

85 коп.

КИБЕРНЕТИЧЕСКАЯ
ЛИНГВИСТИКА



ИЗДАТЕЛЬСТВО • НАУКА •

АКАДЕМИЯ НАУК СССР
НАУЧНЫЙ СОВЕТ
ПО КОМПЛЕКСНОЙ ПРОБЛЕМЕ «КИБЕРНЕТИКА»

**КИБЕРНЕТИЧЕСКАЯ
ЛИНГВИСТИКА**



ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА, 1983

Книга посвящена комплексному освещению кибернетической лингвистики. Статьи отражают результаты, полученные в фундаментальных и прикладных исследованиях, связанных с автоматической обработкой текстовой информации, в том числе с автоматическим вводом текста, проблемами понимания ЭВМ текстов на естественном языке, вопросами нейролингвистики и др.

Книга предназначена для специалистов в области теоретических и прикладных вопросов кибернетической лингвистики.

Ответственные редакторы:

Вяч. Вс. ИВАНОВ, В. П. ГРИГОРЬЕВ

Книга посвящена комплексному освещению проблем кибернетической лингвистики — междисциплинарной области, в рамках которой рассматриваются теоретические и прикладные вопросы, связанные с представлением, переработкой и пониманием текстов на естественном языке автоматическими системами. В более узком смысле к кибернетической лингвистике примыкает круг проблем, возникающих при моделировании коммуникационных процессов в человеко-машинных системах. На современном этапе наряду с разработкой методов формального представления и анализа естественного языка на всех его структурных уровнях большое внимание уделяется вопросам понимания компьютером структуры текстовых и речевых сообщений.

Статьи, представленные в книге, отражают основные направления исследований, ведущиеся в нашей стране и за рубежом, и группируются вокруг следующих научных проблем: механизмов понимания естественного языка и нейролингвистики, проблем моделирования слуховых и зрительных сообщений и моделирования коммуникационных процессов.

Цикл статей посвящен исследованию механизмов понимания естественного языка. В статье Вяч. Вс. Иванова предлагается интересная постановка задачи рассмотрения мозга как двухмашинного комплекса, порождающего речь. Ставится вопрос о соотношении лингвистической модели порождения речи и функций полушарий головного мозга. В статье Ю. М. Лотмана обсуждается гипотеза об «интеллектуальном устройстве» и его влиянии на индивидуальные интеллектуальные возможности. В качестве такого «интеллектуального устройства» рассматривается культура как механизм, реализующийся в различных текстах, в том числе и в естественном языке.

Проблемам моделирования акустических и оптических сообщений посвящены статьи Л. А. Чистович, В. А. Кожевникова, Ж. А. Першина «Проблемы моделирования слуховой обработки речевого сигнала» и Е. К. Гусевой «К проблеме понимания ЭВМ графематического уровня языка». В статьях содержатся результаты экспериментальных исследований, значительное внимание уделяется психофизиологическим механизмам идентификации акустических и оптических сигналов, обсуждаются перспективы их решения в связи с проблемой ввода устной и письменной речи.

Ряд статей посвящен различным аспектам моделирования тек-

стов на естественном языке. Проблемам моделирования грамматических категорий средствами тензорной алгебры посвящена статья Б. В. Сухотина, в которой предлагается координатный метод выражения содержания различных языковых уровней. Результаты работы могут быть использованы для решения прикладных проблем. В статье Ю. С. Мартемьянова предложен способ связанного описания синтаксиса средствами порождающей и трансформационной грамматик, который может быть использован при решении различных прикладных задач.

В статье Д. Г. Лахути, Е. Ф. Федорова, И. С. Добронравова, В. Р. Пархоменко «Автоматическое индексирование текстов в документальных ИПС» изложена схема ИПЯ, ориентированного на автоматическое индексирование документов.

Современный этап развития вычислительной техники характеризуется развитием многомашинных комплексов, объединенных сетями связи. Вопросы моделирования коммуникационных процессов в диалоговых человеко-машинных системах рассматриваются в статье Я. П. Выставкина, Е. К. Гусевой «Формальные языковые модели диалоговых взаимодействий». Предлагается формальная языковая модель, основанная на использовании теории формальных грамматик. Проводится аналогия между естественным языком как порождающим устройством, выполняющим коммуникативные функции, и процедурами управления диалоговыми взаимодействиями в распределенных вычислительных системах.

Цель сборника привлечь внимание специалистов широкого профиля (искусственный интеллект, лингвистика, вычислительные системы и др.) к исследованию языка как устройства, порождающего речевые тексты, и к вопросам общения с ЭИМ на естественном языке.

I. ПСИХОФИЗИОЛОГИЧЕСКИЕ И ТЕХНИЧЕСКИЕ АСПЕКТЫ ПОРОЖДЕНИЯ И РАСПОЗНАВАНИЯ РЕЧИ ЧЕЛОВЕКОМ И МАШИНОЙ

УДК 616.831—081.84.07

Вяч. Вс. Иванов

ЕСТЕСТВЕННЫЙ ЯЗЫК — МОЗГ, ИСКУССТВЕННЫЙ ЯЗЫК — МАШИНА

В современной кибернетической литературе много раз ставилась проблема соотношения мозга человека и электронно-вычислительной (или абстрактной логической) машины, переплетающаяся с вопросом о соотношении естественного и искусственного языков. Следовательно, целесообразно рассмотреть взаимосвязь этих проблем.

1. Кибернетические модели двух полушарий мозга. Доминантное по языку (речевое) полушарие головного мозга (левое у подавляющего большинства «правшей» в норме) по характеру основных видов деятельности, на которых оно специализируется, и по предполагаемому характеру производимых им логических операций над последовательностями любых дискретных символов может быть сопоставлено с вычислительной машиной (или с любой машиной, представляющей собой реализацию машины Тьюринга). К таким видам деятельности можно отнести: синтез, анализ или распознавание последовательностей дискретных звуковых единиц речи — фонем или цепочек слов, образующих грамматически правильный текст, логический вывод, счет, логическую классификацию и т. д.

Явно иной характер носит деятельность субдоминантного по языку (неречевого в норме) полушария, которое осуществляет образование и распознавание целостных единств-образов в широком смысле слова, не членищихся на составные части (музыкальные мелодии, изобразительная картина наглядного пространственного мира в его целостности и т. п.) [12]. По гипотезе, принимаемой в настоящее время рядом ученых [24], деятельность этого полушария можно было бы описать голографической моделью (в отличие от модели вычислительной машины, принимаемой для левого полушария). При некоторой эвристической ценности наглядного образа, ими подсказываемого, такие технические аналогии определяются целиком уровнем техники, достигнутым на сегодняшний день (типа известного сравнения мозга с автоматической телефонной станцией, заинтересовавшего Павлова). Большее значение имеют не конкретные технологические (бионические) усовершенствования — мозг и ЭВМ, мозг и голографическое устройство,

а подсказываемые ими модели более общего характера (не зависящие от деталей технической реализации). В частности, для левого полушария существенной представляется аналогия с любой логической машиной, эквивалентной машине Тьюринга; менее продуктивными в настоящее время кажутся предпринятые около 30 лет назад опыты сравнения нервных сетей с автоматами, хотя формализация по схеме автомата с конечным числом состояний направляется по отношению к системам с очень малым числом элементов типа обнаруживаемых у беспозвоночных [50, 51].

Для моделирования деятельности правого полушария мозга, а возможно и исходных принципов в работе левого полушария, более существенный интерес представляют современные работы по теории сложности вычислений, в которых показана роль преобразований Фурье для построения эффективных алгоритмов [1, с. 284—310]. Поэтому обнаруженная в экспериментах В. Д. Глезера и его сотрудников возможность описания системы переработки зрительной информации, которая осуществляет преобразования Фурье [7, 8], может получить объяснение из соответствующей модели.

Ряд вопросов, относящихся к интерпретации тех экспериментальных данных, на основании которых предполагается осуществление зрительной системой преобразований Фурье, еще нуждается в дополнительном обсуждении [19]. Если эта последняя гипотеза подтвердится, ее можно будет связать с теоретико-информационными данными о сложности зрительной информации, обрабатываемой мозгом человека. Характер задач, во-первых, связан с исключительной сложностью организации самой системы переработки зрительной информации, отмечавшейся, в частности, в работе Дж. фон Неймана [25], во-вторых, может предопределить выбор таких эффективных алгоритмов, которыми могут быть дискретные преобразования Фурье. В этом случае можно было бы сохранить предложенное ранее сопоставление мозга с логической машиной, но при этом целесообразно было бы представить мозг (или хотя бы его кору) как аналог комплекса, состоящего по меньшей мере из двух машин.

Отмечавшиеся в предыдущих публикациях [12, 13, 14] конкретные аналогии с осуществленными за последнее время реальными двухмашинными комплексами нуждаются в той же оговорке, которая была сделана выше по отношению к другим конкретным технологическим уподоблениям. Тем не менее для исследования принципов работы парных полушарий мозга значительный интерес представляет то, что уже сейчас в системах управления манипуляционными роботами используются двухмашинные вычислительные комплексы. В этом случае процессы планирования и построения движений реализует одна, более мощная машина, а управление аппаратурой робота в режиме реального времени, требующего большого числа прерываний процессора и преобразований локальной информации — другая, как правило, мини-ЭВМ [28]. Так, в экспериментальном образце интегрального робота

ЛПИ-2, имеющего два манипулятора (аналог двух рук) и устройство для речевого ввода команд, управляющая система построена на базе вычислительного комплекса, включающего ЭВМ «Минск-32» [30]. Само собой разумеется, конкретное разделение функций внутри машин, входящих в двухмашинный комплекс, может не иметь ничего общего с тем, как проводится разделение функций между полушариями мозга. Более существенно принципиальное сходство, позволяющее предположить возможность моделирования мозга посредством такого двухмашинного комплекса, внутри которого разделяются машина, осуществляющая собственно логические операции с дискретными последовательностями символов, и специализированное устройство, рассчитанное на обработку таких сложных массивов информации, где целесообразно осуществление преобразований Фурье.

Можно думать, что Фурье-преобразования правое полушарие осуществляет и тогда, когда оно совершает операции над акустическими образами — не только музыкальными, но и другими неречевыми звуками, а также гласными, в восприятии которых правое полушарие принимает существенное участие [29]. С точки зрения конкретных технических аналогий обнаруживается следующий парадокс. Современные устройства, распознающие речь, сходны скорее с правым, а не с левым полушарием; они могут распознавать относительно небольшое число слов, произносимых определенным диктором, причем наибольшую техническую трудность представляет разделение непрерывного звукового образа на отдельные сегменты.

Из этого не следует, что мозг можно свести к такому двухмашинному комплексу. Из проблем, оживленно обсуждавшихся в нейрофизиологии и смежных науках в последние годы и не находящихся, по всей видимости, решения в рамках двухмашинной модели, можно было бы отметить роль таламуса, латерализация которого указывает на определенное соотношение левой (речевой) части таламуса с левым (доминантным) полушарием [63]. В свете высказываемых в последнее время допущений о характере обмена сигналами между таламусом и корой [27] можно было бы думать, что получение сигналов (типа радиосигнализации) от передатчика (таламуса) индуцирует такие паттерны в соответствующих зонах коры, которые и ведут в дальнейшем к работе одной из описанных в общих чертах частей «двухмашинного» комплекса. Однако детали этой «управляющей» функции таламуса в связи с его латерализацией остаются до настоящего времени неясными.

Другим спорным вопросом, который рассматривается в последних работах Экклза [39] и других ученых, является проблема считывающего устройства по отношению к любой из моделей разделения функций полушарий. Предполагаемая Экклзом [39] связь сознания с доминантным (левым) полушарием, в известной мере согласующаяся и с другими данными [14, 49], не обязательно еще ведет к приурочению считывающего устройства к доминантному полушарию. В любом случае остается неясным, какой может

быть локализация считывающего устройства в коре одного из полушарий или за ее пределами. Наконец, необходимо сделать особую оговорку и по поводу того, как соотносится с машиной каждый нейрон, в определенном смысле являющийся ее аналогом.

Современные представления о работе нейронов связаны с допущением, что некоторые из них являются специализированными детекторами для определенного рода стимулов или образов [19].

Хотя гипотеза о существовании гностических нейронов, реагирующих на известный класс изображений, и вызывает известные сомнения [19, с. 335—337], тем не менее индивидуальные нейроны в настоящее время привлекают все большее внимание. Поэтому наряду с «двухмашинной» моделью мозга (на макроскопическом уровне) вполне возможным (и обязательным для организмов на низшем этапе эволюции) [50, 51] является представление каждого нейрона как «машины-автомата» с конечным числом состояний. На этом пути открывается ряд возможностей, в равной мере существенных для моделирования мозга как системы нейронов и общества — как системы организмов, обладающих мозгом.

2. Проблема двоичного кодирования информации. Из результатов предварительного исследования семантических словарных ассоциаций при правосторонних и левосторонних электросудорожных шоках следует, что левое полушарие переходит на стратегию двоичного поиска ассоциаций. Ответы на заданное слово при относительной инактивации (или торможении) работы правого (субдоминантного) полушария (усталым или истощенным больным, т. е. в экстремальной ситуации) находятся посредством двоичного поиска. Это можно предположить на основании очень часто встречающихся ассоциаций типа антонимов, в частности прилагательных с противоположными значениями (*тяжелый — легкий, жесткий — мягкий* и т. п.); существительных, которые в данном семантическом поле могут выступать как пара, противопоставленная по двум рядам семантических противоположностей (*солнце — луна* и др.). По отношению к речевым зонам левого полушария та же стратегия давно уже выявлена [12] на основании фонемных замен (смещение фонем, отстоящих друг от друга на один различительный признак). Этот вывод согласуется с широким использованием аналогичного механизма в символике бессознательного, изучавшегося Фрейдом на материале сновидений [12]. То же гипотетическое психофизиологическое приурочение можно предположить для мифопоэтического мышления. Механизм его работы, общий с фрейдовским пониманием амбивалентности символов бессознательно, можно пояснить на примере полустигии Катуллы: *Odī et amo «И ненавижу и люблю»* (по-видимому, отразившегося и в строке Блока: «И с ненавистью и с любовью», написанной в то время, когда Блок начинал изучать Катуллу). В этой строке Катуллы раскрывается внутренняя связь нерасчленяемых элементов эмоционального комплекса значений. Сами по себе слова *ненавидеть* и *любить* по грамматическим своим характеристикам и по их аб-

страктности относятся скорее к левому полушарию. Но особенности мифопоэтического мышления сказываются в построении единого образа. По отношению к мифологии и ритуальным системам можно говорить о медиации в смысле Леви—Стросса, т. е. о поиске промежуточного звена между противоположностями, найденными при двоичном поиске. Соответствующие модели деревьев двоичного поиска хорошо разработаны в современной теории сложности вычислений [1, с. 135—146]. На нейрофизиологическом уровне аналога этому двоичному принципу работы, возможно, следует искать не столько в предполагавшейся ранее схеме бинарного кодирования в сети нейронов (что и дало основание для осуществления первых логических моделей нейронных сетей), сколько в гораздо более общем принципе двоичной организации работы мозга. Как формулирует эволюционные выводы из этого принципа Крик в своей недавней работе о мозге, «если новые функциональные области мозга возникают в ходе эволюции, они возникают попарно» [38, с. 228]. Целый ряд существенных для мозга операций может выполняться соотношением определенных групп нейронов (или отдельных центральных нейронов?) в таких попарно связанных областях. Поэтому двоичный принцип поиска может быть связан с некоторыми существенными характеристиками мозга.

3. Проблемы нейросемантики цветовых обозначений. В качестве первого примера (для лингвистов хорошо известного) можно рассмотреть нейросемантический подход к цветовым обозначениям. На примере цветовых обозначений и восприятия цветов можно продемонстрировать существенные сходства и различия между нейросемантическими вопросами исследования лингвистических систем и отличных от них систем неязыковых знаков. Согласно гипотезе, которую оказалось возможным высказать на основании сопоставления данных языка и культуры, в цветовых системах выявляются определенные общие закономерности [15].

Системы цветовых обозначений и цветовых символов в разных языках и культурах явились предметом многочисленных исследований в связи с обсуждением гипотезы лингвистической относительности, предполагавшей влияние родного языка на характер цветовой классификации. Эти исследования привели к выводу о существенном сходстве всех исследованных языков. Для значительного числа языков мира, цветовая терминология которых подвергалась исследованию, установлена иерархия, определяющая наличие цветовых обозначений: если в языке есть обозначение для красного, то в нем заведомо есть термины для белого (светлого) и черного (темного), которые, однако, могут образовываться сами по себе и минимальную двучленную систему. Эти цвета входят в основной общечеловеческий треугольник словесных названий цветов, совпадающий с треугольником, который выявлен в исследованиях цветовых символов в ритуалах и мифологии различных народов. Вхождение в этот треугольник красного цвета

представляет особый интерес с эволюционной точки зрения. Наличие охры, использовавшейся в сакральных целях, известно уже в шалашах из Терра Амата, датировка которых, $230\ 000 \pm 40\ 000$ лет, недавно подтверждена новыми методами (определение хронологии «последсвечения» камней, побывавших в огне). Трехчленная система цветов (красный — черный — белый) является определяющей для наскальной живописи Верхнего Палеолита и поэтому с надежностью может считаться характерной для *Номо sapiens* с самого начала его археологически засвидетельствованного существования. Эта система, в которой красный цвет (охра) был выделен как особо значимый, резко отлична от известной у антропоидов преобладающей роли синего цвета. В истории культуры отдельные изолированные случаи образования культурных систем, в которых черный и белый цвета противопоставляются не красному, а синему и желтому (например, как в древнем Китае), стоят особняком. Но для индивидуального цветового восприятия особая роль синего или голубого (светло-синего) цвета выявилась как в экспериментальной психологии [6], так и в произведениях писателей (преимущественно романтиков: голубой цветок Новалиса, стихи о синем цвете Бараташвили и т. п.).

По данным лингвистической типологии цветовых обозначений синий (голубой) цвет существует только в таких шестичленных системах, в которых, кроме трех цветов, входящих в основную систему (черный — белый — красный), обязательно есть также желтый и зеленый цвета.

Существуют четырехчленные системы с одним из этих последних цветов и пятичленные системы с ними обоими, но без синего.

Согласно модели описания основных цветовых обозначений в терминах теории нечетких (размытых или расплывчатых) множеств (*fuzzy sets*) [12] предполагается, что шесть цветов образуют основные цветовые категории. Остальные пять лингвистически существенных цветов — коричневый (образующий вместе с шестью вышеперечисленными цветами семичленную систему), лиловый, розовый, оранжевый, серый — образуются соответственно благодаря пересечению нечетких (расплывчатых) множеств (ср. традиционное понятие смешения цветов). Соответственно коричневый цвет представляет собой пересечение черного и желтого, лиловый — красного и синего, розовый — красного и белого, оранжевый — красного и желтого, серый — белого и черного. Хотя для каждого из этих цветов, входящих вместе с шестью основными в одиннадцатичленный универсальный набор цветов, во многих языках есть специальные названия (не являющиеся производными от других названий), тем не менее название соответствующих пяти «сложных» («смешанных») цветов оказывается в таком языке, как английский, существенно менее быстрым. Для типологии систем цветовых обозначений существен рус. *голубой*, представляющий пересечение нечетких множеств для синего и белого. Соответственно можно предположить наличие двенадцатичленной системы, отраженной в русском языке.

Новейшие исследования по нейросемиотике цвета [15, 26] обнаружили не только единство общечеловеческой системы названия цветов, но и наличие соответствия между этой системой и в основном изоморфной ей системой основных цветовых эталонов, определяющих цветовое восприятие. Но вместе с тем экспериментально психологические данные позволили предположить, что названия цветов преимущественно связаны с доминантным по речи (обычно левым) полушарием мозга и соответственно с правым зрительным полем. Само же различие цветов осуществляется субдоминантным (обычно правым) полушарием, причем по клиническим данным наибольшую роль играет его затылочная и теменная области, меньшую — лобная, еще меньшую — височная, функции которой в этом отношении сравнимы с функциями левого полушария. Развитие соответствующих способностей правого полушария у детей сопоставимо с данными об антропоидах.

В свете этих данных нейросемиотики особенности названия цветов, соотносимые с левым полушарием, и особенности опознавания цветовых оттенков, производимого преимущественно правым полушарием, следует понимать как два отдельных механизма, которые, хотя в норме и согласованы друг с другом, могут иногда оказываться в конфликте.

4. Нейросемиотика музыки и нейролингвистика. Различие между двумя стратегиями мозга при решении знаковых задач (при различных операциях над знаками) и между соответствующими знаковыми системами лучше всего выявляются по отношению к языку и музыке, в их сравнении друг с другом. При семиотическом изучении музыки одной из центральных проблем является выяснение специфических отличий музыки от других знаковых систем, в частности от естественного языка. Многочисленные труды по этому вопросу, появившиеся в последние годы, предлагают решения, казалось бы, противоположные. С одной стороны, предполагается, что внутреннее функционирование музыки должно быть изучено по образцу лингвистики [62], что соответствует общему для различных течений современной науки пониманию лингвистики как модели структурного (морфологического в смысле Гёте) описания.

С другой стороны, отмечается существенное различие в числе и характере уровней естественного языка и музыки и в заданной длине некоторого первичного элемента морфемы мотива: в языке его длина всегда ограничена, в музыке же он может быть продолжен (в принципе неограниченно) [35]. Последнее различие, однако, становится менее значимым, если учесть, что во многих языках основной единицей, сопоставимой с мотивом, является не морфема, а слово, длина которого может варьировать в очень широких пределах (сложные слова в санскрите и немецком, фразы — атрибуты в современном английском, инкорпорирующие комплексы — в американских, индейских и чукотско-корякских языках и т. п.). Более существенно то, что непосредственная семантизация определенных отрезков (в естественном языке обязательная и отсут-

ствующая лишь в специфических типах речи, например, заушной) в музыке обнаруживается, наоборот, лишь в исключительных случаях.

Поэтому более оправданным представлялось бы не соположение музыковедения и лингвистики, предполагающее непосредственное перенесение методов лингвистического анализа в анализ музыкальный [36, 46, 47, 58, 61, 64], а разработка таких семиотических методов описания музыки, которые учитывали бы не только опыт лингвистики, но и достижения других семиотических дисциплин, в частности логической семантики и математической логики (так, понятие некоторого абстрактного языка, интерпретируемого на определенном множестве, может более непосредственно помочь решению вопросов музыкальной семантики, чем опыт лингвистической семантики). Исключительное, «модельное» значение лингвистики как науки, выработавшей определенные методы анализа знаковых систем естественных языков, не следует понимать как необходимость и неизбежность применения именно лингвистических приемов к исследованию такой существенно отличной от естественного языка системы, как музыка.

Одним из возможных аспектов исследования, проясняющего соотношение между музыкой и другими типами знаковых систем, является прагматический — в семиотическом смысле, т. е. ориентированный на характер отношений между музыкой и человеком, который ее слушает, сочиняет или исполняет. Этот подход отличается от исследованного в последнее время синтаксического и семантического подходов, из которых первый ориентируется на исследование знаков определенной системы в их внутренних соотношениях, второй — на возможные соответствия между знаками этой системы и некоторыми другими знаками или некоторым фрагментом текста (частным случаем является синтаксическая семантика, при которой изучается смысл, задаваемый преобразованиями внутри данной системы).

Большие возможности для прагматического (в указанном семиотическом смысле) исследования открывает интенсивно проводящееся в последние годы изучение связи музыки (и других систем знаков) с функционированием различных частей головного мозга человека. По образцу термина «нейролингвистика», обозначающего исследование речевых функций мозга (и их расстройств при афазии), используется термин «нейросемиотика», обозначающий дисциплину, исследующую семиотические функции мозга. В таком случае нейролингвистика — часть нейросемиотики.

В настоящей статье рассмотрены различные стороны нейросемиотики музыки способствует выявлению некоторых специфических свойств музыки как знаковой системы. На протяжении почти столетия основным средством исследования размещения в мозгу человека музыкальных функций было клиническое изучение больных с поражениями одного из полушарий. Великий английский невролог Джексон еще в 60-х годах XIX в. установил различие функций двух полушарий: левого — логического и

грамматического (пропозиционального — propositional) и правого, передающего пространственными операциями [22].

Джексон же обнаружил и явное различие в функциях двух полушарий по отношению к речи и пению: больные, которые не могли говорить из-за поражения речевых зон левого полушария, сохраняют способность к пению [41]. Вывод о сохранении способности к пению и музыкальному выражению при нарушениях речи, вызванных расстройствами работы левого полушария, был подтвержден и другими исследованиями [41, 44].

Основным результатом многочисленных исследований об амнезии, основанных на клиническом материале, явилось установление преимущественной роли правого, а не левого полушария по отношению к музыке, особенно к музыкальному выражению [44]. При сохранности речевых способностей поражения правого полушария вызывают нарушение способностей музыкальных [41, 44, 73, 74].

Существенный сдвиг в исследовании этой проблемы обозначился в последние годы благодаря применению новейших клинических, хирургических и экспериментальных методов, позволивших дифференцировать разные аспекты музыки в ее отношении к двум полушариям. Б. Милнер установила, что при удалении правой височной доли обнаруживаются нарушения распознавания тембра и тональной памяти, которых не бывает при аналогичных операциях на левом полушарии. Согласно выводам Милнер и других исследователей, подтвердивших ее результаты [41], с правой височной долей связано восприятие громкости и временных музыкальных различий (длительности), но не различий по высоте и ритму. После удаления правой височной доли больные не могли различать простые мелодии из четырех нот, в которых была заменена одна нота.

При исследовании пробой Вада больных эпилепсией перед операцией перерезания соединительных путей между полушариями после введения амилбарбитонового препарата в левую мозговую артерию лучше сохранилась способность к пению при сильной затрудненности речи и арифметических действий, тогда как при аналогичной пробе на правой артерии были полностью утрачены мелодические (но не ритмические) компоненты пения. Обсуждая результаты своего эксперимента, Боген и Гордон предположили, что правое полушарие ведает музыкальным произведением как целостным вневременным единством, тогда как левое организует построение во времени [41].

Однако под построением во времени имеется в виду только ритмическая структура (но не отношения по длительности); заманчивой представляется возможность соотношения его с другими функциями правого полушария, оперирующего с целостными (глобальными) неразчлененными единствами (структурами или гештальтами) в отличие от левого полушария, имеющего дело с построением любых последовательностей (языковых, числовых и т. д.), в том числе и ритмических. Восприятие музыкального

письмо можно сопоставить с нотной записью. Те пространственно-зрительные образы предметов, которые представляют собой означаемую сторону многих знаков иероглифического письма, хранятся в правом (субдоминантном) полушарии, как и означаемые стороны этих знаков — начертания иероглифов. Поэтому кодирование этих означаемых сторон посредством означаемых может быть осуществлено в пределах самого правого полушария (тогда как для слогового письма требуется обращение к фонемному составу слов, которым ведает левое полушарие).

