

МОДЕЛИ КОГНИТИВНЫХ ПРОЦЕССОВ (Вычислительные системы)

1997 год

Выпуск 158

УДК 519.766.2

АЛГЕБРАИЧЕСКОЕ ОПИСАНИЕ СМЫСЛА ВЫСКАЗЫВАНИЙ ЕСТЕСТВЕННОГО ЯЗЫКА ¹

Д.Е.Пальчунов

В работе исследуется понятие смысла текста и связь смысла с контекстом. Рассматриваются различные осмысливания (т.е. приписывания смысла) высказываний, являющихся логически эквивалентными. Строится алгебра таких высказываний, являющаяся частным случаем алгебры высказываний естественного языка [18]. Построенная алгебра, называемая алгеброй контекстов, является естественным образом изоморфной прямому произведению алгебры Линденбаума-Тарского исчисления предикатов данной сигнатуры и булевой алгебры всех подмножеств сигнатуры.

С помощью алгебры контекстов предложен метод работы с множествами высказываний, содержащими противоречивую информацию. Несмотря на то, что этот метод рассматривает высказывания не как синтаксические, а как семантические объекты, отождествляемые со своим содержанием, он позволяет избежать переноса противоречия из некоторого фрагмента информации на всю информацию.

На языке алгебры контекстов проводится логический анализ метода ГАБЕК [11-17]. Метод ГАБЕК разделяется

¹Работа поддержана РФФИ, грант № 96-06-80970.

на две части — синтаксическую, которая проделывается автоматически компьютером, и семантическую, проделываемую экспертом. В предложенной модели метода ГАБЕК синтаксическая часть метода имеет дело только со вторым прямым сомножителем алгебры высказываний естественного языка, а семантическая часть — только с первым. Текст, описываемый иерархией лингвистических образов, представляется в виде наследственно конечного множества — элемента надстройки наследственно конечных множеств над сигнатурой текста (рассматриваемой как множество символов).

§1. Классификация высказываний естественного языка

В работе [18] автором была предложена классификация высказываний естественного языка, и на её основе была построена алгебра высказываний естественного языка. Кратко изложим основные идеи этой классификации и рассмотрим её с точки зрения смысла предложений и их контекста.

Вслед за Сёрлом и Вандервекеном [2, 5, 6-9] мы рассматриваем высказывания естественного языка с точки зрения цели, с которой они произносятся. В связи с этим высказывания называются речевыми действиями.

Речевые действия мы разделяем на простейшие и составные. Одна фраза естественного языка может выражать как простейшее речевое действие, так и достаточно сложное составное.

Цель речевого действия мы определяем в контексте диалога, в котором участвуют два собеседника. У каждого участника диалога мы выделяем во-первых знания, во-вторых, представление о том, что происходит в момент общения, и, в-третьих, побуждения, намерения, желания. Под знаниями собеседника мы подразумеваем всю его систему представлений, его сведения, мнения, заблуждения и т.д. Представление о том, что происходит в момент общения, — это контекст, в котором человек воспринимает слова собеседника: это смысл происходящего в данный момент, цель диалога и т.д.

Простейшие речевые действия делятся на три типа: сообщения, заявления и побуждения. Цель сообщения — изменение знаний собеседника, цель заявления — изменение смысла диалога (как его воспринимает собеседник), цель побуждения — изменение намерений собеседника. Отметим, что при этом не предполагается, что измененные знания, смысл или намерения должны обязательно отличаться от тех, которые имели место до речевого действия. Так, сообщая информацию, человек может не ведать, что его собеседник эту информацию уже знает; более того, во многих ситуациях (например, на лекции, на которой присутствуют много слушателей разной степени осведомлённости) говорящего может не интересовать, известна его информация или нет. Таким образом, сообщение обращено к прошлому (мы ничего не можем *знать* о будущем, все наши сведения о будущем на самом деле являются сведениями о прошлом), заявление — к настоящему, а побуждение — к будущему.

В качестве иллюстрации трех типов простейших речевых действий можно рассмотреть команду машины Тьюринга: если головка в некотором состоянии видит на ленте ноль (или единицу), то она переходит в новое состояние, пишет нечто в ячейке и движется влево, вправо или остается на месте. Здесь указание перейти в новое состояние можно считать заявлением (поскольку будет меняться состояние головки, которое мы можем рассматривать как представление машины о происходящем в данный момент времени, так как от него зависит, как машина Тьюринга будет воспринимать дальнейшую информацию), требование записать что-то в ячейку — сообщением (действительно, в результате записи меняется лента, которую мы можем представлять как знания машины Тьюринга), а предписание перейти влево, вправо или остаться на месте — побуждением.

Другой формальный пример — программа, написанная на АЛГОЛе, где вначале находятся описания, затем идут команды, а заканчивается программа данными. Описания естественно считать заявлениями, поскольку

они определяют смысл объектов, с которыми далее программа будет иметь дело. Команды — это побуждения, а данные, очевидно, являются сообщениями.

