Г.Е. Крейдлин, Г.Б. Шабат

Формальный язык геометрии как знаковая система I. Элементарный синтаксис

Статья продолжает серию работ, написанных совместно двумя авторами – лингвистом и математиком. Центральное место в ней занимает обсуждение природы и механизмов понимания научных текстов в их соотношении с бытовыми текстами. В предыдущих работах мы охарактеризовали основные способы контроля понимания научных текстов и описали важнейшие когнитивные операции над отдельными их видами. В этих работах были описаны с разной степенью подробности несколько языков геометрии: естественно-подобный язык, язык чертежей, координатный язык и язык геометрических преобразований. Основное внимание в этих описаниях было уделено лингвистическим особенностям рассматриваемых языков, прежде всего, их лексике, грамматике и прагматике. В настоящей статье речь идёт ещё об одном научном языке – о формальном языке геометрии. А именно, представлен один из его уровней, который мы назвали элементарный синтаксис.

Ключевые слова: знак, формальный язык, геометрия, понимание, когнитивный подход, элементарный синтаксис, расширенный синтаксис, терм, предикат, формула, сочетаемость знаков, упорядоченность знаков.

Введение. Постановка задачи

Наряду с повседневными, или, так сказать, бытовыми, языками и невербальными знаковыми кодами, которыми пользуются их рядовые носители, лингвистика и семиотика интересуются и другими языками и знаковыми кодами, в частности, языками разных наук. Когда в сферу внимания лингвистов и специалистов в области семиотики входит язык какой-то науки, они изучают, прежде всего, его строение и содержание и втайне надеются, что лингвистическое и, шире, семиотическое изучение этого языка может помочь и пользователям повседневного языка. При этом, как и для повседневного языка, принято выделять два аспекта владения языком науки, а именно понимание, или анализ, письменных и устных текстов на этом языке и построение , или синтез, таких текстов. При лингвистическом анализе научных текстов центральное место среди уровней языка и языковой структуры занимают лексическая и грамматическая системы, а фонетическая и фонологическая системы остаются в стороне, что не удивительно, поскольку языки науки в основном связаны с письменной формой общения.

***

Настоящая статья является продолжением серии работ, написанных совместно двумя авторами – математиком и лингвистом[1]. Центральное место в нашей совместной научной деятельности занимают природа и механизмы понимания текстов, главным образом научных, в их соотношении с бытовыми текстами. В предыдущих работах мы охарактеризовали основные способы контроля понимания научных текстов и описали важнейшие когнитивные операции над отдельными их видами, а именно, в этих работах речь шла о когнитивных операциях над формулировками теорем. Было показано, что владение такими операциями ведет к более полному и глубокому пониманию смысла рассматриваемых математических текстов. Концепт понимания смысла текстов и некоторые их структурные характеристики составляют центральное место и в данной работе. Сходные проблемы, связанные с пониманием текстов разных тематик, видов и жанров, рассматривались в целом ряде исследований.[2]

***

Как известно, школьная геометрия состоит из двух разделов – планиметрии и стереометрии. Здесь мы остановимся только на планиметрии, которая изучает подмножества *плоскости*. Плоскость в настоящей работе – это особое понятие, точный смысл которого будет раскрыт при описании семантики нашего языка. В первом приближении мы можем сказать, что плоскость понимается как *универсум*, то есть «самое большоеї множество. Это означает, что все геометрические утверждения относятся к элементам плоскости (точкам) или к ее подмножествам (прямым и некоторым другим, более сложным подмножествам).

В предыдущих работах о языках геометрии нами были в общих чертах описаны пять языков геометрии. Это естественно-подобный язык, язык чертежей, формальный язык, координатный язык и язык геометрических преобразований. Основное внимание мы уделили лингвистическим особенностям этих языков, прежде всего, их лексике, грамматике и прагматике. В настоящей статье речь пойдет в основном об одном из этих языков, а именно, о формальном языке геометрии. Мы опишем здесь один из его уровней, а именно – синтаксический, причем сначала остановимся на простейших (элементарных) единицах синтаксиса и отношениях между ними, а затем (в следующей статье) перейдем к сложным единицам и конструкциям.

Формальный язык геометрии является представителем класса так называемых формальных языков, под которым математики понимают по существу то, что лингвисты понимают под языком, создаваемым порождающей грамматикой. Семантика, или возможные интерпретации порождаемых синтаксических единиц, рассматривается и в математике, и в лингвистике как отдельный уровень описания языка (ср. лингвистическое понятие и термин *семантический компонент порождающей грамматики*). Исходной частью формального языка геометрии является синтаксис.

Рассматриваемый в настоящей работе язык геометрии является формализованной частью фрагмента привычного для читателя языка школьной планиметрии. Однако в нашем языке, во-первых, нет способов выразить понятие *лежать между двумя точками* на произвольной прямой, вследствие чего отсутствуют понятия луча, отрезка, полуплоскости и угла, а во-вторых, в нем невыразимы понятия длины и площади. Что же касается еще одного понятия геометрии – окружности, – то оно на нашем языке в принципе выразимо, но довольно сложным образом, и мы его здесь опустим. Несмотря на указанные ограничения, рассматриваемый нами фрагмент языка геометрии достаточно богат. В нем выразимы многие нетривиальные геометрические утверждения, некоторые из которых будут приведены во второй части нашго исследования. Сознательное сокращение числа исходных понятий приводит к тому, что часть привычных геометрических представлений нуждается в корректировке, а именно, все, что можно, надлежит выражать (следуя синтаксическим правилам нашего языка) через, так сказать, синтаксические примитивы, то есть точки и прямые (см. ниже). Основное отношение между точкой и прямой – *инцидентность*, которое в русском языке выражается фразами *точка лежит на прямой, прямая проходит через точку, прямая содержит точку* и т.п. Прежде чем начать описание элементарного синтаксиса, обсудим один из вопросов, относящихся к прагматике рассматриваемого языка, а именно вопрос о его адресатах.

