А.В.Чагров

ДВА ЗАМЕЧАНИЯ О СТРОГО ИМПЛИКАТИВНЫХ ФОРМУЛАХ В МОДАЛЬНОЙ ЛОГИКЕ S3[1]

Abstract. *Some questions about strict implication in modal system* **S3** *are discussed. The first theorem shows that strict implication formulas have the property of equivalent replacement; as well known the logic* **S3** *has not the full theorem of equivalent replacement. The second theorem of the paper is connected with an open question [4] on description of one variable strict implication fragment of* **S3***: this fragment contains at least 29 non-equivalent formulas; in the author's paper [2] the number was 24.*

Данная работа является непосредственным продолжением статьи автора [2], и здесь практически все нужные для понимания детали будут по ходу изложения восстановлены.

В [2] изучался вопрос о характеризации строго импликативных формул от одной переменной в модальной логике **S3** и было показано, что таких попарно не эквивалентных формул не менее двадцати четырех, а именно, попарно не эквивалентными в **S3** являются следующие формулы:

$$
\begin{array}{llll}
\varphi_1 = p; & & \varphi_{13} = \varphi_1 \Rightarrow \varphi_9; \\
\varphi_2 = \varphi_1 \Rightarrow \varphi_1; & & \varphi_{14} = \varphi_3 \Rightarrow \varphi_4; \\
\varphi_3 = \varphi_2 \Rightarrow \varphi_1; & & \varphi_{15} = \varphi_5 \Rightarrow \varphi_4; \\
\varphi_4 = \varphi_1 \Rightarrow \varphi_3; & & \varphi_{16} = \varphi_5 \Rightarrow \varphi_6; \\
\varphi_5 = \varphi_4 \Rightarrow \varphi_1; & & \varphi_{17} = \varphi_5 \Rightarrow \varphi_9; \\
\varphi_6 = \varphi_1 \Rightarrow \varphi_5; & & \varphi_{18} = \varphi_6 \Rightarrow \varphi_1; \\
\varphi_7 = \varphi_5 \Rightarrow \varphi_1; & & \varphi_{19} = \varphi_6 \Rightarrow \varphi_4; \\
\varphi_8 = \varphi_7 \Rightarrow \varphi_4; & & \varphi_{20} = \varphi_8 \Rightarrow \varphi_1; \\
\varphi_9 = \varphi_8 \Rightarrow \varphi_6; & & \varphi_{21} = \varphi_8 \Rightarrow \varphi_4; \\
\varphi_{10} = \varphi_1 \Rightarrow \varphi_2; & & \varphi_{22} = \varphi_9 \Rightarrow \varphi_1; \\
\varphi_{11} = \varphi_1 \Rightarrow \varphi_7; & & \varphi_{23} = \varphi_9 \Rightarrow \varphi_4; \\
\varphi_{12} = \varphi_1 \Rightarrow \varphi_8; & & \varphi_{24} = \varphi_9 \Rightarrow \varphi_7.
\end{array}
$$

Оказывается, этот список можно пополнить еще пятью формулами:

[1] Работа выполнена при поддержке РГНФ, грант № 98-03-04220.

$\varphi_{25} = (((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p) \Rightarrow p) \Rightarrow (p \Rightarrow (p \Rightarrow p))$,

$\varphi_{26} = (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow ((p \Rightarrow p) \Rightarrow p))$,

$\varphi_{27} = (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow ((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p))$,

$\varphi_{28} = (p \Rightarrow (p \Rightarrow p)) \Rightarrow ((((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p) \Rightarrow p) \Rightarrow ((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p))$,

$\varphi_{29} = (p \Rightarrow (p \Rightarrow p)) \Rightarrow (((((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p) \Rightarrow p) \Rightarrow ((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p)) \Rightarrow (p \Rightarrow ((p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow p)))$,

или, если продолжить выписанную выше последовательность взаимосвязей формул,

$$\varphi_{25} = \varphi_7 \Rightarrow \varphi_{10};\quad \varphi_{26} = \varphi_{10} \Rightarrow \varphi_4;\quad \varphi_{27} = \varphi_{10} \Rightarrow \varphi_6;$$

$$\varphi_{28} = \varphi_{10} \Rightarrow \varphi_8;\quad \varphi_{29} = \varphi_{10} \Rightarrow \varphi_9.$$

Поскольку и такой, теперь уже пополненный новыми формулами, список формул φ_i $(1 \le i \le 29)$ не является окончательным решением задачи, стоит пояснить ход возникновения (и, тем самым, некоторые свойства) этого списка.