Сложное речевое действие может быть комбинацией нескольких простейших речевых действий одного, двух или всех трёх типов. Приведём примеры.

Приветствие обычно является чистым заявлением. Человек демонстрирует, что в данный момент он собеседника видит и хорошо к нему относится (или, наоборот, прохладно — в зависимости от конкретного содержания приветствия).

Просьба — это комбинация заявления (что имеет место просьба) и побуждения (сделать требуемое действие), либо комбинация заявления, побуждения и сообщения (например, почему нужно сделать эти действия, или почему их должен сделать именно собеседник). Просьба может быть чистым заявлением, когда просящий знает, что его просьба не будет удовлетворена, но хочет зафиксировать отказ (в этом случае имеет принципиальное значение наличие заявления, т. е. явное указание на смысл фразы — это не что-нибудь, а именно просьба). Мы считаем, что нельзя назвать просьбой побуждение сделать что-либо или сообщение о неприятностях, — в случае, если собеседник "не побудится", нельзя его обвинить в отказе выполнить просьбу, поскольку она не была сформулирована в явном виде.

Чистое побуждение отличается от приказа, который является либо комбинацией заявления и побуждения, либо чистым заявлением (и в одном, и в другом случае может быть добавлена составляющая часть сообщения). Например, когда майор и капитан играют в карты, слова одного из них "Раздай карты" являются чистым побуждением. Но фраза "Пойди проверь караулы и доложи мне" является приказом — комбинацией побуждения и заявления. Её может сказать майор капитану, но не может сказать капитан майору. Здесь сама фраза явным образом подразумевает, что происходит ни что иное, как отдача приказа старшим по званию. В этот момент смысл диа-

лога меняется. До этого общение было праздным времяпровождением, теперь оно превратилось в деловое, служебное, подчиненное определённым формальным правилам. Заметим, что если следующей фразой майора будет: "Ещё есть в запасе пять минут, давай доиграем партию", то эта фраза безусловно также будет комбинацией заявления и побуждения. Действительно, смысл диалога опять меняется, он опять превращается в дружеское времяпровождение, и побуждение, которое высказал, майор уже не является приказом.

Обратим внимание, что наша трактовка приказа принципиально отличается от формулировки Сёрла и Вандервекена [7-9]. Они определяют приказ, как и другие речевые действия, чисто синтаксически. Например, фраза "Я приказываю тебе пойти проверить караулы и доложить мне" является приказом, так как содержит словосочетание "Я приказываю", а фраза "Пойди проверь караулы и доложи мне" не является, т.к. слова "приказываю" не содержит; фраза же "Я сообщаю, что тебе нужно пойти проверить караулы и доложить мне", с точки зрения Сёрла и Вандервекена, является сообщением. Вообще, попытка чисто синтаксическими средствами дать классификацию речевых действий заранее обречена на неудачу хотя бы потому, что само понятие речевого действия очевидно содержит не только синтаксическую, но и семантическую, и, более того, прагматическую компоненты.

Кроме того, чисто синтаксическая интерпретация предложений естественного языка подразумевает полную фиксацию смысла предложений, контекста диалога. Однако, насколько речевое действие не является бессодержательным, настолько же оно меняет контекст следующих за ним предложений и, следовательно, их смысл. Поэтому мы не пытаемся свести классификацию предложений естественного языка к их синтаксическому виду.

Тем не менее, есть способ связать синтаксическую классификацию предложений естественного языка с классификацией речевых действий. Так же, как это сделали Р.Бэндлер и Д.Гриндер в книге "Структура

магии. Г" [3], можно рассматривать предложение естественного языка как поверхностную структуру, связанную с некоторой глубинной структурой (в смысле Н.Хомского [1]). Таким образом, по конкретной фразе естественного языка можно восстановить его глубинную структуру — предложение, синтаксический вид которого полностью (в определённых рамках, естественно) определяет его смысл и, в частности, его место в классификации. Например, в качестве глубинной структуры фразы "Пойди проверь караулы и доложи мне" можно рассмотреть предложение "Я, как твой непосредственный начальник, приказываю тебе пойти проверить караулы и доложить мне". Последнее предложение будет синтаксическим приказом в смысле Сёрла и Вандервекена.

Восстановление глубинной структуры по поверхностной не может быть сделано синтаксически. Данной поверхностной структуре может соответствовать много глубинных, и наоборот, данной глубинной структуре соответствует много поверхностных. Глубинная структура — это в точности способ восприятия, смысл данного предложения, который зависит от конкретного воспринимающего человека. Гриндер и Бэндлер [3] показали, что искажения и потеря информации при переходе от глубинной структуры к поверхностной являются причиной большинства психологических проблем. Более того, с их точки зрения, сущность любого "волшебства" — правильное восстановление глубинной структуры по поверхностной.