В ряде предыдущих наших статей о языках геометрии[3] мы говорили о том, что адресаты геометрических текстов с точки зрения владения соответствующими языками могут различаться. Есть люди, которые в равной степени владеют четырьмя языками геометрии, ориентированными на человека, а не на компьютер. Однако они предпочитают, чтобы в разных ситуациях им сначала излагали геометрический текст на одном языке, а потом на каком-то другом – это кажется им психологически более комфортным. Рассмотрим один пример. Предположим, что мы преподаем геометрию в старших классах математической школы и повторяем теорему Пифагора. Мы можем подойти к ее формулировке, используя каждый из четырех языков. И то, с чего мы начинаем изложение, зависит от очень многих факторов, в частности, от желания акцентировать какие-то аспекты, связанные с теоремой или ее доказательством, от нашего психологического настроя, от обстановки в классе и т.д. Однако при формулировке теоремы Пифагора, чтобы обеспечить полное ее понимание, мы пользуемся по возможности всеми четырьмя языками. Дело в том, что обычно при изложении геометрического текста языки не конкурируют между собой, а взаимно дополняют друг друга[4]. Иными словами, полное понимание геометрического текста предполагает владение всеми четырьмя языками.[5]

Что же касается формального языка геометрии, то он, как мы уже говорили, не предназначен для обычных людей[6]: его основным адресатом является абстрактное логическое устройство, осуществляющее автоматическую обработку информации (в настоящее время воплощением такого устройства является компьютер).

«Необычными полиглотами» представляются те адресаты, которые интересуются языками науки и их соотношением с повседневным языком. Они пытаются овладеть всеми пятью языками и научиться переводить с каждого из них на любой другой – как известно, умение переводить является важнейшим критерием понимания и языков, и тех предметных областей, которую они обслуживают. Что касается формального языка геометрии, то наличие параллельных текстов, один из которых написан на этом языке, а другой – на более «человеческом», является настоятельно необходимым для понимания геометрического текста. В свою очередь, перевод с «обычного» языка геометрии на формальный язык тоже представляется весьма важным, поскольку такой перевод позволяет эксплицировать многие неявные смыслы, содержащиеся в тексте на «обычномї языке и выявлять возможные точки непонимания. Таким образом, один из основных тезисов, который мы выдвигаем в этой (да и во всех наших работах), состоит в следующем. Полное понимание человеком данного геометрического текста обеспечивается умением человека переводить его на все пять языков геометрии и установлением всех возможных соответствий между фрагментами данного текста, написанными на любом из них.

1. Основные виды геометрических текстов

Основные виды рассматриваемых нами геометрических текстов известны со школы; это – определения, аксиомы и теоремы. Все они относятся к жанру повествовательных текстов, нарративов. Кроме того, существуют и другие виды геометрических текстов, многие из которых имеют неутвердительную модальность. Это – задачи, упражнения, проверочные тесты и др. В дальнейшем мы будем рассматривать только основные виды геометрических текстов и их формализацию. Синтаксис формального языка геометрии, о котором пойдет речь ниже, – это синтаксис таких текстов. Описать синтаксис любого языка, будь то естественный или формальный, означает описать, во-первых, множество правильных единиц данного языка, во-вторых, множество синтаксических отношений между данными единицами и, в-третьих, множество синтаксических преобразований одних синтаксических единиц в другие.

Начнем с описания множества синтаксически правильных единиц формального языка геометрии. Среди них выделяются три группы, которые называются *термы*, *предикаты* и *формулы*. Если термы и предикаты являются синтаксически не зависящими друг от друга единицами, то формулы – это такие единицы, для которых термы и предикаты служат строительным материалом.

Описание элементарного синтаксиса формального языка геометрии мы разобьем на две части. В главной части перечисляются синтаксические единицы и конструкции данного языка, некоторые синтаксические отношения и отдельные синтаксические преобразования. В дополнительной части, которую мы назвали «Комментарии», даются содержательные пояснения вводимых синтаксических понятий вместе с указанием полезных аналогий, связывающих формальный язык с естественным (русским) языком. Мы полагаем, что поиск таких аналогий является важным инструментом и критерием понимания текстов не только на формальных, но и на естественных языках. Некоторые из комментариев призваны оградить читателя от ложных аналогий, которые, к сожалению, часто возникают в процессе преподавания и обучения геометрии в школе.

2. Элементарный синтаксис формального языка геометрии

Ниже представлены синтаксические единицы, конструкции и правила так называемого <u>элементарного синтаксиса</u>; в последующей работе речь пойдет о строении <u>расширенного</u> <u>синтаксиса</u>. Это разделение аналогично тому, которое используется в курсах преподавания иностранных языков. Формальных языков геометрии существует много; наиболее известные языки в среде математиков были подробно исследованы ранее[7]. Здесь мы опишем наш собственный вариант формального языка, адаптированный, прежде всего, для лингвистов, которые интересуются языками науки, их соотношением с естественными языками и механизмами понимания текстов, написанных на языках науки. Этим объясняется обилие в данной работе содержательных комментариев, сопровождающих введение формальных конструкций. Наш формальный язык обладает следующим важным свойством: в его описании не используются никакие другие формальные языки.

