrightarrow 0 \in v(A)$ и $0 \in v(B)$,
7. $1 \in v(A \to B) \Leftrightarrow$ если $1 \in v(A)$, то $1 \in v(B)$,
8. $0 \in v(A \to B) \Leftrightarrow$ если $0 \notin v(A)$, то $0 \in v(B)$.

Используя ту интерпретацию терминов «истина» и «ложь», которую мы описали выше, эти семантические условия могут быть прочитаны естественным образом. Например, условие под номером 1 говорит о том, что формула вида $\neg A$ является истинной, если и только если формула вида A является ложной. Условие под номером 6 говорит о том, что формулы дизъюнктивного вида $A \vee B$ являются ложными тогда и только тогда, когда A ложна и B ложна. Читатель может продолжить этот перевод самостоятельно и обнаружить, что практически все условия, которые тут представлены, согласуются с тем, как эти же самые логические связки интерпретируются в классической логике. Единственным исключением является условие под номером 8 (ложность импликации), и его мы подробно обсудим в следующем параграфе, потому что именно оно позволяет отнести импликацию Фаррелла к особой категории неклассических импликаций.

3. ИМПЛИКАЦИЯ ФАРРЕЛЛА И КОННЕКСИВНОСТЬ

Цель данного параграфа состоит в демонстрации того, что импликация Фаррелла, независимо от оригинальной мотивации к ее созданию, имеет тесную связь с одним из актуальных разделов современной неклассической логики — коннексивной логикой[2].

[2]Более подробно об этом направлении см. Wansing, 2023.

Коннексивная логика — это раздел современной символической логики, направленный на изучение некоторых особенностей интерпретации условных высказываний (кондиционалов), которые не могут быть адекватно эксплицированы средствами классической логики. В этом контексте коннексивная логика тесно связана с такими направлениями, как релевантная и условная логики.

В основе коннексивной логики лежит идея о том, что связка импликации должна в том или ином виде отражать следующие вполне естественные интуитивные представления об условной связи:

- ни одно высказывание не должно быть логическим следствием собственного отрицания,
- ни одно высказывание не должно иметь в качестве логического следствия собственное отрицание,
- никакое высказывание не должно иметь в качестве логического следствия оба противоречащих друг другу высказывания.

Эти принципы могут быть формализованы по-разному, но, пожалуй, наиболее традиционный способ, предложенный впервые в работе С. МакКолла (McCall, 1966), заключается в том, чтобы «закодировать» их в виде следующих пропозициональных формул.

(Тезис Аристотеля I)	$\neg(\neg A \to A)$
(Тезис Аристотеля II)	$\neg(A \to \neg A)$
(Тезис Боэция I)	$(A \to \neg B) \to \neg(A \to B)$
(Тезис Боэция II)	$(A \to B) \to \neg(A \to \neg B)$

Стандартное на сегодняшний день определение подразумевает, что коннексивная логика — это логическая теория, законами которой являются все четыре из упомянутых выше формулы (Wansing, 2023). Соответственно, связка $\to$ в такой логике называется коннексивной импликацией. Обычно вдобавок выдвигается требование о том, чтобы $\to$ не обладала свойством симметричности, то есть чтобы формула $(A \to B) \to (B \to A)$ не была законом соответствующей теории.

Примечательно, что ни одна из этих формул не является законом классической логики. А поскольку классическая пропозициональная логика обладает свойством синтаксической полноты, ее нельзя расширить за счет простого добавления к ней тезисов Аристотеля и Боэция. Это означает, что построение коннексивной логики всегда подразумевает отказ от тех или иных законов классической логики и добавление каких-то формул, которые по умолчанию не являются классически приемлемыми.

Одним из наиболее существенных вкладов в развитие коннексивной логики является идея Х. Вансинга (Wansing, 2004), которая заключается в том, что построение коннексивной логики возможно за счет использования нестандартного условия ложности для импликации. В неформальном виде идею Х. Вансинга можно сформулировать так:

$$\text{высказывание с логической формой } A \to B \text{ является ложным,}$$
$$\text{если и только если истинность } A \text{ влечет ложность } B. \tag{W}$$

Это условие ложности не соответствует классическому пониманию ложности импликации в двух аспектах.

