Московский физико-технический институт
Факультет инноваций и высоких технологий
Математическая логика и теория алгоритмов, осень 2015
Лекция 1: формальный язык записи логических утверждений

Краткое содержание

Необходимость в формальной системе записи математических утверждений. Основные понятия теории формальных языков: алфавит, слово, язык. Пустое слово. Операции со словами: конкатенация, обращение. Операции с языками: теоретико-множественные, конкатенация, звезда и плюс Клини. Отношения на словах: префиксы, суффиксы, подслова, подпоследовательности. Три определения правильной скобочной последовательности: через скобочный итог, через разбиение скобок на пары и рекурсивное. Их эквивалентность. Неоднозначность разбора при рекурсивном определении. Алфавит логики высказываний: пропозициональные переменные, символы логических операций, скобки. Правила построения пропозициональных формул. Лемма о скобочном итоге и теорема об однозначности разбора для пропозициональных формул.

1 Зачем нужен формальный язык

По мере развития математики её язык совершенствовался. Постепенно появились различные обозначения: буквы в качестве переменных, символы операций и т.д. Однако собственно математические (в частности, логические) рассуждения долго проводились на естественном языке. К концу XIX – началу XX вв. возникла потребность в формализации рассуждений: с одной стороны, формальная система записи облегчала коммуникацию между разными математиками, с другой стороны, были открыты парадоксы, связанные с рассуждениями на естественном языке.[1] От этих парадоксов нужно было как-то избавляться.

Примером такого парадокса может служить парадокс Рассела. Назовём *правильным* множество, которое не является собственным элементом: например, множество натуральных чисел само не является натуральным числом, а множество всех треугольников не является треугольником. С другой стороны, есть и *неправильные* множества. Например, множество всех множеств является множеством, а множество всех бесконечных множеств является бесконечным множеством. Вопрос: каким является множество правильных множеств? Если оно правильное, то оно входит в множество правильных множеств, т.е. является собственным элементом, т.е. является неправильным. Если же оно неправильное, то является собственным элементом, т.е. входит в множество правильных множеств, т.е. является правильным. Таким образом, в любом случае получается противоречие. Впрочем, этот парадокс разрешается не формализацией рассуждений, а аксиоматизацией теории множеств, ограничивающей способы формирования множеств. Зато без формализации нельзя даже сформулировать, например, вторую проблему Гильберта: являются ли аксиомы арифметики непротиворечивыми?

[1] О сведéнии любого рассуждения к механической процедуре, вычислению говорил ещё Лейбниц, но даже с самой постановкой вопроса он значительно опередил своё время.

Формальные системы записи математических утверждений и рассуждений были в целом построены в первой половине XX в. и сыграли очень большую роль в последовавшем развитии вычислительной техники. Во-первых, компьютеры не могут понимать никаких языков, кроме формальных, поэтому знание теории формальных языков очень помогает при программировании. Во-вторых, формальная запись математических утверждений и доказательств позволяет искать и проверять доказательства машинными методами. Пожалуй, самым известным примером такого доказательства является решение проблемы четырёх красок: верно ли, что любую географическую карту можно покрасить в 4 цвета, так чтобы области одного цвета не граничили друг с другом. Первоначально проблема была положительно решена в 1976 году Аппелем и Хакеном путём компьютерного перебора. Впоследствии доказательство было несколько раз упрощено, хотя и оставалось переборным, а в 2005 году Вернер и Гонтье записали полное доказательство на формальном языке и верифицировали его при помощи программы Coq.

В этой лекции мы коснёмся самых основ теории формальных языков, более подробно она будет освещена в курсе «Формальные языки и трансляции».

2 Основные понятия теории формальных языков

Определение 1. *Алфавитом* называется любое конечное непустое множество. Элементы алфавита называются *символами.*

Как правило, алфавиты мы будем обозначать большими греческими буквами: Σ, Γ и т.д., а символы алфавита — маленькими греческими: σ, τ и т.д., либо конкретными значками. Алфавит $\{0, 1\}$ будем называть *бинарным* или *двоичным.* Теоретически можно рассмотреть и бесконечный алфавит (и даже несчётный), но по умолчанию мы этого делать не будем.