Эта проблема стала в центре внимания японской научной (и даже более широкой) общественности [70]. Высказанная еще Н. Бором мысль о японской культуре как выражении принципа дополнительности [12] получила неожиданное нейрофизиологическое подтверждение. Японская культурная традиция, обращающаяся к двум разнородным типам систем знаков — левополушарной слоговой азбуке и устному языку, с одной стороны, правополушарной иероглифике — с другой, занимает в этом отношении промежуточное место между двумя крайними полюсами.

На одном полюсе крайней правополушарности находятся культурные традиции типа китайской, где правополушарность прослеживается в различных семитических системах [12, 45], в частности крайне осложняющих процедуру и сроки обучения, растягивавшегося иногда в старом Китае на целую жизнь взрослого, всю жизнь готовившегося к сдаче экзамена. Хотя иероглиф можно воспринимать, запоминать и воспроизводить не только как глобальное целое с помощью правополушарной стратегии, но и как состоящий из набора элементов («черт» в китайской педагогической традиции) и, следовательно, возможно оперирование с иероглифами обоих полушарий [45], тем не менее основной является установка на глобальность (правополушарность). Другой тип культуры в истории формируется с опорой на алфавитное письмо и другие дискретные наборы стимулов (аналоги им обнаруживаются уже в графике Верхнего Палеолита) [12]. Однако в современных европейских и американских научных текстах значительную часть составляют специальные символы, функционально соответствующие иероглифике (простейшие примеры — обозначения чисел цифрами; из более сложных можно упомянуть знаки интегралов и вообще всю систему символов, используемую в анализе и т. п.). Поэтому некоторые черты, характерные для «мозга японца», с полным правом могут быть сопоставлены и с характеристиками «европейского ученого».

Разнесенность по двум полушариям оптических систем знаков выявляется при афазии, вызванной поражением левого полушария, когда у глухонемого страдает пальцевая азбука (соотносящаяся с дискретными единицами речи, функционально соответствующими фонемам), но не язык иероглифических (образных) знаков [60, 70]. Однако иная ситуация (видимо, с другим исходным распределением функций доминантного и субдоминантного полушарий) встречается в ряде случаев [60]. Им **следует обратить**

особое внимание потому, что и по другим, недавно опубликованным данным не только правое, но и левое полушарие может оперировать (видимо, посредством левополушарной дискретной стратегии) с иероглифическими (образными) знаками (но обратное неверно: правое не может оперировать с расчлененными элементами типа пальцевой азбуки) [60, 70].

А. Р. Лурия впервые установил различие в локализации в мозге операций над различными системами письма, сравнив случаи поражения алфавитных письменных систем, сопряженных с анализом речи на дискретные единицы и поэтому соотносящихся с речевыми (передне-речевыми-височными) зонами левого полушария, и случаи поражения систем иероглифических (в обычном понимании — как китайская иероглифика, в расширенном — как французское письмо, в котором отсутствуют взаимнооднозначные соответствия между буквами и фонемами, и слово должно поэтому запоминаться как глобальная единица — целостное единство) [20, 21, 22].

Замечательные для своего времени открытия А. Р. Лурия, сделанные на материале французского письма, были заново «переоткрыты» на материале того же письма в нейропсихологическом исследовании, напечатанном через двадцать с лишним лет после работ Лурия, без ссылок на него [60].

Работы по расщепленному мозгу приводят к выводу, что «правое полушарие выполняет задачи посредством трансформаций: ОРФОГРАФИЯ → ЗНАЧЕНИЕ → КАРТИНКА и ЗВУК → ЗНАЧЕНИЕ → ОРФОГРАФИЯ» [75, с. 187]. Правое полушарие приобретает зрительный словарь слов (в этом смысле — иероглифов, что при уточнении локализации подтверждает основную идею А. Р. Лурия, приведенную выше) независимо от звукового словаря (в котором слова также хранятся и вызываются из памяти правым полушарием в виде целостных образов) [75, с. 186].

При экспериментальном изучении японских пациентов с частично расщепленным мозгом удалось установить, что правое полушарие научилось читать иероглифы вслух, а знаки слоговой азбуки воспринимать как целостные зрительные символы [70]. При этом следует иметь в виду, что количество информации, приходящейся на один японский иероглиф, в 500 раз больше, чем количество информации, приходящейся на одну английскую букву.

Исследование проблемы соотношения алфавитных и иероглифических систем письма у японцев, помимо общего интереса, который она представляет для моделирования (в том числе и на ЭВМ) процессов распознавания образов, имеет непосредственное отношение к вопросам функционирования двух полушарий в среде, пользующихся двумя разнофункциональными знаковыми системами. В случае японцев речь идет о двух разных системах письма; в случае же полиглотов — о двух разных языках.

По данным, полученным в последнее время, при социальном двуязычии обычно распределение функций такого рода: язык, связанный с логической мыслью европейского типа, ориентирован

на левое полушарие; язык, связанный с мифопоэтической или сказочной (образной) традицией, — на правое. Так распределены соответственно навахо и английский [66]; холи и английский у американских индейцев в США, туркменский и русский и т. п. [16, 31]. Правое полушарие в начале используется для усвоения нового (неродного) языка, но затем постепенно этот последний может по социальным и культурным причинам переместиться в левое полушарие [16].

6. Две стратегии мозга по отношению к решению задач. Различие двух стратегий мозга и соответствующих им морфологических систем (которыми соответственно могут быть левое и правое полушария) при операциях над знаками можно показать на сопоставлении языка и музыки.

Приведенные факты, касающиеся сохранения музыкальных способностей у композиторов, речевые возможности которых оказываются нарушенными из-за поражения левого полушария, аналогичны сходным клиническим данным, относящимся к исполнителям, сохраняющим в подобных условиях музыкальные способности. Описан случай, когда больной с поражением левого полушария был способен руководить оркестром, петь с точным соблюдением ритма, модуляции мелодий, мог замечать неправильности в исполняемой мелодии и сохранял способность запоминания нот [5; 10, с. 47].

Если запоминанием и творческим (синтетическим) воспроизведением музыкальных произведений занимается в основном правое полушарие, то с левым полушарием могут быть связаны преимущественно аналитические операции, необходимые при восприятии музыки людьми со специальным образованием. Судя по результатам недавних экспериментов, с правым полушарием (и левым ухом) связаны преимущества в различении мелодий у музыкально необразованных людей, тогда как у лиц с музыкальным образованием выявилось явное преобладание левого полушария и правого уха [34, 60]. Эти результаты, для которых предложено несколько разных интерпретаций [37, 45, 60, 70], позволяют внести существенные уточнения в высказывавшиеся ранее суждения о характере связи левого полушария с восприятием музыки. Утверждение о преимущественной связи музыки с левым полушарием [55] или же более отчетливые предположения о связи левого полушария именно с музыкальным восприятием (соответственно сенсорной или рецептивной амузии с поражением левого полушария [73, 74]), а также допущение связи амузий разного типа как с левым полушарием, так и с правым, по-видимому, нуждаются в дополнительной экспериментальной проверке, при которой учитывалась бы степень музыкальной образованности исследуемых взрослых (при этом следует иметь в виду, что, как правило, музыкальное образование получают преимущественно люди с генетически переданными музыкальными способностями, с которыми в некоторых случаях, в частности у исполнителей,

может быть связана и амбидекстрия, предполагающая особое соотношение между функциями полушарий).

Что касается новорожденных детей, то недавно проведенное электроэнцефалографическое исследование показало, что у них четко дифференцируется реакция на музыку, локализуемая в правом полушарии, и реакция на речевые стимулы, соотносимая с левым полушарием [40]. Хотя авторы этого исследования находят сходство в своих результатах и в данных аналогичных экспериментов на взрослых [59, 65, 71], тем не менее необходимо подчеркнуть существенные различия в характере работы больших полушарий головного мозга у нормального взрослого и младенца.

Поскольку у новорожденного ребенка межполушарные связи еще не сложились, функционирование каждого из его полушарий происходит вне воздействия другого, как у взрослого при рассечении межполушарных связей или при инактивации одного из полушарий. Наблюдающееся при инактивации доминантного (левого) полушария улучшение распознавания субдоминантным (правым) полушарием коротких музыкальных фраз (в частности, сокращение латентного периода ответов) [2, с. 83—93; 3] вызывается именно отсутствием тормозящего воздействия противоположного полушария. В этом смысле можно говорить о врожденных музыкальных способностях правого полушария («правого мозга»), противостоящих речевым способностям левого полушария («левого мозга»). Очевидно это противопоставление характерно для ранних этапов онтогенеза (и возможно филогенеза), затем оно осложняется воздействием демпфирующего механизма парных полушарий, соединенных комиссурами.

Особенно интересны данные, по которым доминантное (речевое) полушарие в условиях инактивации правого (субдоминантного) полушария у взрослого тяготеет к замене распознавания мелодий их словесной классификацией [2, 3, 4]. В этом смысле языковые (не специальные) описания музыки, частично уже разбирившиеся с точки зрения лингвистической семантики, интересны именно как результат функционирования механизмов, в норме соотносящихся с доминантным полушарием. Так, при инактивации правого полушария (височный электрошок) в экспериментах, проведенных Л. Я. Балоновым и В. Л. Деглиным [2, 4], наблюдаются следующие показательные смысловые ассоциации со словом музыка: «бывает печальная, веселая, грустная, добрая, радостная музыка»; после дополнительного настояния экспериментатора на продолжении ассоциаций «что еще к музыке? веселая музыка настроение поднимает... — хорошо»; «музыка тихая, громкая, веселая, грустная»; «грустные звуки»; «музыка может быть красивой, может быть некрасивой, бывает хорошая, бывает плохая. Музыка нервы успокаивает, даже излечивает людей хорошая музыка»; «музыка громкая, тихая музыка, красивая музыка, некрасивая музыка». Характерно для левого полушария здесь использование бинарных семантических оппозиций (тихий—громкий, веселый—грустный).

ЛИТЕРАТУРА

1. Азо А., Хопкрофт Дж., Ульман Дж. Построение и анализ вычислительных алгоритмов/Пер. с англ. А. О. Слисенко. Под ред. Ю. В. Матиясевича. М.: Мир, 1979.
2. Балонов Л. Я., Деглин В. Л. Слух и речь доминантного и субдоминантного полушарий. Л.: Наука, 1976.
3. Балонов Л. Я., Деглин В. Л. Восприятие звуковых неречевых образов (слуховой и музыкальный глосис) в условиях инактивации доминантного и недоминантного полушарий.— Физиология человека, 1977, т. 3, № 3, с. 415—433.
4. Балонов Л. Я., Деглин В. Л. Унилатеральный электросудорожный припадок (нейрофизиология, клиника, лечебное действие при психозах). Л.: Наука, 1979.
5. Брагина Н. Н., Доброхотова Т. А. Функциональные асимметрии человека. М.: Медицина, 1981.
6. Вейн А. М., Родштадт И. В., Соловьева А. Д. Вопросы полушарной асимметрии головного мозга и проблема бессознательного в свете анализа функционально-органных клинических симптомов.— В кн.: Бессознательное: Природа, функции, методы исследования. Тбилиси: Мецниереба, 1978, с. 314.
7. Глезер В. Д., Иванов В. А., Щербач Т. А. Ответы рецептивных полей зрительной коры кошки на сложные стимулы.— Физиол. журн. СССР, 1972, т. 58, № 3.
8. Глезер В. Д., Иванов В. А., Щербач Т. А. Исследование рецептивных полей нейронов зрительной коры кошки как фильтров пространственных частот.— Физиол. журн. СССР, 1973, т. 59, № 2.
9. Глезер В. Д., Дудник К. Н., Куперман А. Н. и др. Зрительное опознавание и его нейрофизиологические механизмы. Л.: Наука, 1975.
10. Доброхотова Т. А., Брагина Н. Н. Функциональная асимметрия и психопатология очаговых поражений мозга. М.: Медицина, 1977.
11. Иванов В. В. Об одном типе архаических знаков искусства и пиктографии.— В кн.: Равные формы искусства/Под ред. Е. М. Мелетинского. М.: Искусство, 1972.
12. Иванов В. В. Чет и нечет: Асимметрия мозга и знаковых систем. М.: Сов. радио, 1978. (Сер. Кибернетика).
13. Иванов В. В. Машинные комплексы и полушария мозга.— Вопр. философии, 1979, № 4.
14. Иванов В. В. Структура гомеровских текстов, описывающих психические состояния.— В кн.: Структура текста. М.: Наука, 1980.
15. Иванов В. В. Цветовая символика в географических названиях в свете данных типологии.— В кн.: Балто-славянские исследования, 1980. М.: Наука, 1981.
16. Иванов В. В. Нейролингвистический аспект порождения двуязычных текстов у полиглотов.— В кн.: Структура текста-81: Тез. симпозиум. М.: Ин-т славяноведения и балканистики АН СССР, 1981, с. 27.
17. Кауфман Д. А. Экспериментальное изучение особенностей функциональной специализации полушарий мозга при шизофрении.— Физиология человека, 1979, т. 5, № 6, с. 1007—1019.
18. Кауфман Д. А., Траченко О. П. Исследование межполушарной асимметрии у здоровых и больных шизофренией методом дихотического тестирования.— Физиология человека, 1981, т. 7, № 6, с. 1034—1040.
19. Позин Н. В., Любинский И. А., Левашов О. В. и др. Элементы теории биологических анализаторов/Под ред. Н. В. Позина. М.: Наука, 1978.
20. Лурия А. Р. Травматическая афазия. М.: Изд-во АМН, 1947.
21. Лурия А. Р. Очерки психофизиологии письма. М.: Изд-во АН РСФСР, 1950.
22. Лурия А. Р. Основы нейропсихологии. М.: Изд-во МГУ, 1973.
23. Лурия А. Р., Цветкова Л. С., Футер Д. С. Афазия у композитора.— В кн.: Проблемы динамической локализации функций мозга: Сб. тр., посвящ. С. А. Саркисову в связи с 70-летием со дня рождения. М.: Медицина, 1968, с. 328—333.
24. Манин Ю. И. Физика и математика. М.: Знание, 1980, с. 23.
25. Нейман Дж. Теория самовоспроизводящихся автоматов/Пер. с англ. В. А. Стефанюка. Под ред. В. И. Варшавского. М.: Мир, 1971.
26. Николаенко Н. Н. О роли доминантного и недоминантного полушарий мозга в восприятии и обозначении цвета.— Физиология человека, 1981, т. 7, № 3, с. 443—448.
27. Осовец С. М., Гинзбург Д. А., Гурфинкель В. С., Латаш Л. П., Малкин Б. В. и др. Ритмы мозга.— В кн.: Нелинейные волны. Стохастичность и турбулентность. Материалы V Всесоюз. школы по нелинейным волнам. Горький: Ин-т прикл. физики АН СССР, 1980, с. 172—199.
28. Попов Е. П., Верещагин А. К., Зенкевич С. Л. Манипуляционные роботы: Динамика и алгоритмы. М.: Наука, 1978.
29. Симеоничная Э. Г. Доминантность полушарий. М.: Изд-во МГУ, 1978.
30. Тимофеев А. В. Роботы и искусственный интеллект. М.: Наука, 1978.
31. Черниговская Т. В., Балонов Л. И., Деглин В. Л. Функциональная асимметрия мозга и построение текстов при билингвизме; (Нейрофизиологическое исследование).— В кн.: Структура текста-81. М.: Ин-т славяноведения и балканистики АН СССР, 1981, с. 163—165.
32. Alajouanine T. Aphasia and artistic realization.— Brain, 1948, vol. 71, p. 229—241.
33. Barlow H. B. Is neurobiology totally non-mathematical? — Lect. Math. Life Sci., 1976, vol. 8, p. 164—182. (Some mathematical questions in biology VII).
34. Bever T. G., Chiarello R. J. Cerebral dominance in musicians and nonmusicians.— Science, 1974, vol. 185, p. 537—539.
35. Cazimir B. Qu'est-ce que la linguistique mathématique peut faire pour la sémiotique musicale (II).— Rev. roum. linguist., 1977, t. 22, p. 171—174.
36. Chandola A. Some system of musical scales and linguistic principles.— Semiotica, 1970, N 2, p. 135—160.
37. Corballis M. C., Beale I. L. The psychology of left and right. Hillsdale (N. J.), 1976.
38. Crick F. H. C. Thinking about the brain.— Sci. Amer., 1979, vol. 241, N 3 (The Brain), p. 219—232.
39. Eccles J. C., Popper K. R. The self and its brain. N. Y. etc.: Spring. Intern., 1977.
40. Gardiner M. F., Walter D. O. Evidence of hemispheric specialization from infant EEG.— In: Lateralization in the nervous system/Ed. by S. Harnad. N. Y.: Acad. press, 1977, p. 481—502.
41. Gates A., Bradshaw J. L. The role of the cerebral hemispheres in music.— Brain and Language, 1977, vol. 4, N 3, p. 403—431.
42. Goodglass H., Baker E. Semantic field, naming and auditory comprehension in aphasia.— Ibid., 1976, vol. 3, p. 359—374.
43. Goodglass H., Kaplan E. The assessment of aphasia and related disorders. Philadelphia: Lea a. Febiger, 1972.
44. Gordon H. W. Hemispheric asymmetries in the perception of musical chords.— Cortex, 1970, vol. 6, p. 387—398.
45. Hardyck C., Tzeng O. J. L., Wang W. S.-Y. Cerebral lateralization of function and bilingual decision processes: is thinking lateralized? — Brain and Language, 1978, vol. 5, N 1, p. 56—71.
46. Herndon M. Le modèle transformationnel en linguistique: ses implications pour l'étude de la musique.— Semiotica, 1974, vol. 15 N 1, p. 71—82.
47. Hirbour-Paquette L. La linguistica: una nuova via per l'analisi musicale?— Versus, 1973, vs/5, p. 2—18.
48. Ingegnieros J. Les aphasies musicales.— Nouvelles iconographie de la Salpetriere, 1906, vol. 19, p. 362—381.
49. Jaynes J. The origin of consciousness in the breakdown of the bicameral mind. Boston: Houghton Mifflin, 1976.
50. Kandel E. R. Cellular basis of behaviour: an introduction to behavioural neurobiology. W. H. Freeman a. Co, 1976.

51. *Kandel E. R.* Cellular insights into behaviour and learning.— Harvey Lect., 1979, Ser. 73, p. 29—92.
52. *Kandel E. I., Peresedov V. V.* Stereotaxic clipping of arterial aneurysms and arteriovenous malformations.— J. Neurosurgery, 1977, vol. 46, N 1, p. 12—23.
53. *Kimura D.* Left-right differences in the perception of melodies.— Quart. J. Exp. Psychol., 1964, vol. 16, p. 355—358.
54. *Kimura D.* Functional asymmetry of the brain in dichotic listening.— Cortex, 1967, vol. 3, N 2, p. 163—178.
55. *Kleist K.* Sensory aphasia and amusia. The myeloarchitectonic basis. Oxford; London: Pergamon press, 1962.
56. *Levy J., Trevarthen C., Sperry R. W.* Perception of bilateral chimeric figures following hemispheric deconnection.— Brain, 1972, vol. 95, p. 61—78.
57. *Lurçat L.* Completément de formes et symétrie.— J. psychol. norm. et pathol., 1978, N 3, p. 309—320.
58. *Mache F. B.* Méthodes linguistiques en musicologie.— Musique en jeu, 1971, vol. 5, p. 75—91.
59. *Mc Kee G., Humphrey B., McAdam D.* Scaled lateralization of alpha during linguistic and musical tasks.— Psychophysiology, 1973, vol. 10, p. 441—443.
60. *Moskovitch M.* On the representation of language in the right hemisphere of the right-handed people.— Brain and Language, 1976, vol. 3, N 1, p. 47—71.
61. *Moutard N.* L'emploi des termes linguistiques en musicologie actuelle.— Linguistique, 1979, vol. 15, p. 85—93.
62. *Nattiez J. J.* Trois modes linguistiques pour l'analyse musicale.— Musique en jeu, 1973, vol. 10, p. 3—11.
63. *Ojemann G. A.* Asymmetric function of the thalamus in man.— Ann. N. Y. Acad. Sci., 1977, vol. 299, p. 380—396.
64. *Ruwet N.* Musicologie et linguistique.— Rev. Intern. sci. social., 1967, vol. 19, N 1, p. 85—93.
65. *Schwartz G. E., Davidson R., Maer F., Bromfield F.* Patterns of hemispheric dominance during musical, verbal and spatial tasks.— Psychophysiology, 1974, vol. 11, p. 227.
66. *Scott S., Hynd G. W., Hunt L., Weed W.* Cerebral speech lateralization in the native America Navaho.— Neuropsychologia, 1979, vol. 17, p. 89—92.
67. *Searleman A.* A review of right hemisphere linguistic capabilities.— Psychol. Bull., 1977, vol. 84, N 3, p. 502—528.
68. *Spellacy F.* Lateral preferences in the identification of pattern stimuli.— J. Acoust. Soc. Amer., 1970, vol. 47, N 2, p. 574—578.
69. *Spreen O., Spellacy F. T., Reid J. R.* The effect of interstimulus interval and intensity on ear asymmetry for nonverbal stimuli in dichotic listening.— Neuropsychologia, 1970, vol. 8, N 2, p. 245—250.
70. *Sugishita M., Iwata M., Toyokura Y.* et al. Reading of ideograms and phonograms in Japanese patients after partial commissurotomy.— Ibid., 1978, vol. 16, p. 417—426.
71. *Taub J. M., Tanguay P. E., Doubleday C. N., Clarkson D.* Hemisphere and ear asymmetry in the auditory evoked response to musical chord stimuli.— Physiol., Psychol., 1976, vol. 4, p. 11—17.
72. *Van Lancker D.* Heterogeneity in language and speech: neurolinguistic studies. Los Angeles: Univ. Cal., 1975. (Working papers in phonetics; 29).
73. *Wertheim N.* The amusias.— In: Handbook of clinical neurology. Amsterdam: North Holland, 1969, vol. 4.
74. *Wertheim N., Botz M. I.* Receptive amusia: a clinical analysis.— Brain, 1961, vol. 84, p. 19—30.
75. *Zaidel E.* Lexical organization in the right hemisphere.— In: Cerebral correlates of conscious experience/Ed. by J. Eccles. Amsterdam: North-Holland Biomed. press, 1978, p. 177—179.
76. *Zurif E. B., Blumstein S. E.* Language and the brain.— In: Linguistic

theory and psychological reality. Cambridge (Mass.); The MIT press, 1978, p. 228—275.

77. *Zurif E. B., Caramazza A., Myerson R., Galvin J.* Semantic feature representation in normal and aphasic language.— Brain and Language, 1974, vol. 1, p. 167—187.
78. *Zurif E. B., Mendelson M.* Hemispheric specialization for the perception of speech sounds: the influence of intonation and structure.— Perception and Psychophys., 1972, vol. 11, p. 329—332.

УДК 519.95:7

Ю. М. Лотман

КУЛЬТУРА И ТЕКСТ КАК ГЕНЕРАТОРЫ СМЫСЛА

Моделирование искусственного интеллекта наталкивается на ряд трудностей, наиболее существенные из которых связаны с неясностью самого понятия «интеллект». Это вызвано тем, что понятие «интеллект» и «интеллектуальная деятельность» безусловно отождествляются с интеллектом отдельного человека. Предметом изучения становится некоторый уникальный объект, который мы определяем через противопоставление объектам, лишенным признака интеллектуальности (часто это противопоставление имеет чисто априорный характер, как, например, при абсолютизации оппозиции «человек ↔ животное»), и не можем определить основным для науки способом — методом включения в некоторый ряд подобных по основному признаку явлений. В результате неясно, что же мы определяем: интеллект вообще или человеческий интеллект. Подмена одного понятия другим ясно проявилась, например, в определении Тьюринга, согласно которому к интеллектуальным реакциям следует относить такие, которые в процессе длительного общения не могут быть отличены от человеческих.

В основе настоящей работы лежит гипотеза о том, что человечество в процессе цивилизации создало интеллектуальные устройства, дополняющие соответствующие возможности отдельного индивида. Включение индивидуального сознания в единый ряд с объектами этого рода позволит построить модель интеллектуальности и, следовательно, уяснить сущность искомого в процессе моделирования искусственного интеллекта.

Другой исходной гипотезой является предположение, что в качестве таких «интеллектуальных устройств» могут быть названы культура как семиотический механизм и художественный текст¹. В дальнейшем будет сделана попытка показать, что элементы ряда индивидуальный интеллект — художественный текст —

¹ В определенной мере интересующие нас свойства обнаруживает всякий текст, но мы будем рассматривать тексты искусства, так как в них эти черты проявляются наиболее ярко.

культура обнаруживают черты изоструктурности и изофункционализма, что позволяет говорить о подобии этих объектов. Не ставя перед собой задачи дать исчерпывающее определение и ограничиваясь целью выработки практически удобной формулы, можно было бы определить мыслящий объект как такой, который осуществляет следующие действия:

1) хранит и передает информацию; обладает языком и механизмом порождения правильных текстов на этом языке; 2) осуществляет алгоритмизованные операции по правильному преобразованию этих сообщений; 3) образует новые сообщения.

Мы будем исходить из определения, согласно которому сообщения, образуемые в результате операций, предусмотренных п. 2, новыми не являются, выступая лишь как законные трансформации исходных текстов в соответствии с некоторыми правилами, и могут рассматриваться как один и тот же текст. К п. 3 будут относиться лишь такие тексты, которые не могут быть выведены с помощью однозначных операций из каких-либо уже имеющих текстов. Следовательно, «мыслящим объектом» мы можем считать такой, который обладает памятью и способностями выше некоторого примитивного порога.

Исключительный прогресс всего цикла «наук о человеке», характеризующий исследования последних пятидесяти лет, позволяет одновременно подвергнуть рассмотрению и уточнению некоторые из его принципов, считавшихся до сих пор неизменными.

Нами будут рассмотрены две предпосылки. Первая исходит из того, что наука рассматривает лишь повторяющиеся явления и инвариантные их модели и строится на снятии индивидуальных различий между текстами, с одной стороны, лишенных значимости для коллектива, а с другой — недоступных для научного описания. Метод структурного изучения интерпретировался при этом как ориентированный в принципе на моделирование инвариантного ядра объекта, которое вместе с тем рассматривалось как носитель смысла.

Вторая заключается в утверждении, что целью всякой коммуникации является максимально точная передача некоторого смыслового инварианта, что всякая возникающая в процессе коммуникации трансформация смысла должна рассматриваться как искажение и утрата — результат несовершенства канала связи. Следовательно, в идеальной коммуникационной цепи смысл сообщения отправленного и сообщения полученного будут идентичными.

Предпосылки эти явно способствовали развитию семиотики на определенном ее этапе. Однако в настоящее время уместно обратить внимание на те противоречия, к которым они привели.

Рассматривая всю совокупность порождаемых культурой сообщений, мы обнаруживаем, что только определенная часть из них ориентирована на сосредоточение смысла в инвариантной структуре и, следовательно, на переводимость смысла. Значитель-

ная часть текстов имеет противоположную ориентацию. Факт этот нуждается в объяснении. В равной мере привлекает внимание постоянная тенденция культуры увеличивать внутри себя число языков и умножать разнообразие текстов. Является ли подобная «расточительность» конструктивным дефектом или в ней скрыт глубокий смысл?