Напомним, что первые девять формул – φ_i $(1 \le i \le 9)$ – взяты из работы [4], где доказано, что в **S4** они составляют полную и независимую совокупность строго импликативных формул от одной переменной: между собой они попарно не эквивалентны в **S4**, а всякая другая строго импликативная формула от одной переменной в **S4** эквивалентна одной из них. При этом под эквивалентностью разумно понимать такую эквивалентность, относительно которой справедлива теорема о замене эквивалентных. Для **S4** этому условию удовлетворяют оба следующих определения эквивалентности формул φ и ψ: эти формулы считаются эквивалентными а) если **S4** принадлежит формула $\varphi \leftrightarrow \psi$ и б) если **S4** принадлежит формула $\varphi \Leftrightarrow \psi$. На самом деле оба определения эквивалентности просто-напросто равносильны для **S4**.

Иная ситуация в **S3**. Здесь с точки зрения теоремы о замене эквивалентных годится только следующее определение: формулы φ и ψ считаются эквивалентными в **S3**, если **S3** принадлежит формула $\varphi \Leftrightarrow \psi$. То есть по существу речь идет о строгой эквивалентности. Именно это имелось в виду перед формулировкой теоремы 1 в [2] (в пункте б) этой теоремы говорится, конечно, о логике **S3**).

Отмеченный факт хорошо известен в общем случае. Например, логике **S3** принадлежит формула $\top \leftrightarrow \Box\top$ и формула $\Box\top$, однако попытка замены в последней формуле $\top$ на $\Box\top$ приводит нас к недоказуемой в **S3** формуле $\Box\Box\top$, и более того: присоединение к **S3** формулы $\Box\Box\top$ дает в результате логику **S4**. Но ведь нас интересуют строго импликативные формулы, а для этого случая справедлива

Теорема 1. *Пусть φ и ψ такие строго импликативные формулы, что логике* **S3** *принадлежит $\varphi \leftrightarrow \psi$. Тогда из того, что* **S3** *принадлежит χ, а формула χ' получена из χ заменой некоторых вхождений φ на ψ, следует, что* **S3** *принадлежит и формула χ'.*

Доказательство несложно. Мы воспользуемся полнотой **S3** относительно моделей Крипке вида M = <F,V>, где F = <W,N,D,R> – **S3**-шкала, т.е. W – множество миров, N (N $\subset$ W) – множество нормальных миров, D (D $\subset$ N) – множество действительных миров (можно всегда полагать, что множество D одноэлементно), R – транзитивное и рефлексивное бинарное отношение (отношение достижимости) на W, V – оценка на F. Детали определений и формулировок можно найти в [1], напомним только получающееся из определений [1] определение истинности формул, имеющих вид строгой импликации: $\varphi \Rightarrow \psi$ истинна в мире модели тогда и только тогда, когда этот мир нормален, и во всяком достижимом из него мире либо опровергается формула φ, либо истинна формула ψ. Истинность эквивалентностей (материальной и строгой), естественно, понимается как истинность двух импликаций: *слева направо* и *справа налево*. Истинность формулы в модели означает ее истинность в действительных мирах этой модели.

Итак, пусть φ и ψ такие строго импликативные формулы, что логике **S3** принадлежит $\varphi \leftrightarrow \psi$.

Прежде всего заметим, что если одна из формул φ или ψ является простейшей строго импликативной формулой, т.е. переменной, то вторая с ней совпадает, а потому в этом случае применение замены эквивалентных тривиально. В самом деле, допустим, что $\varphi = p$. Рассмотрим различные варианты для формулы ψ. Если это переменная, то она должна совпасть с p, поскольку в **S3** недоказуема формула $p \to q$. Пусть теперь ψ – строго импликативная формула, отличная от переменной. Если она принадлежит **S3**, то очевидным образом формула $\varphi \to p$ не принадлежит **S3** – достаточно взять одноэлементную **S3**-модель, в которой p ложна. Если же формула ψ не принадлежит логике **S3**, то и формула $p \to \psi$ не принадлежит **S3**. Для обоснования этого предположим, что ψ опровергается в действительном мире d некоторой **S3**-модели. Преобразуем эту модель, добавив новый действительный мир d', полагая его рефлексивным и считая, кроме того, что из d' достижимы все те миры, которые достижимы из d; оценку в старых мирах сохраняем, а в d' считаем переменную p истинной. Ясно, что теперь в d' опровергается ψ (здесь используется транзитивность отношения достижимости, более точно - тот факт, что из d'