Понимание дальнейшего текста, в котором излагается синтаксис данного языка, предполагает некоторое знакомство читателя с понятиями и терминами математической логики и теории множеств. Дело в том, что правила построения основных синтаксических единиц существенно используют аппарат математической логики, в основном ее раздела, называемого исчислением предикатов первого порядка. С этим, в частности, связано введение логических связок и кванторов в алфавит нашего формального языка геометрии (см. ниже).

2.1. Алфавит

Как и любое порождающее описание синтаксиса естественного языка, описание синтаксиса формального языка геометрии начинается с задания <u>алфавита</u>. Алфавит формальных языков состоит из букв и специальных значков. В нашем языке есть три вида букв:
- заглавные латинские AB ... Z;
- строчные латинские ab ... z;
- строчные латинские маркированные $\vec{a}\,\vec{b}$ ... $\vec{z}$.

Специальные значки нашего языка делятся на несколько видов:
- насечка ˌ
- два вида скобок ( ) [ ]
- запятая ,
- пять значков геометрических операций • • • ! $^{\rightarrow}$
- пять логических символов ∀ ∃ ¬ ∧ ∨

Различие между буквами и специальными значками состоит в том, что для каждой из букв существует соответствующее ей однобуквенное слово. Таким образом, в нашем языке есть три множества однобуквенных слов: $\{A; B; \ldots ; Z\}$; $\{a; b; \ldots ; z\}$ и $\{\vec{a}\,\vec{b} \ldots \vec{z}\}$.

Комментарий

В соответствии со сложившимися в математике традициями словоупотребления некоторые значки называются *знаками*,а некоторые – *сим-*

волами. Правила употребления всех знаков и букв и правила чтения содержащих их последовательностей будут описаны по отдельности.

(a) Насечка – это особый значок, который может приписываться к любым буквам в качестве нижнего индекса (это и есть правило употребления данного значка). В результате приписывания насечки тоже образуются (неоднобуквенные) последовательности или, иначе, *простые термы* нашего формального языка геометрии (см. о простых термах подробнее в следующем разделе). Помимо букв, насечка может приписываться к простому терму, в результате чего образуется буква с последовательностью насечек. Количество насечек в такой последовательности читается как номер при букве или при простом терме, например, $b_{|||}$ читается как «*b* три» или «*b* третье».

Один из аналогов насечки в русском языке – это морфема *пра*, которая, будучи многократно присоединенной к словам *бабушка* и *дедушка*, дает слова *прабабушка*, *прадедушка*, *прапрабабушка*, и т.д.

(b) В состав специальных значков нашего языка входят также два вида скобок. Они различаются своими употреблениями (соответствующие правила даны ниже) и не являются синонимичными знаками.

(c) Запятая как значок нашего языка используется в качестве разделителя элементов конечного множества. Заметим, что она нетождественна запятой естественного языка.

(d) В алфавит входят также пять знаков геометрических операций, не являющиеся общепринятыми в математике.

Знак ˙ называется *верхний кружок*, знак • называется *средний кружок*, а знак . называется *нижний кружок*. Четвертый знак ! за отсутствием лучшего названия мы будем называть *восклицательным знаком*, хотя с естественноязыковым восклицательным знаком он не имеет ничего общего. Компоненты этого знака графически напоминают о точке и прямой. Наконец, пятый знак $^{\rightarrow}$ называется *верхняя стрелка*.

(e) Знаки ∀ и ∃ – это *кванторы общностии существования*, а ¬, ∧ и ∨ – это знаки отрицания, дизъюнкции и конъюнкции.

Замечание.

Импликация определяется через дизъюнкцию и отрицание, а логическая эквивалентность – через пару импликаций. В настоящем тексте импликация и логическая эквивалентность рассматриваться не будут.

2.2. Правила построения термов

Естественно-языковыми аналогами термов являются существительные и синтаксические группы с существительным в качестве главного слова. Такие группы обычно называют *именными группами*. Естественноязыковые аналоги предикатов – это глаголы, краткие формы прилагательных (*предикативы*) и *глагольные* (*предикативные*) *группы*, или *предикации*.

Термы делятся на два вида – *простые термы* и *сложные термы*, то есть не простые. *Простыми термами* мы будем называть, во-первых, однобуквенные слова, соответствующие буквам нашего алфавита, а, во-вторых, слова, состоящие из букв с приписанными к ним одной или несколькими насечками. Примерами простых термов служат термы A; b; $\vec{v}$. Простые термы называются также *переменными*.

Определение сложных термов дается в два этапа: сначала определяется понятие *терма*, а *сложным термом* называется терм, не являющийся простым. Определение термов основывается на важном понятии *сорта терма*.

Множество термов нашего языка разбивается на три непересекающихся подмножества – *точки*, *прямые* и *векторы*. В так называемом *многосортном исчислении предикатов первого порядка* им соответствуют три сорта термов. Простые термы соответствующих сортов обозначаются латинскими буквами (без насечек или с насечками) – заглавными, строчными и строчными маркированными. Поскольку простые термы называют также переменными, сорт простого терма называют <u>сортом переменной</u>.