В предыдущем параграфе мы выписывали условие ложности для импликации в классической логике. Зафиксируем его еще раз, только уже в неформальном виде:

$$\text{высказывание с логической формой } A \to B \text{ является ложным,}$$
$$\text{если и только если } A \text{ истинно и } B \text{ ложно.} \tag{C}$$

Главное отличие между классическим подходом и подходом Х. Вансинга состоит в том, что определяющая часть в классическом условии ложности импликации по сути является конъюнктивным мета-языковым утверждением, а определяющая часть в условии Х. Вансинга является условным мета-языковым утверждением. Если в классической логике для ложности импликации достаточно и необходимо лишь одновременной истинности антецедента и ложности консеквента, то в случае с условием Х. Вансинга ложность консеквента должна возникать вследствие истинности антецедента импликации.

Эта простая и естественная идея, будучи конкретизированной в рамках той или иной формальной семантики, позволяет получить целое семейство логических теорий, которые удовлетворяют описанному выше определению коннексивной логики. К ним относятся, например, логики Х. Вансинга **C** и **MC** (Wansing, 2004; 2023), логика Х. Омори и Х. Вансинга **C3** (Omori & Wansing, 2020), логика Дж. Кэнвелла **CN** (Cantwell, 2008) и многие другие.

Второе интересное отличие заключается в том, что поместив условие Х. Вансинга в контекст формальной семантики для классической логики, мы получим логическую связку, которая просто будет эквивалентна функции истинности, соответствующей отрицанию конъюнкции. Таким образом, неверно было бы утверждать, что (W) есть некий альтернативный способ определить классическую импликацию. Действительно, нетрудно убедиться, что условие

$$V(A \to B) = 0 \Leftrightarrow \text{если } V(A) = 1, \text{ то } V(B) = 0, \qquad (W')$$

будучи эквивалентным условию

$$V(A \to B) = 0 \Leftrightarrow V(A) = 0 \text{ или } V(B) = 0, \qquad (W'')$$

приводит к тому, что формулы $A \to B$ и $\neg(A \,\&\, B)$ выражают одну и ту же функцию истинности (если они интерпретируются при помощи классической оценки V, определенной в параграфе 2). Таким образом, использование условия (W) в качестве формальной семантики для импликации в контексте классической логики высказываний оказывается безуспешным, ведь это не позволяет смоделировать связку импликации в каком бы то ни было виде.

Отличительной особенностью классической логики является то, что в ней истинность и ложность являются взаимоисключающими свойствами. Для любой формулы в классической логике верно, что она является ложной, если и только если она не является истинной, а также, что она является истинной, если и только если она не является ложной. Семантика классической логики не позволяет отличить ложность от не-истинности и истинность от не-ложности. С этой точки зрения, семантика классической логики не является в достаточной степени выразительной, чтобы адекватно эксплицировать следующую модификацию условия (W).

высказывание с логической формой $A \to B$ является ложным,
если и только если не-ложность A влечет ложность B. (F)

Действительно, переформулировав условие (F) средствами классической логики мы придем к тому же результату, что и в случае с условием (W).

$$V(A \to B) = 0 \Leftrightarrow \text{если } V(A) \neq 0, \text{ то } V(B) = 0, \qquad (W''')$$

может быть преобразовано в

$$V(A \to B) = 0 \Leftrightarrow V(A) = 0 \text{ или } V(B) = 0, \qquad (W'')$$

что влечет эквивалентность формул $A \to B$ и $\neg(A \,\&\, B)$.

Этот эффект возникает как раз в силу того, что семантика классической логики не позволяет выразить различие между истинностью и не-ложностью. Однако это различие может быть выражено в контексте семантики обобщенных истинностных значений. И здесь уместно напомнить то условие ложности, которое нам удалось выявить при формулировке импликации Фаррелла.

$$0 \in v(A \to B) \Leftrightarrow \text{если } 0 \notin v(A), \text{ то } 0 \in v(B). \qquad (\mathcal{F})$$

Учитывая неформальную интерпретацию обобщенных истинностных значений из предыдущего параграфа, нетрудно убедиться в том, что условие ложности импликации Фаррелла полностью совпадает с условием (F).