Исследования в области семиотики культуры убеждают нас, что все семиотические процессы, совершающиеся в ее пространстве, можно разделить на две группы: функционирующие в пределах одного какого-либо языка или в пространстве двух и более языков. Наиболее изучена первая группа, которая долгое время воспринималась как модель всех семиотических процессов. Изучение второй только начинается.

В пределах одного языка осуществляются: 1) операции по передаче информации (имеют целью передать «другому» в предельном объеме то сообщение, которым располагаю «я»); 2) операции по алгоритмическим трансформациям тех или иных данных текстов; 3) метаязыковые операции; 4) операции по очерчиванию границ личности (смысл их будет раскрыт ниже). В пределах двух или более языков совершаются операции по выработке новых сообщений, связанных с творческим мышлением.

Если выстроить цепочку объектов индивидуальное сознание человека — художественный текст — культура, то мы обнаружим общую структурную черту: соединение в одной структуре двух различных по семиотическим характеристикам структур.

Гипотеза Сепира-Уорфа, имеющая лишь ограниченное значение применительно к языкам, оперирующим условными и дискретными знаками, приобретает весьма серьезный характер, когда речь заходит о переводе с таких языков на принципиально иначе устроенные языки (например, оперирующие иконическими знаками или имеющие недискретную, континуальную природу), и обратно. В этих случаях структурное различие столь велико, а моделирующее воздействие на их сообщение столь значительно, что вопрос о возможности однозначного перевода вообще снимается. Одновременно соотносимость этих сообщений с какими-то общими фактами вне семиотической реальности позволяет устанавливать между ними отношения соответствия. Все перечисленные выше объекты являются текстопорождающими устройствами, состоящими, в свою очередь, из более чем одного текстопорождающего устройства, в которые заложены языки с предельно удаленными структурами.

Так, например, открытия последних двух десятилетий в области функциональной асимметрии больших полушарий головного мозга человека рисуют ясную картину семиотического механизма, единого и биполярного одновременно. Если семиотические процессы, руководимые правым полушарием, характеризуются недискретностью (например, интонации), то левое полушарие связывает значимость с границами признаков, классификационными и подобными отчетливо дискретными механизмами: «в норме левое

полушарие осуществляет классификацию фонем по дифференциальным признакам и поддерживает иерархию этих признаков, обеспечивая устойчивость фонологической системы языка»¹. Мир иконических знаков, по-видимому, тяготеет к правополушарности в отличие от словесного семиозиса левого полушария. Правое полушарие, судя по современному состоянию изучения вопроса, связано с синкретизацией семиотических систем и образованием значений, «размазанных» в некотором семиотическом пространстве, в то время как левое — образует разграниченные между собой и сегментированные внутри семиотические конструкции. В результате работы такого механизма порождаются гетерогенные тексты, представляющие амальгамы противоположных семиотических реализаций (например, в естественной речи — дискретных грамматико-лексических единиц и интонаций). Если учесть, например, для устной речи паралингвистику, для письменной — неизбежность зрительных впечатлений от шрифта, цвета бумаги, то не покажется преувеличением утверждение, что всякий текст на естественном языке представляет собой текст на нескольких языках, вернее, на амальгаме языков со сложной системой отношений между ними. Естественный язык в этом отношении находится в некоторой срединной позиции — от него возможны движения в двух направлениях: в сторону искусственной моноязычности (эта тенденция ощущается в письменной нехудожественной речи и преобладает в искусственных языках и метаязыковых образованиях) или в сторону углубления языковой полярности в механизме текста. Примером второго являются художественные тексты.

На каком уровне ни взяли бы мы художественный текст — от такого элементарного звена, как метафора, и до сложнейших построенных целостных художественных произведений — мы сталкиваемся с соединением несоединимых структур.

Тропы обычно рассматриваются как определенные замещения или соположения знаков, в результате чего между замещенным и замещающим возникают отношения смыслового сдвига (определенное число сем совпадает при несовпадении остальных, что приводит к активизации семантических полей). В результате троп рассматривается как риторическая фигура в пределах одного языка. Однако если отказаться от представления о естественном языке как об однородной семиотической системе и признать его неустранимую гетерогенность и гетероструктурность, то возможно будет представление о тропе как об элементарном тексте на двух языках, взаимная переводимость между которыми возможна лишь условно и с известным напряжением.

Нетрудно было бы показать, что любой художественный текст не является последовательным и однолинейным развертыванием какой-либо одной структуры, а представляет собой переплетение

¹ *Балочов Л. Я., Деглин В. Л.* Слух и речь доминантного и недоминантного полушарий. Л., 1976, с. 182.

и взаимное перекодирование двух или более текстовых структур с разной степенью взаимной непереводимости. Существеннее выяснить, почему многоязычность структуры проявляется на разных уровнях интеллектуальных устройств.

Представим себе некоторое устройство, разделенное на два (или более) кодирующих субустройства, коды которых построены принципиально столь различно, что точный перевод с одного на другой вообще невозможен. Предположим, что один из них будет представлен языком с дискретными знаковыми единицами, имеющими стабильные значения, с линейной последовательностью синтагматической организации текста, а другой будет характеризоваться недискретностью и пространственной (континуальной) организацией элементов. Соответственно и планы содержания этих языков будут построены принципиально различным образом. Если нам, например, потребуется передать текст на языке L_1 средствами L_2 , ни о каком точном переводе речи быть не может. В лучшем случае возникнет текст, который в отношении к некоторому культурному контексту сможет рассматриваться как адекватный первому. Если потом произвести обратный перевод на L_1 , то мы, естественно, не получим исходного текста: по отношению к исходному это будет новым сообщением. Имея такое устройство и подавая ему на вход некоторые тексты, мы будем получать на выходе новые тексты. Следовательно, устройство такого типа будет способно осуществлять акты творческого сознания.

Однако для того чтобы даже относительная эквивалентность между семиотическими единицами одного и другого текста могла быть установлена, необходимы, во-первых, некоторая совместимость кодов, достигаемая отнесенностью их к некоторой внешней реальности и памятью о предшествующих перекодировках, и, во-вторых, наличие некоторой метаязыковой системы, описывающей оба текста как единый текст, а оба механизма — как один механизм. Как правило, такую метаязыковую функцию берет на себя дискретный кодирующий механизм, который одновременно выступает и как часть устройства, и как образ устройства в целом (ср. тенденцию левополушарного сознания представлять сознание как таковое или письменной формой языка подменять собой язык). Творческое сознание основной своей нагрузкой падает на биполярную структуру, а логическое — на моноязыковую метаструктурную область сознания.

Если прибавить этому механизму функцию памяти, рассредоточенную и на различных уровнях текста (ср., например, «память жанра» М. М. Бахтина) и культуры, то мы получим цепь объектов, обладающих интеллектуальными способностями.

Таким образом, мы можем определить всю цепь объектов от элементарного художественного текста типа метафоры («образа») до культуры как текстопроизводящие устройства. Может показаться, что определение это страдает преувеличением, поскольку для того, чтобы любой из этих механизмов выполнил текстопроизводящую работу, необходим вне его находящийся человек, который

бы вводил в него тексты. Вопрос этот можно перефразировать так: можно ли считать интересующие нас объекты «искусственными интеллектами *sui generis*», если для того, чтобы запустить их в действие, необходим на входе системы вводимый извне текст?

Ни одна машина не может работать «сама собой». Это вытекает из второго закона термодинамики и является опровержением возможности перпетуум-мобиле. Человеческий мозг не является исключением из этого правила. Он тоже не работает «сам собой». И не только в том смысле, что он получает из организма физиологическое обеспечение, а из внешнего мира через сенсорные рецепторы — впечатления. Мышление в одиночку невозможно в том смысле, что для его «запуска» необходим вне его находящийся интеллект. Наглядной иллюстрацией этому являются случаи с нормальными детьми, полностью изолированными с младенческого возраста от общества людей: материальная сторона интеллектуального механизма у них в полном порядке, но из-за отсутствия поступающих извне текстов она «не запущена» и остается в бездействии. Таким образом, мышлящее устройство не может заполнять весь свой универсум — оно нуждается в другом мышлящем устройстве и организованных актах обмена текстами. Это закон не только для искусственного интеллекта, но и для всякого интеллекта вообще.

По мере усложнения внутренней семиотической организации и насыщения памяти возникает способность конструировать «другого» внутри собственного семиотического пространства и экстраполировать его во вне. Внутри устройства возникают замкнутые цепи коммуникации, которые позволяют черпать тексты не извне, а из резервов памяти и трансформировать их в новые тексты. Однако, видимо, это лишь частично может заменить необходимость коммуникации с «другим», лежащим вне данного механизма семиотическим образованием.

Сказанное вносит некоторые коррективы в привычные представления о смысле и характере коммуникативных связей.

Распространенное представление о задачах коммуникации основано на мысли о том, что целью ее является полная и свободная от утрат передача сообщения от адресанта к адресату. Если в процессе передачи возникает изменение смысла, то оно трактуется как искажение, нежелательное последствие помех в канале связи. В идеальной схеме, в которой помехи условно приравниваются нулю, текст на входе и выходе коммуникационной цепи берется как абсолютно идентичный. Конечно, такая тенденция в системе коммуникации реально существует, она обеспечивает циркуляцию в коллективе уже выработанных сообщений. Не является ли она единственной и даже доминирующей?

Полная идентичность сообщений на входе и выходе возможна лишь при абсолютной идентичности системы кодирования и декодирования. Если говорить не об идеальных схемах, а о реальных человеческих коммуникациях и определять систему культурных кодов, свойственную той или иной человеческой индивидуальности как семиотической личности, то абсолютная неизменность

сообщения может быть гарантирована лишь при условии, что передающий и принимающий, по существу, являются одной семиотической личностью. В этом случае возникают два соображения. Первое состоит в том, что идеальная (или приближенная к ней) циркуляция сообщений возможна только в условиях, когда она практически не нужна, так как передавать неизменное сообщение от одной семиотической личности к другой невозможно, а в пределах единой семиотической личности — бессмысленно. Второе соображение имеет более принципиальный характер: общее направление зоосемиотического и культурно-семиотического развития ведет в сторону усложнения и индивидуализации кодов каждой отдельной особи.

Ценность сообщения для адресата пропорциональна трудности декодировки, т. е. различию между кодом отправителя и получателя. Ценность эта заключается в том, что в процессе декодировки исходное сообщение трансформируется, выполнив свою главную работу, включив интеллектуальную машину получателя. Как сознание индивидуального человека стереоскопично, образуясь взаимодействием двух типов семиотического преломления текстов в разных полушариях (два сознания) и их слиянием в гетерогенном едином текстообразовании человеческого ума, так и на другом уровне мышление не бывает индивидуальным, оно подразумевает партнера, который в одном аспекте должен быть «не то, что я» (два сознания), а в другом — «то же, что я» (одно сознание).

В связи с этим необходимо уточнить понятие диалога. Диалог включает момент семиотической сопротивляемости получателя информации, который переживает столкновение языка сообщения с языками своей памяти. Роль адресата в диалоге активна: он сопротивляется текстовому вторжению, испытывает такое вторжение, возбуждающее его текстовые потенции и передает новый текст адресанту, который, превращаясь в получателя информации, ведет себя аналогичным образом. Там, где происходит пассивная передача информации, получение команд, диалога нет. В системе «команда — исполнение», даже если она направлена в обе стороны (отдающий и принимающий команду меняются местами), диалога не происходит. Следовательно, так называемые «диалоговые системы» типа «человек — машина» на современной стадии конструирования еще диалоговыми не являются, и превратить их в диалоговые методом количественного усложнения команд и программ не удастся.

Таким образом, мы можем выделить два типа перекодировки: с языка на язык и одной личности (индивидуального многокодового образования) на другую. Разнообразие текстов и разнообразие личностей обеспечивают информационную емкость культуры как интеллектуального текстопорождающего устройства. В механизм культуры заложена тенденция к умножению ее языков и к умножению комбинаций языков, а также к постоянному росту относительно самостоятельных смыслогенерирующих образований типа «текст» и «личность».

Конструктивной особенностью, реализующей принцип семиотического обмена и перекодировки внутри мыслящей единицы как основы порождения новых текстов, является принцип, согласно которому каждое мыслящее устройство одновременно является соединением двух (минимально) мыслящих устройств и частью мыслящего устройства более высокого уровня. На одном конце этой цепи стоит индивидуальное сознание человека; культура, накопленная человечеством, — на другом, если брать синхронный срез. Но поскольку в интеллектуальном механизме особо важную роль играет память, диахронная ось также исключительно активна.

Моделирование искусственного интеллекта до сих пор развивалось в двух направлениях: освоение логических способностей и построение искусственной памяти. Однако без проникновения в механизм творческого мышления построение интеллекта невозможно. На этом пути мы встречаем исторически созданные человеком формы надличностного интеллекта. Изучение их в какой-то мере более доступно, чем проникновение в механизм индивидуального сознания: протекающие в коллективе семиотические процессы доступны прямому наблюдению и широко документированы. Пренебрегать этим источником было бы неблагоприятно. Именно на этом пути семиотика может оправдать возлагавшиеся на нее надежды — синтезировать наиболее отдаленные сферы научного знания человека.

УДК 518.53

Л. А. Чистович, В. А. Кожевников, Ж. А. Першин

ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ СЛУХОВОЙ ОБРАБОТКИ РЕЧЕВОГО СИГНАЛА

Проблема автоматического распознавания устной речи была поставлена перед инженерами примерно 20—25 лет назад. Она довольно быстро разделилась на две задачи. Одна из них — распознавание изолированных слов-команд из ограниченного словаря — довольно успешно решается формальными математическими методами и не требует исследования ни процессов восприятия, ни речеобразования.

Другая задача — автоматическое распознавание естественной слитной устной речи без ограничений, наложенных на словарь, — претерпела значительную эволюцию. Она стала формулироваться как задача понимания смысла устного сообщения и вылилась в более крупную проблему общения человека с машиной на естественном человеческом языке. Важность и значение ее решения ясны. Например, возможность речевого общения с информационной системой, осуществляемого по телефону, открыла бы совер-

шенно новые перспективы использования ЭВМ в самых разных областях.

Уже лет десять тому назад стало очевидным, что единственная корректная формулировка проблемы автоматического понимания слитной речи состоит в определении ее как проблемы моделирования процессов восприятия и понимания речи человеком. Неизбежность такого подхода впервые была показана советскими исследователями [2]. Сейчас этот подход принят повсеместно: во всем мире резко усилились фундаментальные исследования, направленные на выяснение того, что такое процесс восприятия и понимания речи человеком.

Известно, что за последние годы реализованы первые варианты систем, которые можно с некоторой натяжкой назвать системами понимания связной речи [7]. Хотя на условия, при которых системы работают, наложены пока еще серьезные ограничения, уже ясно, что задача перешла из области научной фантастики в область реальности. Другой важный результат состоит в том, что принципы организации системы в значительной мере определились: выяснилось, какие последовательные блоки обработки информации эта система обязательно должна включать, какие вопросы, касающиеся самих этих отдельных блоков, наиболее существенны [4]. Образовалось некоторое новое научное направление, в котором не только объединились, казалось бы, далекие области знания, но и стало ясным, какие задачи оказываются на попечении акустиков, физиологов, фонетиков, математических лингвистов.

Обработка речевого сигнала у человека начинается со слухового анализа этого сигнала. Еще Дарвином, Сеченовым, Павловым и др. высказывались соображения, что воспринимая звуковой речевой сигнал, человек определяет по нему соответствующие этому звуку речевые движения. В 60-х годах наш коллектив провел цикл работ, направленных на экспериментальную проверку этой гипотезы [13].

Использовался очень простой прием. Человеку предъявлялись речевые сигналы (естественные или искусственные — синтетические). Испытуемый должен был как можно точнее и быстрее их повторять. Была разработана система датчиков, позволяющая параллельно звуковой речи регистрировать комплекс артикуляторных явлений, т. е. увидеть, какие движения и в какой последовательности человек делает. Результаты показали прежде всего, что перекодирование в движение происходит очень быстро — задержки ответных артикуляторных реакций относительно начала предъявляемого испытуемому звука составляют 150—200 мс.

Данные принципиальны: из них следует, что требования к объему слуховой оперативной памяти модели не очень высокие — модель должна одновременно обрабатывать сравнительно короткий отрезок сигнала. Более длинные отрезки могут запоминаться уже не в виде слухового изображения, а в сжатом виде (программы артикуляторных движений). Другой результат состоит в том, что пе-

рекодирование принимаемого речевого сигнала в программу артикуляторных движений эквивалентно переходу к описанию сигнала в виде последовательности дискретных элементов, которые уже реально можно обозначить фонетическими знаками.

Это очевидно из экспериментов, в которых испытуемые имитируют синтетические сигналы. Оказывается, что группа физически отличающихся звуковых сигналов, но находящихся в пределах некоторого диапазона значений параметров, вызывает одну и ту же ответную реакцию, а группа сигналов, попадающих в другой диапазон — другую. Число различающихся реакций достаточно ограничено.

Перекодирование устного речевого сигнала в последовательность дискретных элементов — фонетических образов — необходимое условие для того, чтобы в сообщении были выделены и проанализированы слова, т. е. найдено семантическое значение слова и его грамматические характеристики. Иначе говоря, как только речевой сигнал превращается в последовательность дискретных элементов, он поступает на вход следующего уровня обработки информации. Разработка модели этого уровня — ее называют моделью анализирующей части языка — находится в руках математических лингвистов, психолингвистов, нейролингвистов.

В настоящем сообщении будет говориться только о более низких уровнях обработки сигнала, т. е. тех, которые кончаются превращением устного речевого сигнала в дискретные фонетические элементы. Эти уровни обозначаются как блок первичной обработки сигнала, выделяющий в нем полезные признаки, и блок фонетической интерпретации [12]. Стержневым вопросом разработки этих блоков является природа полезных признаков, так как правила принятия решений практически зависят от того, какая система признаков будет выбрана.

Исходный устный речевой сигнал — это просто звуковое давление, определенным образом изменяющееся во времени. Если его записать, получим осциллограмму. На рис. 1 приведена для примера такая запись в общем очень простого слова «какой». Осциллограммы речевых сигналов весьма изменчивы, и узнать прямо по ним, что произносилось, практически невозможно. Неизбежна обработка этих сигналов. От эмпирических поисков методов обработки исходного сигнала все серьезные исследователи речи уже давно отказались. Осталось два пути: ориентированный на модель речеобразования и ориентированный на модель восприятия.

Первый путь вначале казался очень привлекательным, так как акустическая модель речеобразования в достаточно хорошем приближении уже разработана [10]. При определенных упрощениях модель описывается с помощью довольно небольшого числа параметров. Остается, таким образом, найти такие математические методы обработки звукового сигнала, которые позволили бы восстанавливать текущие значения этих параметров и уже по ним осуществлять фонетическую интерпретацию. Однако сейчас у тех исследовательских коллективов, которые раньше всех пошли по

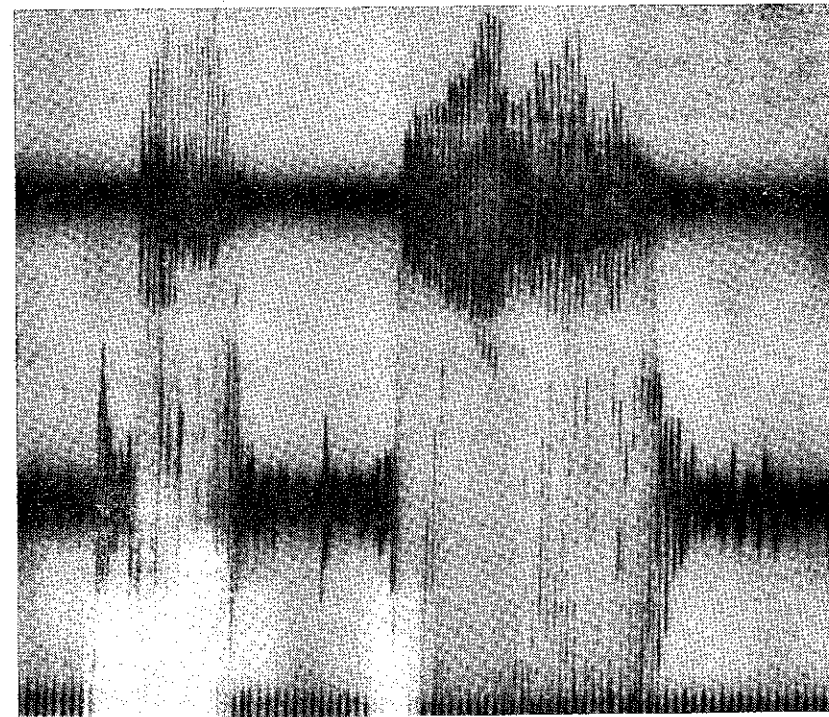


РИС. 1. Осциллограф слова «Какой»

этому пути и лучше всех его разработали, наблюдается существенное разочарование в этом подходе [17]. Теперь все более выражена тенденция пойти по второму пути, т. е. обратиться к механизмам слуха.

Слуховую систему можно условно рассматривать как состоящую из двух частей: периферической и центральной. Периферическая часть — это улитка внутреннего уха, сложная гидродинамическая система, приводимая в колебания поступающим сигналом. В ней, собственно, и осуществляется спектральный анализ звука. Здесь же во внутреннем уже имеется система рецепторов, своего рода датчиков, расположенных вдоль улитки. Рецепторы эти преобразуют механические колебания соответствующих областей улитки в нервную импульсацию, текущую по соответствующим волокнам слухового нерва.

Центральная же слуховая система представлена целым набором определенным образом упорядоченных скоплений нейронов — нервных узлов, расположенных в разных отделах мозга и связанных между собой нервными путями.

Периферическая часть слуховой системы в настоящее время исследуется чрезвычайно интенсивно [15]. Хотя физические про-

цессы, происходящие в улитке, еще не полностью поняты, данных, характеризующих улитку как спектральный анализатор, практически уже достаточно для составления требований к ее функциональной модели. Процесс преобразования механических колебаний в нервную импульсацию значительно менее изучен. Однако и здесь в первом приближении понятно, какие формальные преобразования сигнала должны обеспечиваться и какие эффекты должны воспроизводиться моделью.

Хуже обстоит дело с исследованием центральной части слуховой системы. Здесь вся полученная информация исчерпывается набором эффектов и зависимостей, полученных, с одной стороны, нейрофизиологами, исследующими реакции на звук нейронов центральных образований, и, с другой стороны, психоакустиками, изучающими восприятие звуков. На основании этих данных пока можно делать только некоторые догадки о принципах преобразования сигнала. Вместе с тем именно эта центральная слуховая обработка сигнала наиболее интересна для исследователей речи [9].

Начав экспериментальные исследования в этом направлении, мы быстро убедились, что нам необходима модель периферической части слуховой системы. Без нее невозможно представить, какое изображение поступает на вход центральных уровней. Практически эту периферическую модель нужно было иметь в виде прибора, чтобы с его помощью можно было бы получать изображения любых звуковых сигналов. Необходима также возможность дополнять эту модель более «центральными блоками», которые мы разрабатываем.

На следующей стадии в работу включались французские специалисты.

Исследования по функциональному моделированию в основном выполнялись в Институте физиологии им. И. П. Павлова АН СССР В. С. Шупляковым. В настоящее время модель выполнена в виде двух цифровых вариантов и действующего аналогового прибора [15, 16].

На рис. 2 приведена блок-схема модели. Блок СУ корректирует входной сигнал согласно частотной характеристике среднего уха. Блоки 1 — это резонансные фильтры второго порядка с малой добротностью. Они включены последовательно и развязаны буферными усилителями. Их частотные характеристики умножаются, в результате формируется характеристика с очень крутым склоном в сторону высоких частот. Блоки 2 — аналогичные резонансные фильтры, собственные частоты которых систематически понижаются. Они образуют длинную линию, воспроизводящую прохождение волн вдоль по улитке. Блоки 3 — резонансные фильтры более высокой добротности, обеспечивающие дополнительное формирование частотных характеристик. Выходной сигнал каждого из фильтров представляет колебания определенной области вдоль оси улитки. Таким образом, вся система осуществляет преобразование «частота — координата» и является спектральным анализатором параллельного действия.

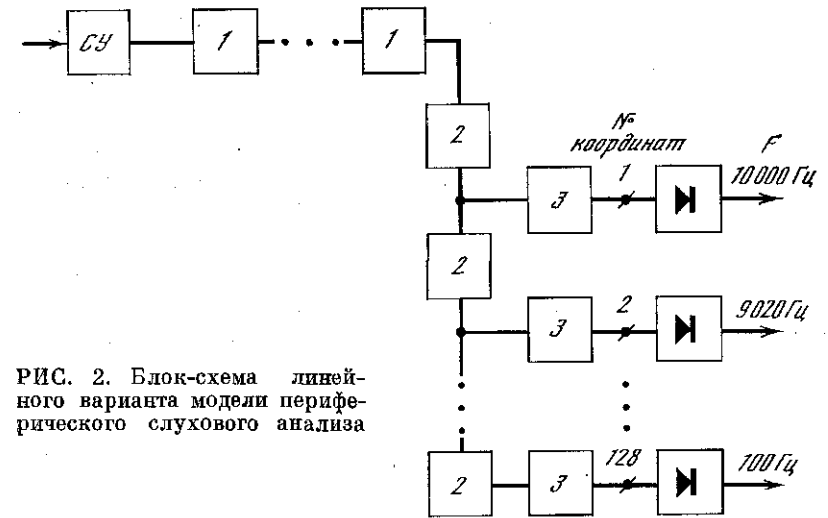


РИС. 2. Блок-схема линейного варианта модели периферического слухового анализа

На рис. 3 приведен отклик модели на слово КАКОЙ. Такое изображение можно условно назвать слуховой спектрограммой [3]. По оси ординат — текущее время, по оси абсцисс — координаты улитки, или, что то же самое, — частота; почернение отражает интенсивность колебаний.

Одним из ключевых вопросов в автоматическом распознавании является вопрос о сегментации сигнала. Если сигнал сначала разделяется на отрезки (на рис. 3 такие отрезки, соответствующие звукам слова КАКОЙ, легко могут быть выделены), то можно использовать при его распознавании такие признаки, как длины отрезков, «начальный спектр», «конечный спектр» и т. п. Однако, как правило, большая часть звуков в нормальной слитной речи непосредственно не разделяется, даже между большей частью слов во фразе паузы отсутствуют.

Если сегментации нет, то приходится основываться на текущих значениях спектра или иных параметров сигнала. Типичен подход, когда берутся пробы спектра каждые 10 мс и каждый раз решается вопрос, на какой гласный или согласный эта проба больше всего похожа. Именно этот подход используется техниками в большинстве существующих сейчас систем автоматического распознавания [7].