Комментарий

Если понятия *точка* и *прямая* являются неопределяемыми, то понятие *вектор*, которое у нас введено как неопределяемое, на самом деле может быть определено. Поясним, как это делается, для читателей, знакомых с понятиями плоскость и отображение. Вводятся понятие и термин *преобразование плоскости*. Под преобразованием плоскости понимается взаимно однозначное отображение плоскости в себя, при котором все прямые отображаются в прямые. *Вектором* называется такое преобразование плоскости, при котором либо никакая точка плоскости не переходит в себя, либо все точки плоскости переходят в себя (в последнем случае преобразование называется тождественным). Поскольку под вектором мы понимаем некоторую трансформацию, имеет смысл говорить о результате этой трансформации, то есть имеет смысл выражение вектор, примененный к точке. Формула $\vec{v} \bullet P = Q$ читается как *вектор v, примененный к точке P, дает точку Q* или как *вектор $\vec{v}$ переводит точку P в точку Q*.

Термы определяются рекурсивными правилами, а именно:
- Простые термы всех трех сортов суть термы
- Если $\mathfrak{A}$ и $\mathfrak{B}$ – две различные точки, то $(\mathfrak{A}^{\bullet}\mathfrak{B})$ – прямая
- Если $\vec{\mathfrak{b}}$ – вектор, а $\mathfrak{A}$ – точка, то $(\vec{\mathfrak{b}} \bullet \mathfrak{A})$ – точка
- Если $\mathfrak{a}$ и $\mathfrak{b}$ – две различные точки, то $(\mathfrak{a} \bullet \mathfrak{b})$ – точка
- Если $\mathfrak{A}$ – точка, а $\mathfrak{b}$ – прямая, то $(\mathfrak{A}!\mathfrak{b})$ – прямая
- Если $\mathfrak{A}$ и $\mathfrak{B}$ – точки, то $(\mathfrak{A}^{\rightarrow}\mathfrak{B})$ – вектор

Две части сложного терма, стоящие внутри круглых скобок и разделенные знаком операции, называются *аргументами сложного терма*.

Например, аргументы сложного терма $(\mathfrak{A}!\mathfrak{b})$ – это $\mathfrak{A}$ и $\mathfrak{b}$.

Сложные термы нашего языка имеют один из следующих пяти видов: $(\mathfrak{A}^{\bullet}\mathfrak{B})$, $(\vec{\mathfrak{b}} \bullet \mathfrak{A})$, $(\mathfrak{a} \bullet \mathfrak{b})$, $(\mathfrak{A}!\mathfrak{b})$ и $(\mathfrak{A}^{\rightarrow}\mathfrak{B})$. Два из них – точки, два – прямые и один – вектор.

Комментарий

1. В приведенных правилах готическими буквами кодируются последовательности знаков нашего алфавита, которые обозначают точки, прямые и векторы (т.е. сложные термы). К буквам готического алфавита применимо то же соглашение, что и к буквам латинского алфавита (см. о нем в разделе «Правила построения термов»).

2. Во введенных определениях термов неявно учтены три положения о существовании и единственности, которые соответствуют двум аксиомам и одной теореме евклидовой планиметрии.[8] Эти аксиомы таковы: *Через две точки можно провести прямую, причем единственную* и *Через точку вне данной прямой можно провести прямую, параллельную ей, причем единственную*. А теорема гласит: *Любые две непараллельные прямые пересекаются в единственной точке*.

3. Два слова – *точка* и *прямая*, – выбранные для обозначения сортов термов, соответствуют интуиции, сложившейся у читателя с ранних лет, а с третьим словом – *вектор* – современные люди знакомятся в школе. Сами понятия и термины «точка», «прямая» и «вектор» у нас считаются неопределяемыми.

4. Точку как сорт термов необходимо отличать от точки как знака препинания естественного языка и от слова, омонимичного данному слову. Слово *точка* в другом значении встречается, например, во фразе *Сказал, что сделаю, – и точка*.

Поскольку простые термы иначе называются переменными (см. выше), а сложные термы состоят из простых, можно говорить, что с каждым сложным термом связано *множество переменных*, входящих в него. Это множество представляет собой теоретико-множественное объединение переменных, входящих в составляющие сложного терма.

В перечисленных выше правилах построения термов заглавными и строчными готическими буквами обозначаются, соответственно, любые из латинских заглавных и строчных букв и, кроме того, сложные термы всех трех сортов. Например, выражение $((P!a_{||}) \bullet b)$, представляющее собой сложный терм, построено по правилу «Если a и b – прямые, то $(a \bullet b)$ – точка». Данное выражение означает «точка пересечения прямой, проходящей через точку P параллельно прямой a-два, с прямой b».

Круглые скобки являются маркерами сложных термов. Иными словами, в круглых скобках стоят сложные термы, и наоборот, сложные термы всегда заключаются в круглые скобки. Круглые скобки в синтаксических образованиях нашего языка геометрии не читаются.

Верхняя стрелка соответствует привычному со школы обозначению вектора $\vec{AB}$, где точку A называют началом вектора, а B – концом, и при этом никакого смысла стрелке как знаку не придается. Для нас же верхняя стрелка – это семантически нагруженный знак, обозначающий результат операции, которая сопоставляет паре точек некоторый сдвиг плоскости. Тем не менее мы предлагаем читать запись $\vec{AB}$ не только как «А верхняя стрелка В», но и более привычно как «вектор АВ». Особо отметим, что в результате применения правил синтеза сложных термов из простых образуются термы тех же трех сортов, причем сорта результатов применения операций отличаются от сортов операндов.

Примеры термов

<u>Термы сорта «точка»</u>	<u>Термы сорта «прямая»</u>
P	ℓ
$(l \bullet m)$	$(P^{\bullet} Q)$
$((\mathrm{A}!a_{\|\|}) \bullet m_{\|\|})$	$(Z_{\|\|\|}!(P_{\|}(A!a_{\|\|})))$

Комментарии

1. Каждая из приведенных групп термов содержит три примера. Первые примеры в каждой группе – это простые термы сорта «точка» (соответственно, сорта «прямая»), а второй и третий примеры – это сложные термы сорта «точка» («прямая»). Сложные термы являются аналогами синтаксически сложных номинативных сочетаний русского языка.