Обратим внимание на то, что $(\mathcal{F})$ также может быть эквивалентным образом переписано в виде $(\mathcal{F}')$.

$$0 \in v(A \to B) \Leftrightarrow 0 \in v(A) \text{ или } 0 \in v(B). \qquad (\mathcal{F}')$$

Однако несмотря на то, что $(\mathcal{F}')$ по смыслу совпадает со стандартным условием ложности для конъюнкции, это не приводит к эквивалентности $A \to B$ и $\neg(A \,\&\, B)$ в контексте логики $\mathbf{F_3^*}$. Другими словами, импликация Фаррелла не является «вырожденной» связкой, которая совпадает с отрицанием конъюнкции и более того, она обладает некоторым минимальным набором свойств, которые позволяют охарактеризовать ее именно как импликацию (подробнее об этом будет сказано в параграфе 5).

Учитывая связь импликации Фаррелла с подходом Х. Вансинга, возникает вопрос: в какой мере она является коннексивной импликацией? Отвечая на этот вопрос, можно с уверенностью сказать, что импликация Фаррелла является коннексивной, поскольку для нее выполняется определение, упомянутое нами в начале данного параграфа. Нетрудно проверить, что в контексте $\mathbf{F_3^*}$ общезначимыми формулами являются все тезисы Аристотеля и Боэция, а принцип симметричности для $\to$ не является общезначимым.

Однако у импликации Фаррелла есть еще одно любопытное свойство, которое отличает ее от множества других коннексивных импликаций. Дело в том, что большинство коннексивных теорий, полученных по методу Х. Вансинга, помимо уже упомянутых выше тезисов Боэция, содержат в качестве общезначимых следующие формулы.

$$\neg(A \to B) \to (A \to \neg B) \qquad \text{(HB1)}$$

$$\neg(A \to \neg B) \to (A \to B) \qquad \text{(HB2)}$$

Такие теории, вслед за Р. Силваном (Sylvan, 1989), принято называть гипер-коннексивными. Существует мнение, что общезначимость этих формул не согласуется с некоторыми интуитивными представлениями о том, что коннексивная импликация должна выражать связь по содержанию между антецедентом и консеквентом. Действительно, по мнению С. МакКолла, мы вполне можем принять истинность высказывания «неверно, что если снег белый, то трава зеленая», однако из этого вряд ли можно заключить, что «если снег белый, то трава не является зеленой», как предписывает нам одна из приведенных выше схем.

Нетрудно убедиться в том, что импликация Фаррелла успешно избегает гипер-коннексивности, поскольку ни (HB1), ни (HB2) не являются общезначимыми в $\mathbf{F}_3^*$.

Подводя итог данному параграфу, еще раз отметим, что именно условие ложности для импликации Фаррелла позволяет классифицировать ее как разновидность коннексивной импликации, которая при этом не является гиперконнексивной. Последнее свойство роднит импликацию Фаррелла с другими теориями, способными избежать гипер-коннексивности (Belikov, 2024; Belikov & Zaitsev, 2022).

4. ИМПЛИКАЦИЯ ФАРРЕЛЛА И КОГНИТИВНАЯ ПСИХОЛОГИЯ МЫШЛЕНИЯ

В предыдущем параграфе мы отмечали, что условие ложности для импликации Фаррелла совпадает со стандартным условием ложности для конъюнкции, однако в силу специфики семантики обобщенных истинностных значений, это не приводит к тому, что $A \to B$ и $\neg(A \mathbin{\&} B)$ эквивалентны в $\mathbf{F}_3^*$. Вместе с тем в $\mathbf{F}_3^*$ эквивалентными оказываются формулы $\neg(A \to B)$ и $\neg(A \mathbin{\&} B)$. Получается, что с точки зрения $\mathbf{F}_3^*$ отрицание условного высказывания попросту совпадает с отрицанием конъюнктивного высказывания.

На первый взгляд, эта особенность импликации Фаррелла действительно может показаться чем-то противоестественным. С другой стороны, если мы посмотрим на эту проблему в более общем контексте, то мы сможем найти интересную связь между импликацией Фаррелла и исследованиями условных высказываний в когнитивных науках.