Определение 2. *Словом* называется любая конечная последовательность символов. Число символов в этой последовательности называется *длиной* слова.

При записи слов, в отличие от последовательностей, мы не будем использовать скобки и запятые, просто записывая символы один за другим. Как правило, слова мы будем обозначать маленькими латинскими буквами: u, v, x и т.д. Особое значение играет *пустое слово*, т.е. слово нулевой длины, последовательность нуля символов. Оно ровно одно и не зависит от алфавита, поэтому для используем для него специальное обозначение ε (иногда также пишут λ или Λ). Длину слова u мы будем обозначать через $|u|$ (также используют обозначения $l(u)$ или $\mathrm{len}(u)$).

Определение 3. *Языком* называется любое множество слов в некотором алфавите.

Языки могут быть и конечными, и бесконечными. Языки мы будем обозначать большими латинскими буквами: L, M и т.д. Язык может быть пустым, в таком случае его обозначают $\varnothing$, как пустое множество (иногда используют обозначение Φ).

Над словами и языками можно проводить различные операции.

Определение 4. Пусть $u = \sigma_1 \dots \sigma_k$ и $v = \tau_1 \dots \tau_l$ суть два слова. Тогда *конкатенацией* слов u и v называется слово $u \cdot v = \sigma_1 \dots \sigma_k \tau_1 \dots \tau_l$.

Суть определения заключается в том, что слово v приписывается справа к слову u. Также можно отметить, что конкатенация очень похожа на умножение. Если каждый символ рассматривать как переменную, а слова — как произведения переменных, то перемножаться они должны именно так. Поэтому знак $\cdot$ для конкатенации, как и для умножения, часто опускают. Также конкатенация подобно умножению порождает понятие степени.

Определение 5. Пусть u — некоторое слово, а n — натуральное[2] число. Тогда его n-ой степенью называется слово $u^n = \underbrace{u \cdot \dots \cdot u}_{n \text{ раз}}$.

При этом нулевая степень любого слова, т.е. конкатенация нуля слов, по определению считается пустым словом. Можно было бы определить степень итеративно: $u^0 = \varepsilon$, а при $n > 0$ выполнено $u^n = u^{n-1} \cdot u$.

Можно отметить следующий простой факт:

Утверждение 6. *Для всех слов u, v и w выполнены следующие свойства:*

а) $(u \cdot v) \cdot w = u \cdot (v \cdot w)$ *(ассоциативность конкатенации);*

б) $u \cdot \varepsilon = \varepsilon \cdot u = u$ *(пустое слово является нейтральным элементом);*

в) $u^n \cdot u^m = u^{n+m}$*;*

г) $(u^n)^m = u^{nm}$*.*

В алгебраических терминах можно сказать, что слова образуют моноид, или полугруппу с нейтральным элементом относительно конкатенации. (И именно поэтому она похожа на умножение). Более того, эта полугруппа свободная.

Определение 7. *Обращением* или *зеркальным образом* слова $u = \sigma_1 \dots \sigma_k$ называется слово $u^{\mathsf{R}} = \sigma_k \dots \sigma_1$, т.е. слово u, записанное задом наперёд.

Утверждение 8. *Для всех слов u и v выполнены следующие свойства:*

а) $\left(u^{\mathsf{R}}\right)^{\mathsf{R}} = u$*;*

б) $(u \cdot v)^{\mathsf{R}} = v^{\mathsf{R}} \cdot u^{\mathsf{R}}$*;*

в) $\varepsilon^{\mathsf{R}} = \varepsilon$*.*

Доказательство. Во всех случаях пусть $|u| = k$, $|v| = l$. Заметим, что i-ый символ слова u^{R} есть $(k + 1 - i)$-ый символ слова u.