Нейрофизиологические данные показывают, что в первой системе существует какой-то механизм выделения краев звука. Значительное число нейронов отвечает только в моменты «начал» или «концов» звуковой посылки. Был проведен цикл психоакустических работ, направленных на выяснение этих механизмов [11, 12]. Данные позволяют утверждать, что выделяются резкие изменения огибающей плотности импульсации в отдельных частотных каналах слуховой системы, т. е. в группах нейронов, связанных с узкими участками мембраны улитки.



РИС. 3. Изображение на выходе модели слухового спектрального анализа

Экспериментальные данные в первом приближении описываются моделью [9], состоящей из полосового фильтра с центральной частотой 25 Гц, на вход которого поступает выпрямленный и подвергнутый логарифмированию сигнал с выхода фильтра улитки. На выходе полосового фильтра находятся устройства, вырабатывающие марки слуховой сегментации [5]. Рис. 4 показывает реакцию модели на то же слово «какой». Видно, что марками отмечаются те явления, которые обозначаются как «начала» звуков (черные марки), «концы» звуков (белые марки), и частотные «переходы» — белые марки, идущие в косом направлении.

Наша модель, как и человек, начинает обнаруживать паузы в сигнале длительностью всего 4—5 мс. Один из недостатков большинства существующих систем автоматического распознавания состоит в том, что они теряют быстрые и кратковременные акустические события.

Классический вопрос автоматического распознавания касается выбора фонетически значимых параметров спектра. Его удобно пояснить, показав (рис. 5) слуховой спектр естественных стационарных гласных *U, O, A*. Эти спектры получены с помощью цифрового варианта модели улитки [8]. Видно, что изображения довольно сложные и описать их можно разнообразными способами.

Ряд исследований с синтетическими гласными направлен на поиск минимальных различий в звуках (т. е. в их спектральных изображениях), при которых они относятся к двум разным гласным, и на поиск условий, при которых имеется максимальное сходство при восприятии звуков с явно разными спектральными изображениями [6, 14]. Результаты этих экспериментов позволяют считать, что первым этапом центральной обработки периферического слухового спектра является выделение на нем локальных неоднородностей (по оси частот). Простейшая и физиологически

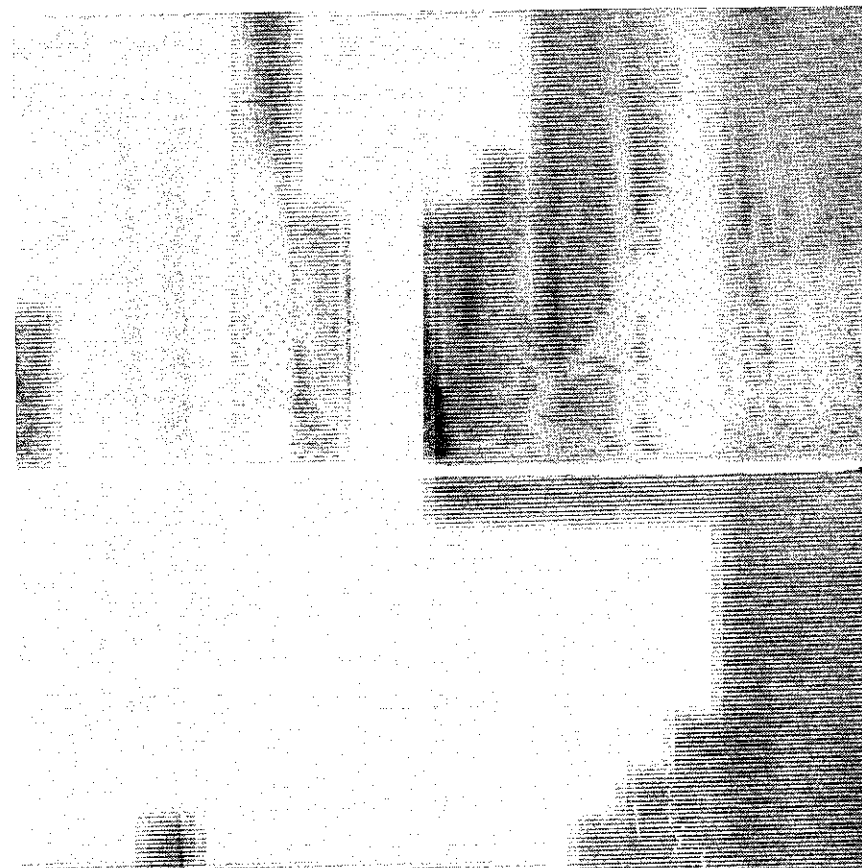


РИС. 4. Изображение слова «Какой» на выходе модели обработки амплитудных изменений в «Слуховых каналах»

правдоподобная модель, способная это делать — так называемая процедура латерального торможения. Она вычисляет нечто близкое ко второй производной от кривой, описывающей спектр. Параметры этой модели в первом приближении уже определены [1].

На втором этапе обработки происходит интегрирование изображения по довольно широким областям шкалы частот: область усреднения не менее октавы. В настоящее время нам приходится рассматривать несколько вариантов модели этого этапа; проводятся эксперименты для выбора между вариантами и определения количественных параметров блока модели.

В заключение остановимся на двух принципиальных вопросах.

В задаче автоматического распознавания речи исследователь все время сталкивается с противоречивыми требованиями. С одной стороны не терять подробностей изображений как временных, так

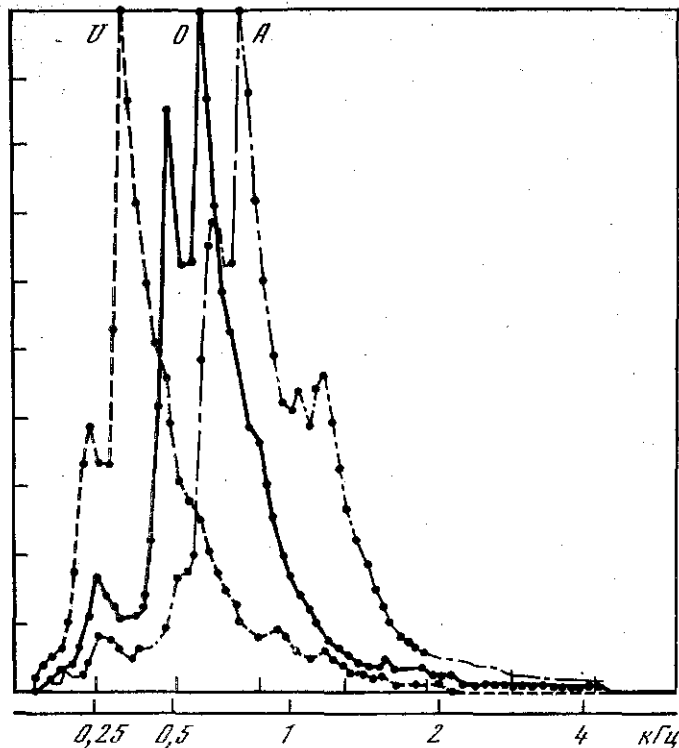


РИС. 5. «Слуховые спектры» стационарных гласных И, О, А

и спектральных, с другой — производить укрупнение информации как по оси времени, так и по оси частот.

Можно достаточно твердо сказать, что в слухе эта задача решается благодаря параллельной обработке. Периферическое изображение сигнала размножается в центральной слуховой системе в ряд параллельных версий. Но существенно, что оси изображений (оси времени и частот) остаются теми же самыми во всех версиях. Таким образом, наблюдается принцип, существенно отличающийся от использованного техниками, пытающимися с самого начала резко сократить обрабатываемую информацию.

Вопрос второй касается того, практически ли стремиться к копированию в технической системе принципов обработки сигналов мозга? По мнению многих специалистов — да. Затраты человеческого труда и машинное время в этом случае непосредственно уйдут на исследования и разработку модели, а как только будут разработаны приемлемые, хотя бы и громоздкие, алгоритмы, их можно будет реализовать с помощью микропроцессоров.

ЛИТЕРАТУРА

1. Бедров Я. А., Чистович Л. А., Шейкин Р. Л. Частотное положение «центра тяжести» формант как полезный признак при восприятии гласных. — Акуст. журн., 1978, т. 24, с. 480—486.
2. Бондарко Л. В., Загоруйко Н. Г., Кожеевников В. А. и др. Модель восприятия речи человеком. Новосибирск: Ин-т математики СО АН СССР, 1968.
3. Бородин А. Н., Головешкин В. Т., Кожеевников В. А., Шупляков В. С. Особенности изображения речевых сигналов моделью спектрального слухового анализа. — Физиол. журн. СССР, 1980, т. 66, № 1, с. 125—131.
4. Загоруйко Н. Г. Методы распознавания и их применение. М.: Сов. радио, 1972.
5. Кожеевников В. А., Столярова Э. И. Особенности реакций на речевые сигналы модели слухового выделения амплитудных неравномерностей. — Физиол. журн. СССР, 1980, т. 66, № 1, с. 132—138.
6. Люблинская В. В., Чистович Л. А. Факторы, определяющие субъективную близость гласноподобных стимулов. — Там же, 1978, 64, 12, с. 1782—1789.
7. Редди. Машинное распознавание речи: Обзор. — Тр. Ин-та инж. по электротехнике и радиоэлектрон., 1976, т. 64, № 4, с. 95—130.
8. Родионов В. Д. Минимально-фазовая модель динамического спектрального анализа звуков на периферии слуха. — Физиол. журн. СССР, 1980, т. 66, № 1, с. 119—124.
9. Столярова Э. И., Чистович Л. А. Амплитудно-частотные характеристики и пороговые устройства модели слуховой обработки огибающей. — Физиология человека, 1977, т. 3, № 1, с. 72—76.
10. Фант Г. Акустическая теория речеобразования. М.: Мир, 1964.
11. Чистович Л. А. О возможной природе акустических событий. — Физиология человека, 1977, т. 3, 6, с. 1103—1112.
12. Чистович Л. А., Венцов А. В., Гранстрем М. П. и др. Физиология речи: Восприятие речи человеком. — В кн.: Руководство по физиологии. Л.: Наука, 1976.
13. Чистович Л. А., Кожеевников В. А., Алякринский В. В. и др. Речь Артикуляция и восприятие. М.; Л.: Наука, 1965.
14. Чистович Л. А., Шейкин Р. Л. Полезные признаки спектра стационарных гласных. — В кн.: Сенсорные системы Л.: Наука, 1979, с. 116—129.
15. Шупляков В. С., Долмазон Ж.-М., Бастэ Л. Исследование функциональной модели периферического спектрального слухового анализа: (Линейный вариант). — Физиол. журн. СССР, 1978, т. 64, № 12, с. 1796—1802.
16. Dolmazon J. M., Shupljakov V. S., Bastet L. Etude d'un modèle fonctionnel du système auditif périphérique. — Rev. acoust., 1977, vol. 10, N 41, p. 169—175.
17. Lindblom B. Motor control mechanisms. — Prepr. of the Speech commun. Sem., Stockholm, 1974, Chairman review of papers in session 2, p. 1—14.

К ПРОБЛЕМЕ ПОНИМАНИЯ ЭВМ ГРАФЕМАТИЧЕСКОГО УРОВНЯ ЯЗЫКА

Система письменности с кибернетической точки зрения может быть рассмотрена в двух планах: как средство коммуникации, осуществляющее смысловое кодирование сообщений, поступающих на вход зрительного анализатора, и как системно-структурное образование, состоящее из конечного набора графических единиц, составляющих графематический уровень языка. Будучи низшей в иерархии языковых уровней, система письменности является в некотором смысле внешней по отношению к языку; отношения между планом выражения и планом содержания здесь наиболее детерминированы. В зависимости от типа графической системы содержательным референтом графемы могут быть любые значимые единицы языка — фонемы или сочетания фонем в слоговых и буквенно-слоговых системах, морфемы или сочетания морфем в идеографических системах письма.

При разработке проблем машинного понимания графематического уровня языка необходимо различать два аспекта — психофизиологический и моделирования. Психофизиологический аспект предполагает исследование механизмов работы мозга и решение круга задач, связанных с классификацией и идентификацией зрительной системой структуры входного сообщения. Существенная особенность этого процесса — наличие механизмов, осуществляющих прием сообщения, его идентификацию, классификацию по наиболее общим признакам и пересылку в соответствующие центры головного мозга. Способность группировать входные сообщения и различать их по элементам чрезвычайно важна: без этого мы не могли бы классифицировать и опознавать эти сообщения, каждое из которых не является точной копией ранее встречающегося.

Вопрос о том, как зрительная система выделяет и измеряет признаки зрительного сигнала, изучен недостаточно глубоко. Но ряд данных, свидетельствует о том, что в сетчатке глаза, которая выполняет роль зрительного анализатора, происходит определение контуров изображения, выделение дискретных элементов, их идентификация и т. д. На следующем этапе входное сообщение в закодированном виде передается в мозг, где вступают в действие другие механизмы и происходит собственно распознавание зрительного сообщения. Результаты исследований подтверждают тот факт, что полная слепота наступает не только от повреждения сетчатки или нервных путей, но и от выпадения функций определенного участка коры головного мозга. Повреждение отдельных участков коры головного мозга приводит к нарушению обработки зрительных сообщений и связано с расстройством зрительного восприятия (явление зрительной диагностики) [1—7].

С другой стороны, зрительное сообщение может быть воспринято при многих искажениях и даже при отсутствии составляющих его элементов. Данные экспериментальных исследований свидетельствуют о том, что при восприятии неполного или искаженного сообщения привлекается контекстуальная информация, которая записана и хранится в соответствующих участках коры головного мозга. В процессе восприятия графического сообщения происходит анализ его структуры и оптико-графических характеристик составляющих его элементов¹.

Другой круг проблем возникает при моделировании процесса зрительного восприятия и состоит в последовательном решении комплекса задач, связанных с интерпретацией и пониманием машиной сообщений естественного языка в графической форме. Процесс письма может быть рассмотрен как коммуникационный процесс, целью которого является смысловое кодирование сообщений естественного языка. Для того чтобы эти сообщения были восприняты машиной, они должны быть декодированы, т. е. представлены в виде набора некоторых структурных зависимостей составляющих его элементов. В общем виде эта задача сводится к поиску методов и созданию устройств, позволяющих автоматически классифицировать и анализировать различные оптические сообщения, разработка которых в конечном итоге связывается с проблемой автоматического чтения текстов на естественном языке [9—13].

При решении практических задач, связанных в большей или меньшей степени с автоматическим вводом текста в машину, возникает ряд трудностей объективного и субъективного характера, зависящих, в частности, от учета степени сложности графического сообщения и его размерности, учета контекста, ранга соподчиненности различных типов шрифтов, от наличия шумов и степени загрязненности графического сообщения и др.

В современной отечественной и зарубежной литературе наиболее распространенные принципы, положенные в основу построения различных экспериментальных систем понимания графических сообщений, — это последовательное сравнение элементов заданного класса с эталоном, кластеризация и структуризация.

Методы анализа, в основу которых положено последовательное сравнение элементов заданного класса с эталоном, предполагают запоминание системой последовательно перечисленных элементов заданного класса. Система относит каждое новое сообщение к новому классу, который совпал с находящимся в ее памяти эталоном.

Методы анализа, использующие принципы кластеризации, предполагают разбиение пространства заданного набора призна-

¹ Применительно к графематическому уровню русского языка нами было проведено экспериментальное исследование, в результате которого были намечены комплекс структурных и количественных характеристик, существенных при восприятии зрительного сообщения [8].

ков на кластеры. Построение системы определяется исследованием геометрических особенностей многомерного кластерного пространства и количественными оценками кластерной структуры.

Принцип структуризации предусматривает построение систем, основанных на выделении структурных признаков, общих для класса в целом, которые вводятся в память системы. Для каждого входного объекта система выделяет набор описывающих его признаков, а затем сопоставляет с хранящимся в ее памяти набором характерных признаков различных классов. Принцип структуризации лежит в основе построения некоторого класса систем автоматического понимания графических сообщений, который известен как лингвистический (синтаксический) подход. В рамках этого подхода может быть составлено семиотическое описание заданного класса графических объектов, использованием некоторого содержательного набора знаков и их комбинаций, для моделирования которых применяется аппарат формальных грамматик. В основе подхода лежит использование аналогии между структурой изображения и синтаксисом языка. Графический объект структурируется с помощью некоторого набора элементарных «слов», представляющих собой вполне определенные графические элементы. Далее графический объект сегментируется и преобразуется в линейную строку. Процедура идентификации машиной входной структурной зависимости осуществляется проверкой соответствия или несоответствия входной структуры грамматике структур, заданных в машине. Подобно фрагменту текста естественного языка, структура графического сообщения рассматривается как совокупность иерархически соподчиненных структурных уровней, включая уровень непроектируемых элементов [9—21].

Система анализа состоит из трех основных блоков: предобработки, представления объекта и синтаксического анализа. Первые два блока выполняют функции предпроцессора. Исходный графический объект сегментируется, кодируется и представляется в виде цепочки символов (структуры объекта). В целях снятия искажений используются методы фильтрации. Блок процессора осуществляет грамматический разбор структуры входного объекта и принимает решение о соответствии его грамматике структур, заложенной в системе. Если описание объекта синтаксически неправильно, он либо исключается из рассмотрения, либо анализируется на основе других заданных грамматических структур. Проблема снятия неоднозначностей решается путем введения статистических критериев. Этот метод эффективен при анализе объектов, богатых структурной информацией: идентификации портретов, аэрофотосъемки, отпечатков пальцев, слитной речи и т. д. Особый класс задач составляют вопросы, связанные с анализом символьной информации. Поскольку для различных типов символьных сообщений, представленных различными графическими системами письменности, не найдено универсальных непроектируемых элементов, подобных, например, фонемам в речевых сообщениях,

анализ каждого класса графических структур представляет чрезвычайно сложную и трудоемкую задачу. Алфавиты различных естественных языков представлены графическими единицами, составленными из разных компонент — линейных, нелинейных, комбинированных и др., которые могут следовать друг за другом, образуя простые линейные последовательности из одной или из двух графем, либо составлять сложные двухуровневые конфигурации из сочетаний двух и более графем и т. д.

С точки зрения разнообразия и богатства структурной информации, заложенной в комбинации графических единиц, представляет интерес графическая система корейской письменности, которая воплощает особенности, присущие разным графическим системам¹.

Рассмотрим в качестве иллюстративного примера структуру алфавита корейской письменности. Алфавитная система включает 24 основных знака — 10 гласных и 14 согласных графем и, кроме того, 9 производных графем, которые порождаются последовательным наращиванием или дополнением линейных компонент и последовательно выводятся один из другого.

Комбинации этих графем могут порождать любые значения слоговые структуры и, кроме того, большое число «пустых» графических структур слога, не имеющих смысла. Возможность образования большого числа таких комбинаций — одна из самых сложных проблем автоматического понимания графических структур, возникающая, в частности, в связи с проблемой ввода. Если посмотреть на этот алфавит с точки зрения его формальной структуры, нетрудно определить соединение двух противоположных принципов — алфавитного и иероглифического. С одной стороны, каждая графема предназначена для обозначения одной и только одной фонемы. С другой — группировка графем в слоги происходит в пределах воображаемого квадрата, как это имеет место в иероглифических системах письменности — пример наложения китайского компонента на неазиатскую основу. Структура алфавита примечательна еще и тем, что графемы образуют конфигурации от одного до четырех элементов, которые образуют слоги. Таким образом, по структуре этот алфавит не буквенный и не слоговой², а буквенно-слоговой. Нами был проведен анализ графических структур слога и установлено, что слоги образуют многоуровневые структуры, которые записываются комбинациями графем в пределах одного, двух и трех ярусов. В основе гласных графем горизонтального ряда V_1 лежит горизонтальный компонент (для обозначения фонем типа \perp , \top и их производных $\perp\perp$, $\top\top$ и др.).

¹ Будучи произведением средневековой культуры, что нашло отражение в первом письменном памятнике «Хунмин Чоном», корейская алфавитная письменность и сейчас привлекает к себе внимание исследователей своей уникальной простотой, четкостью, однозначностью и изяществом структуры.

² Типичным примером слоговых систем письма являются японские азбуки хирагана и катакана.

В основе гласных графем вертикального ряда V_2 лежит вертикальный компонент $|$ (для обозначения фонем типа $|-$, $-|$ и их производных).

Слог записывается в один ярус, когда имеют место следующие структуры:

- V_1C [гласная графема вертикального ряда + согласная графема].
- V_2C [гласная графема горизонтального ряда + согласная графема].

Слог записывается в два яруса, если имеет место структура

- CV_1C [согласная + гласная вертикального ряда + согласная].

Слог записывается в три яруса, если имеет место следующая структура:

- CV_2C [согласная + гласная горизонтального ряда + согласная].

При этом порядок распределения графем по ярусам следующий:

- верхний ярус — согласная, гласная вертикального ряда;
- средний ярус — гласная горизонтального ряда;
- нижний ярус — согласная.

В языке реально функционируют те слоговые структуры, которые отражают деление лексики на генетически разнородные пласты — исконно корейский, китайский и заимствования из других языков.

Следует также обратить внимание на некоторые закономерности употребления слоговых структур. Наиболее употребительные в языке 350 грамматических аффиксов и синтаксем записываются всего лишь 60 различными слоговыми структурами. Слоговые структуры типа $-кетт$, $-атт$, $-бтт$, $-йотт$ и некоторые другие, обозначающие грамматические форманты различных грамматических аффиксов, используются только в позиции временных аффиксов и ни в каких других. Слоговые структуры, которыми записываются слова китайского происхождения, не имеют в конечной позиции исконно корейских графем, обозначающих гортанно-взрывные, придыхательные и другие фонемы типа $-тх$, $-чх$, $-кх$, $-кк$ и др.

Основные закономерности корейской системы письменности могут быть представлены в виде структуры, показанной на рис. 1.

Она включает четыре компонента: верхний левый (C_1), верхний правый (V_1), средний (V_2) и нижний (C_2). На основании анализа выделяются две основные структуры: C_1V_1 и C_1V_2 (для открытых слогов) и три производные структуры: $C_1V_1C_2$, $C_1V_1V_2C_2$ (для закрытых слогов).

Процедура анализа сводится к представлению и преобразованию исходной слоговой структуры в инвариантную ей кодовую, которая используется на дальнейших стадиях ее обработки. Для этого должен быть проведен детальный формальный анализ графических структур слога и составлен словарь исходных производных элементов и характерных точек, составляющих структуру графемы. Для набора корейских графем характерными точками являются следующие:

↑ терминальная точка; > ↗ точки излома; × точки пересечения; + точки разветвления.

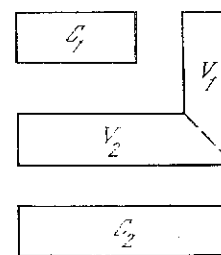


РИС. 1. Фреймовая структура слога



РИС. 2. Элементы графического образа слога

Входной буквенный образ при использовании этих элементов может быть представлен цепочкой символов и преобразован в кодовую структуру — его графический образ (рис. 2).

Графический образ представляет собой совокупность отрезков, связывающих характерные точки, которые аппроксимируются и рассматриваются как частичные образы графического образа. Сформированный графический образ кодируется в форме, подходящей для машинной обработки, и представляется в виде графической матрицы, включающей координаты узлов графического образа. Затем из графического образа формируется дерево буквы. Определяется начальный узел (S), в качестве которого выбирается верхний левый (UL) или верхний правый (UR) узел графического образа. Формирование дерева заканчивается, когда все основные элементы графического образа получают древовидное выражение.

Из набора деревьев, выражающих структуру буквенного образа, может быть сконструирована древовидная грамматика (G_t), порождающая деревья некоторой буквы, принадлежащей к данному классу:

$$G_t(V_N, V_T, P, S),$$

где V_N — конечное множество нетерминальных символов; V_T — конечное множество терминальных символов; P — правила вывода; S — начальный символ.

Непроизводные элементы и соответствующие им терминальные символы приведены в табл. 1.

ТАБЛИЦА № 1

Непроизводные элементы	Терминальные символы	Непроизводные элементы	Терминальные символы	Непроизводные элемент	Терминальные символы
→	0	↓	6	↘	7
↑	2	↖	3	↗	1
←	4	↙	5	.	S

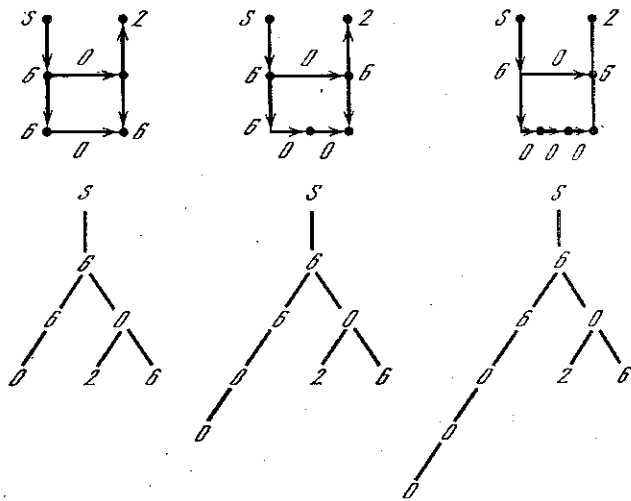
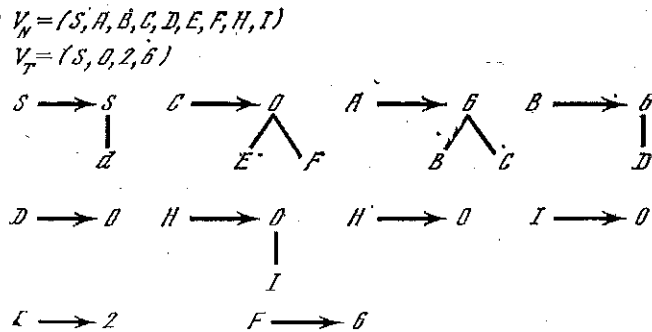


РИС. 3. Древоподобная грамматика буквы «п»

На рис. 3 показан пример древоподобного представления буквы «п». Далее приводится древоподобная грамматика, сконструированная на основе деревьев, показанных на рис. 3.



Процедура собственно «понимания» машиной входной структуры сводится к ее грамматическому разбору и сопоставлению кодового описания структурной зависимости с эталонным описанием, заданным в системе.

Этот метод может быть применен для анализа и машинного представления других графических систем письма, например японской слоговой азбуки катакана и др.