В современной когнитивной психологии одним из доминирующих подходов к объяснению механизма понимания высказываний естественного языка является «теория ментальных моделей» (Johnson-Laird, 1983). Согласно этой теории, понимание человеком смыслового значения высказывания достигается за счет того, что он конструирует в своей рабочей памяти так называемые «ментальные модели», которые репрезентирую различные комбинации положений дел, описываемых в высказывании.

Авторами статей Barrouillet & Lecas, 1999 и Barrouillet & Lecas, 1998 был проведен ряд экспериментальных исследований, которые подтверждают следующие две гипотезы. Во-первых, это гипотеза о том, что процесс конструирования ментальных моделей у детей ограничен возможностями их рабочей памяти. Вторая гипотеза утверждает, что возможности рабочей памяти ребенка определяют то, как он интерпретирует (то есть оценивает истинность или ложность) условных высказываний.

Согласно Barrouillet & Lecas, 1999 и Barrouillet & Lecas, 1998, в теории принято выделять три типа ментальных моделей, которые потенциально могут быть использованы человеком для понимания условных высказываний:

1. человек конструирует одну ментальную модель, репрезентирующую одновременное наличие A и B (конъюнктивная интерпретация),
2. человек конструирует две ментальные модели: первая репрезентирует одновременное наличие A и B, а вторая — одновременное отсутствие A и B (би-кондициональная интерпретация),
3. человек конструирует три ментальные модели: первая репрезентирует одновременное наличие A и B, вторая — одновременное отсутствие A и B, а третья — отсутствие A и наличие B (кондициональная интерпретация).

В работе ibid. был проведен эксперимент, где в качестве испытуемых принимали участие школьники трех возрастных групп. Первая группа включала в себя детей, имеющих средний возраст 8.2 лет, вторая — 11.3 лет, и третья — 15 лет. Участникам демонстрировалась серия высказываний с логической формой «если A, то B» и четыре возможных комбинации положений дел, описываемых в этом высказываний: «A, B», «A, не-B», «не-A, B» и «не-A, не-B». Задание заключалось в том, что участники должны выбрать случаи, которые, по их мнению, нарушают (фальсифицируют) продемонстрированное им условное высказывание[3].

С точки зрения авторов эксперимента, выбор определенной комбинации случаев указывает на имплицитное использование человеком того или иного типа ментальной модели. В частности, если человек выбирает все случаи, кроме «A, B», то это говорит о том, что он интерпретирует условное высказывание в соответствии с конъюнктивной интерпретацией, поскольку считает, что для его фальсификации достаточно отсутствия хотя бы одного из двух положений дел, описываемых в антецеденте и консеквенте.

Результаты этого эксперимента свидетельствуют в пользу того, что переход от использования более простой (конъюнктивной интерпретации)

[3]Более подробное описание эксперимента можно найти в Barrouillet & Lecas, 1998.

к более сложной (кондициональной интерпретации) сопряжен с возрастом ребенка, а именно — зависит от объема его рабочей памяти. Было установлено, что при выполнении заданий, схема фальсификации, соответствующая конъюнктивной интерпретации, наиболее распространена в ответах самых младших участников (63% ответов), фальсификация, соответствующая би-кондициональной интерпретации, чаще всего встречается среди второй возрастной категории (38% ответов), и наконец схема фальсификации, соответствующая кондициональной интерпретации, наиболее распространена в ответах 15-летних подростков (52% ответов).

Итак, опираясь на результаты описанного выше эксперимента, можно сказать, что отождествление отрицания условных высказываний с отрицанием конъюнктивного высказывания не является чем-то противоестественным. Более того, это наблюдение открывает интересные перспективы для логики как теории рассуждений. Если мы хотим иметь более комплексный подход к построению теории дедукции, то очевидно, что импликация Фаррелла позволяет сделать шаг в совершенно новом направлении. Она дает нам возможность привычными формальными средствами моделировать рассуждения с учетом динамики развития когнитивных способностей человека. Фактически, импликация Фаррелла позволяет формализовать то понимание условных высказываний и, как следствие, рассуждений, содержащих эти высказывания, которое присуще человеку на ранних этапах развития его когнитивных способностей.

5. АКСИОМАТИЗАЦИЯ ЛОГИКИ $\mathbf{F}_3^*$

В этом параграфе мы проведем синтаксический анализ логики $\mathbf{F}_3^*$. Это позволит ответить на фундаментальные вопросы о мета-теоретических свойствах $\mathbf{F}_3^*$.