---

[2] Мы предполагаем, что натуральные числа начинаются с нуля, т.е. $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. Это соответствует пониманию, что натуральное число отвечает на вопрос «Сколько?» и обозначает количество. В средней школе говорят, что натуральные числа начинаются с единицы, и, таким образом, отвечают на вопрос «Какой по счёту?» и обозначают порядковый номер. Выбор одного из двух подходов является делом вкуса, но в логике и информатике удобнее первый.

а) i-ый символ слова $\left(u^{\mathsf{R}}\right)^{\mathsf{R}}$ есть $(k+1-i)$-ый символ слова u^{R}, т.е. $(k+1-(k+1-i))$-ый символ u, т.е. i-ый, что и требовалось.

б) Докажем, что i-ый символ $(u \cdot v)^{\mathsf{R}}$, т.е. $(k+l+1-i)$-ый символ $u \cdot v$, равен i-ому символу $v^{\mathsf{R}} \cdot u^{\mathsf{R}}$. Рассмотрим случай $i \leqslant l$. В таком случае $k+l+1-i \geqslant k+1$, и $(k+l+1-i)$-ый символ $u \cdot v$ равен $(l+1-i)$-му символу v, т.е. i-ому символу v^{R}, т.е. i-ому символу $v^{\mathsf{R}} \cdot u^{\mathsf{R}}$, что и требовалось.

   Теперь рассмотрим случай $i > l$. В таком случае $k+l+1-i \leqslant k$, и $(k+l+1-i)$-ый символ $u \cdot v$ равен $(k+l+1-i)$-му символу u, т.е. $(k+1)-(k+l+1-i) = (i-l)$-му символу u^{R}, т.е. i-му символу $v^{\mathsf{R}} \cdot u^{\mathsf{R}}$, что и требовалось.

в) Поскольку пустое слово не содержит ни одного символа, то и его зеркальный образ не содержит символов, т.е. также пуст.

$\square$

Операция обращения позволяет определить одно широко известное понятие:

Определение 9. Слово u, для которого выполнено соотношение $u^{\mathsf{R}} = u$, называется *палиндромом.*

Примерами палиндромов могут служить слова ТОПОТ, РОТАТОР в русском языке, CIVIC, RACECAR, REDIVIDER в английском языке или SAIPPUAKIVIKAUPPIAS в финском. Чаще всего рассматривают не слова, а фразы-палиндромы, которые становятся палиндромами в строгом смысле, если исключить из них все пробелы и знаки препинания, а также не различать регистры букв, например «А роза упала на лапу Азора» или «Madam, I'm Adam».

Далее перейдём к операциям над языками. Во-первых, над языками можно проводить все теоретико-множественные операции: объединение, пересечение, разность, дополнение. Наиболее часто встречается объединение, поэтому для него могут использовать специальный знак $+$ вместо $\cup$. Во-вторых, возникают специальные операции, связанные с конкатенацией.

Определение 10. Пусть L и M суть два языка. Тогда их *конкатенацией* называется язык $L \cdot M = \{uv \mid u \in L, v \in M\}$. Если n натурально, то n-ой степенью языка L называется язык $L^n = \underbrace{L \cdot \dots \cdot L}_{n \text{ раз}} = \big\{u_1 \dots u_n \mid u_i \in L \text{ при всех } i \in \{1, \dots, n\}\big\}$.

Например, если $L = \{0, 01\}$, а $M = \{0, 10\}$, то $L \cdot M = \{00, 010, 0110\}$. Обратите внимание, что слово 010 получается двумя способами: как $0 \cdot 10$ и как $01 \cdot 0$. Степень языка вновь можно определить итеративно: $L^0 = \{\varepsilon\}$, $L^n = L^{n-1} \cdot L$. Можно заметить выполнение нескольких простых свойств:

Утверждение 11. *При всех L, M и K выполнены следующие соотношения:*

а) $(L \cdot M) \cdot K = L \cdot (M \cdot K)$;

б) $L \cdot \{\varepsilon\} = \{\varepsilon\} \cdot L = L$;

в) $L \cdot \varnothing = \varnothing \cdot L = \varnothing$;

г) $L^0 = \{\varepsilon\}$, $L^1 = L$;

д) $L^n \cdot L^m = L^{n+m}$;

е) $(L^n)^m = L^{nm}$.