достижимо все, что достижимо из d), а потому и формула $p \to \psi$ опровергается в мире d'.

В результате нам нужно рассмотреть лишь случай, когда $\varphi = \varphi_1 \Rightarrow \varphi_2$ (здесь обозначения φ_1 и φ_2 локальны и никак не связаны с формулами φ_i, вводившимися выше) и $\psi = \psi_1 \Rightarrow \psi_2$, т.е. когда обе формулы являются строгими импликациями.

Предположим, что формула $(\varphi_1 \Rightarrow \varphi_2) \leftrightarrow (\psi_1 \Rightarrow \psi_2)$ принадлежит **S3**, и покажем, что формула $(\varphi_1 \Rightarrow \varphi_2) \Leftrightarrow (\psi_1 \Rightarrow \psi_2)$ тоже принадлежит **S3** от противного. Непринадлежность этой формулы **S3** по теореме о полноте означает, что в некоторой **S3**-модели в некотором мире a, достижимом из действительного, истинна одна из формул $(\varphi_1 \Rightarrow \varphi_2)$ и $(\psi_1 \Rightarrow \psi_2)$, но ложна другая. Поскольку истинная формула начинается с необходимости, этот мир нормален. Значит, мы можем образовать из имеющейся **S3**-модели новую **S3**-модель, объявив a действительным миром. Поскольку в a формулы $(\varphi_1 \Rightarrow \varphi_2)$ и $(\psi_1 \Rightarrow \psi_2)$ имеют разные значения, то в нем опровергается эквивалентность $(\varphi_1 \Rightarrow \varphi_2) \leftrightarrow (\psi_1 \Rightarrow \psi_2)$, что противоречит ее выбору.

Теорема 1 доказана.

Вернемся к основной теме – описанию совокупности строго импликативных формул от одной переменной в **S3**. Понятно, что начинать следует с выяснения взаимоотношений формул φ_i при $1 \leq i \leq 9$, коль скоро они точно описывают соответствующий фрагмент **S4**. На этом пути и были обнаружены указанные выше попарно неэквивалентные в **S3** формулы. Далее, конечно, следует рассматривать взаимоотношения формул из уже расширенного списка. Вовсе не факт, что процедура пополнения списка и внесения в него строгих импликаций формул из него, которые не эквивалентны ни одной из формул из него же, когда-нибудь завершится; т.е. вполне возможно, что рассматриваемый фрагмент бесконечен. В любом случае, громоздкость проводимых рассмотрений с каждым новым пополнением списка существенно возрастает и он (список) становится все менее регулярным.

Теперь уместно сформулировать точно высказанное выше утверждение.

Теорема 2. *Строго импликативные формулы* φ_i *при* $1 \leq i \leq 29$ *попарно не эквивалентны в* **S3**.

Доказательство проводится построением соответствующих **S3**-моделей, опровергающих проверяемые эквивалентности. Само по себе построение этих моделей не очень интересно, но отметим тот факт, что для всех эквивалентностей оказалось достаточным рассматривать модели, содержащие не более трех миров. Это тем более любопытно, что при прямом использовании финитной аппроксимируемости **S3** количество миров в контрмоделях оценивается экспонентой от длины формулы (эта оценка не понижаема, здесь легко применимы методы работы [5]); другими словами, рассмотренные формулы, несмотря на их громоздкость, имеют достаточно простой семантический «смысл». С другой стороны, это вызывает некоторую надежду решить задачу описания попарно не эквивалентных в **S3** строго импликативных формул от одной переменной компьютерными методами, главным препятствием для которых был бы именно экспоненциальный рост перебираемых контрмоделей в зависимости от длины опровергаемой формулы. Конечно, в том случае, когда контрмодели для проверяемой формулы нет, алгоритм перебора следует дополнить эвристиками. И вот поиск подходящих эвристик составляет здесь самую трудную и интересную проблему. Надежду на то, что они должны быть, выражает следующая

Гипотеза 1. *Существует лишь конечное число строго импликативных формул от одной переменной, попарно не эквивалентных в* **S3**.