Для начала определим аксиоматическое исчисление для $\mathbf{F}_3^*$. Оно задается следующим списком аксиомных схем (запись $A \leftrightarrow B$ используется как сокращение для $(A \to B) \mathbin{\&} (B \to A)$)

$$A \to (B \to A) \tag{A1}$$

$$(A \to (B \to C)) \to ((A \to B) \to (A \to C)) \tag{A2}$$

$$((A \to B) \to A) \to A \tag{A3}$$

$$(A \to C) \to ((B \to C) \to ((A \vee B) \to C)) \tag{A4}$$

$$A \to (A \vee B) \tag{A5}$$

$$B \to (A \vee B) \tag{A6}$$

$$(A \to B) \to ((A \to C) \to (A \to (B \,\&\, C))) \tag{A7}$$

$$(A \,\&\, B) \to A \tag{A8}$$

$$(A \,\&\, B) \to B \tag{A9}$$

$$\neg\neg A \leftrightarrow A \tag{A10}$$

$$\neg(A \,\&\, B) \leftrightarrow (\neg A \vee \neg B) \tag{A11}$$

$$\neg(A \vee B) \leftrightarrow (\neg A \,\&\, \neg B) \tag{A12}$$

$$\neg(A \to B) \leftrightarrow (\neg A \vee \neg B) \tag{A13}$$

$$A \vee \neg A \tag{A14}$$

и единственным правилом вывода *modus ponens*

$$\frac{A \to B, \quad A}{B} \tag{MP}$$

Понятие вывода и доказуемой формулы в этом исчислении определяется стандартно (см. например, Mendelson, 2015). Отношение выводимости будем обозначать с помощью символа $\vdash$.

Сделаем несколько важных замечаний касаемо этого исчисления.

Поскольку позитивный фрагмент этого исчисления (а значит и чистый импликативный фрагмент тоже) является позитивным фрагментом классической логики, ясно, что для этого исчисления будет иметь место стандартная теорема о дедукции.

Теорема 1. Если $\Gamma, A \vdash B$, то $\Gamma \vdash A \to B$.

Доказательство. Общая схема доказательства этой теоремы, которая существенно опирается на использование аксиом (A1) и (A2), может быть найдена в ibid. □

Наличие среди правил вывода *modus ponens* и теоремы о дедукции в системе $\mathbf{F_3^*}$ позволяют судить о том, что импликация Фаррелла обладает неким минимальным набором свойств, по которым ее можно идентифицировать именно как импликацию, а не как какую-то другую логическую связку.

Отличительной особенностью исчисления для $\mathbf{F_3^*}$ является наличие аксиомной схемы (A13), которая и кодирует в себе идею о том, что отрицание условных высказываний эквивалентно отрицанию конъюнктивных высказываний. Нетрудно построить доказательство формулы

$$\neg(A \to B) \leftrightarrow \neg(A \,\&\, B),$$

используя схемы (A13) и (A11).

Для доказательства семантической непротиворечивости и полноты нашего исчисления, введем ряд технических определений[4].

Множество формул $\mathscr{U}$ называем *теорией*, если и только если оно замкнуто относительно отношения выводимости в исчислении $\mathbf{F}_3^*$, то есть если $\mathscr{U} \vdash A$, то $A \in \mathscr{U}$. Теория $\mathscr{U}$ называется *нетривиальной*, если и только если $\mathscr{U}$ не совпадает со множеством всех формул языка $\mathbb{L}$. Теория $\mathscr{U}$ называется *простой*, если и только если выполняется следующее условие: если $A \vee B \in \mathscr{U}$, то $A \in \mathscr{U}$ или $B \in \mathscr{U}$. Теория $\mathscr{U}$ называется *$\neg$-полной*, если и только если для любой формулы A верно, что $A \in \mathscr{U}$ или $\neg A \in \mathscr{U}$.

Нам потребуется следующая вспомогательная лемма.

Лемма 1. Для всякой простой теории $\mathscr{U}$ верно, что $A \to B \in \mathscr{U}$, если и только если $A \notin \mathscr{U}$ или $B \in \mathscr{U}$.