Таким образом, конкатенация языков тоже ведёт себя подобно умножению, причем помимо единицы ($\{\varepsilon\}$) там есть ноль ($\varnothing$). На языки можно распространить и операцию обращения:

Определение 12. *Обращением* языка L называется язык $L^{\mathsf{R}} = \{u^{\mathsf{R}} \mid u \in L\}$.

Утверждение 13. *При всех L и M выполнено:*

а) $(L \cdot M)^{\mathsf{R}} = M^{\mathsf{R}} L^{\mathsf{R}}$;

б) $(L^{\mathsf{R}})^{\mathsf{R}} = L$.

Определение 14. *Итерацией* (замыканием Клини) языка L называется язык $L^* = \cup_{n=0}^{\infty} L^n$. Эту операцию также называют *звездой Клини*. *Плюсом Клини* называют операцию, переводящую язык L в $L^+ = \cup_{n=1}^{\infty} L^n$.

Например, если $L = \{\text{парам}, \text{пам}\}$, то $L^* = \{\varepsilon$, парам, пам, парампарам, пампарам, парампам, пампам, парампампарам, парампампам, $\ldots\}$.

Утверждение 15. *Для всех языков L выполнены свойства:*

а) $L^* = L^+ \sqcup \{\varepsilon\}$, $L^+ = L \cdot L^*$;

б) $(L^*)^* = L^*$, $(L^+)^+ = L^+$ *(идемпотентность)*;

Полезно понять, что получится в результате применения звезды и плюса Клини к пустому языку и к языку из пустого слова. Для языка из пустого слова можно по индукции доказать, что $\{\varepsilon\}^n = \{\varepsilon\}$, поэтому $\{\varepsilon\}^* = \{\varepsilon\}^+ = \{\varepsilon\}$. Для пустого языка по определению $\varnothing^0 = \{\varepsilon\}$, а при $n > 0$ по индукции можно доказать, что $\varnothing^n = \varnothing$. Таким образом, $\varnothing^* = \{\varepsilon\}$, а $\varnothing^+ = \varnothing$.

Если в качестве языка L взять алфавит Σ (или, если угодно, множество однобуквенных слов), то операции звезды и плюса позволяют ввести удобные обозначения: Σ^* — множество всех слов, Σ^+ — множество всех непустых слов.

Завершим раздел определением нескольких отношений на множестве слов.

Определение 16. Слово u является *префиксом* (началом) слова v, если для некоторого слова w выполнено $v = uw$. Обозначение: $u \sqsubset v$. Слово u является *суффиксом* (концом) слова v, если для некоторого слова w выполнено $v = wu$. Обозначение: $u \sqsupset v$. Слово u является *подсловом* слова v, если для некоторых слов w и z выполнено $v = wuz$. Обозначение: $u \sqsubset\!\sqsupset v$. Слово u называется *подпоследовательностью* слова v, если $v = \sigma_1 \ldots \sigma_n$, а $u = \sigma_{i_1} \ldots \sigma_{i_k}$ для некоторых $1 \leqslant i_1 < \cdots < i_k \leqslant n$. Префиксы, суффиксы, подслова и подпоследовательности, которые короче самого́ слова, мы будем называть *собственными*.