В целях практического использования и автоматизации процессов ввода иероглифических текстов (на японском, корейском и китайском языках) создаются специальные системы, способные воспринимать и кодировать подаваемые на вход графические

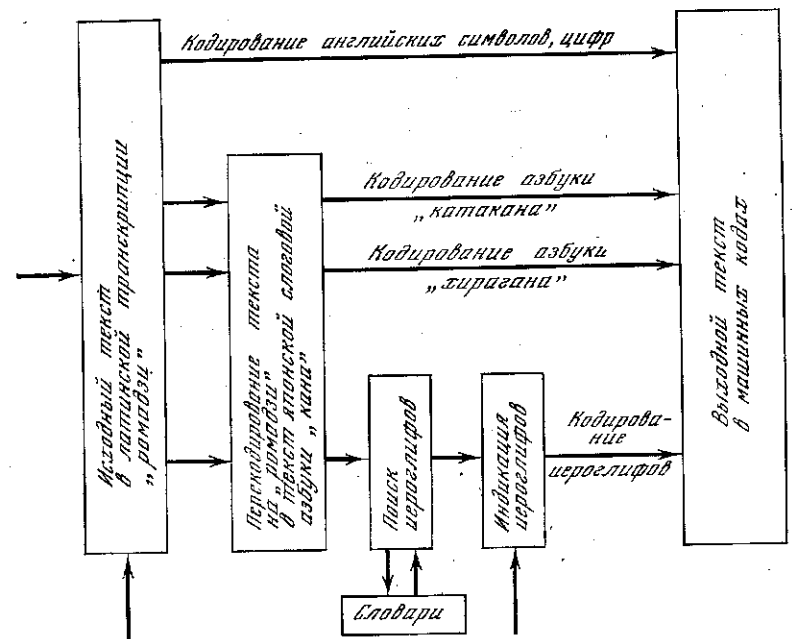


РИС. 4. Блок-схема диалоговой системы ввода в ЭВМ иероглифического текста

структуры. Ввод иероглифических текстов сопряжен с серьезными трудностями, связанными с особенностями графики, обработкой и хранением в памяти машины больших массивов иероглифической информации, поиском эффективных методов ее распознавания, решением проблемы омонимии иероглифических слов и др.

В связи с разработкой интеллектуального интерфейса высокопроизводительных компьютеров, осуществляющих интеллектуальные функции (ввод/вывод графической текстовой информации и др.), особую актуальность приобретает создание систем автоматического и полуавтоматического ввода иероглифической информации. Наиболее простыми для практической реализации являются диалоговые системы ввода в ЭВМ японских смешанных иероглифических текстов с помощью оператора¹. Блок-схема такой системы показана на рис. 4.

¹ Для записи японского текста используются три типа знаков — две слоговые азбуки (катакана и хирагана) по 71 буквенному знаку в каждой и иероглифические знаки. Как правило, корень слова записывается одним или несколькими иероглифами, а окончание — азбукой хирагана, которая также используется для записи служебных слов. Катакана служит для записи слов иностранного происхождения.

ТАБЛИЦА 2. Соответствие знаков «ромадзи» русской транскрипции

Начальный символ	Конечный символ						Начальный символ	Конечный символ					
	A	I	U	E	O	N		A	I	U	E	O	N
	а	и	у	э	о	н		а	и	у	э	о	н
K	ка	ки	ку	кэ	ко		B	ба	би	бу	бэ	бо	
S	са	си	су	сэ	со		P	па	пи	пу	пэ	по	
T	та	ти	ту	тэ	то		KY	кя		кю		кё	
N	на	ни	ну	нэ	но		SY	ся		сю		сё	
H	ха	хи	ху	хэ	хо		TU	тя		тю		тё	
M	ма	ми	му	мэ	мо		NY	ня		ню		нё	
Y	я		ю		ё		MY	мя		мю		мё	
R	ра	ри	ру	рэ	ро		RY	ря		рю		рё	
W	ва						GY	гя		гю		гё	
G	га	ги	гу	гэ	го		ZY	дзя		дзю		дзё	
Z	дза	дзи	дзу	дзэ	дзо		DY	дя		дю		дё	
D	да	ди	ду	дэ	до		BY	бя		бю		бё	
							PY	пя		пю		пё	

Соответствие знаков «ромадзи» русской транскрипции приведено в табл. 2.

Поиск иероглифических знаков и их кодирование осуществляются специальной индикаторной системой, состоящей из базы данных (словари фонетического звучания однозначных иероглифов и иероглифических сочетаний), блока формирования образов иероглифов и блока управления системой отображения образов иероглифов на экране дисплея. После ввода исходного текста и поиска требуемых иероглифов группа иероглифов, имеющих одинаковые фонетические звучания (чтения), последовательно выдается на дисплей (в количестве 9—11 омонимичных слов), и оператор визуально определяет требуемый иероглиф.

Для более эффективного извлечения требуемого иероглифа могут быть использованы дополнительные признаки его идентификации: китайское фонетическое звучание (оное чтение) и японское звучание (кунное чтение), ограничения, накладываемые вхождением иероглифа в данное словосочетание (лексический контекст), а также составление специального файла иероглифов, имеющих обширную омонимию.

Структурная схема поиска иероглифов показана на рис. 5.

Входной текст на «ромадзи» вводится с терминала и преобразуется в коды японской слоговой азбуки «кана». В соответствии с чтением иероглифа с помощью справочного файла производится поиск иероглифов в словаре. Если в результате поиска определена группа омонимов, то изображения их последовательно вы-

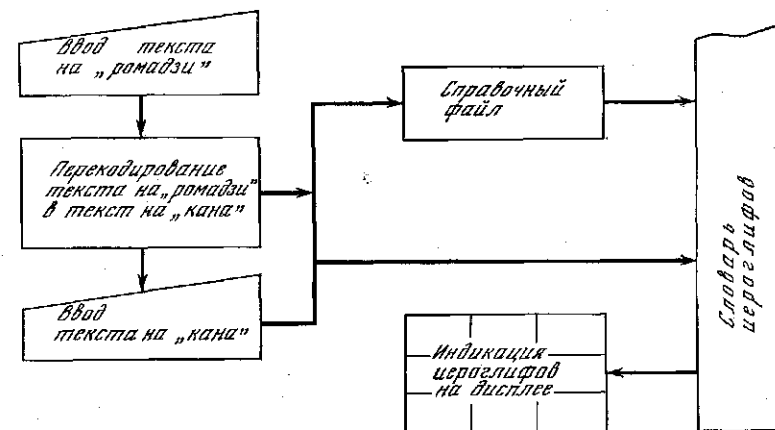


РИС. 5. Структурная схема поиска иероглифов

даются на дисплей для визуальной индикации оператором требуемого иероглифического сочетания. Далее изображения передаются в блок кодирования и формирования выходного текста.

В целях более эффективного поиска иероглифические сочетания, состоящие из двух и более иероглифов, можно классифицировать следующим образом в соответствии с их чтением: 1) иероглифические сочетания, чтения которых начинаются с графем *ya, yi, yo* (группа С); 2) иероглифические сочетания, чтения которых начинаются с графем *и, у, ки, ку, ту, н* (группа В); 3) остальные иероглифические сочетания, чтения которых не вошли в группы В и С (группа А).

На основании этой классификации можно определить распределение иероглифических сочетаний по группам чтения, которые приведены в табл. 3.

Общая конфигурация диалоговой системы ввода иероглифического текста, которая может быть реализована на мини-ЭВМ, приведена на рис. 6.

Блок памяти 1 предназначен для запоминания и генерации образов иероглифов.

В блоке памяти 2 размещена словарная база данных, включающая словари чтений одиночных иероглифов и сочетаний из двух и более иероглифов.

Пульт оператора 1 используется при вводе текста на «ромадзи».

Пульт оператора 2 используется при вводе текста на слоговой азбуке «кана».

На экране дисплея выводятся изображения иероглифов с одинаковым чтением для визуальной индикации оператором. Выходной текст выдается на дисплей для редактирования и распечатывания на АЦПУ. Для более эффективного ввода и поиска

ТАБЛИЦА 3. Распределение иероглифических сочетаний по группам чтений

Распределение групп чтений иероглифических сочетаний по числу знаков наны		Группы чтений иероглифических сочетаний	
		Чтение первого иероглифа	Чтение второго иероглифа
Из двух знаков	Чтение первого иероглифа из одного знака	A	A
	Чтение второго иероглифа из одного знака	B	B
Из трех знаков	Чтение первого иероглифа из одного знака Чтение второго иероглифа из двух знаков	A	AB
		A	AC
		A	BB
		A	BC
		B	AB
		B	AC
	Чтение первого иероглифа из двух знаков Чтение второго иероглифа из одного знака	B	BB
		B	BC
		AB	A
		AC	A
BB	A		
BC	A		
AB	B		
AC	B		
BB	B		
BC	B		
Из четырех знаков	Чтение первого иероглифа из одного знака Чтение второго иероглифа из трех знаков	A	ACB
		A	BCB
		B	ACB
		B	BCB
	Чтение первого иероглифа из двух знаков Чтение второго иероглифа из двух знаков	AB	AB
		AB	AC
		AB	BB
		AB	BC
		AC	AB
		AC	AC
		AC	BB
		AC	BC
BB	AB		
BB	AB		
BB	AC		
BB	BB		

ТАБЛИЦА 3 (окончание)

Распределение групп чтений иероглифических сочетаний по числу знаков наны		Группы чтений иероглифических сочетаний	
		Чтение первого иероглифа	Чтение второго иероглифа
Из четырех знаков		BB	BC
		BC	AB
		BC	AC
		BC	BB
Из четырех знаков	Чтение первого иероглифа из трех знаков Чтение второго иероглифа из одного знака	BC	BC
		ACB	A
Из пяти знаков	Чтение первого иероглифа из двух знаков Чтение второго иероглифа из трех знаков	ACB	A
		BCB	B
	Чтение первого иероглифа из трех знаков Чтение второго иероглифа из двух знаков	ACB	A
		BCB	B
		BCB	B
		BCB	B
Из шести знаков	Чтение первого иероглифа из трех знаков Чтение второго иероглифа из трех знаков	AB	ACB
		AB	BCB
	Чтение первого иероглифа из трех знаков Чтение второго иероглифа из двух знаков	AC	ABC
		CB	BCB
		BB	ACB
		BB	BCB
		BC	ACB
		BC	BCB
		BC	BCB
		BC	BCB
BC	BCB		

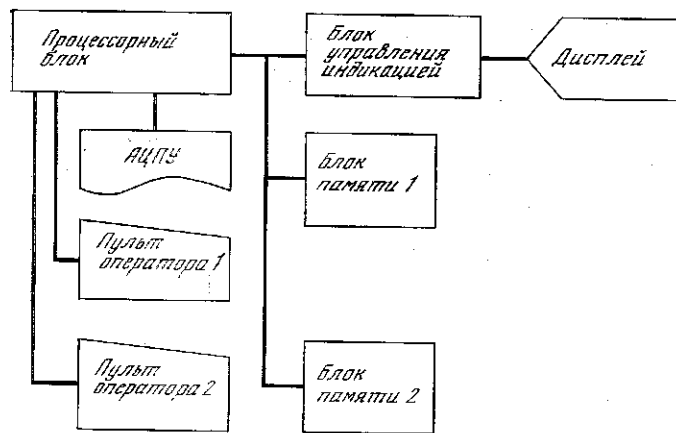


РИС. 6. Общая конфигурация диалоговой системы ввода

иероглифов целесообразно использовать специальные знаки-указатели, которые проставляются оператором при вводе иероглифического текста в ЭВМ и до и после чтения иероглифического сочетания, а также хранятся в машинном словаре иероглифов в памяти ЭВМ.

Поскольку графическая структура иероглифов при всем ее разнообразии имеет определенные закономерности, иероглифы можно классифицировать в соответствии с типом их графической структуры на две или более групп.

Если разбить иероглифы на две группы, то в первую войдут иероглифы, имеющие «слоистую» структуру — сверху вниз из горизонтальных или слева направо из вертикальных слоев. Во вторую — все остальные типы иероглифов. Соответственно иероглифы, вошедшие в первую группу, можно отметить знаком ≠, а во вторую — знаком X.

В результате машинного поиска иероглифических сочетаний по их чтению из машинного словаря устраняются омонимичные слова и на дисплей оператору поступает значительно меньше омонимов, т. е. только те, которые, кроме одинакового чтения, имеют еще и одинаково проставленные знаки — указатели. При таком подходе количество ошибок и скорость идентификации иероглифических слов значительно сокращаются.

Ожидается, что в ближайшем будущем устройства автоматической обработки иероглифической информации в специальных системах будут обладать высокими возможностями. Такие системы будут использовать микро- и мини-ЭВМ и входить в состав автоматизированных комплексов для обработки информации различного назначения. Предлагается увеличение возможностей средств программного обеспечения систем рассматриваемого класса.

Вместе с тем по-прежнему сложной будет оставаться задача

полной автоматизации процессов ввода, которая возможна при создании специальных вводных устройств, способных воспринимать и кодировать входные графические структуры. Эти автоматические устройства, получившие название читающих автоматов (OCR), предназначены для оптического считывания с первичных документов и автоматического ввода алфавитно-цифровой информации в ЭВМ. Каждому входному символу ставится в соответствие код этого символа на внутреннем языке ЭВМ. Алфавитно-цифровая информация представляется в основном в виде машинописных или рукописных знаков.

В ходе работ по созданию читающих устройств выяснили, что форма буквенно-цифровых символов плохо приспособлена для достоверного автоматического считывания. Поэтому для повышения надежности и помехоустойчивости знаки нормализуются и представляются в виде стилизованных шрифтов.

За последние годы опубликован ряд интересных данных о сферах применения читающих автоматов [22, 23]. Эти данные свидетельствуют, что основными в обработке символьной информации можно считать три категории читающих устройств. К первой относятся устройства, воспринимающие цифровые знаки и несколько специальных символов одного шрифта. Такие автоматы применяются для ввода символьной информации в основном экономического характера в производственной сфере. Ко второй — устройства более широкого назначения, предназначенные для считывания алфавитного текста в пределах одного шрифта, например машинописного. Такие устройства находят применение в деловой и производственной сферах. Устройства третьей категории более универсального назначения и предназначены для чтения информации, представленной различными шрифтами. Потребность в создании таких автоматов возникает в связи с решением задач автоматической обработки больших массивов текстовой информации, в частности автоматического перевода с одного языка на другой и пр. Подобные автоматы могли бы найти широкое применение, поскольку они имитируют функцию человеческого чтения. Однако при построении практических систем автоматического понимания символьной информации, помимо трудностей чисто технического характера, возникает ряд серьезных задач, связанных с проблемой чтения и ранжирования отрезков входного текста в соответствии с различными типами типографских шрифтов (на одной странице текста встречается иногда до 20 типографических единиц), с преодолением низкого качества печати, загрязненности текста и пр.

Как показали результаты исследований, проводимые в нашей стране и за рубежом на протяжении двух десятилетий, самая большая трудность состоит в том, что проблема автоматического понимания символьной информации рассматривается как часть более общей задачи понимания естественного языка автоматическими системами. В более общем виде эта задача может быть сформулирована как задача воссоздания зрения роботов и об-

учения их чтению и пониманию текста. Пути к решению этой проблемы лежат в исследовании и моделировании зрительных и речевых механизмов мозга, которые выработались в процессе длительной эволюции человека и его адаптации к внешней среде. Интенсивные исследования в области психологии и нейрофизиологии мозга, направленные на моделирование процессов человеческого восприятия и раскрытие алгоритмов работы мозга, позволяют надеяться, что в недалеком будущем можно будет взглянуть на весь этот комплекс проблем с иных позиций и приблизить результаты машинного восприятия к человеческому.

ЛИТЕРАТУРА

1. Прибрам К. Языки мозга. М.: Прогресс, 1978.
2. Уинстон И. Искусственный интеллект. М.: Мир, 1980.
3. Сорокин Ю. А., Тарасов Е. Ф., Шахнарович А. М. Теоретические и прикладные проблемы речевого общения. М.: Наука, 1979.
4. Mann W. C., Moore J. A. A comprehension model for human dialogue.— In: Proc. 5th Intern. Joint Conf. on Artificial Intelligence. Boston: MIT Press, 1977.
5. Цемель Г. И. Опознавание речевых сигналов. М.: Наука, 1971.
6. Л. Л. Мясников, Е. Н. Мясникова. Автоматическое распознавание звуковых сигналов. М.: Энергия, 1970.
7. Распознавание образов: Исследование образов живых и автоматических систем. М.: Мир, 1976.
8. Гусева Е. К. О количественном подходе к анализу буквенных знаков: Проблемы грамматического моделирования. М.: Наука, 1973.
9. Ярославский Л. П. Введение в цифровую обработку изображений. М.: Сов. радио, 1979.
10. Фролов С. А. Кибернетика и инженерная графика. М.: Машиностроение, 1974.
11. Розенфельд А. Распознавание и обработка изображений. М.: Мир, 1972.
12. Р. Гонзалес, Ту. Дж. Принципы распознавания образов. М.: Мир, 1977.
13. Pappalard T. H. Trends in graphics.— Ergonomics, 1980, vol. 23, N 9, p. 921—933.
14. Хомский Н. Аспекты теории синтаксиса. М.: Изд-во МГУ, 1972.
15. Фу К. Структурные методы в распознавании образов. М.: Мир, 1977.
16. Fu K. S. On mixed approaches to pattern recognition.— In: Proc. Intern. Conf. Cybern. and Soc. Boston; New York, 1980, p. 930—934.
17. Fu K. S. Recent developments in pattern recognition.— IEEE Trans. Comput., 1980, vol. 29, N 10, p. 845—854.
18. Fu K. S. Syntactic pattern recognition and application, Compcon spring 80.— In: 20th IEEE Comput. Soc. Intern. Conf., San Francisco, Calif. N. Y., 1980, p. 391—397.
19. Stallings W. W. Recognition of printed chinese characters by automatic pattern analysis.— Comput. Graph. and Image Process., 1972, vol. 1, p. 47—65.
20. Chang S. K. An interactive system for Chinese character generation and text editing.— In: Proc. IEEE Intern. Conf. Cybernet. and Soc. Wash., 1973, Oct., vol. 9, N 12.
21. Гусева Е. К. Об одном подходе к машинному пониманию графематического уровня языка.— В кн.: Тез. докл. Междунар. семинара по машин. переводу. М.; ГИИТ, ВЦП, 1979.
22. Переверзев-Орлов В. Г. Модели и методы автоматического чтения. М.: Наука, 1976.
23. Киркис З. А. Технические предпосылки развития оптических читающих устройств. Из истории естествознания и техники Прибалтики. Вильнюс, 1980.

II. ОБЩЕНИЕ С ЭВМ И МОДЕЛИРОВАНИЕ ЕСТЕСТВЕННОГО ЯЗЫКА

УДК 519.76

Б. В. Сухотин

МОДЕЛИРОВАНИЕ ГРАММАТИЧЕСКИХ КАТЕГОРИЙ СРЕДСТВАМИ ТЕНЗОРНОЙ АЛГЕБРЫ

В основе идей данной статьи лежит убеждение в эффективности аналитического метода в той его разновидности, которая принята в аналитической геометрии. Главный прием аналитической геометрии — изображение точек пространства n — чисел, координат. Этот прием не только позволил заменить исследование геометрических объектов, представляемых чертежами, исследованием соответствующих алгебраических выражений, но и дал возможность ввести в рассмотрение объекты, не допускающие наглядного представления.

Использование аналитического метода в геометрии встретились с проблемами, внутренне ему присущими, но чуждыми изучаемым геометрическим объектам. Основная из этих проблем легко формулируется лингвистически: это проблема синонимии. Дело в том, что аналитический метод не обеспечивает единообразного изображения геометрического объекта в алгебраическом виде. Поэтому алгебраические выражения могут различаться, несмотря на изображаемое ими общее геометрическое содержание. Для отделения собственно геометрических свойств объектов от свойств, связанных с их алгебраическим выражением, был разработан специальный математический аппарат — тензорное исчисление.

Проблема синонимии — едва ли не центральная проблема лингвистики. Будем понимать ее широко: выражения языка могут выглядеть различно, но иметь сходное содержание. При этом сопоставляемые выражения могут принадлежать, а могут и не принадлежать к одному и тому же языку. В последнем случае лингвисты обычно не говорят о синонимии.

Решение проблемы синонимии применительно к лингвистике означает умение распознать одно и то же содержание под маской различного внешнего оформления. Наш подход к решению этой проблемы основывается на предположении о том, что лингвистическая и соответствующая математическая проблематика имеют глубокое существенное сходство. Однако лингвистическая сторона дела, скорее всего, сложнее математической: выразительные средства аналитической геометрии являются плодом со-

знательного творчества и потому регулярнее, чем выражения естественных языков. Поэтому аппарат, необходимый для решения проблемы синонимии в естественных языках, во всяком случае, не может быть проще аппарата тензорного исчисления. Тем не менее мы надеемся показать, что тензорное исчисление может служить основой соответствующей лингвистической теории.

Проблема синонимии в геометрии, иначе называемая проблемой инвариантов, наиболее глубоко разработана для особых объектов, так называемых тензоров. Мы постараемся показать, что грамматические категории во многих случаях могут быть поняты как тензоры, а их взаимные преобразования (деривация) с успехом изображаются гибким формализмом тензорной алгебры.

С о д е р ж а н и е и м е н и с у щ е с т в и т е л ь н о г о .
К о о р д и н а т ы. Исходным для большинства лингвистических моделей является понятие текста. Текст T есть множество пар $\{v_i, v_j\}$, где $v_i \in V$ (V — алфавит слов, т. е. словарь) и $v_j = 1, 2, \dots, N$ есть порядковый номер слова от начала текст. Пара v_i, v_j есть вхождение слова v_i в текст T . Выражение «слово v встретилось в тексте q раз» означает, что $|\{v, v_j\}| = q$, т. е. количество вхождений слова v в текст T равно q .

Число слов, активно употребляющихся пользователем языка, невелико по сравнению с количеством явлений действительного мира, нуждающихся в наименовании. Поэтому многозначность слов следует считать не досадным недостатком, а необходимым и существенным свойством слов.

Пусть задан текст $T = \{v_i, v_j\}$, $v_i \in V$ и список предметов $\Pi = \{\pi_x\}$, которые так или иначе упоминаются в тексте T . Выберем некоторое существительное S ($S \in V$) и сопоставим его со списком Π . Допустим, имя S встретилось в тексте q раз, причем в значении предмета π_x встретилось s_x раз. Мы имеем

$$\sum_{x=1}^{x=n} s_x = q \quad (q \geq s_x \geq 0).$$

Таким образом, существительному можно сопоставить n -ку чисел s_x ($x = 1, 2, \dots, n$), где $n = |\Pi|$ и где s_x показывает, сколько раз в тексте T существительное S встретилось в значении предмета π_x .

Представим, что текст T описывает обстановку некоторой комнаты, содержащей два стола и три стула. Пусть слово стул встречается в тексте два раза в значении первого стула, три раза называет второй и один раз — третий. Содержание слова *стул* описывается в этом случае следующей пятеркой чисел:

Π	S		
1-й стол	0	1-й стул	2
2-й стол	0	2-й стул	3
		3-й стул	1

Эту пятерку чисел можно понимать как координаты точки пятимерного пространства, соотнесенные с начальной точкой $O = 0,0,0,0,0$.

Допустив в качестве координат не только частоты существительных из текста T , но и любые другие целые неотрицательные числа, мы сможем описать значение любого существительного в произвольных текстах «на ту же тему», т. е. содержащих информацию о тех же предметах из списка Π . При этом содержание существительного будет выражаться пятеркой (в общем случае n -й) целых неотрицательных чисел. Множество таких n -ок назовем пространством R_N (наличным пространством). В пространстве R_N выделим точки, изображающие существительные из данного текста T или данной совокупности текстов $\Theta = \{T_j\}$. Множество таких точек назовем выделенным множеством $W(\Theta)$ или W .

Большее число лингвистических ситуаций можно описать с помощью «пространства возможностей» R_B , координаты точек которого имеют строение

$$\tilde{s}_x = \sum_j p_{xj} s_{xj},$$

где s_{xj} — значение частоты имени; S, p_{xj} — вероятность данного значения в $\{T_j\}$; следовательно, \tilde{s}_x — среднее значение (математическое ожидание) частоты \tilde{s}_x в множестве текстов Θ . В R_B координаты могут иметь любые неотрицательные значения.

Аналитические возможности модели станут значительно богаче, если мы введем в рассмотрение разности точек пространства R_B . Разности точек \tilde{S} и \tilde{S}' из R_B соответствует n -ка разностей координат этих точек:

$$\bar{S} = \tilde{S} - \tilde{S}' = \langle \tilde{s}_i - \tilde{s}'_i \rangle \quad (i = 1, 2, \dots, n).$$

Разность точек $\bar{S} = \tilde{S} - \tilde{S}'$ характеризует заменимость существительного \tilde{S}' на \tilde{S} . Положительность координаты разности свидетельствует о том, что вхождений \tilde{S} «хватает» для замены \tilde{S}' на \tilde{S} в соответствующем значении; отрицательность свидетельствует о нехватке. \tilde{S} и \tilde{S}' необязательно входят в один и тот же текст.

Точки пространства R_B можно рассматривать как разности существительных и «нулевого существительного», которое не встретилось в Θ ни разу. Пространство R_B , дополненное произвольными разностями, назовем «относительным пространством» R_V . Разности точек, составляющие R_V , представляют собой векторы, т. е. направленные отрезки. Координаты векторов R_V можно считать координатами точек, если все векторы исходят из нулевой точки. R_V отличается от R_B точками, содержащими отрицательные координаты.

Пространство R_V выгодно отличается от R_N и R_B своей линеаризуемостью. Это значит, что над векторами R_V можно ввести операции сложения, вычитания и умножения вектора на число. Сложение и вычитание векторов можно определить как соответ-

ствующие по координатным действиям, поскольку математическое ожидание обладает аддитивностью: математическое ожидание суммы случайных величин равно сумме их математических ожиданий даже в случае зависимых случайных величин. Введение отрицательных и нецелых координат обеспечивает замкнутость операций.

Лингвистически сложение векторов выражает равенство последовательных замен $\bar{0}$ на \bar{S}' и затем на \bar{S}'' непосредственной замене $\bar{0}$ на $\bar{S} = \bar{S}' + \bar{S}''$.

Наличие списка предметов Π наряду с текстом (или совокупностью текстов) представляет известное неудобство и как бы отрицает возможность описания содержания выражений языка без обращения к действительности. Это затруднение сравнительно легко устраняется заменой списка Π списком «собственных имен». Собственным именем мы назовем такое существительное ϵ_x , которое встречается только в значении фиксированного предмета, имеющего номер x . Таким образом, содержание собственного имени описывается n -кой чисел, одно из которых отлично от нуля, а остальные равны нулю. Такое представление собственного имени лишь немного отличается от понимания, установившегося в лингвистике. Единичным именем мы будем называть собственное имя, ненулевая координата которого равна 1 (обозначение — ϵ_x).

Операциями сложения и умножения на число каждое имя можно представить в виде линейной комбинации собственных или единичных имен, например:

Π	$B = \{\epsilon_x\}$	\bar{S}
1-й стол	2 0 0 0 0 0	
2-й стол	0 3 0 0 0 0	
1-й стул	0 0 2 0 0 1	
2-й стул	0 0 0 3 0 2	
3-й стул	0 0 0 0 1 1	

Здесь имя \bar{S} представляется с помощью списка $B = \{\epsilon_x\}$ следующим образом:

$$\bar{S} = 0 \cdot \epsilon_1 + 0 \cdot \epsilon_2 + 1/2 \cdot \epsilon_3 + 2/3 \cdot \epsilon_4 + 1 \cdot \epsilon_5.$$

Векторы, линейная комбинация которых дает \bar{S} , необязательно должны быть изображениями собственных имен. Достаточно, чтобы они были линейно независимыми. Тем самым обеспечивается большая свобода в выборе слов, через которые описывается содержание других слов.