Доказательство. Докажем часть утверждения «слева-направо». Допустим, что $A \to B \in \mathscr{U}$, $A \in \mathscr{U}$ и $B \notin \mathscr{U}$. Применяя (MP), получаем противоречие.

Чтобы доказать утверждение «справа-налево», нужно рассмотреть два случая.

Сначала допустим, что $A \notin \mathscr{U}$ и $A \to B \notin \mathscr{U}$. Опираясь на тот факт, что в $\mathbf{F}_3^*$ доказуема формула $A \vee (A \to B)$ (используя (A3) и теорему дедукции) и то, что $\mathscr{U}$ является простой, мы снова получаем противоречие.

Теперь допустим, что $B \in \mathscr{U}$ и $A \to B \notin \mathscr{U}$. Используя схему (A1), мы снова получаем противоречие. □

Доказательство семантической полноты исчисления $\mathbf{F}_3^*$ может быть получено адаптацией стандартного доказательства по методу Хенкена (см. Wójcicki, 1984). Прежде всего, необходимо ввести понятие канонической оценки.

Определение 3. Пусть $\mathscr{U}$ есть нетривиальная простая $\neg$-полная теория. Для всякой пропозициональной переменной p языка $\mathbb{L}$, определим $\mathbf{F}_3^*$-каноническую оценку $v_{\mathscr{U}}$ следующим образом:

$$1 \in v_{\mathscr{U}}(p) \;\Leftrightarrow\; p \in \mathscr{U}, \qquad 0 \in v_{\mathscr{U}}(p) \;\Leftrightarrow\; \neg p \in \mathscr{U}.$$

[4]Для того чтобы не перегружать текст статьи техническими деталями, все последующие доказательства будут даны лишь схематично. Однако используя указанные по ходу изложения ссылки, определения и замечания, читатель сможет успешно восстановить все необходимые детали самостоятельно.

С помощью следующей леммы мы демонстрируем, что определение канонической оценки может быть расширено с множества всех пропозициональных переменных на множество всех формул языка $\mathbb{L}$.

Лемма 2. Пусть $v_{\mathscr{U}}$ есть $\mathbf{F}_3^*$-каноническая оценка. Для всякой формулы A в языке $\mathbb{L}$ верно, что

$$1 \in v_{\mathscr{U}}(A) \;\Leftrightarrow\; A \in \mathscr{U}\,, \qquad 0 \in v_{\mathscr{U}}(A) \;\Leftrightarrow\; \neg A \in \mathscr{U}\,.$$

Доказательство. Индукцией по сложности формулы A. Доказательство ведется по стандартной для многозначных логик методологии, см., например, доказательства аналогичных лемм в работах Belikov, 2022: лемма 5.5, Dunn, 2000: лемма 10. □

Следующая лемма также является модификацией стандартной леммы Линденбаума (см., например, доказательства аналогичных лемм в работах Belikov, 2022: лемма 5.3, Carnielli & Coniglio, 2016: теорема 2.2.6).

Лемма 3 (Лемма Линденбаума)**.** Если $\Gamma \nvdash A$, то существует нетривиальная $\neg$-полная простая теория Γ', такая, что $\Gamma \subseteq \Gamma'$ и $\Gamma' \nvdash A$.

Теорема 2 (Семантическая полнота)**.** Если $\Gamma \vDash_{\mathbf{F}_3^*} A$, то $\Gamma \vdash A$.

Доказательство. Методом «от противного», используя лемму 3 и лемму 2. □

Теорема 3 (Семантическая непротиворечивость)**.** Если $\Gamma \vdash A$, то $\Gamma \vDash_{\mathbf{F}_3} A$.

Доказательство. Методом индукции по длине вывода. Необходимо показать, что все аксиомные схемы сохраняют свойство общезначимости, а правило вывода *modus ponens* удовлетворяет определению семантического следования из определения 2. □

6. ЗАКЛЮЧЕНИЕ

Подводя итог нашей работе, еще раз перечислим основные полученные результаты и наметим перспективы дальнейших исследований.

Нами был предложен логико-философский анализ трехзначной импликации Фаррелла, в результате которого была выявлена тесная связь этой импликации с двумя, на первый взгляд, независимыми областями исследований.