Например, слово ТУН является префиксом слова ТУНИС, слово ОЛА является суффиксом слова АНГОЛА, слово ГЕНТ является подсловом слова АРГЕНТИНА, а слово РИМ является подпоследовательностью слова СУРИНАМ.[3]

Обратите внимание, что запись $u \sqsupset v$ означает, что u есть суффикс v, а не что v есть префикс u. Нетрудно заметить, что любые префиксы и суффиксы являются подсловами, а любое подслово — подпоследовательностью. Также верны следующие свойства:

Утверждение 17. *При всех u, v и w выполнено:*

а) $u \sqsubset u$, $u \sqsupset u$, $u \sqsupset\!\sqsubset u$ (рефлексивность);

б) Если $u \sqsubset v$ и $v \sqsubset u$, то $u = v$ (антисимметричность, аналогично для $\sqsupset$ и $\sqsupset\!\sqsubset$);

в) Если $u \sqsubset v$ и $v \sqsubset w$, то $u \sqsubset w$ (транзитивность, аналогично для $\sqsupset$ и $\sqsupset\!\sqsubset$).

Говорят, что слова образуют частично упорядоченные множества относительно всех трёх операций. (Более того, это упорядоченное множество будет решёткой).

Понятие префикса позволяет определить важное понятие:

Определение 18. Язык L называется *беспрефиксным*, если ни для каких различных u и v из L не выполняется $u \sqsubset v$.

Беспрефиксные языки очень важны: таким свойством обладают множество пропозициональных формул, множество формул первого порядка (т.е. формул с кванторами), множества корректных программ во многих языках программирования и т.д.

3 Правильные скобочные последовательности

В любых математических формулах важную роль играют скобки. Поскольку мы анализируем формальный язык, то разумно сначала проанализировать язык, состоящий только из скобок.

Определение 19. *Скобочной последовательностью* назовём любое слово в алфавите $\{(,)\}$, т.е. любой элемент $\{(,)\}^*$.

Интуитивное понятие *правильной скобочной последовательности* можно формализовать тремя различными способами:

Определение 20. Скобочная последовательность называется правильной, если все входящие в неё скобки можно разбить на пары, так что в каждой паре есть открывающая и закрывающая скобки, причём открывающая встречается раньше.

Определение 21. *Скобочным итогом* (балансом) скобочной последовательности называется разность числа закрывающих и открывающих скобок. Скобочная последовательность называется правильной, если её скобочный итог равен нулю, а у любого её префикса скобочный итог неотрицательный.

[3] В географическом смысле Тун расположен в Швейцарии, Ола — в Магаданской области России, Гент — в Бельгии, а Рим, разумеется, в Италии.

Определение 22. Пустая скобочная последовательность (т.е. ε) является правильной. Если s является правильной скобочной последовательностью, то (s) тоже является правильной скобочной последовательностью. Если s_1 и s_2 являются правильными, то $s_1 \cdot s_2$ тоже является правильной. Никак иначе образовать правильную скобочную последовательность нельзя.

Теорема 23. *Определения 20, 21 и 22 эквивалентны.*

Доказательство. Проведём доказательство по циклу: докажем, что из определения 20 следует 21, из определения 21 следует 22, а из 22 — 20.

Действительно, если все скобки разбиты на пары, то в любой префикс от каждой пары либо не входит ни одной скобки, либо ровно одна, либо обе. Пары, от которых не входят ни одной скобки или входят обе, дают нулевой вклад в скобочный итог префикса, а пары, от которых входит ровно одна скобка, дают вклад $+1$, т.к. эта скобка обязательно открывающая, раз встречается раньше. Значит, у любого префикса скобочный итог неотрицательный. Ну а вся последовательность содержит все пары, поэтому её скобочный итог нулевой.