Другим аспектом вопроса о мощности контрмоделей строго импликативных формул является надежда, что с ограничением числа используемых переменных экспоненциальная оценка, получаемая в общем случае, может быть снижена до полиномиальной. Во всяком случае все линейно растущие последовательности формул, требующих для своего опровержения (и действительно опровержимых!) экспоненциального роста контрмоделей, автору удавалось строить лишь без какого-либо ограничения числа используемых переменных. С этим связана

Гипотеза 2. *Фрагменты* **S3** *и* **S4**, *состоящие из строго импликативных формул, в которых используется не более некоторого заранее фиксированного числа n переменных, линейно финитно аппроксимируемы, то есть всякая строго импликативная формула φ от n переменных, не принадлежащая рассматриваемому фрагменту, опровергается на подходящей модели не более чем с $cl(\varphi)+d$ мирами, где c, d – константы, а $l(\varphi)$ – длина формулы φ.*

К слову, попарно неэквивалентных в **S4** строго импликативных формул от двух переменных бесконечно много, см. [2]. Однако это в принципе не мешает подтверждению гипотезы 2. Даже в классической логике высказываний количество попарно не эквивалентных чисто импликативных формул от n переменных не меньше 2^{n-1}, примером являются формулы вида

$$p^*_1 \to (p^*_2 \to (\ldots \to (p^*_{n-1} \to p^*_n)\ldots)),$$

где p^*_i есть либо p_i, либо $p^*_i \to p^*_n$. Мы здесь не задаемся целью точно описать количество импликативных булевых функций. Только что отмеченная экспоненциальная оценка достаточна для наших иллюстративных целей. Однако «справедливости ради» и «полноты для» напомним известный факт: число попарно неэквивалентных в классической логике импликативных формул от n переменных по порядку равно двойной экспоненте от n (т.е. это «почти все» булевы функции). Обратим внимание, что с 60-х годов активно интересовались соответствующим числом в случае интуиционитской логики и ее фрагментов. Упомянем лишь совсем немногие результаты из большого их количества, собранного и полученного в [6]: интуиционистских попарно не эквивалентных чисто импликативных формул от двух переменных две, от двух – 14 (попарно не эквивалентных в классической логике чисто импликативных формул шесть), от трех – 25165802 (всех булевых функций от трех переменных «всего» $2^8 = 256$), от четырех – $2^{623\ 662\ 965\ 552\ 393} - 50\ 331\ 618$ (это число не поддается никаким естественным сравнениям, его невозможно даже назвать астрономическим[2]). Говоря об интуиционистских формулах, уместно напомнить, что аналог высказанной выше гипотезы 2 для интуиционистских формул (не чисто импликативных, а в полной сигнатуре) тоже до сих пор остается не доказанным утверждением.

ЛИТЕРАТУРА

1. *Фейс Р.* Модальная логика. М., Наука, 1974.
2. *Чагров А.В.* Строго импликативные формулы в модальных логиках, близких к интуиционисткой // Логические исследования. Выпуск 6. М.: РОССПЭН, 1999. С. 69-75.
3. *Чагров А.В.* О сложности пропозициональных логик // Сложностные проблемы математической логики. Калинин: КГУ, 1985. С. 80-90
4. *Byrd M.* Single variable formulas in S4 // J.Phil.Log., 1976, 5, No. 4, 439-456.
5. *Urquhart A.* Implicational formulas in intuitionistic logic // Journal of Symbolic Logic, 1974. Vol. 39. P. 661-664.
6. *Hendrix A.* Computations in Propositional Logic. PhD. thesis, Amsterdam, 1996.

---

[2] Можно вспомнить известную притчу о создателе шахматной игры, запросившем в оплату за свою игру одно зернышко с первой клетки, два со второй, четыре с третьей и т.д., всего $2^{64} - 1$ зерен, это зерно довольно толстым слоем покрыло бы весь земной шар.