2. Простые термы читаются очевидным образом, а сложные читаются при помощи следующего приема: в структуре сложного терма ищется «главныйї знак, и чтения начинается с этого знака. Например, терм $((\mathrm{A}!a_{||}) \bullet m_{||})$ читается как *точка пересечения двух прямых, одна из которых – прямая, проходящая через точку А параллельно прямой а-два, а вторая прямая – это прямая т-два.*[9]

<u>Термы сорта «вектор»</u>

Поскольку для нашего читателя геометрическое понятие вектора, по всей видимости, не столь привычно, как понятия прямой или точки, мы выделяем примеры термов – имен векторов в отдельную группу:

$$\vec{a} \qquad (P^{\rightarrow} R) \qquad (Z^{\rightarrow}_{|||}\,(P^{\bullet}_{||}\,(A!a_{||})))$$

2.3. Правила употребления и особенности чтения термов нашего языка

Правила употребления термов применительно к разным видам знаков читателю известны в разной степени. Прежде всего, отметим, что в составе группы сложных термов есть единицы, которые можно переставлять местами. Перестановочность единиц оговаривается отдельными правилами.

Хотя заглавные буквы у нас используются для обозначения точек, те же точки, как можно увидеть из правил образования сложных термов, могут обозначаться и более громоздко. Например, точка пересечения прямых a и b обозначается как $(a \bullet b)$. В отличие от простых термов сорта «точка» для сложных термов того же сорта используются знаки «нижний кружок» и круглые скобки. То же касается обозначений прямых и векторов. Правило прочтения букв с насечками таково: $A_{|}$ читается как «*a* с одной насечкой», как «*a* один» или как «*a* первое»; $b_{||}$ – как «*b* с двумя насечками», «*b* два» или «*b* второе»; $\vec{c}_{|||}$ – как «вектор *c* с тремя насечками», «вектор *c* три» или «вектор *c* третье» и т.п.

Верхний кружок $^{\bullet}$ разрешается ставить между несовпадающими точками, ср., например, обозначение прямой $(P^{\bullet}Q)$, которое читается неидиоматично как «P – верхний кружок – Q», а идиоматично – как «прямая, проходящая через точки *P* и *Q*» или «прямая *PQ*». Здесь буквы *P* и *Q* перестановочны.

Средний кружок • ставится между вектором и точкой. Так, запись $(\vec{v} \bullet P)$ читается буквально как «*v* – средний кружок – *P*», а, с учетом данного выше определения вектора (см. выше), – как «вектор *v*, примененный к точке *P*». Вектор и точка в записи $(\vec{v} \bullet P)$ неперестановочны.

Нижний кружок $_{\bullet}$ ставится между несовпадающими прямыми. Например, обозначение точки $(l \bullet m)$ читается неидиоматичным образом как «*l* – нижний кружок – *m*», а идиоматично – как «точка пересечения прямых *l* и *m*». Здесь единицы *l* и *m* перестановочны.

Восклицательный знак ! ставится между точкой и прямой в указанном порядке. Например, запись $(P!l)$ читается как «*P* – восклицательный знак – *l*».

Стрелка $^{\rightarrow}$ ставится между двумя точками. Запись $(P^{\rightarrow}Q)$ читается неидиоматично как «*P* – стрелка – *Q*», или как «вектор из *P* в *Q*», а идиоматично – как просто «вектор *PQ*».

Скобки называются обычным образом: *круглые* и *квадратные*. При чтении они обычно опускаются.

Комментарий (о мнемонике)

Мы старались по возможности обозначать точки, прямые и вектора начальными буквами соответствующих английских слов *point*, *line*, *vector* и близкими к ним.

В нашем алфавите есть три вида основных значков – кружки, восклицательный знак и стрелка. При этом значимым является такой признак кружка, как его расположение. Оно бывает трех видов: верхнее, среднее и нижнее. Запомнить, какой сорт терма образуется с их помощью, можно так: верхний кружок ставится между точками, и этим он напоминает вектор в записи, близкой к традиционной, нижний кружок – по некоторому принципу дополнительности – ставится между двумя прямыми. Остается запомнить лишь расположение среднего кружка – он используется в комбинации знаков, наименее привычной для читателя (вектор, примененный к точке, есть точка).

2.4. Предикаты и формулы

Двумя другими видами синтаксических единиц нашего формального языка геометрии являются *предикаты* и *формулы*. Естественно-языковыми аналогами предикатов являются глаголы и краткие формы прилагательных, или *предикативы*, а аналогами формул – предикаты с заполненными аргументами (местами), или *предикации*. В частности, предикации могут быть полными фразами. Предикаты можно представлять себе как предикации с незаполненными местами.

Формулы, так же как предикаты и как термы, бывают *простыми* и *сложными* (не простыми). Простых предикатов всего два – это *равенство* и *инцидентность*, обозначаемые, соответственно, = и ∈. Им соответствуют два вида простых формул, причем одна из них имеет два варианта. Простые формулы строятся из простых предикатов путем заполнения их мест и имеют один из трех видов:
$\mathfrak{A} = \mathfrak{B}$ (равенство точек),
$\mathfrak{l} = \mathfrak{m}$ (равенство прямых) и
$\mathfrak{P} \in \mathfrak{l}$ (инцидентность точки и прямой).