Пусть для последовательности t теперь выполнено определение 21. Будем доказывать индукцией по длине t, что она соответствует определению 22. Если $t = \varepsilon$, то определение 22 выполнено, и это база индукции. Иначе рассмотрим кратчайший непустой префикс $t' \sqsubset t$, скобочный итог которого равен нулю. Заметим, что первым символом t' обязательно является открывающая скобка, а последним — закрывающая, иначе условие на неотрицательность скобочного итога будет нарушено либо для однобуквенного префикса t', либо для $(|t'| - 1)$-буквенного. Пусть $t' = t$. Тогда в силу сказанного выше $t = (s)$. При этом скобочный итог s такой же как у t, т.е. нулевой, а скобочный итог любого префикса $s' \sqsubset s$ неотрицательный, поскольку $(s'$ есть префикс t', и по определению t' скобочный итог $(s'$ положительный. Значит, для s' выполнено определение 21, и по предположению индукции выполнено определение 22. Значит, для $t = (s)$ определение 22 также выполнено, что и требовалось. Если же t' короче t, то $t = t' \cdot s$. Для t' выполнено определение 21, поскольку все его префиксы являются префиксами t. Для s оно тоже выполнено, поскольку у префикса $s' \sqsubset s$ скобочный итог такой же, как и у $t' \cdot s'$, которое есть префикс t. По предположению индукции определение 22 выполнено для последовательностей t' и s, значит оно выполнено и для $t = t' \cdot s$, что и требовалось.

Наконец, докажем, что из определения 22 следует определение 20. Вести доказательство будем индукцией по построению. Для пустой последовательности всё верно: никаких скобок нет и ничего разбивать не надо. Если последовательность s разбита на пары скобок, то последовательность (s) тоже разбивается: надо добавить пару внешних скобок. Наконец, если последовательности s_1 и s_2 разбиты на пары скобок, то объединение этих разбиений даёт разбиение последовательности $s_1 \cdot s_2$. Таким образом, все три импликации и вся теорема доказаны. $\square$

Рекурсивные определения, похожие на 22 будут часто нам встречаться. Теория формальных языков подробно их рассматривает, называя контекстно-свободными грамматиками. Так, определение 22 могло бы быть записано как

$$pmp \to \varepsilon \mid (pmp) \mid pmp \cdot pmp,$$

где *pmp* означает “properly matched parantheses”. Мы не будем подробно расшифровывать эту запись, оставляя читателю возможность догадаться о её значении самостоятельно.

Также можно отметить, что правильные скобочные последовательности важны для комбинаторики: количество правильных скобочных последовательностей для n пар скобок называется n-ым числом Каталана и имеет много интересных свойств.

Заметим, что для правильных скобочных последовательностей не соблюдается *однозначность разбора*: по последовательности невозможно однозначно сказать, по какому правилу и из каких подпоследовательностей она образована. Например, последовательность $()()()$ может возникнуть как $()() \cdot ()$ и как $() \cdot ()()$. (Впрочем, несколько вариантов разбиения на составляющие по третьему правилу это единственный источник неоднозначности).

4 Построение пропозициональных формул

После проведённой подготовки перейдём к описанию языка пропозициональной логики. Слово «пропозициональная» означает «относящаяся к высказываниям», т.е. утверждениям, которые могут быть истинны или ложны. Прежде всего опишем алфавит.

Алфавит будет содержать пропозициональные переменные, т.е. символы, обозначающие высказывания. Обычно мы будем обозначать их буквами p, q, r и т.д. Мы договорились, что алфавит конечен, но на практике удобно иметь хотя бы счётное число переменных. Нужно либо отказаться от требования конечности алфавита, либо разрешить обозначать переменные более, чем одним символом, например латинской буквой с десятичным числом в индексе. Помимо переменных, язык будет содержать символы логических операций $\wedge$ (конъюнкция, «и»), $\vee$ (дизъюнкция, «или»), $\to$ (импликация, «влечёт») и $\neg$ (отрицание, «не»), а также скобки.

Пропозициональные формулы определяются рекурсивно:

Определение 24. Если p — пропозициональная переменная, то p есть формула.[4] Если φ есть формула, то $\neg\varphi$ — также формула. Если φ и ψ являются формулами, то $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$ и $(\varphi \to \psi)$ также являются формулами.