Важно также отметить следующее. Предположим, что некоторый человек произнёс в диалоге несколько фраз подряд. Предположим, что можно (конечно, достаточно условно) каждой фразе поставить в соответствие её глубинную структуру. Поскольку в момент произнесения этого монолога контекст, смысл (точнее, способ приписывания смысла) у говорящего не менялся, то смысл каждой фразы однозначно восстанавливается по её глубинной структуре. Следовательно, при таких ограничениях

синтаксическая классификация Сёрла и Вандервекена будет работать достаточно хорошо.

Итак, один из способов представления смысла и контекста высказывания естественного языка — синтаксическая запись его глубинной структуры. Восстановление глубинной структуры предложения может быть сделано только человеком, носителем этого языка. При этом от эксперта, его способа понимания предложения и ситуации в целом, полностью зависит, какая глубинная структура будет получена.

Приведённое разделение простейших высказываний естественного языка на сообщения, заявления и побуждения позволяет нам ввести алгебру высказываний естественного языка. Высказывание состоит из трёх частей — сообщения, заявления и побуждения, каждая из которых является комбинацией простейших высказываний. Таким образом высказывание представляется в виде тройки предложений, каждое из которых состоит из слов одного и того же языка. Удобно рассматривать не весь язык, а только некоторый его фрагмент, соответствующий данному классу задач. Если мы, например, рассмотрим только сообщения, то им соответствует (при определённых ограничениях) булева алгебра Линденбаума-Тарского чистого исчисления предикатов сигнатуры, состоящей из всех понятий, входящих в данный фрагмент естественного языка. Поскольку заявления и побуждения содержат те же самые понятия, что и сообщения, мы можем считать, что части высказывания, являющиеся заявлениями, принадлежат алгебре, изоморфной указанной алгебре Линденбаума-Тарского, то же самое будем предполагать и для побуждений. Поэтому высказывание из данного фрагмента естественного языка будет представлено тройкой элементов алгебры Линденбаума-Тарского, или, иными словами, элементом куба алгебры Линденбаума-Тарского. Таким образом, алгебра высказываний данного фрагмента естественного языка — это куб булевой алгебры Линденбаума-Тарского чистого

исчисления предикатов первого порядка сигнатуры этого фрагмента.

§2. Предложения естественного языка, определяющие смысл текста

В этом параграфе мы рассмотрим вторую составляющую речевого действия — заявление. Первая составляющая — сообщение — очень хорошо описана и формализована в классической логике. Заметим, что смысл предложений классической логики (например, исчисления предикатов) всегда точно определён и фиксирован. Смысл предложения естественного языка, как мы уже отмечали выше, не может быть восстановлен по его синтаксической форме, а следовательно не может быть и зафиксирован. Поэтому, первым шагом для расширения классической логики до логики высказываний естественного языка является введение возможности изменения смысла предложений, т.е. введение в рассмотрение заявлений.

Здесь мы остановимся на одном частном случае заявлений. Частью контекста, во многом определяющей смысл предложения, является набор понятий — объектов, их свойств, действий над ними и пр., — которые в данный момент времени находятся в рассмотрении, "в локсе внимания".

Проиллюстрируем это на примере. Для объяснения, что такое немонотонная логика, часто рассматривают следующую ситуацию [10, стр. 3]. Допустим, мы знаем, что Твити — птица. В обычном контексте мы считаем, что утверждение "Птицы умеют летать" верно. Поэтому мы делаем вывод, что Твити также умеет летать. Далее мы узнаём, что Твити — домашний страус, и после этого, естественно, делаем противоположный вывод — что Твити не умеет летать. Рассмотрим теперь эту же ситуацию с точки зрения контекста предложений и его изменения. Каждую фразу естественного языка будем представлять как речевое действие, содержащее, кроме всего прочего, составляющую — заявление о том, что в рассмотрении

вводятся все понятия, содержащиеся в этой фразе. Каждое такое заявление будет расширять контекст, в котором воспринимаются предложения. Первая фраза — "Твити является птицей" вводит в рассмотрение два понятия — "Твити" и "птица". Следующая фраза — вопрос: "Может ли Твити летать?", вводит в рассмотрение понятие "летать" (вопрос в нашей классификации является комбинацией побуждения — ответить, и заявления — что задан вопрос; здесь мы видим, что вопрос может содержать и другие составляющие заявления). В контексте, состоящем только из понятий "Твити", "птица" и "летать" мы безусловно считаем, что все птицы умеют летать и, соответственно, делаем вывод что и Твити умеет летать. Следующая фраза — "Твити — это домашний страус" вводит в рассмотрение понятие "домашний страус". В контексте, содержащем это понятие, мы, очевидно, знаем, что страусы не летают, страусы — это птицы, и, следовательно, во-первых, утверждение, что все птицы летают, является ложным, и, во-вторых, верно, что Твити летать не умеет.