Список векторов, через которые выражается любой другой вектор пространства, называется базисом данного пространства. При использовании базиса список предметов Π не нужен. Это позволяет описывать содержание слов, не зная языка, т. е. исключая экстралингвистический подход. В настоящей работе вопрос о выборе какого-нибудь «наилучшего» базиса мы оставляем открытым.

Для дальнейшего изложения нам понадобятся еще следующие понятия:

1) максимальное собственное имя G_i , G_i в качестве i -й (ненулевой) координаты имеет i -ю координату суммы всех имен из W ;

2) максимальное имя $G = \sum_i G_i$;

3) безразличное имя U , которое есть сумма всех базисных имен;

4) частичное имя U_i — собственное имя, имеющее в качестве ненулевой координаты i -ю координату вектора U . Максимальное имя моделирует «глобальное» местоимение, замещающее любое имя из W .

Размерность пространства R_0 будем считать постоянной, но достаточно большой. Во всяком случае, нет нужды считать ее большею, чем число вхождений существительных в тексты, определяющие W .

Формализация понятия «сказуемое». Одноместный предикат. При попытках формализовать лингвистические категории в терминах математической логики представление о сказуемом обычно моделируется с помощью понятия «предикат».

Предикат есть функция, переводящая m -ку имен-предметов (аргумент) в истинностное значение, т. е. «истину» или «ложь», изображаемые числами 1 и 0. В геометрической модели мы также будем сопоставлять m имен-векторов некоторое число (скаляр), являющееся аналогом истинностного значения в математической логике. В простейшем случае одноместного предиката его геометрический аналог — это функция, переводящая вектор в число.

Чтобы выяснить арифметический смысл этой функции, мы будем исходить именно из аналогии математико-логического и геометрического представлений о предикате. Обратим внимание на то, что тривиально истинный предикат имеет значение 1 для любого аргумента; можно добавить: в любых условиях, в любой момент времени. В наших предпосылках это будет значить: для любого вхождения имени в текст.

Напротив, нетривиальность предиката означает, что некоторые имена в некоторых вхождениях в текст отображаются предикатом в истину, а остальные — в ложь. Поэтому в качестве аналога истинностного значения можно использовать частоту употребления данного имени с выражением, обозначающим предикат. Возможно, нет нужды стремиться к слишком полной аналогии, т. е. в качестве истинностного значения принимать число лишь тех вхождений имени и сказуемого, которые образуют истинное высказывание. Достаточно учитывать лишь число (математическое ожидание) совместных появлений и говорить не об истинности, а об инцидентности аргументов и предиката. Такой подход избавляет от метафизических рассуждений о том, какие высказывания и в каком смысле следует считать истинными.

Итак, предикативная функция выбирает часть вхождений

имени и фиксирует число этих вхождений. Диапазон ее значений — от нуля до частоты (математического ожидания) данного имени. Для того чтобы ее значение не зависело от содержания аргумента, ее следует определять относительно одного и того же имени. Удобно использовать максимальное имя G или безразличное имя U .

При этом (в случае одноместного предиката) предикативная функция естественно задается коэффициентами v_i , показывающими, например, какая доля вхождений G в смысле предмета π_i встречается совместно с предикатом V :

$$v_i = \frac{M(V, G_i)}{M(G_i)},$$

где G_i — максимальное собственное имя, описывающее вхождение G в смысле предмета π_i ; символ M означает математическое ожидание.

В случае отсутствия в текстах однородных (в грамматическом смысле) сказуемых число v_i может быть понято как условная вероятность появления предиката V , если появилось слово, обозначающее предмет π_i . При наличии однородных сказуемых коэффициент v_i может быть больше единицы.

Значительное удобство представляет определение коэффициентов v_i относительно безразличного имени U , поскольку в этом случае предикат непосредственно связывается с базисом:

$$v_i = \frac{M(V, U_i)}{M(U_i)}.$$

Принятое определение коэффициента v_i отражает предположение, в силу которого употребительность глагола зависит не от имени, а от содержания имени. Произведение именной координаты s_i на v_i интерпретируется как математическое ожидание числа появлений сказуемого на все появления определяемого имени в смысле данного предмета π_i , а значение предикативной функции определяется как скалярное произведение вектора-имени на последовательность коэффициентов $\langle v_i \rangle$:

$$\varphi(S, V) = \sum_i s_i v_i.$$

Последовательность коэффициентов $\langle v_i \rangle$, изображающая одноместный предикат, также представляет собой вектор. Покоординатное сложение векторов-предикатов частично напоминает дизъюнкцию. Умножение вектора-предиката на число, как и сложение неограниченного числа предикатов, возможно в силу допущения однородных предикатов (т. е. предикатов, относящихся к одному и тому же вхождению имени). Безразличный предикат U в тексте без однородностей представляется вектором, все координаты которого равны 1.

Существует, однако, важное отличие векторов-предикатов от векторов-имен, заставляющее признать пространство векторов-

предикатов необъединяемым с пространством имен. Следующий параграф посвящен рассмотрению этого вопроса.

Инвариантность скалярного произведения, ковариантность предикатов и контравариантность имен в преобразованиях, изменяющих степень грамматической однородности.

Два или несколько предложений с общей частью могут быть объединены в одно, при этом различающиеся части прежних предложений образуют так называемые однородные члены предложения. Не вызывает сомнения, что такое объединение не изменяет содержания текста. Рассмотрим, как реагируют на эти преобразования предложения составляющие их имена, предикаты и скалярные произведения имен и предикатов.

Возьмем два небольших текста, носящих несколько искусственный характер ввиду того, что до сих пор нами разобрано лишь весьма небольшое число грамматических категорий.

Т. 1. Небо сереет. Небо светлеет. Море светлеет.

Т. 2. Небо сереет, светлеет. Море светлеет.

Список предметов, соответствующий как первому, так и второму тексту, содержит лишь два предмета: *небо*, *море*. Однако векторы имен *небо* и *море* для первого и второго текстов неодинаковы. Для первого текста имеем:

$$\text{небо} = 2,0; \quad \text{море} = 0,1;$$

для второго:

$$\text{небо} = 1,0; \quad \text{море} = 1,0.$$

Неодинаковы и векторы-предикаты:

Первый текст	Второй текст	
	светлеет	сереет
1/2	1	1
1	0	0

Напротив, скалярные произведения остаются неизменными:

первый текст:

второй текст:

	светлеет	сереет
светлеет	1	1
сереет	1	0

	светлеет	сереет
светлеет	1	1
сереет	1	0

Очевидно, что скалярные произведения отражают более существенные свойства лингвистических единиц, чем те, которые отражаются частотой.

Векторы-имена и векторы-предикаты второго текста получают умножением векторов из первого текста на взаимно обратные матрицы P и P^{-1} :

$$P = \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} \quad P^{-1} = \begin{vmatrix} 1/2 & 0 \\ 0 & 1 \end{vmatrix}.$$

Действительно, имена преобразуются так:

$$2,0 \times \begin{vmatrix} 1/2 & 0 \\ 0 & 1 \end{vmatrix} = 1,0,$$

$$0,1 \times \begin{vmatrix} 1/2 & 0 \\ 0 & 1 \end{vmatrix} = 0,1.$$

Предикаты преобразуются следующим образом:

$$\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} \times \begin{vmatrix} 1/2 & 1 \\ 2 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} \times \begin{vmatrix} 1/2 & 1 \\ 0 & 0 \end{vmatrix} = 0.$$

Как первый, так и второй тексты предполагались связанными с одним и тем же списком предметов или с одним и тем же базисом, состоящим из единичных собственных имен. Преобразование координат векторов мы получили, заменив первый текст вторым. Однако такое же изменение координат можно получить, не изменяя текста *T.1.*, а всего лишь изменив базис.

Пусть новый базис равен

$$B' = \begin{cases} \varepsilon_1' = 2, 0 \\ \varepsilon_2' = 0, 1. \end{cases}$$

В этом базисе координаты имен будут такими:

$$\text{небо} = 1,0; \quad \text{море} = 0,1.$$

Для того, чтобы скалярные произведения остались неизменными, координаты предикатов должны быть равны следующим числам:

$$\text{светлеет} = 1,1; \quad \text{серееет} = 1,0.$$

Таким образом, при преобразованиях базиса предикаты изменяются так же, как и векторы базиса, т. е. в данном случае умножением на матрицу P ; напротив, векторы-имена получаются из прежних умножением на P^{-1} . По этой причине векторы-предикаты называются ковариантными, а векторы-имена — контравариантными. Приведенные примеры показывают: 1) что преобразование текстов в некотором смысле эквивалентно преобразованию базиса; 2) что основное отличие предикатов от имен заключается в противоположном характере их изменения относительно преобразований, сохраняющих скалярное произведение.

Л и н е й н ы е п р е о б р а з о в а н и я. Векторы, составляющие пространство, могут подвергаться преобразованиям. Наиболее простыми среди них являются так называемые линейные преобразования. Мы постараемся показать, что ряд линейных преобразований имеет ясный лингвистический смысл и что большинство (может быть, все) грамматических категорий представляется тензорами — геометрическими объектами, координаты которых изменяются при линейных преобразованиях по особым законам.

Линейным называется такое преобразование Φ , для которого

справедливо следующее равенство:

$$\Phi(x + y) = \Phi(x) + \Phi(y).$$

Имеются два источника преобразований векторов: преобразование векторов при неизменном базисе и преобразования базиса при неизменном пространстве. Первый источник соответствует в конечном счете реальным преобразованиям текстов, второй — лишь затрагивает способы их описания. Закон преобразования векторов задается в обоих случаях одинаково.

Рассмотрим некоторые примеры линейных преобразований с краткой характеристикой их лингвистического смысла.

1. Тожественное преобразование переводит вектор $\langle x_i \rangle$ в вектор $\langle x_i \rangle$ с теми же координатами. Это преобразование задается умножением вектора $\langle x_i \rangle$ на единичную матрицу E .

2. Параллельный перенос преобразует вектор $\langle x_i \rangle$ в вектор $\langle x_i \rangle = \langle x_i + h_i \rangle$.

3. Преобразование подобия переводит вектор $\langle x_i \rangle$ в вектор $\langle x_i \rangle = \langle \lambda x_i \rangle$ и задается матрицей λE .

4. Неравномерное растяжение (сжатие) сопоставляет вектору $\langle x_i \rangle$ вектор $\langle x_i \rangle$, где $x_i = \lambda_i x_i$ и задается умножением на диагональную матрицу с элементами λ_i на главной диагонали.

5. Проекция имеет место в том случае, когда некоторые λ_i равны 0.

6. Поворот определяется умножением вектора на матрицу $\| a_{ij} \|$, такую, что

$$\sum_j a_{ij}^2 = 1; \quad \sum_j a_{ij} a_{kj} = 0 \quad (i \neq k).$$

Преобразования 1, 2, 3, 4, 6 носят название аффинных. При этих преобразованиях в общем случае не сохраняется длина векторов и углы между ними, но сохраняются скалярные произведения для тех векторов, для которых эти произведения определены, параллельность прямых и последовательность расположения точек на прямой. Аффинные преобразования задаются уравнениями следующего вида:

$$x^i = p_{\alpha}^i x^{\alpha} + h_i,$$

где определитель матрицы $\| p_{\alpha}^i \|$ не равен нулю. В данном выражении, как и в дальнейшем, считается, что по повторяющемуся индексу производится суммирование.

Преобразования 1, 2, 6 носят название эвклидовых: они сохраняют длины векторов, углы, площади фигур. Эвклидовы преобразования характеризуются тем, что они оставляют неизменным так называемый метрический тензор, т. е. систему положительных чисел g_{ij} , с помощью которых определяется скалярное произведение для любой пары векторов по формуле

$$xy = g_{ij} x^i y^j.$$

Преобразование 2 легко интерпретируется лингвистически. Ему соответствует присоединение текста T , на основе которого вычисляются координаты, к другому тексту T' или исключение текста T' из T .

Преобразования 3, 4 и 5 соответствуют ранее рассмотренным преобразованиям, изменяющим степень грамматической однородности, а также реферированию (комментированию) исходного текста. При этом преобразование 5 соответствует реферированию с потерей информации.

Важный случай поворотов имеет место, когда матрица $\|a_{ij}\|$ содержит в каждой строке и в каждом столбце по одной единице, тем самым задавая подстановку. Такие преобразования соответствуют буквальному переводу с одного языка на другой. В этом случае список предметов Π' для языка L' отличается от списка Π для языка L порядком предметов. Матрица $\|a_{ij}\|$ задает переход к описанию имен относительно списка Π' . Если Π и Π' совпадают, то $\|a_{ij}\| = E$. Вообще повороты связаны с переименованиями предметов.

Весьма важен вопрос, какие преобразования считать более приспособленными для целей лингвистического моделирования: аффинные или эвклидовы? Ранее мы показали, что при преобразованиях, изменяющих степень грамматической однородности, поразному изменяются векторы-имена и векторы-предикаты. В то же время, как известно, при эвклидовых преобразованиях все векторы изменяются одинаково, в силу чего пропадает разница между ковариантными и контравариантными векторами, а значит, между именами и предикатами.

Далее, в аффинном пространстве по той же причине скалярное произведение определено лишь для разнородных, дуальных векторов, в эвклидовом — для любых векторов. Это обстоятельство благоприятствует аффинной модели, поскольку высказывание можно считать всегда содержащим аргумент и предикат, а предложение — сказуемое и подлежащее. Отсутствие в предложении подлежащего или сказуемого лучше считать фактом наличия «невидимого», нулевого элемента.

Недостатком аффинного подхода является то, что в аффинном пространстве вектору не сопоставляется никакого инварианта (например, длины), в силу чего оказывается, что существительное или глагол не имеют никакого постоянного содержания. Постоянное содержание могут иметь лишь высказывание и некоторые другие конструкции.

Если заняться поисками инвариантов векторов, то невольно вспоминаются такие характеристики, как частота $F = \sum_i s_i$; дисперсия

$$D = \frac{1}{n} \sum_i (s_i - s^0)^2 = M(s - Ms)^2$$

(здесь n — число координат, M — математическое ожидание,

s^0 — среднее значение координаты); энтропия

$$H = - \sum_i \frac{s_i}{F} \log \frac{s_i}{F}.$$

Однако эти интересные характеристики инвариантны относительно линейных преобразований. Нелинейная же модель, не имея преимуществ простоты, вряд ли имела бы более адекватный характер.

В самом деле, можно заметить, что аффинная и эвклидова модели правильны каждая по-своему. Для этого надо вспомнить, что, помимо грамматически оформленных слов, существуют и неоформленные элементы языка: основы или корни слов. Очевидно, что для изучения функционирования слов в предложении важно различие имен и глаголов, а следовательно, здесь предпочтительнее аффинный подход. Напротив, для выявления лексического значения естественнее ограничиваться основами или корнями; здесь правомернее эвклидова модель.

Для уяснения эвклидовой модели следует остановиться на лингвистическом смысле эвклидовых инвариантов: длины вектора и угла между векторами.

Обратим внимание на то, что обычное понятие длины вектора лишь немного отличается от дисперсии. Как известно, дисперсию можно представить как разность среднего квадрата и квадрата средней данной случайной величины:

$$D = \frac{1}{n} \sum_i s_i^2 - s^0^2.$$

Таким образом, квадрат длины вектора равен (с точностью до постоянного множителя) дисперсии плюс квадрат средней. Отличаясь от дисперсии на s^0^2 , длина вектора имеет преимущество в инвариантности относительно линейных преобразований. Со статистической точки зрения длина вектора есть начальный момент и, подобно дисперсии, может служить мерой «рассеянности», а значит, и содержательности координат.

Еще интереснее статистическая (а следовательно, и лингвистическая) интерпретация угла между векторами. В силу одного из определений коэффициент корреляции между двумя системами чисел $\langle x_i \rangle$ и $\langle y_i \rangle$ равен

$$K = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}}.$$

Этой же формулой определяется косинус угла между соответствующими векторами. Коэффициент K , измеряющий степень линейной зависимости между векторами, равен 1 в том случае, когда сравниваемые векторы равны или пропорциональны. Лингвистически его нельзя интерпретировать иначе, как степень смыс-

ловой близости сравниваемых слов, конечно, если согласиться с предложенной моделью содержания слов.

Таким образом, эвклидовы преобразования не изменяют степени смысловой близости слов.

Дальнейшее изложение будет базироваться на аффинной модели, поскольку основная цель статьи есть изложение методов геометрического представления грамматических категорий.

Геометрическая модель атрибута. Формальное истолкование атрибута (определения) в рамках математической логики наталкивается на серьезные трудности. Напротив, геометрическая модель этого понятия вполне очевидна. Для построения этой модели достаточно указать на два определяющих факта. Во-первых, атрибут переводит одно имя в другое, например: *стул* → *большой стул*.

Во-вторых, атрибут сужает значение исходного имени. Это значит, что лишь в некоторых вхождениях слова *стул* реализуется значение выражения *большой стул*. Иначе говоря, атрибут *большой* выбирает некоторые вхождения имени *стул*.

Преобразование имени, переводящее его в другое имя, как известно, задается матрицей, определяющей преобразование соответствующих векторов. Сужающий характер атрибута говорит о том, что это матрица стяжения. Если вектор соотнесен с базисом из собственных имен, то такая матрица имеет диагональный вид. Диагональные элементы матрицы вычисляются по формуле

$$a_{ii} = \frac{M(G_i, A)}{M(G_i)},$$

где $M(G_i, A)$ — математическое ожидание числа совместных появлений максимального собственного имени и данного атрибута. Некоторые диагональные элементы матрицы-атрибута могут быть равными нулю. Поэтому правильнее говорить, что атрибут определяет проекцию. Рассмотрим пример, введенный на с. 91.

Если слово *большой* всегда сопровождает слово *стол* в значении *второй стол* и в двух случаях из трех сопровождает слово *стул* в значении *второй стул*, то матрица атрибута *большой* будет иметь следующий вид:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2/3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Имя *большой стул* будет иметь содержание, определяемое умножением вектора $0, 0, 2, 3, 1$ на матрицу A :

$$0, 0, 2, 3, 1 \times A = 0, 0, 0, 2, 0.$$

Неоднородные определения изображаются последовательным умножением вектора имени на соответствующие матрицы. Инте-

ресен случай последовательного применения одного и того же атрибута, являющийся частным случаем неоднородной (в грамматическом смысле) цепочки определений. Он моделируется умножением вектора-имени на степень матрицы-атрибута и, естественно, оказывает более сильное сужающее действие, чем однократный атрибут: *большой-большой стул* →

$$0, 0, 2, 3, 1 \times A^2 = 0, 0, 0, 4/3, 0.$$

Подобная ситуация в рамках математико-логического представления совершенно необъяснима.

Каким бы сложным выражением не являлся реальный атрибут, его содержание в силу приведенных выше аргументов должно изображаться матрицей. Этот факт имеет большое эвристическое значение.

Тензоры. Формальное описание наречия. Тензором называется система чисел (компонент), пронумерованных верхними и нижними индексами. Компоненты тензора при линейном преобразовании, определяемом матрицей p_{α}^i , должны изменяться следующим образом:

$$*g_{\alpha_1 \dots \alpha_n} \beta_m = g_{i_1 \dots i_n} p_{\alpha_1}^{i_1} p_{\alpha_2}^{i_2} \dots p_{\alpha_n}^{i_n} q_{j_1}^{\beta_1} q_{j_2}^{\beta_2} \dots q_{j_m}^{\beta_m}.$$

В этом выражении по повторяющимся индексам производится суммирование; матрица q_j^{β} — обратная относительно p_{α}^i ; звездочка отмечает новое значение тензора g .

Общее число индексов называется валентностью, или рангом тензора. Тензоры, имеющие только верхние или только нижние индексы называются соответственно контравариантными и ковариантными; остальные тензоры суть смешанные.

Тензоры с одинаковым числом верхних и нижних индексов можно складывать покомпонентно. Два тензора любого строения можно перемножать, при этом каждая компонента первого тензора умножается на каждую компоненту второго, в результате чего получается тензор, валентность которого равна сумме валентностей сомножителей.

Важнейшей операцией является свертка. Смешанный тензор можно свернуть по любой паре разнородных индексов. Индексы, по которым происходит свертка, именуется общей переменной, затем строится новый тензор сложением всех компонент, отличающихся лишь значением этой переменной, например:

$$a_{jk}^i \rightarrow a_{ik}^i = \left(\sum_i a_i^i \right)_k = b_k.$$

Сверткой можно назвать также умножение двух тензоров с последующей сверткой, например:

$$a_{jk}^i b_t^s \rightarrow a_{ik}^i b_t^t = c_{kt}^i$$

есть свертка тензоров a_{jk}^i и b_t^s по индексам j, s .

В этой терминологии вектор-имя есть контравариантный тензор валентности 1, а одноместный предикат — ковариантный тензор валентности 1. Скалярное произведение есть свертка имени с предикатом; атрибут — смешанный тензор валентности 2; определение имени атрибутом есть свертка имени с атрибутом.

Тензорный характер имен и предикатов обнаруживается в силу способа их изменения в зависимости от изменения базиса. При изменении базиса по закону

$$e_{\alpha} = p_{\alpha}^i e_i$$

предикатные векторы изменяются ковариантно:

$$v_{\alpha} = p_{\alpha}^i v_i$$

а именные — контравариантно по формуле

$$s^{\alpha} = q_j^{\alpha} s^j$$

где произведение матриц

$$p_{\alpha}^i q_j^{\alpha} = \delta_j^i$$

есть единичная матрица.

Тензорный характер атрибута легко уяснить, если учесть, что формулировку атрибута в новых координатах можно получить, перейдя к изображению определяемого вектора в старых координатах, осуществить преобразование-атрибут A и затем представить результат в новых координатах:

$$S = \tilde{S} \times P^{-1} \times A \times P.$$

Здесь P — матрица преобразования базиса. Тем самым коэффициенты атрибута изменяются по формулам, характеризующим изменение координат смешанного тензора валентности 2:

$$a_{\alpha}^{\beta} = p_{\alpha}^i q_j^{\beta} a_i^j$$

Интуитивно очевидна близость атрибутов и одноместных предикатов. Смысловое сходство этих категорий подчеркивается экономностью средств языка, преобразующих атрибут в предикат. В русском языке для этого достаточно изменить порядок слов: *большой дом* → *дом — большой*. С тензорной точки зрения одноместный предикат получается из атрибута умножением последнего на безразличный предикат y_i с последующей сверткой:

$$v_j = a_j^i y_i.$$

Безразличный предикат моделирует одноместный глагол-связку. Получаемый предикат подсчитывает именно те вхождения имени, которые выбираются атрибутом $A = a_j^i$.

В разнообразных синтаксических ролях может выступать наречие. Наиболее естественная его роль — приглагольное определение, например: *он быстро работает*. Однако наречие может

определять и прилагательное, причастие или другое наречие: *быстро летящие гобы, очень красивый дом, очень быстро протекающий процесс*. Возможно также определение при существительном, выраженное наречием: *сапоги всмятку, полностью банкрот*. Однако сочетаемость наречий с существительными все же ограничена. Например, из правильности фразы *очень высокий дом* не следует правильность фразы *очень дом*.

Мы исходим из аналогии роли наречия в качестве приглагольного определения и прилагательного — приименного определения. Разница в том, что приименное определение модифицирует контравариантный вектор, а приглагольное — ковариантный. Таким образом, $d^i v_j = v_i$, иначе говоря, наречие d^i представляет собой смешанный тензор валентности 2, т. е. матрицу. Рассматривая матрицу наречия саму по себе, в ней нельзя заметить отличия от матрицы прилагательного (атрибута). Однако функционирование этих матриц различно. При умножении имени на прилагательное происходит свертка по нижнему индексу прилагательного и верхнему индексу имени, тогда как при умножении глагола и наречия свертка осуществляется по верхнему индексу наречия и нижнему индексу глагола. Если наречие и прилагательное связаны отношением производности, то матрица наречия является транспонированной по отношению к матрице прилагательного: $d_i^j = a_j^i$.

Матричный характер наречия позволяет понять возможность применения его к прилагательному или другому наречию:

$$A \times D = A'; \quad D' \times D = D''$$

Здесь A, A' — прилагательные, D, D', D'' — наречия. Возможность применения наречия к существительному объясняется тем, что наречие может быть «понято как прилагательное». Одним из факторов, мешающих такому отождествлению, является наличие у прилагательных форм словоизменения.

Многоместные предикаты. Некоторые деривационные соотношения. Номинализация. Подобно тому, как однесместный предикат выражается линейной функцией, выбирающей некоторые вхождения имени, двухместный может быть изображен билинейной функцией, выбирающей некоторые вхождения пар имен и подсчитывающей число этих вхождений.

Таким образом, билинейная функция переводит пару контравариантных векторов в скаляр и, следовательно, представляет собой ковариантный тензор валентности 2, обозначаемый v_{ij} . Значение соответствующего высказывания определяется сверткой по индексам i и j : $\varphi(V, S_1, S_2) = v_{ij} s^i s^j$.

В этом выражении координаты одного из имен следует заменить нормированными, т. е. поделенными на сумму вхождений имени, например, заменить s^j на

$$\dot{s}^j = s^j / \sum_j s^j.$$

Тогда произведение $s^i s^j$ (индексы фиксированы) будет представлять собой математическое ожидание появления имени S_1 в значении предмета π_i , а S_2 — в значении предмета π_j ; коэффициент v_{ij} может быть интерпретирован как условная вероятность того, что данная пара предметов обладает данным свойством (в том случае, если однородные сказуемые отсутствуют).

Итак, коэффициент v_{ij} имеет следующую структуру:

$$v_{ij} = \frac{M(V, G_i, G_j)}{M(G_i, G_j)}.$$

Здесь G_i, G_j — максимальные собственные имена.

Коэффициенты безразличного двухместного предиката для текстов без однородностей должны быть все равны единице, коэффициенты тождественно ложного, «пустого» предиката — нулю. Предикат отождествления изображается билинейной функцией, все элементы которой вида v_{ij} равны единице, а остальные — нулю. Предикат, симметричный в логическом смысле, моделируется симметричной билинейной функцией.

Преобразование двухместного предиката в одноместный моделируется сверткой двухместного предиката с зависимым именем: $v_i = v_{ij} s^j$. Так получают соотношения типа *бомбить* = *бросать бомбы*. Преобразование имени в глагол может быть представлено как свертка двухместного предиката с именной частью сказуемого: *он (есть) учитель* = *он учительствует* $\rightarrow s^i v_{ij} s^j = s^i v_i$, где $v_i = v_{ij} s^j$.