Например, формулой будет являться выражение $((p \vee q) \to \neg(q \wedge \neg\neg r))$. Видно, что по данному определению нужно ставить излишнее количество скобок. Например, внешние скобки вокруг всей формулы не нужны, а расстановка приоритетов операций позволяет ещё уменьшить число скобок. Однако, данное нами определение наиболее просто и в формулировке, и при анализе. Мы докажем, что в отличие от правильных скобочных последовательностей пропозициональные формулы разбираются однозначно, что исключает разночтения в трактовке их смысла.

Лемма 25 (о скобочном итоге)**.** *Скобочным итогом любой последовательности назовём разность количеств открывающих и закрывающих скобок в ней (безотносительно*

[4] Если переменная обозначается одним символом, то формулой будет однобуквенное слово из этого символа, если многими, то формулой будет соответствующее слово.

прочих символов). Пусть φ есть пропозициональная формула, а s — некоторый её префикс. Тогда скобочный итог s неотрицателен, причём равен нулю, только если $s = \varphi$ или $s \in \{\neg\}^*$ (т.е. либо пуст, либо представляет собой несколько знаков отрицания).[5]

Доказательство. Будем доказывать лемму индукцией по построению формулы. Для переменной всё верно, поскольку у неё только два префикса: пустой и она сама. Если для формулы φ лемма верна, то и для формулы $\neg\varphi$ тоже верна: любой префикс формулы $\neg\varphi$ получается дописыванием $\neg$ в начало префикса φ, отчего не меняется ни скобочный итог, ни заключение леммы. Наконец, если для формул φ и ψ формула верна, то она верна и для $(\varphi * \psi)$, где $*$ обозначает один из символов $\wedge$, $\vee$, $\to$. Действительно, рассмотрим собственный непустой префикс $(\varphi * \psi)$. Тогда он не подходит под заключение леммы. Посчитаем его скобочный итог. Если он имеет вид $(\varphi'$, где $\varphi' \sqsubset \varphi$, то его скобочный итог на единицу больше скобочного итога φ', т.е. положительный. Если же он имеет вид $(\varphi * \psi'$, где $\psi' \sqsubset \psi$, то его скобочный итог на единицу больше скобочного итога ψ', т.е. тоже положительный. Таким образом, утверждение леммы доказано. □

Теорема 26 (об однозначности разбора)**.** *Пусть φ есть пропозициональная формула. Тогда можно однозначно сказать, по какому правилу и из каких подформул она образована.*

Доказательство. Посмотрим на первый символ φ. Если это переменная, то формула сама является переменной. Если это символ $\neg$, то формула образована по второму правилу из подформулы, получающейся зачёркиванием этого символа. Если же это открывающая скобка, то априори возможны различные варианты. Докажем от противного, что вариант единственный. Пусть $(\varphi_1 * \psi_1) \mathrel{=\!\!=} (\varphi_2 * \psi_2)$, где знак $\mathrel{=\!\!=}$ (графическое равенство, посимвольное совпадение) подчёркивает равенство выражений как слов, а не в каком-то другом смысле. Знаки $*$ могут обозначать разные символы, это не повлияет на наши рассуждения. Если $\varphi_1 \mathrel{=\!\!=} \varphi_2$, то $\psi_1 \mathrel{=\!\!=} \psi_2$, т.е. это то же самое представление. Пусть $\varphi_1 \not\mathrel{=\!\!=} \varphi_2$. Тогда одна из этих формул является префиксом другой, без ограничения общности $\varphi_1 \sqsubset \varphi_2$. Но, с одной стороны, скобочный итог φ_1 нулевой, а с другой стороны, скобочный итог собственного префикса φ_2 может быть нулевым, только если он лежит в $\{\neg\}^*$. Такие слова не являются формулами, поэтому мы пришли к противоречию. Значит, при образовании формулы по третьему правилу также можно однозначно сказать, из каких подформул она образована, так что теорема доказана. □

[5]Здесь мы считаем, что переменная это один символ, иначе пришлось бы добавить вариант, когда φ есть префикс переменной.