Приведённый пример, с одной стороны, показывает, что смысл и, следовательно, истинность утверждений (т.е., речевых действий, являющихся сообщениями) зависит от контекста, в котором они воспринимаются. Предложение "Все птицы умеют летать" в изначальном контексте имело смысл суждения об обычных птицах и, поэтому, было истинно; в обогащённом контексте смысл этого предложения меняется, оно говорит и о страусах тоже, и, в силу этого, является ложным.

С другой стороны, мы видим, что каждая фраза естественного языка может быть представлена как речевое действие, имеющее составляющую — заявление о том, что в рассмотрение вводятся все понятия, содержащиеся в этом предложении. Напомним, что сообщения и побуждения не предполагают, что сообщаемая информация не была известна, или что собеседник сам не собирался сделать требуемое действие. Точно так же и указан-

ное заявление не предполагает, что вводимые им понятия ещё не находятся в рассмотрении.

Введение таких заявлений соответствует предложенной выше связи между фразой естественного языка как поверхностной структурой и её глубинной структурой. Действительно, вопрос "Может ли Твити летать?" можно расширить до глубинной структуры "Рассмотрим Твити и свойство — уметь летать. Я хочу, чтобы вы ответили на вопрос — может ли Твити летать?". Такая глубинная структура содержит две фразы, первая из которых является чистым заявлением, а вторая — комбинацией побуждения и заявления.

Заметим следующее. Мы можем каждому предложению исчисления предикатов некоторой сигнатуры (рассматриваемому как синтаксический объект) сопоставить заявление о том, что в рассмотрение вводятся все сигнатурные символы, содержащиеся в этом предложении. Такое заявление однозначно представляется конечным подмножеством сигнатуры — множеством сигнатурных символов, входящих в предложение. Очевидно, что логически эквивалентным предложениям могут соответствовать различные заявления, т.е. различные подмножества сигнатуры. В силу интерполяционной теоремы Крейга [4, с.105], среди предложений, логически эквивалентных данному, найдётся предложение, содержащее *наименьшее по включению* множество символов.

Действительно, рассмотрим два логически эквивалентных предложения чистого исчисления предикатов первого порядка, содержащих разный набор сигнатурных символов. Из первого предложения следует второе, поэтому, по теореме Крейга, найдётся интерполянт — предложение, все символы которого входят и в первое предложение, и во второе, оно следует из первого предложения и из него следует второе предложение. Значит, во-первых, интерполянт логически эквивалентен каждому из двух эквивалентных предложений, и, во-вторых, множество символов интерполянта содержится в пересечении множества символов, входящих в первое предло-

жение, и множества символов, входящих во второе. Стало быть, в силу конечности множества символов, входящих в предложение, найдётся ему эквивалентное предложение, содержащее наименьшее по включению множество символов.

Таким образом, каждому предложению исчисления предикатов мы можем поставить в соответствие это наименьшее множество сигнатурных символов, т.е., множество сигнатурных символов, о которых предложение говорит *по-существу*. Тождественно истинному и тождественно ложному предложениям (и не только им, например, предложению, утверждающему, что существуют n различных элементов) будет соответствовать пустое множество; предложению и его отрицанию будет соответствовать одно и то же множество сигнатурных символов; конъюнкции и дизъюнкции двух предложений будет соответствовать некоторое подмножество объединения множеств, соответствующих этим предложениям.

Рассмотрим прямое произведение двух булевых алгебр — алгебры Линденбаума-Тарского исчисления предикатов данной сигнатуры и алгебры всех подмножеств сигнатуры (будем считать, что сигнатура конечна). Алгебру подмножеств сигнатуры будем рассматривать как булеву алгебру, двойственную булевой алгебре множеств: частичный порядок будет обратным теоретико-множественному включению, объединение и пересечение поменяются местами. Прямое произведение будем рассматривать как булеву алгебру, операции на которой определены покомпонентно. Таким образом, на первой компоненте прямого произведения будет естественный частичный порядок, а на всём прямом произведении — покомпонентный. Элемент этого прямого произведения будем понимать как композицию сообщения и заявления. Сообщение представлено первой компонентой — элементом алгебры Линденбаума-Тарского, заявление — второй — подмножеством сигнатуры. Смысл сообщения — это обычный смысл предложения исчисления предикатов, смысл заявления — введение в рассмотрение данного множества

сигнатурных символов. Такое прямое произведение будет частным случаем описанной выше алгебры высказываний естественного языка. Далее будем называть его алгеброй контекстов.