Примеры простых формул

$\mathrm{P} = \mathrm{Q}$
$(l_\bullet m) = p \ (l_\bullet m) \in l \ (l_\bullet m) \in (P^\bullet Q)$

Правила чтения и употребления знаков = и ∈ достаточно прихотливы. Сначала о знаке = . Он может читаться *равен*; *равно*; *равняется* или при помощи других родственных слов, и он же может читаться иначе: *совпадает*; *совмещается*; *является тем же <самым> и под.*, а также при помощи слов и выражений, синонимичных указанным словам. Теперь о знаке ∈. Он тоже может читаться двумя способам: при помощи несимметричного и симметричного предикатов. В первом случае запись $P \in l$ канонически читается как «точка P принадлежит прямой l» или «точка P лежит на прямой l»; но ее же можно читать «инвертированно», а именно как «прямая l проходит через точку P».[10] Во втором случае запись $P \in l$ читается с помощью симметричного предиката инцидентность, а именно как «точка P и прямая l инциденты», или как «прямая l и точка P инциденты».

Сложные предикаты и сложные формулы строятся из простых применением к ним *синтаксических операций* двух классов.

Первый класс таких операций носит название *булевых*[11], среди которых мы выделяем три основные: унарную операцию *отрицание* $\neg$ и две бинарные операции – *дизъюнкция* $\wedge$ и *конъюнкция* $\vee$.[12] При применении булевых операций их аргументы (то, к чему они применяются) заключаются в квадратные скобки.

Приведем примеры сложных формул, построенных при помощи булевых операций. Формула $\neg[P = Q]$ читается как «неверно, что точки P и Q совпадают».[13] Формула $[p \in l] \vee [\neg[P = Q]]$ читается неидиоматично как «точка P лежит на прямой l, или неверно, что точки P и Q совпадают», а идиоматично как «одно из двух: либо точка P лежит на прямой l , либо неверно, что точки P и Q совпадают».

Комментарий

Некоторые сложные формулы, построенные с помощью перечисленных выше операций, допускают полезные сокращения, или *редукции*.[14] Так, формулу $\neg[P = Q]$ можно записать в виде $P \neq Q$, которая читается как «точки P и Q не совпадают», или, более идиоматично, «точки P и Q – разные», а формулу $\neg[P \in l]$, которая читается как «неверно, что точка P и прямая l инцидентны» – в виде $P \notin l$; последняя формула может читаться как «точка P и прямая l неинцидентны», или, иначе, «точка P не лежит на прямой l» или, наконец, «прямая l не проходит через точку P».

Второй класс операций, с помощью которых сложные предикаты и сложные формулы строятся из простых, носят название *кванторных*. В математической логике применение этих операций называют *навешиванием кванторов*, а в лингвистике соответствующие операции называют *квантификацией*. Напомним, что в алфавите нашего формального языка геометрии есть два знака $\forall$ и $\exists$, обозначающих, соответственно, кванторы общности и существования. Для определения операций навешивания кванторов, для установления равносильности кванторных выражений, а также для определения одного вида редукции формул, нам понадобятся некоторые вспомогательные понятия. Первое из них – это понятие *множества переменных, входящих в формулу*. Оно опирается на введенное ранее понятие множества переменных, входящих в терм.

Комментарий

Хотя в формальном языке геометрии простые термы суть переменные, в нем обычно не принято использовать понятия «терм, входящий в данную формулу» и «множество термов, входящих в данную формулу».

Множество переменных, входящих в простую формулу, определяется как объединение множеств аргументов всех термов, входящих в нее. Например, множество переменных, входящих в формулу $P \in l$, – это $\{P; l\}$ а множество переменных, входящих в формулу $P = P$ – это $\{P\}$.

Множество переменных, входящих в сложную формулу, определяется отдельно для унарной операции отрицания и отдельно для бинарных операций дизъюнкции и конъюнкции. Операция отрицания не меняет множества переменных формулы, к которой она применяется, а при использовании бинарных операций множества переменных, входящих в аргументы операций, объединяются. Переменные, входящие в данную формулу, делятся на *свободные* и *связанные*. Соответствующие определения даются рекурсивно. Приведем их. Все переменные, входящие в простые формулы или в сложные формулы, получаемые из простых булевыми операциями, называются *входящими в них свободно*, или *свободными в них*.

Операции навешивания каждого из кванторов могут применяться к паре $< \mathcal{F}, \mathcal{E} >$, состоящей из формулы $\mathcal{F}$ и свободной переменной $\mathcal{E}$ в ней; здесь $\mathcal{E}$ обозначает переменную любого из трех рассматриваемых нами сортов, то есть точек, прямых и векторов.

В результате применения операций навешивания кванторов возникают формулы $\forall\mathcal{E}[\mathcal{F}]$ и $\exists\mathcal{E}[\mathcal{F}]$.

Комментарий

Квадратные скобки служат здесь для формирования так называемых *областей действия кванторов* и отделения их от кванторных приставок, то есть последовательности «квантор-переменная». Таким образом, квадратные скобки у нас используются в двух случаях – как маркеры операндов булевых операций и как маркеры областей действия кванторов.

2.5. Примеры формул с кванторами

$\forall P[\forall Q[[P = Q] \vee [[P \in l] \wedge [P \in l]]]]$ – «каковы бы ни были две точки, либо они совпадают, либо через них проходит некоторая прямая» (это часть первого постулата Евклида без требования единственности прямой, проходящей через две точки).

$\exists P[\neg[P \in l]]$ – «относительно данной прямой верно, что существует точка вне ее» (в этой формуле квантором связана переменная P, а переменная l – свободная).