Во-первых, мы установили, что импликация Фаррелла удовлетворяет стандартному определению коннексивной импликации, то есть, будучи помещенной в контекст логической теории $\mathbf{F}_3^*$, эта импликация позволяет нам получить общезначимость тезисов Аристотеля и Боэция, а также избежать свойствами симметричности. Мы показали, если импликацию Фаррелла представить в контексте семантики обобщенных истинностных значений, то ее условие ложности может быть воспринято, как абсолютно естественная модификация условия ложности Вансинга, используемого во многих современных теориях коннексивной импликации.

Во-вторых, нами было показано, что импликация Фаррелла может быть использована для формального моделирования некоторых особенностей понимания условных высказываний, присущих человеку на ранних этапах развития его когнитивных способностей. Привлекая результаты экспериментальных исследований проблемы интерпретации детьми и подростками условных высказываний, нам удалось установить, что отождествление отрицания условных высказываний с отрицанием соединительных высказываний, распространенное среди детей младшего возраста, может быть смоделировано с помощью импликации Фаррелла, поскольку формулы $\neg(A \to B)$ и $\neg(A \,\&\, B)$ логически эквиваленты в $\mathbf{F}_3^*$.

Наконец, нами было предложено аксиоматическое исчисление, которое адекватно формализует логику $\mathbf{F}_3^*$, то есть обладает свойствами семантической полноты и непротиворечивости.

В качестве перспектив для будущего исследования можно отметить следующее. На наш взгляд, интересным представляется продолжение исследований по «натурализации» импликации Фаррелла. Другими словами, исследование возможностей для конструирования такой связки импликации, которая бы помимо моделирования конъюнктивного подхода к построению ментальных моделей каким-то образом могла бы моделировать и остальные два (би-кондициональный и кондициональный), действительно представляет интерес для будущих исследований.

Литература

Barrouillet P., *Lecas J. F.* How Can Mental Models Theory Account for Content Effects in Conditional Reasoning? A Developmental Perspective // Cognition. — 1998. — Vol. 67, no. 3. — P. 209–253.

Barrouillet P., *Lecas J. F.* Mental Models in Conditional Reasoning and Working Memory // Thinking & Reasoning. — 1999. — Vol. 5, no. 4. — P. 289–302.

Belikov A. On Bivalent Semantics and Natural Deduction for Some Infectious Logics // Logic Journal of the IGPL. — 2022. — Vol. 3, no. 1. — P. 186–210.

Belikov A. A Simple Way to Overcome Hyperconnexivity // Studia Logica. — 2024. — Vol. 112. — P. 69–94.

Belikov A., *Zaitsev D.* A Variant of Material Connexive Logic // Bulletin of the Section of Logic. — 2022. — Vol. 51, no. 2. — P. 227–242.
Cantwell J. The Logic of Conditional Negation // Notre Dame Journal of Formal Logic. — 2008. — Vol. 49, no. 3. — P. 245–260.
Carnielli W. A., *Coniglio M. E.* Paraconsistent Logic : Consistency, Contradiction and Negation. — Springer : Dordrecht, 2016. — (Logic, Epistemology, and the Unity of Science ; 40).
Dunn J. M. Partiality and Its Dual // Studia Logica. — 2000. — Vol. 66. — P. 5–40.
Farrell R. J. Material Implication, Confirmation, and Counterfactuals // Notre Dame Journal of Formal Logic. — 1979. — Vol. 20, no. 2. — P. 383–394.
Johnson-Laird P. N. Mental Models. — Cambridge : Cambridge University Press, 1983.
McCall S. Connexive Implication // The Journal of Symbolic Logic. — 1966. — Vol. 31, no. 3. — P. 415–433.
Mendelson E. Introduction to Mathematical Logic. — New York : Chapman & Hall, 2015.
Omori H., *Wansing H.* An Extension of Connexive Logic C // Advances in Modal Logic. — 2020. — Vol. 13. — P. 503–522.
Sylvan R. Bystanders' Guide to Sociative Logics. — Canberra : Australian National University, 1989. — (Research Series in Logic and Metaphysics ; 4).
Wansing H. Connexive Modal Logic // Advances in Modal Logic. — 2004. — Vol. 5. — P. 387–399.
Wansing H. Connexive Logic / The Stanford Encyclopedia of Philosophy ; ed. by E. N. Zalta. — 2023. — URL: `https://plato.stanford.edu/archives/sum2023/entries/logic-connexive/`.
Wójcicki R. Lectures on Propositional Calculi. — Wroclaw : Ossolineum, 1984.