Каждому предложению исчисления предикатов (рассматриваемому как синтаксический объект — слово в соответствующем алфавите) поставим в соответствие элемент алгебры контекстов. Его первой компонентой будет элемент алгебры Линденбаума-Тарского — класс эквивалентности данного предложения. Вторая компонента — множество сигнатурных символов, входящих в предложение. Очевидно, что логически эквивалентным предложениям могут соответствовать разные элементы алгебры контекстов. Рассмотрим два предложения и два соответствующих им элемента алгебры контекстов. Легко видеть, что первый элемент меньше либо равен второму тогда и только тогда, когда, во-первых, из первого предложения следует второе, и, во-вторых, множество сигнатурных символов первого предложения содержит множество сигнатурных символов второго. При таких условиях будем говорить, что из первого суждения естественного языка (представленного первым элементом алгебры контекстов) следует второе.

Завершая этот параграф, введём три определения логического следования из множества элементов алгебры контекстов. *Первое.* Следствием назовём любой элемент фильтра, порождённого данным множеством элементов. *Второе.* Множество следствий будем строить из исходного множества элементов используя две операции: а) логический вывод из множества предложений с одинаковой второй компонентой (определяющей контекст); б) увеличение (т.е. теоретико-множественное уменьшение) второй компоненты данного элемента. *Третье.* Множество следствий строится используя только п."а" предыдущего определения.

Первое определение соответствует пониманию второй компоненты элемента алгебры контекстов как заявления, вводящего в рассмотрение новые понятия. Третье опре-

деление — как заявления, констатирующего данный контекст (глубинной структурой такого заявления будет: "Сейчас мы будем рассматривать следующие понятия и только их ..."). С помощью третьего определения можно описать немонотонное логическое следование в приведённом выше примере о Твити — домашнем страусе. Очевидно, что предложение "Все птицы летают" истинно во всех контекстах, не содержащих понятий "страус", "курица", "пингвин", "жареный гусь" и т. д. (т.е., истинна пара — данное предложение и соответствующий контекст — множество символов, обозначающих понятия). Поэтому, если мы рассмотрим речевое действие "Твити — птица; может ли Твити летать?", то в его контексте — множестве {Твити, птица, летать}, высказывание "Все птицы летают" истинно, и мы можем сделать вывод, что Твити также может летать. Речевое действие "Твити — птица; Твити — домашний страус; может ли Твити летать?" имеет другой контекст — {Твити, птица, домашний, страус, летать}, в котором предложение "Все птицы летают" ложно, а истинно другое — "Страусы не умеют летать", поэтому мы делаем вывод, что Твити летать не умеет.

Таким образом приведённые определения, в особенности третье, дают возможность иметь дело с текстами, содержащими противоречивую информацию. Третье определение, например, не позволяет выносить противоречие за пределы контекста, в котором оно получено.

§3. Логический анализ метода ГАБЕК

В этом параграфе мы, с помощью введённых выше понятий, проделаем анализ метода ГАБЕК, предназначенного для обработки текстов естественного языка.

Метод ГАБЕК [11-17] (GABEK — Ganzheitliche Bewältigung sprachlich erfaßter Komplexität — Целостное преодоление языковых сложностей) был предложен Йозефом Цельгером в 1990 году [11]. Мы будем использовать обзор этого мето-

да, приведённый в настоящем сборнике [19]. Кратко изложим основные идеи метода ГАБЕК.

Метод ГАБЕК имеет дело с текстами, состоящими из неупорядоченного, ничем не организованного набора предложений. Каждое предложение состоит из одной или нескольких фраз обыденного языка. Предложение представляет собой суждение о некоторых предметах; например, предложения, из которых состоит текст, могут быть ответами различных людей на некоторый вопрос; при этом весь ответ полностью (состоящий, быть может, из нескольких фраз) считается одним предложением. Предложение мыслится как некая самостоятельная, целостная единица текста.

Текст рассматривается на трех уровнях — синтаксическом, семантическом и прагматическом. На синтаксическом уровне предложение отождествляется со множеством так называемых ключевых понятий, входящих в это предложение. Набор ключевых понятий определяется экспертом для каждого предложения на самом первом этапе работы с текстом. Ключевые понятия — это наиболее существенные понятия (т.е. объекты, их свойства и т. д.), о которых говорит данное предложение.

На семантическом и прагматическом уровнях предложение рассматривается как вполне конкретное осмысленное суждение, которое может понять только человек. На семантическом уровне от предложения или текста требуется наличие идеального примера ситуации, где оно может быть реализовано, на прагматическом уровне — реального примера ситуации (но не такой ситуации, которая была на самом деле, а которую может помыслить себе эксперт). Например, описание животного с четырьмя ногами, копытами и двумя крыльями имеет идеальный пример — пегас, но не имеет реального примера. Описание же животного, имеющего четыре ноги, копыта и черный цвет, очевидно, имеет реальный пример — вороной конь.