***

На этом описание основных единиц элементарного синтаксиса нашего формального языка геометрии закончено. Перейдем теперь к описанию синтаксических отношений на множестве формул построенного языка.

3. Синтаксические отношения

Не претендуя на полное описание синтаксических отношений, дадим определения двух важнейших. Они играют особую роль при сопоставлении формального языка с естественным и при установлении соответствий между их фрагментами.

3.1. Отношение «быть подформулой»

Формула $\mathcal{B}$ называется подформулой формулы $\mathcal{A}$, если $\mathcal{A}$ построена из $\mathcal{B}$ и, возможно, других формул, по одному из перечисленных выше правил построения формул. Например, формула $\exists P[P \in l]$ содержит подформулу $P \in l$, формула $\neg[(\vec{v} \bullet A) = B]$ содержит подформулу $(\vec{v} \bullet A) = B$, а формула $[P \in l] \wedge [P \in m]$ содержит две подформулы: $[P \in l]$ и $[P \in m]$.

Комментарий

В естественно-подобном языке геометрии подформулам соответствуют такие фрагменты предложений, как причастные и деепричастные обороты, простые предложения (в составе сложного) и вставные предложения и конструкции, как во фразе «Он сказал: “Вернусь поздно” и ушел».

3.2. Отношение синтаксической эквивалентности

Не приводя определения отношения синтаксической эквивалентности[15], опишем идеи, положенные в его основу. Определение строится в несколько шагов. На первом шаге каждая из двух формул, синтаксическая эквивалентность которых устанавливается, приводится к так называемой предваренной нормальной форме, то есть такой, у которой все кванторы предшествуют всем предикатам. Это всегда возможно[16]. Далее у двух формул сравниваются кванторные приставки. Подпишем их одну под другой (переменная под переменной, скобка под скобкой). Если хотя бы на одном из соответствующих мест кванторы не совпадают, то формулы синтаксически неэквивалентны. Пусть теперь все кванторы на соответствующих местах совпали. В этом случае существует процедура переименования переменных, превращающая кванторные приставки двух формул в тождественные. Эта же процедура применяется к оставшейся (бескванторной) части формулы. Отбрасывая ставшие одинаковыми кванторные приставки, мы сводим задачу установления синтаксической эквивалентности исходных формул к задаче так называемой *алгебры высказываний*. Она решается многими хорошо известными в логике способами, в частности, выяснением возможности преобразования одной формулы в другую по законам булевой алгебры, то есть действуя чисто синтаксически.

4. Синтаксические операции

К формулам, состоящим из элементарных синтаксических единиц, применяются некоторые синтаксические операции.

Один их класс состоит из операций замен формул на синтаксически эквивалентные им, обладающие какими-либо заранее заданными свойствами. Это принадлежность к каким-либо стандартным формам (предваренная нормальная формы, вид «посылка ⇒ заключение» и др.), краткость, понятность, легкость перевода на другие языки и т.п. Некоторые из этих операций носят не формализованный характер, между тем, они широко используются в разного рода лингвистических исследованиях[17].

Другой класс операций – это преобразование линейной записи формул в соответствующие им синтаксические деревья.

Третий класс операций состоит в расширении лексики исходного языка за счет замены его фрагментов новыми знаками. Такие операции мы подробно рассмотрим в следующей работе. В ней, также посвященной синтаксису формального языка геометрии, мы приведем примеры более сложных синтаксических единиц с целью показать их разнообразие и возможность выразить с их помощью нетривиальные геометрические смыслы.

Заключение

Несмотря на кажущуюся сложность или, по крайней мере, непривычность нашего изложения для лингвиста элементарного синтаксиса формального языка геометрии, оно имеет параллели в целом ряде формальных лингвистических теорий. К ним относятся такие инструменты описания, как порождающие грамматики и семантики в духе грамматик Н. Хомского и его последователей[18]. Это также интерпретативная семантика Катца и Фодора[19], порождающая семантика Лакоффа[20] и др. К упомянутым исследованиям примыкают также работы Е.В. Падучевой[21] и ее сотрудников, А. Вежбицкой[22] и ее коллег, и еще многие другие работы.

Отличительной особенностью описанного нами фрагмента языка геометрии является предельная минимальность набора исходных синтаксических понятий и выражающих их слов (*точка, прямая; равенство, инцидентность*). Благодаря принятой в геометрии еще со времен Евклида установке на выбор как можно меньшего числа неопределяемых понятий все многообразие пространственных категорий и явлений описывается крайне экономными средствами. Вместе с тем такая минимальность на практике приводит к большей сложности восприятия синтаксических единиц и их анализа. Лингвистам это хорошо известно, например, из работ А. Вежбицкой и ее учеников и последователей. А именно, осознанное стремление к минимизации числа семантических атомов (примитивов) приводит к весьма сложным по своей синтаксической структуре текстам толкований.

Альтернативный подход к построению элементарного синтаксиса языка геометрии мог бы быть связан с увеличением числа исходных понятий. В качестве таковых к перечисленным выше понятиям можно было бы добавить количественные понятия и термины *расстояние, площадь, величина угла* и качественные концепты и термины *отрезок, луч, окружность* и др. Помимо имен, в качестве основных могли бы быть введены предикаты равенства = и неравенства <, аргументами которых являются не геометрические объекты, а числа.

Именно такой подход к изложению оснований геометрии был принят со времен Евклида. Он позволяет соотнести синтаксические единицы языка геометрии с синтаксическими конструкциями естественного языка. Обнаружение синтаксического параллелизма между конструкциями этих языков, несомненно, облегчает понимание текстов на языке геометрии и их перевод на естественно-подобный язык.