Преобразование одноместного глагола в имя сложнее в том смысле, что требует введения особой категории — номинализатора, который есть функция, не соответствующая какой-либо совокупности слов. Номинализатор представляет собой тензор валентности 2, контравариантный:

$$(некто) работает \rightarrow рабочий \rightarrow n^{ij} v_j = s^i.$$

Подобно тому, как двухместный предикат формализуется в виде скалярной билинейной функцией, многоместный — задается полилинейной функцией, например:

$$\varphi(s^i, r^j, z^k) = v_{ijk} s^i r^j z^k.$$

Здесь символами r, z , в отличие от s , выделены нормированные векторы, соответствующие косвенным падежам. Представление косвенных падежей именно нормированными векторами находит интуитивное оправдание в том, что нормированное имя с одноместным предикатом не дает законченного высказывания, значением которого должна быть абсолютная частота имени с предикатом.

Свертка предиката с одним или несколькими косвенными падежами производит предикат меньшей валентности:

$$\text{он рубит дрова} \text{ топором} \rightarrow \text{он рубит дрова} \rightarrow \text{он рубит};$$

$$s^i v_{ijk} r^j z^k \rightarrow s^i v_{ij} r^j \rightarrow s^i v_i.$$

При этом свертка с потерей информации соответствует замене дополнения глобальным местоимением (максимальным именем).

Представление о валентности предиката осложняется тем, что некоторые «условия» (например, места, времени) характеризуют почти любой предикат. Мало того, обычно в естественных языках существует набор категорий сказуемого (время, наклонение), которые непременно должны квалифицироваться.

Дело обстоит так, что у каждого предикатного тензора как бы существует собственный набор индексов (i_1, i_2, \dots, i_r) и дополнительный (j_1, j_2, \dots, j_r) :

$$V = v_{i_1 i_2 \dots i_r j_1 j_2 \dots j_r},$$

причем дополнительные наборы индексов различных предикатов в некотором смысле совпадают. Это совпадение отражает независимость дополнительного набора от собственного; иначе говоря, каждый предикат можно считать мультипликативным и представимым в виде произведения двух тензоров: собственно предиката $v_{i_1 \dots i_r}$ и тензора дополнительных категорий $h_{j_1 \dots j_r}$:

$$V = v_{i_1 \dots i_r} h_{j_1 \dots j_r}.$$

Тензор $h_{j_1 \dots j_r}$ можно назвать тензором абстрактной предикации. Его, по-видимому, в свою очередь, полезно разложить на тензор обязательных условий и тензор возможных условий: $h_{j_1 \dots j_r} = v_{k_1 \dots k_m} f_{l_1 \dots l_n}$. Различие тензоров v и k в том, что первый из них не допускает свертки с местоимениями и содержит обязательные, т. е. грамматические, категории.

Наконец, следует допустить, что существует особая валентность факта (события). Смысл ее в том, что каждое повествовательное предложение соответствует некоторому событию, что можно пояснить такой искусственной фразой:

Ветер дует, (тем самым) производя событие.

Обычно эту валентность следует считать свернутой с местоимением, но она приобретает важное значение при номинализации.

Номинализация в многоместном предикате осуществляется (как и в одноместном) умножением на тензор n^{ij} с последующей сверткой по одному из его индексов. Ввиду многоместности предикатов при этом получают имена различных типов, например *победитель, побежденный, победа*. Слова типа *победа* получают в результате номинализации по валентности факта. Такая номинализация важна для понимания рекурсивного построения фраз, когда предложение занимает место аргумента в предикате более высокого порядка, например:

я знаю, город будет и т. д.

В настоящей статье освещались главным образом основания геометрической модели языка. Больше число примеров можно

найти в работе [2], в том числе такие важные, как изложение тензорного представления о падежах. В той же работе отмечена связь данной модели с K -грамматиками.

ЛИТЕРАТУРА

1. Гуревич Г. Б. Основы алгебраической теории инвариантов. М.; Л.: ОГИЗ, 1948.
2. Сухотин Б. Ю. Основные проблемы грамматики и семантики в тензорном изложении. — В кн.: Проблемы структурной лингвистики. М.: Наука, 1979.

УДК 802.03:681.3+002.5:681.

Ю. С. Мартемьянов

О СВЯЗНОМ ЗАДАНИИ СИНТАКСИСА

В системе ФЛОС (формализованное лингвистическое обеспечение синтеза) осуществляются следующие представления о синтаксисе естественного языка.

Синтаксис упорядочен: в нем есть только одна такая часть — набор иерархизованных синтаксических отношений, — которая составляет необходимое ядро для всех остальных, от них не зависит и может быть задана достаточно простой контекстно-свободной грамматикой (§ 1); остальные части усложняют или расширяют это ядро в том или ином аспекте; ядро с введенными усложнениями (§ 2) соответствует базовому синтаксису; заданием определенных преобразований создаются расширения базы (§ 3) вплоть до полного синтаксиса;

синтаксис связан: максимальное расширение синтаксиса может выводиться из минимальной базы и сводиться к ней путем непрерывных приближений; для достижения такой связности все элементы базы, способные быть предметом усложняющих и расширяющих операций, должны быть введены в явном и расчлененном виде (§ 1);

синтаксис единствен: более сильная, «семантическая», а вернее, «метасинтаксическая» правильность предложения (ср. [2]) может быть достигнута без смены ядерных синтаксических отношений, одним только заданием иерархии метасинтаксических категорий, распределяемых по отношениям синтаксической структуры при ее усложнении и сохраняемых при ее расширениях (§ 2);

синтаксис самоуправляем: каждый очередной шаг вывода, приближающий базовую или производную структуру к ее воплощению в текст — фразу или сверхфразовое единство, — должен всецело определяться формой и составом этой структуры, в частности присутствием в ней тех или иных факторов усложнения и/или расширения; особый мотивационный компонент синтаксиса учитывает

все специфические характеристики структуры, к которым данный естественный язык «чувствителен» в том смысле, что требует их преобразования с привлечением определенных операций (см. § 3—4);

синтаксис самодостаточен: при выводе реляционных структур базового и расширенного синтаксиса должны сами собой получиться все такие узлы категорий, которые в совокупности составят полный набор известных синтаксических частей речи, а также более дробных синтаксических единиц, составляющих основу для грамматической классификации слов; для вывода таких частей речи деревья иерархизованных отношений должны переводиться из расчлененной формы в стянутую, т. е. в дерево подчинительных связей (§ 1).

1. Ядерный синтаксис. Ядро синтаксиса содержит набор категориальных пар (групп), связанных некоторым бинарным отношением (1), из которых по определенным правилам (2) строятся иерархические структуры — реляционные деревья (3),

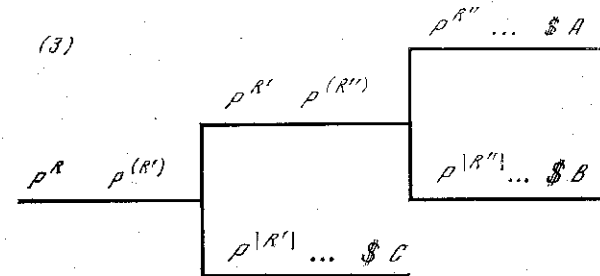


РИС. 1. Дерево отношений (реляционная структура)

допускающие синтаксическую конкретизацию категорий (4) и располагающие особой, «стянутой» формой (5).

(1) Обозначение категориальной группы с отношением имеет вид $P^{(R)}$, где при категориальной характеристике свернутой группы P имя отношения указывается в круглых скобках.

(2) Правил образования структур — два: заполнение (II) и развертывание (I):

$$I. P^{(R)} \rightarrow P^R + P^{[R]},$$

$$II. P^R \text{ (или } [R]) \rightarrow P^{(R)}.$$

По первому правилу символ свернутой группы развертывается на два элемента соответствующей категориальной характеристики с сохранением при них имени отношения в прямом смысле R у левого элемента и в конверсном $[R]$ у правого; очевидная избыточность обозначений допущена из содержательных лингвистических соображений.

Согласно второму правилу элементы развернутой группы могут

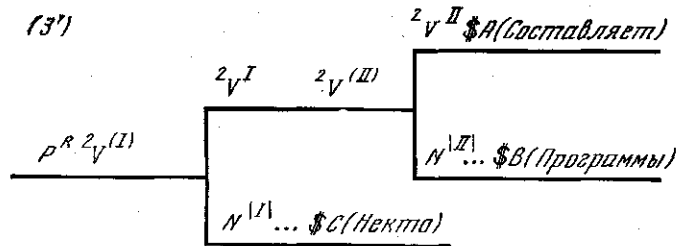


РИС. 2. Дерево вербальной группы

заполняться либо новой свернутой группой с отношением R' , либо символом словарного входа $\$$.

(3) Общий вид структуры — дерево отношений, которое для удобства строится слева направо, так что заполняющая группа располагается правее заполняемого элемента, а правый элемент развернутой группы становится ниже левого (рис. 4).

Дерево отношений, как и структура непосредственных составляющих, позволяет выразить взаимное расположение элементов отношения (A, B) , $C \neq C$, (A, B) , а также иерархию («ранг») самих отношений по порядку заполнения, или по «тесноте»: (A, B) , $C \neq A$, (B, C) . Сохранение в реляционной структуре объединяющего символа группы $P^{(R)}$ позволяет вносить в имена синтаксических отношений дополнительные (метакатегориальные) различия (§ 2) и упрощает задание и осуществление операций, изменяющих реляционную структуру (§ 3).

(4) Конкретная категоризация формулы: группа $P^{(R)}$ может получить категорию высказывательной (вербальной) группы ${}^nV^{(M)}$ или номинальной (юнктивной) группы $W^{(J)}$.

Для вербальной группы правила (2) принимают конкретный вид (2') с результирующим деревом (3') (рис. 2).

$$(2'). \text{ I. } {}^nV^{(M)}, 1 \leq M \leq n \rightarrow {}^nV^M + N^{|M|}$$

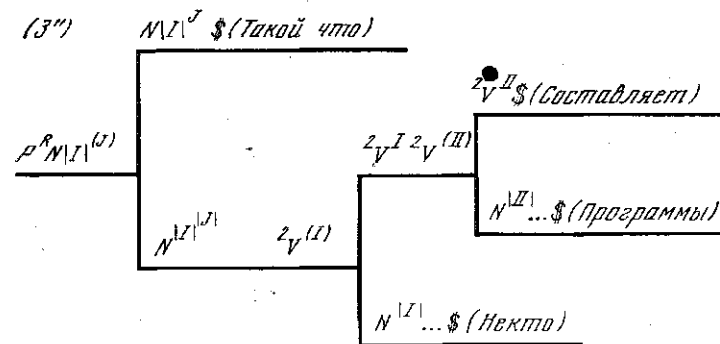
т. е. вербальная группа с отношением M развертывается на две составляющие: вербальную V^M и номинальную $N^{|M|}$, с прямым и конверсным повтором имени отношения M ; при этом вербальная составляющая сохраняет левый индекс n , показывающий число разрешенных заполнений («объем») вербальной составляющей, одно из которых и имеет номер M .

$$\text{II. } {}^nV^M, M < n \rightarrow {}^nV^{(M+1)}$$

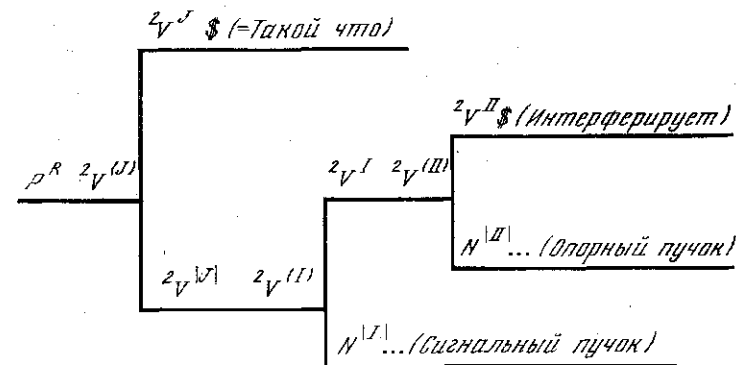
$$\text{III. } {}^nV^M, M = n \rightarrow \$$$

т. е. вербальная составляющая, не полностью реализовавшая свой объем ($M < n$), заполняется вербальной группой очередного номера, а составляющая, реализовавшая свой объем полностью, отсылается на вход словаря.

Объем n реляционной структуры ${}^nV^{(I)}$ со словом $\$A$ из



= $N^{|Z|}^{(J)}$, Такой что Составляет Программы $N^{|Z|}$ (=Некто)



= $(\text{Факт})^2 V^{(J)}$, Такой что Интерферирует Опорный пучок Сигнальный пучок

РИС. 3. Дерево юнктивной группы

вербальной составляющей самой тесной группы ${}^nV^{(M)}$, $M=n$ можно считать реализацией присоединительных способностей слова $\$A$, отчего все отношения этих структур называются валентными, точнее, валентностями слова $\$A$, а имена $N^{|M|}$ — дополнениями слова $\$A$ по соответствующим валентностям ${}^nV^{(M)}$, $1 \leq M \leq n$. Уместно подчеркнуть, что в пределах валентной структуры одного слова различие синтаксических отношений только по возрастающему номеру M , или по «тесноте», оказывается вполне достаточным (ср. § 2).

Особую конкретизацию общие правила (2) получают для группы с отношением номинализации (или юнкции) по элементу — см. правила (2''), реализуемые в виде (3''), рис. 3.

(2'')

$$\text{IV. } W^{(J)} \rightarrow W^J + W^{|J|} (W = V, {}^nV, N^{|M|});$$

V. $N|M|^{M1} \rightarrow {}^nV^{(I)}, n \geq M;$

VI. $W^{M1} \rightarrow {}^nV^{(I)};$

VII. $W^J \rightarrow \$ (= \text{ТАКОЙ, ЧТО}).$

Согласно IV юнктивная группа развертывается на юнкт — левая составляющая, — и адъюнкт — правая составляющая, причем категориальной характеристикой юнкции, сохраняющейся у ее составляющих, является отсылка к элементу W той вербальной группы, которая по правилам V и VI заполняет адъюнкт. Этим элементом может быть валентное слово V , вербальная группа в целом nV или дополнение по любой из валентностей $N|M|$; в последнем случае общее число валентностей n подадъюнктивной группы должно быть не меньше номера M у этого дополнения. Согласно правилу VII юнкт всегда осложняется лексическим оператором ТАКОЙ, ЧТО.

Можно сказать, что юнкция по элементу адъюнктивного высказывания есть способ обогащения этого элемента за счет высказывания. При этом сам элемент W , подлежащий обогащению, непосредственно упоминается только в составе адъюнктивного высказывания на правах его рядового члена: $N|M|$ (НЕКТО) СОСТАВЛЯЕТ ПРОГРАММЫ; обогащенное состояние этого элемента выражается отдельно — при категории юнкта, и опосредованно — в виде адресной отсылки к его адъюнктивному вхождению ($N^J|M| \$$ ТАКОЙ ЧТО...).

Главное достоинство расчлененного представления номинальных конструкций — сохранение в полной форме, без местоименных замен и иных сокращений, того высказывания, которое лежит в основе номинализации.

Расчлененное представление реляционных структур, как оно создается правилами (2') и (2''), отличается элементарностью своих претерминальных категорий и их малочисленностью; это валент V , дополнение $N|M|$ и оператор-юнкт W^J . Названные свойства не только полезны сами по себе, например для введения первичной, категориально-простой лексики (см. [11]), но и удобны для перехода к другой форме представления, где элементарные категории стянуты в узлы и образуют составные категории, такие, как отглагольное прилагательное (*составляющий, интерферирующий, сигнальный*) или имя (*интерференция, программирование, составитель, программист*).

(5) Стянутая форма реляционной структуры достигается простым способом (см. [11]): достаточно в каждом отношении реляционного дерева (рис. [2; 3]) элиминировать групповой символ $P^{(R)}$, а вместо него для заполнения соответствующего элементарного символа использовать непосредственно одну из составляющих, соответственно вербальную или юнктивную, назначив ее для этого главной в своей паре, — другими словами, представив синтаксическое отношение между составляющими подчинительной связью, исходящей из главного элемента (5'), см. рис. [4].

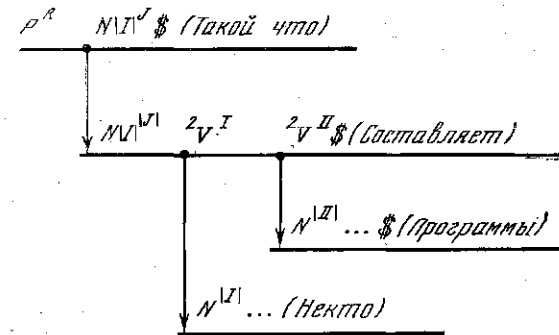


РИС. 4. Стянутое дерево отношений

Очевидно, что стянутая форма реляционной структуры есть дерево подчинения, в котором связь любых двух узлов всегда опосредована парой согласованных синтаксических категорий, концевых для данной связи. В терминах этих концевых категорий не только называется сама связь, но и фиксируется взаимное расположение объединяемых узлов, а также указывается условное направление подчинения. При этом для всякой комбинации категорий, образующейся при данном узле из категории зависимого конца входящей связи и/или главного конца исходящей(их) связи(ей), сам собой складывается определенный порядок по заполнению, или по тесноте; ср. выше узел:

$$\{ N|M|^M1 \cdot 2V^I \cdot 2V^II \$$$

Замечательное свойство стянутой формы, отличающее ее от формы расчлененной, состоит как раз в том, что юнктивные и вербальные символы, объединенные и упорядоченные в составе того или иного узла, четко описывают некоторую составную синтаксическую категорию («часть речи»), такую, как «личный глагол», «отглагольное прилагательное» и др. (рис. 5).

Стяжение реляционных структур, как ядерных, так и особенно производных (см. § 3), создает условия для того, чтобы на основе структур с первичной, категориально-простой лексикой (НЕКТО, СОСТАВЛЯЕТ, ПРОГРАММЫ; ТАКОЙ, ЧТО...) вводить категориально более сложную, вторичную лексику: дериваты (СОСТАВЛЯЮЩИЙ, ПРОГРАММИРОВАНИЕ) и композиты (СОСТАВИТЕЛЬ, ПРОГРАММИСТ) (ср. [3]).

2. Усложнение ядерного синтаксиса. Различаются три вида усложнений ядра:

1) гиперсинтаксическое, надстраивающее над более или менее крупными фрагментами реляционной структуры какие-либо из коммуникативных интенций — «снижение глубины» (по Ингве), «сепарация», «тематизация», «эмфаза» (см. [7, 10]);

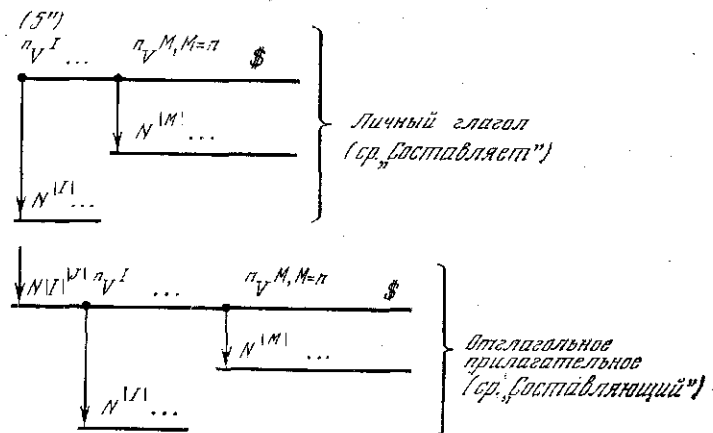


РИС. 5. Составные синтаксические категории

2) метасинтаксическое, вводящее на отдельных синтаксических отношениях особые различительные имена в прямом и конверсном виде: «действие»/«агенса», «состояние»/«пациенса», «восприятие»/«перцепиенса» и т. п.;

3) гипосинтаксическое, снабжающее определенные категории дополнительными лексико-грамматическими (субкатегориальными) признаками: предметная или сигнификативная тождественность для V или N , род — число — лицо для N , предложноподлежащее управление для V^M (ср. [4]).

Остановимся подробнее на метасинтаксическом усложнении. Оно состоит в придании синтаксическим отношениям особых различий, имеющих абсолютный характер, не зависящий от вхождения в тот или иной валентный узел, подобно «глубинным падежам» Ч. Филмора [5]. Однако в отличие от последних метасинтаксические категории:

во-первых, не привязаны к структуре отдельного слова, охватывают реляционную структуру в целом, до ее расчленения на словесные узлы и в определенной мере сами являются поводом для такого расчленения (см. [7, 8]);

во-вторых, они более разнообразны по составу; наряду с актантными связями типа приведенных выше сюда входят отношения для обозначения, с одной стороны, начальных стадий фиксации термов («дефиниция»/«дефинит»: «а — является М»; «квантификация»/«квантификат»: «аМ — всякий», ср. [9]), а с другой — для введения более поздних шагов характеристики предметов и ситуаций («параметризация»/«явление»: «всякий аМ имеет длину»; «этот факт V^M имеет время»);

в-третьих, метасинтаксические категории частично упорядочены между собой и при введении в вербальные узлы выступают в виде правильных комбинаций, соответствующих числу валентностей узла; соответствие может использоваться в обе стороны:

от заранее объявленного числа валентностей вербальной группы к построению комбинации метасинтаксических категорий и обратно — от правильно построенной комбинации метакатегорий к объявлению на символе V^I требуемого числа валентных отношений.

Даже тогда, когда мы отводим метасинтаксическому усложнению направляющую роль в выводе реляционной структуры, чисто синтаксическая природа самого вывода и результирующего дерева отношений не меняется, поскольку целиком определяется правилами вербальной грамматики (2'—2"). Но тем самым реляционная форма ядерного синтаксиса, пока она не получила специфически поверхностных усложнений и расширений, достаточно обща, чтобы, сохраняя синтаксическую правильность, выражать и правильность более «тонкого», метасинтаксического уровня. Поэтому мы не видим необходимости в особом «семантическом» языке с его синтаксисом (ср. [3, 6]) и считаем ядерный реляционный синтаксис, предполагающий усложнения и расширения, единственным, в котором нуждается полное описание естественного языка.

Характерно, что все три названные типа усложнений задаются одинаковым образом: комбинации гипер-, мета- и гипосинтаксических категорий формируются по своим правилам и вводятся в ядерную структуру при объектах определенного вида: узлах, отношениях или отдельных категориях, см. пример (рис. 6).

Таким образом, усложненная базовая структура (6) содержит все виды информации: коммуникативную (гиперсинтаксическую), семантическую (метасинтаксическую) и первично-грамматическую (гипосинтаксическую), но только в автономном, недораспределенном и невоплощенном виде.

Предполагается, что в дальнейшем автономное коммуникативное задание (например, ТНЕМ) с помощью определенных преобразований будет воплощено в самой древесной структуре; окончательная перестройка структуры будет зависеть также от наличия в ней тех или иных метасинтаксических категорий (например, параметрического отношения, вводящего «причину»), от знаков лексической тождественности \oplus и др., а также от характера первичной лексики, точнее, от того, какой вторичной лексикой она обеспечена в словаре конкретного языка, т. е. на какие более емкие композитные слова и более подходящие для изменившегося категориального узла синтаксические дериваты ее принято заменять в этом языке.

Для реализации усложненной базисной структуры необходимы, во-первых, определенные способы преобразования, расширяющие арсенал синтаксиса (§ 3), и, во-вторых, правила избирательного, мотивированного привлечения этих средств (§ 4).

3. Расширения базового синтаксиса. По характеру преобразуемых объектов различаются три типа расширения базового синтаксиса: модуляционное, модификационное и дистрибуционное.

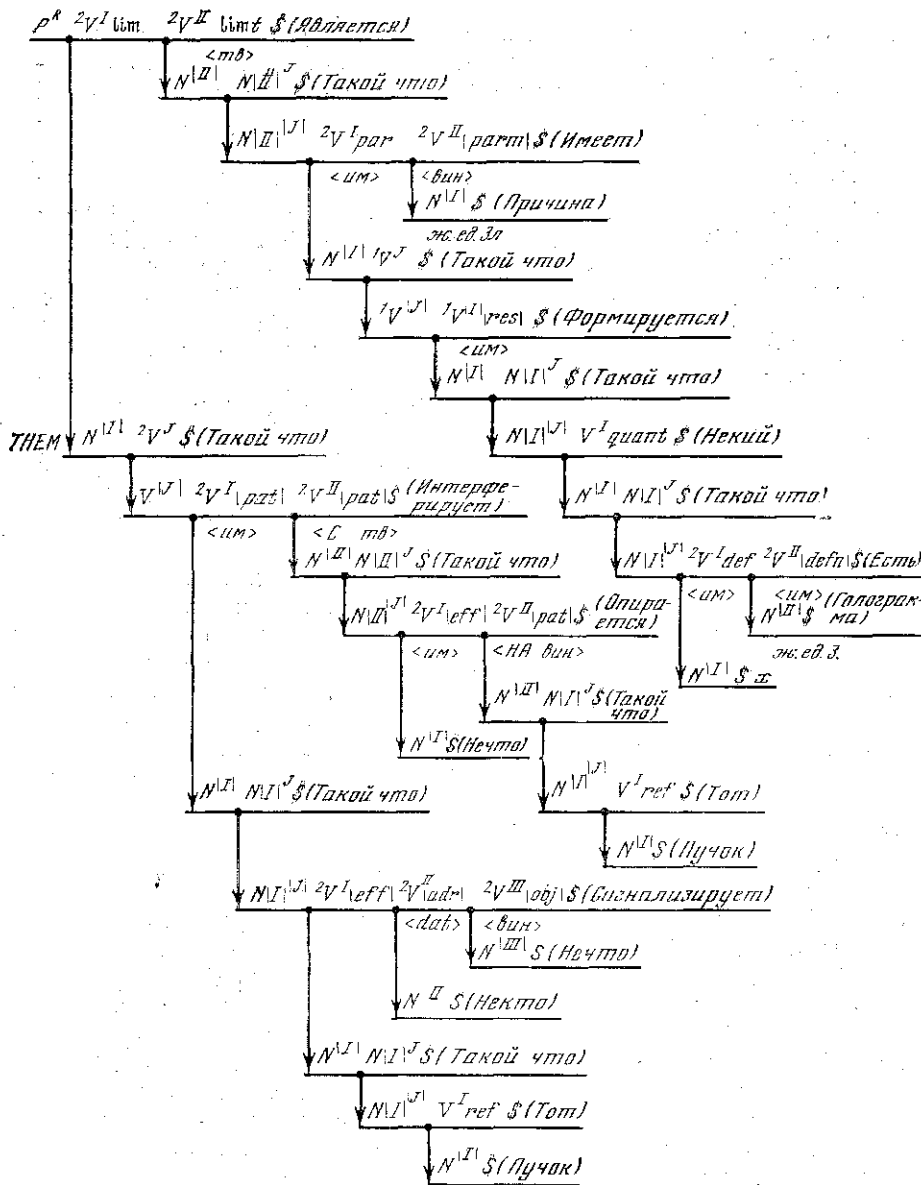


РИС. 6. Стянутое реляционное дерево с усложнениями

А. посредством модуляционных преобразований и создаются производные структуры с новыми синтаксическими отношениями (подчинительными связями). Модуляции имеют дело с узлами, которые вычлениваются либо из внешней, корневой подструктуры, либо из подструктуры внутренней —

валентной или юнктивной. С подструктурой внешней имеют дело модуляции аннексирования (ANX) и прелиминации (PRAEL), а с внутренней — модуляции инкорпорирования (INC) и экспонирования (EXP), причем первая преобразует валентную подструктуру, а последняя — юнктивную.