С нашей точки зрения, семантический и прагматический уровни рассмотрения предложения в ГАБЕКе,

по-существу, являются одним, который имеет черты и семантики, и прагматики. Действительно, с одной стороны, семантический уровень не предполагает наличия формальной семантики и с необходимостью требует наличия эксперта, который со своей точки зрения (т.е., используя своё собственное понимание и толкование предложения) оценивает возможность наличия идеального примера. С другой стороны, реальный пример ситуации — это не то, что было на самом деле, а то, что может себе представить как реальную ситуацию эксперт. Поэтому, с нашей точки зрения, граница между семантическими и прагматическими требованиями в ГАБЕКЕ достаточно условна, и, на наш взгляд, более естественно выделять здесь два главных уровня рассмотрения — синтаксический и семантический (включающий в себя элементы прагматики).

Текст в ГАБЕКЕ рассматривается на синтаксическом и семантическом уровнях поочередно. На синтаксическом уровне текст обрабатывается компьютером, на семантическом — экспертом. Сначала текст рассматривается как множество осмысленных предложений, и эксперт определяет для каждого предложения набор его ключевых понятий. Затем текст как чисто синтаксический объект представляется в виде неупорядоченного графа, в вершинах которого находятся предложения и ключевые понятия, рёбра соединяют ключевые понятия со всеми предложениями, в которые они входят. Такой граф называется сетью выражений. В соответствии с определёнными синтаксическими требованиями текст разбивается на группы предложений, называемые формальными лингвистическими образами.

Далее текст рассматривается на семантическом уровне. Эксперт, работая с данной группой предложений, добивается, чтобы она удовлетворяла не только синтаксическим, но и семантическим, и прагматическим требованиям. Полученная в результате группа предложений называется прагматическим лингвистическим образом. После этого эксперт формулирует предложение, по его

мнению, суммирующее всю информацию, содержащуюся в данной группе предложений.

Такая процедура проделывается со всеми группами предложений, на которые разбит текст. В результате будет получен новый текст, состоящий из резюмирующих предложений для всех групп исходного текста. Новый текст по объёму в 7-10 раз меньше исходного.

Новый текст опять рассматривается как чисто синтаксический объект и разбивается на группы предложений, называемые формальными лингвистическими сверхобразами первого уровня. Затем строятся сверхобразы второго уровня, третьего и т. д., пока не будет получен один сверхобраз. Для него строится резюмирующее предложение, которое и будет итоговым суждением для всего текста.

Таким образом, исходный текст представляется, во-первых, с помощью иерархии лингвистический образов и, во-вторых, своим итоговым предложением.

Рассмотрим теперь метод ГАБЕК на языке введенной выше алгебры контекстов. В качестве сигнатуры мы должны взять множество всех понятий, встречающихся в тексте. Множество ключевых понятий будет подмножеством сигнатуры. Каждому предложению, содержащемуся в тексте, поставим в соответствие элемент алгебры контекстов: первой компонентой будет класс эквивалентности этого предложения, второй — множество входящих в него ключевых понятий.

Как мы отмечали выше, каждому предложению можно сопоставить множество сигнатурных символов (т.е. понятий), входящих в это предложение *по-существу*. Это множество является наименьшим по включению среди множеств символов предложений, принадлежащих классу эквивалентности данного предложения. Мы считаем разумным добавить следующее семантическое условие на лингвистический образ: каждое ключевое понятие, входящее в предложение, должно входить в него *по-существу*. Более подробно его можно сформулировать так: для каждого предложения, входящего в лингвисти-

ческий образ, нельзя построить предложение, содержащее меньший набор ключевых понятий и описывающее то же свойство идеальной ситуации, что и данное предложение. Тот факт, что в ГАБЕКЕ такое требование отсутствует, на наш взгляд говорит только о том, что в реальных текстах оно, как правило, выполнено. Однако, полезно считать, что семантический образ *обязательно* удовлетворяет этому условию.

Таким образом, исходный текст мы можем представить как множество элементов алгебры контекстов (далее для краткости будем называть его просто *множеством элементов*). Вместе с этой алгеброй удобно отдельно рассматривать две её компоненты — её прямые сомножители: алгебру Линденбаума-Тарского исчисления предикатов данной сигнатуры и алгебру подмножеств сигнатуры (с двойственным порядком и двойственными операциями), а также проекции алгебры контекстов на эти прямые сомножители. *Первая* проекция элемента алгебры контекстов даёт его семантический смысл (т.е. его класс эквивалентности в исчислении предикатов), *вторая* — множество его ключевых понятий.

Рассмотрим теперь основные понятия, используемые в методе ГАБЕК [11-17, 19]. Семантическое условие, накладываемое на лингвистическую сеть, равносильно непротиворечивости (или, то же самое, совместности) первой проекции множества элементов (которыми представлен текст). Прагматическое условие равносильно наличию реально существующей модели первой проекции.