Belikov, A. A. 2024. "Zamechaniya ob implikatsii Farrella [Remarks on Farrell's Implication]" [in Russian]. *Filosofiya. Zhurnal Vysshey shkoly ekonomiki [Philosophy. Journal of the Higher School of Economics]* 8 (2), 123–141.

Belikov Aleksandr
PhD in Philosophy
Senior Lecturer
Lomonosov Moscow State University (Moscow, Russia)
Senior Researcher
St. Petersburg State University (Saint-Petersburg, Russia); orcid: 0000–0003–1395–8878

Remarks on Farrell's Implication

Submitted: Aug. 29, 2023. **Reviewed:** Oct. 31, 2023. **Accepted:** Apr. 21, 2024.

Abstract: The paper provides an analysis of the three-valued implication developed by Robert Farrell in his work "Material Implication, Confirmation, and Counterfactuals" (1979).

Our analysis shows that, firstly, Farrell's implication has all the properties sufficient to put it into the category of connexive implications. Secondly, that Farrell's implication, similar in some of its properties to conjunction, can be used for the formal modelling of some peculiarities of understanding of conditional statements which are inherent to the early stages of development of humans' cognitive abilities. The main technical result of the paper is the construction and meta-theoretical study of the three-valued logic $\mathbf{F}_3^*$ containing Farrell's implication.

Keywords: Conditional Sentences, Conjunctive Sentences, Implication, Conjunction, Connexive Logic, Three-Valued Logic, Mental Models.

DOI: 10.17323/2587-8719-2024-2-123-141.

REFERENCES

Barrouillet, P., and J. F. Lecas. 1998. "How Can Mental Models Theory Account for Content Effects in Conditional Reasoning? A Developmental Perspective." *Cognition* 67 (3): 209–253.

———. 1999. "Mental Models in Conditional Reasoning and Working Memory." *Thinking & Reasoning* 5 (4): 289–302.

Belikov, A. 2022. "On Bivalent Semantics and Natural Deduction for Some Infectious Logics." *Logic Journal of the IGPL* 3 (1): 186–210.

———. 2024. "A Simple Way to Overcome Hyperconnexivity." *Studia Logica* 112:69–94.

Belikov, A., and D. Zaitsev. 2022. "A Variant of Material Connexive Logic." *Bulletin of the Section of Logic* 51 (2): 227–242.

Cantwell, J. 2008. "The Logic of Conditional Negation." *Notre Dame Journal of Formal Logic* 49 (3): 245–260.

Carnielli, W. A., and M. E. Coniglio. 2016. *Paraconsistent Logic: Consistency, Contradiction and Negation.* Logic, Epistemology, and the Unity of Science 40. Springer: Dordrecht.

Dunn, J. M. 2000. "Partiality and Its Dual." *Studia Logica* 66:5–40.

Farrell, R. J. 1979. "Material Implication, Confirmation, and Counterfactuals." *Notre Dame Journal of Formal Logic* 20 (2): 383–394.

Johnson-Laird, P. N. 1983. *Mental Models.* Cambridge: Cambridge University Press.

McCall, S. 1966. "Connexive Implication." *The Journal of Symbolic Logic* 31 (3): 415–433.

Mendelson, E. 2015. *Introduction to Mathematical Logic.* New York: Chapman & Hall.

Omori, H., and H. Wansing. 2020. "An Extension of Connexive Logic C." *Advances in Modal Logic* 13:503–522.

Sylvan, R. 1989. *Bystanders' Guide to Sociative Logics.* Research Series in Logic and Metaphysics 4. Canberra: Australian National University.

Wansing, H. 2004. "Connexive Modal Logic." *Advances in Modal Logic* 5:387–399.

———. 2023. "Connexive Logic." Ed. by E. N. Zalta. The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2023/entries/logic-connexive/.

Wójcicki, R. 1984. *Lectures on Propositional Calculi.* Wroclaw: Ossolineum.