Поскольку основным адресатом формального языка геометрии является, как мы говорили, не человек, а компьютер (или подобное ему устройство), для него такой трудности не существует. В отличие от человека, компьютер легко обращается с длинными последовательностями знаков и легко разбивает их на синтаксически правильные подпоследовательности. Это разбиение является ступенью на пути к полному грамматическому разбору синтаксических сочетаний и фраз анализируемого языка. Математика нашла компромисс между установкой на минимальность множества исходных понятий и сложностью синтаксических правил построения геометрических текстов. Некоторое расширение этого множества и введение содержательных производных понятий позволяет упростить изложение языка геометрии и его понимание человеком. Во второй части этой работы, посвященной расширенному синтаксису формального языка геометрии, мы опишем некоторые из таких производных понятий, на основе которых целый ряд содержательных геометрических результатов излагается и проще, и понятнее.

Примечания

1 *Крейдлин Г.Е., Шабат Г.Б.* Теорема как вид текста I. Понятность // Вестник РГГУ. Серия Языкознание/МЛЖ. 2007, № 8. С. 102-112.; *Крейдлин Г.Е., Шабат Г.Б.* Теорема как вид текста II. Когнитивные операции над формулировками теорем // Вестник РГГУ. Серия Языкознание/МЛЖ. 2011, № 11. С. 241-270; *Крейдлин Г.Е., Шабат Г.Б.* Пространство в естественных языках и языках геометрии // Вестник РГГУ. Серия «История. Филология. Культурология. Востоковедение». 2015. №1. С. 116-130. (Московский лингвистический журнал, Том 17 (1)). *Крейдлин Г.Е., Шабат Г.Б.* Естественный язык и язык геометрических чертежей: точки соприкосновения // Znaki czy nie znaki? / Red naukowa J. Piatkowska, G. Zeldowicz. Warszawa, 2016. S. 197-221.

2 *Гладкий А.В., Крейдлин Г.Е.* Математика в гуманитарной школе // Математика в школе. 1991. № 6. С. 6-9; *Звонкин А.К.* Абстракции с языковой поддержкой // Язык и структура знания. М., 1990. С. 86-95; *Корельская Т.Д., Падучева Е.В.* Обратная теорема (алгоритмические и эвристические процессы мышления). М., 1978; *Крейдлин Г.Е., Шмелев А.Д.* Языковая деятельность и решение задач // Математика в школе, 1989, № 3. С. 39-45; *Крейдлин Г.Е., Шмелев А.Д.* Математика помогает лингвистике. М.: Просвещение, 1994; *Манин Ю.И.* Математика как метафора. М., 2008.

3 См. упомянутые ранее статьи авторов «Теорема как вид текста» I, II.

4 Там же.

5 Этот тезис мы сознательно формулируем в сильной форме. Возможные его ослабления связаны не только со структурой и содержанием самого текста, но и с различными экстралингвистическими, в частности, социальными, факторами.

6 Это обстоятельство необходимо учитывать при преподавании тех разделов математики, которые связаны с формализацией рассуждений. Особенно это относится к преподавателям, ведущим занятия в гуманитарной аудитории.

7 *Гильберт Д.* Основания геометрии. М.-Л., 1948; *Tarski A.* What is elementary geometry? // The axiomatic method. With special reference to geometry and physics : Proceedings of an International Symposium / Eds. L. Henkin, P Suppes., A. Tarski. Amsterdam, 1959. P. 16-29 (Studies in Logic and the Foundations of Mathematics, 27).

8 Под *евклидовой планиметрией* здесь подразумевается раздел геометрии, построенный на аксиомах, сформулированных Евклидом. Они формулируются в школьной геометрии и дают представление об аксиоматическом построении научной теории.

9 Поскольку формулы языка геометрии человек, вообще говоря, читать не должен, способов прочтения более длинных термов мы не приводим.

[10] Вообще говоря, фразы «точка P лежит на прямой l» и «прямая l проходит через точку P» различаются актуальным членением. В первой фразе именная группа «точка P» является темой предложения, а группа «прямая l» – ремой, а во второй фразе – наоборот (вторую фразу иногда записывают в виде «$l \ni P$», но мы не ввели значок $\ni$ в наш алфавит). Наш язык не обладает формальными средствами для выражения актуального членения.

[11] Операции названы так в честь Джорджа Буля (1815-1864).

[12] Знаки $\neg$, $\wedge$, $\vee$ называют также *логическими связками*.

[13] Напомним, что, в соответствии с введенным ранее соглашением, заглавными латинскими буквами обозначаются точки, а строчными – прямые.

[14] Подробнее о редукциях разных видов будет говориться во второй части данной работы.

[15] См. *Падучева Е.В.* О семантике синтаксиса. Материалы к трансформационной грамматике русского языка. М., 1974.

[16] См. *Крупский В.Н., Плиско В.Е.* Математическая логика и теория алгоритмов. М., 2013.

[17] *Корельская Т.Д., Падучева Е.В.* Указ. соч.

[18] *Тестелец Я.Г.* Введение в общий синтаксис. М., 2001.

[19] *Katz J J., Fodor J.A.* The structure of a semantic theory // Language. 1963. №39(2). P. 170-210.

[20] *Лакофф Дж.* Когнитивная семантика // Язык и интеллект. М., 1996.

[21] *Падучева Е.В.* Указ. соч.

[22] *Вежбицкая А.* Язык. Культура. Познание. М., 1996.