Далее для краткости будем первую и вторую проекции множества элементов алгебры контекстов, соответствующего данному лингвистическому образу, называть соответственно первой и второй проекциями лингвистического образа.

Синтаксические требования, накладываемые на лингвистический образ, имеют дело с его второй проекцией. Требование формального различия говорит, во-первых, что вторая проекция данного лингвистического образа является взаимно-однозначным отображением (мы здесь

одним словом — "проекция", называем и отображение, и его образ; по смыслу всегда ясно, о чём идёт речь), и, во-вторых, что разность любых двух множеств символов во второй проекции содержит как минимум два элемента.

Требование взаимной однозначности второй проекции лингвистического образа имеет принципиальное значение. Из него следует, что для рассмотрения лингвистического образа на синтаксическом уровне нам достаточно иметь дело *только* с его второй проекцией. С другой стороны, в силу добавленного нами выше семантического требования на лингвистический образ, для работы с ним на семантическом уровне достаточно иметь только его первую проекцию. Поэтому, обработка текста естественным образом разбивается на работу с его первой и второй проекциями по отдельности.

Требование формальной связи состоит в том, что любой символ, входящий в объединение всех элементов второй проекции лингвистического образа, встречается как минимум в двух таких элементах.

Ослабленное требование формальной связи говорит, что для любого элемента второй проекции существуют три различных отличных от него элемента и три входящих в него символа такие, что первый символ входит в первый элемент, второй — во второй, а третий — в один из трёх.

Слабое требование формальной связи означает, что для любого элемента второй проекции существуют два различных отличных от него элемента и два входящих в него символа такие, что первый символ входит в первый элемент, а второй — во второй.

Требование формального расстояния заключается в том, что для любых двух элементов второй проекции лингвистического образа найдётся третий элемент (не обязательно отличный от них), имеющий с каждым из них непустое пересечение.

Семантические и прагматические условия на лингвистический образ, имеют дело с его первой проекцией. Лингвистический образ рассматривается как часть линг-

вистической сети; семантическое и прагматическое условия, накладываемые методом ГАБЕК на лингвистическую сеть, требуют наличия соответственно идеального и реального (с точки зрения эксперта) примеров, на которых реализуются (т.е., являются истинными, имеют место) все предложения, входящие в лингвистическую сеть. В нашем представлении метода ГАБЕК с помощью алгебры контекстов в качестве идеального примера мы рассматриваем алгебраическую систему (модель), на которой истинна первая проекция лингвистической сети, в качестве реального примера — алгебраическую систему, имеющую место в реальном мире, причём с естественной интерпретацией (т.е., сигнатурному символу "человек" должны соответствовать реальные люди, символу "быть красным" — именно это свойство и т. д.).

Требование семантического различия говорит, что никакие два разных элемента первой проекции лингвистического образа не являются эквивалентными в элементарной теории алгебраической системы — идеального примера. Требование прагматического различия означает, что любые два сигнатурных символа, входящие во вторую проекцию, имеют *различные интерпретации* в алгебраической системе — реальном примере (т.е., им соответствуют разные отношения, операции или элементы).

Требования семантической связи, прагматической связи и внешнего различия с нашей точки зрения не могут быть сформулированы в обычных теоретико-модельных терминах, поскольку они формализуют принципиально новое понятие — понятие *образа*. Разработка теоретико-модельных конструкций, описывающих это понятие, была бы, на наш взгляд, важным шагом на пути логической формализации естественного интеллекта (в определенном смысле начатой Джорджем Булем в 1847 г. в его книге "Математический анализ логики").

Образ, выражаясь кратко, это нечто целостное, состоящее из частей и являющееся взаимосвязью, взаимодействием этих частей. С нашей точки зрения, в теории моделей образу наиболее близко соответствует понятие

типа кортежа элементов модели, особенно если этот тип является главным, т.е. если он описывается (аксиоматизируется) одной формулой. Понятие образа имеет много общего с понятием объекта в объектно-ориентированном программировании.

В результате обработки текста с помощью метода ГАБЕК получается резюмирующее предложение и иерархия лингвистических образов. Эта иерархия может быть представлена как множество наследственно-конечных множеств — элементов надстройки наследственно-конечных множеств над сигнатурой, рассматриваемой как множество символов. Сигнатурные символы будут праэлементами, предложения — множествами праэлементов, лингвистические образы первого уровня — множествами множеств праэлементов и т.д. Иерархию также можно рассматривать как множество элементов надстройки наследственно-конечных множеств над второй проекцией текста (разница состоит в том, что в качестве праэлементов в первом случае будут рассматриваться элементы сигнатуры, а во втором случае — подмножества сигнатуры, соответствующие предложениям). С другой стороны, иерархию можно представлять в виде одного элемента надстройки, наибольшего среди указанного выше множества элементов надстройки; этот элемент представляет лингвистический образ самого высокого уровня. Каждый элемент надстройки (кроме праэлементов — в первом случае) представляет некоторое суждение об объектах и их свойствах, описываемых ключевыми понятиями; это суждение — либо исходное предложение, либо резюмирующее предложение для некоторого лингвистического образа. Элемент надстройки, не являющийся праэлементом (а также множеством праэлементов — в первом случае) представляет соответствующий ему лингвистический образ. Наибольший элемент представляет предложение, резюмирующее весь текст.

В заключение важно отметить, что наследственно-конечное множество, являющееся указанным выше наибольшим элементом, полностью описывает структуру

текста, представляемую методом ГАБЕК. К этому наследственно-конечному множеству естественно добавить отображение из его транзитивного замыкания (множества всех наследственно-конечных множеств, являющихся его элементами, элементами его элементов и т.д.) в алгебру Линденбаума-Тарского (содержащую первую проекцию): каждому наследственно-конечному множеству ставим в соответствие класс эквивалентности резюмирующего предложения лингвистического образа (некоторого уровня), описываемого этим множеством; если же это множество является множеством сигнатурных символов, т.е. описывает некоторое исходное предложение, то ставим ему в соответствие класс эквивалентности этого предложения. Таким образом, исходный текст будет представлен парой — наследственно-конечным множеством и отображением из его транзитивного замыкания в алгебру Линденбаума-Тарского.

Л и т е р а т у р а

1. CHOMSKY N. Syntactic Structures, s'-Gravenhage, 1957.
2. SEARLE J.R. Speech acts: an essay in the philosophy of language.- Cambridge: University Press, 1969.
3. BANDLER R., GRINDER J. Structure of magic. Vol.1.- 1975. (Русский перевод — Гриндер Д., Бэндлер Р. Структура магии.- М.: КААС, 1995.)
4. КЕЙСЛЕР Г., ЧЕН Ч.Ч. Теория моделей. - М.: Мир, 1977.
5. SEARLE J.R. Expression and meaning, 1979.
6. SEARLE J.R. Intensionality, 1983.
7. SEARLE J.R., VANDERVEKEN D. Foundations of illocutionary logic.- Cambridge,1985.
8. VANDERVEKEN D. Meaning and speech acts. Vol. 1: Principles of language use.- Cambridge: Cambridge Univ. Press, 1990.
9. VANDERVEKEN D. Meaning and speech acts. Vol. 2: Formal semantics of success and satisfaction.- Cambridge: Cambridge Univ. Press, 1991.

10. MAREK W., NERODE A., REMMEL J. A theory of nonmonotonic rule systems I, Center of Excellence in Mathematical Sciences, Cornell Univ., 1990.

11. ZELGER J. Ein ganzheitliches Verfahren zur Bewältigung sprachlich erfassbarer Komplexität (GABEK). Universität Innsbruck, Institut für Philosophie, 1990, preprint № 10.

12. ZELGER J. GABEK, a new method for qualitative evaluation of interviews and model construction with PC-support, In: Stuhler E., Suilleabhain M.O., (Eds.): Enchanging human capacity to solve ecological and socio-economic problems, Munchen, Mering: Rainer Hampp Verlag, 1993, pp. 128-172.

13. ZELGER J. GABEK I: Constituting conceptual networks. II: Evaluation of conceptual networks. III: From conceptual networks to linguistic Gestalten. Innsbruck, Institut für Philosophie der Universität Innsbruck, 1993, preprints № 15, 16, 17.

14. ZELGER J. Von sprachlichen Gestalten zur Hypergestalt, In: Philosophie und Verfahren kreativer Selbstorganisation, Innsbruck, Institut für Philosophie der Universität Innsbruck, 1993, preprint № 26.

15. ZELGER J. Sprachlichen Gestaltbildung durch das PC-unterstützte Verfahren GABEK, Innsbruck: Universität, Institut für Philosophie, Manuskript 1993.

16. ZELGER J. Constructing Linguistic Gestalten by GABEK.

17. ZELGER J., Qualitative Auswertung sprachlicher Äußerungen. Wissensvernetzung, Wissensverarbeitung und Wissensumsetzung durch GABEK, In: Wille R. & Zickwolff M. (Hrsg.): Begriffliche Wissensverarbeitung: Grundfragen und Aufgaben, Mannheim (B.I. Wissenschaftsverlag), 1994, pp. 239-266.

18. ПАЛЬЧУНОВ Д.Е. О логическом анализе естественного языка//Теория вычислений и языки спецификаций.- Новосибирск.- 1995.- Вып.152: Вычислительные системы.- С. 61-75.

19. ПАЛЬЧУНОВ Д.Е. Анализ текстов естественного языка с помощью метода ГАБЕК//Настоящий сборник.- С.149-166.

Поступила в редакцию
19декабря 1996 года