В результате модуляционных преобразований у производных структур появляются новые, модуляционные отношения (связи) с соответствующими названиями.

Связь аннексивная подчиняет внешний высказывательный узел, лишенный одного внутреннего высказывательного узла, этому последнему узлу (коннексу) в качестве обстоятельства (аннекса) при нем;

Интерференция... является причиной того, что формируется голограм-

ANX ← |ANX|
 ма ⇒. Вследствие интерференции формируется голограмма.

В частном случае внешний высказывательный узел сокращается до одной первой валентности, лишаясь целого комплекса всех остальных валентностей, которым предстоит составить самостоятельное высказывание — коннекс:

Совмещенный луч записывает голограмму в ячейку памяти ⇒ Совмещен-

ANX ← |ANX|
 ным лучом голограмма записывается в ячейку памяти.

Связь прелиминативная подчиняет усеченной версии высказывательного узла (прелиминату) его полную версию («постлиминат»);

надежность хранения информации голограммой превосходит надежность

PRAEL →
 хранения информации обычными средствами. ⇒ Голограмма превосходит
 обычные средства по надежности хранения информации.

Связь инкорпоративная непосредственно подчиняет валентному узлу тот элемент, который от этого же узла прежде зависел опосредованно;

Формирование голограммы имеет то, что интерференция является при-

INC → |INC|
 чиной формирования голограммы ⇒ Формирование голограммы имеет интер-
 ференцию своей причиной.

Связь экспозитная подчиняет высказывательному узлу, вынесенному из юнкции другого высказывания, весь этот другой узел, становящийся аппозитом при вынесенном узле — экспозите.

Вследствие того, что сигнальный пучок интерферирует с опорным, фор-

EXP →
 мируется голограмма ⇒ Сигнальный пучок интерферирует с опорным, след-
 ствие чего формируется голограмма.

Экспозитивная связь может быть слаборазделительной (как в предшествующем примере) или сильноразделительной, ср.:

Голограмма, формирующаяся вследствие интерференции..., записывается в ячейку памяти \Rightarrow Вследствие интерференции... формируется голограмма. $\xrightarrow{\text{EXP}}$
Она (эта голограмма) записывается в ячейку памяти. $\leftarrow \text{EXP}$

Как показано в [7,10], именно с модуляцией экспозиции и ее серийным применением к исходному, достаточно глубинному (в смысле Ингве) высказывательному узлу, связано построение сверхфразового единства (текста) в разных его формах; сверхфразовое единство оказывается частным следствием расширения базового синтаксиса.

Б. Модификационные преобразования имеют дело с различными параметрами подчинительных связей (базисных и модуляционных), меняя взаимное расположение концевых категорий одной и той же связи (конверсия — CONV), иерархию по тесноте у связей одного и того же узла (реклинация — RECL) или глубину узла среди других узлов дерева подчинения с вынесением внутреннего узла наружу (превенция — PRAEV) или введением внешнего узла вглубь (индукция — IND).

Примеры:

CONV: Является причиной формирования голограммы интерференция \Rightarrow Интерференция... является причиной формирования голограммы.

RECL: Интерферирует с опорным пучком сигнальный пучок \Rightarrow Интерферирует сигнальный пучок с опорным пучком.

PRAEV: Интерференция является причиной, которую имеет формирование голограммы \Rightarrow *Формирование голограммы имеет то, что интерференция является причиной, которую имеет формирование голограммы.

IND*: Есть возможность, что сигнальный пучок интерферирует с опорным \Rightarrow Сигнальный пучок может интерферировать с опорным.

Вводятся внутрь дерева, причем на разную глубину, могут узлы чисто категориальные, прежде в дереве отсутствовавшие, например пустой присубстантивный глагол + имя, приаждективный глагол + прилагательное, присирконстантный глагол + сирконстант: Специальная программа сжимает информацию \Rightarrow Происходит сжатие информации специальной программой.

При более глубоком индуцировании субстантивный узел может подчинять только сказуемое исходной структуры или также его дополнение: Специальная программа производит сжатие информации или Специальная программа подвергает информацию сжатию.

В. Дистрибуционные преобразования вообще не затрагивают синтаксической структуры и только перемещают по ней некатегориальные элементы — лексику и грамматические (гипосинтаксические) признаки.

* Интерференция является $N(II)^J$ таким, что имеет $N|II|$ (= причину) формирование голограммы \Rightarrow Интерференция является причиной, которую имеет формирование голограммы.

В приведенных выше примерах, иллюстрирующих результат некоторого преобразования, мы для удобства использовали хорошую поверхностную форму, т. е. на самом деле сопровождали иллюстрируемое преобразование целым рядом сопутствующих. Дело в том, что такие преобразования задаются целыми блоками, поскольку именно в совокупности они служат воплощению какого-либо одного мотива, в силу макроправил мотивационного компонента (§ 4).

4. Избирательное (мотивированное) привлечение средств расширенного синтаксиса. Базовым структурам, предназначенным для выражения максимума значащих и только значащих различий, свойственны сугубая аналитичность, жесткое единообразие и бедность выразительных средств. Разные типы расширений базового синтаксиса представляют собой арсенал средств для перевода базовых структур в нормальную для данного естественного языка идиоматичную и разнообразную форму. Но чтобы, с одной стороны, учесть все те специфические особенности в строении и усложнении базовых структур, которые в данном естественном языке вообще нетерпимы или к которым он просто «чувствителен», а с другой — чтобы организовать избирательное привлечение оптимальной комбинации средств для «реакции» на определенную особенность базовой или не до конца преобразованной структуры, необходим особый мотивационный компонент.

Мотивационные правила приближают некоторую специфичную входную структуру к нормальному, по возможности оптимальному в данном языке (стиле) виду, причем так, чтобы в ходе этого последовательного улучшающего приближения сохранить верность именно этой базисной структуре с ее значением. При изменении формы улучшаемой структуры ее смысл должен остаться тем же.

Четкие различия в реляционных структурах, обеспечиваемые связным заданием базового синтаксиса с его эксплицитностью и расчлененностью (см. § 1, 2), составляют главную предпосылку для все более дифференцированного и мотивированного подхода к воплощению этих структур в поверхностной форме текста. Более подробно это освещено в [12], а приведенные примеры преобразований, фактически составных (т. е. трансформаций), дают достаточное представление об отдельных шагах трансформационного вывода, приближающего базовую структуру (б) к поверхностному виду (Г):

Вследствие интерференции сигнального пучка с опорным формируется голограмма.

ЛИТЕРАТУРА

1. Yngve V. H. A model and hypothesis of language structure. — Proc. Amer. Philos. Soc., 1960, vol. 104, N 5.
2. Chomsky N. Aspects of the theory of syntax. Cambridge (Mass.): MIT press, Mass. Inst. Technol., 1965.
3. Lakoff G. On the nature of Syntactic irregularity. — In: Mathematical linguistics and automatic translation. Harvard Univ., 1965.

4. Weinreich U. Explorations in semantic theory.— In: Current trends in linguistics/Ed. T. A. Sebeok. The Hague, 1966, Bd. 3.
5. Fillmore Ch. J. The case for case.— In: Universals in linguistic theory/Eds. E. Bach, R. T. Harms. N. Y., 1968.
6. McCawley J. D. Meaning and the Description of Language.— In: Linguistische Forschungen/Ed. F. Kiefer. Athenaum, 1970.
7. Мартемьянов Ю. С. Связный текст — изложение расчлененного смысла. М.: Ин-т рус. языка АН СССР, 1973, вып. 40. Предвар. публ.
8. Мартемьянов Ю. С., Московой В. А. Об исчислении словарных входов. 1. Прикладная лингвистика. М.: МГПИИЯ, 1976, вып. 18.
9. Бестужев Э., Делла Ж.-П. ЛИТФАЛЛ — программа исследований по естественному языку.— В кн.: Взаимодействие с ЭВМ на естественном языке. Новосибирск, 1978.
10. Мартемьянов Ю. С. Текст: от канонической структуры к разнообразию форм.— В кн.: Лингвистические проблемы проектирования информационных систем. Киев: Ин-т кибернетики АН УССР, 1978.
11. Мартемьянов Ю. С. Синтез: трансформирование глубинной структуры в поверхностную.— В кн.: Современное состояние теории и практики машинного перевода и автоматизации информационных процессов. М.: ВЦП, 1977.
12. Мартемьянов Ю. С. Об управляемом выводе предложений.— В кн.: Диалоговые системы и представление знаний.— Учен. зап. Тартус. ун-та, 1981, № 594.

УДК 025.4:681.3

Д. Р. Лазути, Е. Б. Федоров, И. С. Добронравов,
В. Ф. Пархоменко

АВТОМАТИЧЕСКОЕ ИНДЕКСИРОВАНИЕ ТЕКСТОВ В ДОКУМЕНТАЛЬНЫХ ИПС*

В работе любой документальной информационно-поисковой системы (ИПС) можно выделить три основных семантических процесса.

1. Индексирование документов, т. е. перевод текста документа на информационно-поисковый язык (ИПЯ) данной системы. Результатом индексирования является поисковый образ текста (ПОТ). Последний мы отличаем от поискового образа документа (ПОД), который может содержать, во-первых, несколько разных поисковых образов текста на разных ИПЯ, во-вторых, формальные поисковые признаки типа библиографических, с индексированием которых не связаны серьезные семантические проблемы.

2. Индексирование запросов. Для некоторых ИПС этот процесс аналогичен индексированию документов, однако в ряде ИПС (и прежде всего в большинстве дискрипторных) — это процесс принципиально отличный: имеет место не просто перевод текста запроса на язык системы, а определение «стратегии поиска». Индексатор, опираясь на свою интуицию и знание структурных особенностей ИПС и ее конкретной реализации, а также на зна-

ние предметной области, по которой предполагается поиск, разворачивает тему запроса в ряд подтем (подзапросов), которых может быть достаточно много (в некоторых случаях до 20 и более), определяет их логическую структуру, быть может, приписывает различные веса терминам и т. п. Ответом на запрос является объединение выдач на отдельные подзапросы, независимо от того, проводился ли поиск отдельно по каждому подзапросу или они объединялись — дизъюнктивно — в одно поисковое предписание (ПП).

3. Собственно поиск, т. е. сравнение по формальным правилам — критерию смыслового соответствия (КСС) или критерию выдачи — заиндексированного запроса с заиндексированным документом и принятие на основании этих правил решения о выдаче или невыдаче документа на запрос¹.

Если процесс поиска алгоритмизован и автоматизирован в любой автоматизированной ИПС (можно сказать, по определению, поскольку без этого вообще бессмысленно говорить об *автоматизированной* ИПС), то процесс содержательного развертывания запроса не автоматизирован на сегодняшний день ни в одной производственно действующей АИПС. Это связано прежде всего с содержательной трудностью алгоритмизации этого наиболее интеллектуального из перечисленных процессов. Пока можно говорить лишь о некоторых намечаемых подходах к его алгоритмизации, и среди них можно выделить два основных.

Первый ориентирован на алгоритмизацию адаптации поискового предписания к информационной потребности пользователя в процессе обратной связи с ним. Это всегда подразумевает принципиально неалгоритмический этап — оценку пользователем релевантности полученных на данный момент документов; задача сводится к алгоритмизации процедур перестройки поискового предписания, обеспечивающей максимальное приближение оцениваемой выдачи к требованиям, выраженным оценкой. Чаще предлагаемые методы основаны на учете сравнительной частоты встречаемости дескрипторов в релевантных и нерелевантных документах. Заметим, что сколько-нибудь проработанные решения предложены пока только для дескрипторных ИПС «без грамматики» (ср., например, [29, 23, 4]). Исходное поисковое предписание служит при этом только начальной точкой итеративного процесса; оно может быть результатом послышного автоматического индексирования текста исходного запроса, представленного пользователем [4], либо алгоритмически порождаться из задаваемого пользователем набора «отмеченных» (т. е. релевантных запросу) документов [23].

Второй подход стремится моделировать работу человека-эксперта при порождении подзапросов на основе исходной формулировки запроса; он существенным образом опирается на использование тезаурусов, а также грамматических отношений в ИПЯ,

* Статья основана на докладе, прочитанном на Всесоюзной конференции по единой системе информационных языков (Юрмала, 1977 г.)

В более общем случае КСС определяет степень близости документа к запросу (ранг, или эшелон, документа по данному запросу).

позволяющих как минимум выделять в тексте фрагменты, соответствующие подзапросам (ср., например, [11, 14]).

Промежуточное положение с точки зрения алгоритмизации занимает процесс индексирования текстов, который и является основным предметом данной статьи. Под индексированием текстов мы понимаем как индексирование документов, так и запросов, уже прошедших стадию предварительного содержательного развертывания.

Для первых всерьез предложенных к практическому применению систем автоматического или механизированного поиска документальной информации задача алгоритмизации индексирования текстов не ставилась. Казалось очевидным, что обеспечить приемлемое качество поиска в них может только индексирование, осуществляемое человеком — интеллектуальное, или ручное, индексирование (в литературе эти термины употребляются как взаимозаменяемые; ниже мы увидим, что эта «условная синонимия» не случайна). Среди предлагавшихся в 50-х годах были как системы с предельно простыми языками (Таубовские унитары [30, 31]), так и с языками довольно сложной структуры, обладающими, в частности, развитой системой указателей связи и роли (язык Перри—Кента, ср. [26]). Для последних представление о неперспективности автоматического индексирования, естественно, объяснялось сложностью задачи — определение указателей роли, например, в общем случае не алгоритмизовано и до сих пор. Для первых оно, помимо соображений технической реализуемости, определялось, по-видимому, тем значением, которое придавали задаче *выбора* существенных поисковых признаков текста. Считалось, что из всех элементарных поисковых признаков текста (унитармов, ключевых слов) необходимо выбрать только ограниченное число наиболее существенных (это ограничение было связано с техническими возможностями доступных в то время средств механизированного или автоматического поиска), что выбор несущественных признаков приведет к серьезным ошибкам при поиске и что правильный выбор может осуществить только достаточно квалифицированный человек. Считалось также, что достаточно важную роль в обеспечении нужного качества поиска может играть приписывание тексту поисковых признаков (дескрипторов), выраженных в нем лишь неявным образом (индексирование «между строк»).

Только в 60-х годах начались серьезные исследования, основанные на гипотезе, которую можно сформулировать так:

Можно заранее фиксировать семантически значимые объекты языка в области применения ИПС, например в виде списка ключевых слов (КС) — дескрипторного словаря, состоящего из отдельных слов и словосочетаний. Если использовать эти списки в системе пословного перевода документов с естественного языка на ИПЯ, то результаты поиска по документам, заиндексированным таким образом, в условиях практической эксплуатации не будут существенно уступать результатам поисков в аналогичных ИПС с индексированием, осуществляемым человеком.

Существенна здесь оговорка «в условиях практической эксплуатации». Разумеется, содержательное интеллектуальное индексирование, осуществляемое подготовленным человеком, может в принципе обеспечить весьма высокое качество ПОТ, так что на экспериментальных массивах в несколько сот документов пословное автоматическое индексирование (АИ) может и уступать в чем-то содержательному индексированию. Однако опыт показывает, что в условиях массовой, рутинной обработки текстов при достаточной жесткости норм среднй индексатор далеко не использует всех своих интеллектуальных возможностей, а ориентируется в основном на поисковые признаки, наиболее легко распознаваемые в тексте документа (в заглавии, в первых фразах). И в этом случае есть основания говорить не столько об интеллектуальном, сколько о ручном индексировании, т. е. о процессе, хотя и выполняемом человеком, но по существу достаточно механическом.

Исследования в области пословного АИ начались независимо в 1962 г. в США в связи с созданием системы SMART [20, с. 22, 35] и в СССР — в связи с созданием ИПС класса «Пусто-Непусто» [15, с. 10, 3]. Но если работы группы SMART долгое время сохраняли чисто исследовательский характер, так что практическое внедрение АИ за рубежом связано с другими коллективами и относится уже к 70-м годам (ср. ниже), то АИ на ИПЯ «Пусто-Непусто» начало практически использоваться в информационной службе «Электротехника» уже с 1967 г.

Экспериментальные исследования и практический опыт первых систем пословного АИ подтвердили сформулированную выше гипотезу (ср., например, [29, с. VIII—IX]); в 70-е годы можно констатировать постепенный рост числа использующих его АИПС: система FAIRS в Евратоме [27], система DIRECT фирмы Филипс [25], система PIAF DOC Национального бюро НТИ Франции [28], система АИ АИДОС, разрабатываемая в ГДР комбинатом «Роботрон» [1, с. 40—41, 6] и др. Можно упомянуть также системы, предназначенные для работы непосредственно с текстами на естественном языке, о которых будет говориться ниже.

Представляется полезным на основе некоторых общих соображений оценить сегодняшнее положение дел и попытаться определить наиболее перспективные на ближайшее будущее направления развития АИ. Мы исходим из того, что наиболее реальной перспективой на обозримый период является сосуществование языков, ориентированных как на автоматическое, так и на ручное индексирование (для краткости можно говорить о языках АИ и соответственно РИ). Еще рано решать вопрос о том, должны ли в конечном счете языки АИ полностью вытеснить языки РИ или же их сосуществование будет постоянным, но уже сейчас можно говорить о неуклонном расширении сферы действия языков АИ.

Под индексированием мы понимаем переход от исходного текста документа, написанного на естественном языке, к непосредственному объекту поиска — ПОТ. Соответственно об алгоритмическом (автоматическом) индексировании можно говорить, если этот пе-

реход с начала до конца осуществляется алгоритмически (автоматически). Если автоматизированы не все стадии этого процесса, можно говорить о частично автоматическом, в основном автоматическом (автоматизированном) индексировании и т. п.

В известных ИПС результат индексирования (ПОТ) более или менее далеко отстоит от исходного текста. Предельно простым — вырожденным — случаем является совпадение ПОТ с исходным текстом: «отсутствие» индексирования, или тривиальное индексирование. Иногда в этом случае говорят об «ИПС, работающих с естественным языком» (критику этого термина см. в [19, с. 6, 24, с. 94—98]). Правильнее было бы говорить о системах, работающих в процессе поиска с текстами, написанными на естественном языке. Заметим, что тривиальное индексирование по определению алгоритмично.

ПОТы, получаемые нетривиальным индексированием, весьма разнообразны: наборы унитермов, ключевых слов, в том числе словосочетаний или дескрипторов, быть может связанных грамматическими средствами, простые или сложные предметные рубрики, классификационные индексы, анкеты (например, типа «стандартных фраз») и т. д.

В это разнообразие типов ПОТов и соответственно процедур индексирования можно ввести хотя бы частичную упорядоченность, исходя 1) из размера контекста в исходном документе и 2) степени сложности анализа этого контекста, минимально необходимых для определения простого индекса (элементарного поискового признака) и построения (сборки) из простых индексов полного ПОТга или, иными словами, из степени локальности (соответственно глобальности) индексирования (ср. [22, с. 4—5]).

Так, например, две процедуры индексирования, опирающиеся соответственно на выбор ключевых *словформ* и *слов*, имеют одинаковый минимальный размер анализируемого контекста (от пробела до пробела), но разную степень сложности анализа: в первом случае — простейший анализ на совпадение, во втором — более или менее сложный (в зависимости от степени флективности языка) морфологический анализ. Аналогично опознание словосочетаний в контексте одного предложения можно производить либо просто на основе проверки вхождения всех его элементов в это предложение, либо требовать проверки дополнительных условий, более или менее сложных, вплоть до совпадения синтаксических структур (ср. [2, 20, с. 50—51; 2, с. 256—257])¹.

¹ Вообще говоря, следовало бы различать контекст, необходимый для опознания возможного поискового признака, и контекст, необходимый для решения вопроса о включении его в окончательный ПОТ. Для большинства методов РИ и некоторых АИ (например, основанных на учете частоты встречаемости КС в тексте документа) последний совпадает со всем текстом, в то время как первый может быть сколь угодно ограничен (до словоформы). Здесь мы отвлекаемся от этого аспекта, поскольку для интересующих нас в первую очередь методов АИ второй контекст может, как правило, быть ограничен первым без существенного снижения качества работы ИПС (ср. [15, с. 9—11]).

Соответственно получаем, условно говоря, некую шкалу процедур индексирования — от максимально локальных до максимально глобальных (условно, так как строго линейный порядок все же установить не удастся). На одном конце — тривиальное индексирование (контекстом можно считать один символ, анализ на совпадение), унитарное индексирование (контекст — от пробела до пробела, анализ — на совпадение либо тот или иной вариант морфологического анализа, полный ПОТ — неупорядоченный или условно упорядоченный набор унитермов) и т. д., а на другом — классификационные индексы (например, сложные индексы УДК или индексы анкетного типа), в идеале предполагающие содержательный анализ всего текста.

Промежуточное положение займут дескрипторные языки словосочетаний (анализируемый контекст, как правило, не превышает предложения), дескрипторные языки с грамматикой (контекст, как правило, не превышает одного абзаца, если не считать учета заглавия документа, во всяком случае, при индексировании вторичных документов), языки предметных рубрик, для которых анализируемым контекстом служат достаточно крупные фрагменты текста — межфразовые комплексы. Дескрипторное индексирование в ИПС пословного типа «без грамматики» на этой шкале совпадает с унитарным, поскольку то, что принято считать его специфическим отличием — объединение унитермов в класс условной синонимии — не использует дополнительного анализа контекста.

При классификации процедур индексирования следует также учитывать тип текстов, на который они ориентируются.

1. Первичные документы — тип текстов, наиболее богатый как по лексическому составу, так и по синтаксическим конструкциям; тексты этого типа отличаются значительным объемом и соответственно наиболее трудно поддаются автоматическому индексированию.

2. Вторичные документы — рефераты. Лексически и синтаксически они лишь немногим беднее текстов первого уровня (например, для них в той же мере характерны предикативные конструкции), но значительно меньше их по объему. Этот класс текстов, порождаемых многочисленными реферативными службами, можно считать основным источником информации для современных документальных АИПС и соответственно основным полем применения методов АИ.

3. Аннотации, т. е. тексты, состоящие только из назывных предложений. Синтаксически более простые, они допускают применение более тонких методов АИ, чем пословное (ср. [12, 18]), но для практически действующих ИПС их пришлось бы в большинстве случаев породить специально. Основными, практически важными подклассами текстов этого уровня являются *заголовки* и *запросы*.

Традиционные ИПЯ — классификационные языки и языки предметных рубрик — ориентированы прежде всего на индексирование первичных документов, тогда как унитарные и дескриптор-

ные — вторичных. В этом одна из причин большей глобальности индексирования для первых, но это различие не безусловно, поскольку можно применять как традиционные языки к индексированию вторичных документов, так и дескрипторные (ср., напр., [7, с. 172—174]) к индексированию первичных.

Нетрудно видеть, что упорядочение процедур индексирования по степени локальности (глобальности) индексирования есть в то же время упорядочение их по степени простоты (сложности) алгоритмизации: чем меньше контекст и проще анализ, тем легче его алгоритмизировать, и наоборот.

На этом этапе рассмотрения уместно подчеркнуть, что между процедурами индексирования и поиска (точнее, между процедурой индексирования документов и выбором стратегии поиска) существует связь дополнительности: те или иные частные задачи или процедуры можно переносить из сферы индексирования в сферу поиска, и наоборот (ср. [13, с. 9—10]).

Простейший, но содержательно важный пример: сравним (1) дескрипторную систему с классами условной синонимии и позитивным словарем для пословного индексирования текстов (например, типа «Пусто-Непусто-4») и (2) систему, «работающую с естественным языком» (типа АСОД [16] ¹). Процедура АИ для (1) состоит в опознании слов (КС) по словоформам и опознании дескрипторов — по КС. Для (2) обе задачи переносятся на процедуру поиска: надо в поисковом предписании задать отсечение окончаний (отвлечемся от возможных при этом ошибочных отождествлений, о чем ниже), словари синонимов и т. п.

Отсюда, в частности, вытекает крайняя условность рассмотрения индексирования и ИПЯ в отрыве от поиска и ИПС в целом: это единый объект, те или иные операции которого могут смещаться с одного структурного подразделения на другое.

Еще один содержательно важный пример: опознание (не жестких) словосочетаний в тексте может производиться, как отмечалось выше, по разным критериям, но в любом случае оно протекает в одном из двух режимов: в процессе индексирования, когда опознанное сочетание слов получает свой код, который и заносится в ПОТ (ср., например, [2, с. 262]), либо в процессе поиска, когда текст представлен цепочкой (или графом) КС или их кодов и опознание заданного сочетания происходит каждый раз заново. В этом случае будем говорить о «дескрипторосочетаниях», понимая дескриптор в широком смысле — как нормализованное КС или его код, необязательно с отождествлением до класса условной синонимии (см. [15, с. 12]).

¹ Здесь затрагивается вопрос о соотношении пакета прикладных программы (ППП) и ИПС. Современный ППП для документального поиска (АСОД, АИДОС и т. п.) может реализовать некоторый класс ИПС: унитарную, дескрипторную без грамматики, дескрипторную с элементами грамматики, с тем или иным тезаурусом и т. д. Но можно понимать под ИПС, соответствующей данному ППП, тип максимально сложной (сильной) ИПС, реализуемой с помощью данного пакета.

Возвращаясь к шкале относительной сложности (и соответственно алгоритмизуемости) различных процедур индексирования, можно выделить на ней следующие градации, соответствующие типу анализируемого контекста:

Словоформы,

Слова,

Словосочетания:

жесткие,

рассеянные по тексту:

внутри предложения,

за пределами предложения,

эллиптические.

Во всех этих случаях задачей индексирования является опознание в тексте тех или иных слов и их сочетаний. Для процедур индексирования более высокого уровня характерна задача опознания в тексте не столько тех или иных выражений как таковых, сколько выражаемых ими понятий. Здесь также возможны градации, связанные с тем, что искомые понятия могут быть выражены в тексте непосредственно в своей «естественной» формулировке (этот случай сливается со случаем словосочетаний);

косвенно, т. е. в неявной форме — случай, наиболее плохо поддающийся алгоритмизации.

Граница практически осмысленной алгоритмизации индексирования со временем естественно сдвигается. Если еще недавно она проходила на уровне автоматического отождествления слов по словоформам и опознания жестких словосочетаний, то теперь она охватывает уже и опознание словосочетаний, свободно распределенных по тексту, по крайней мере в рамках предложения, а в некоторых случаях и вне его (ср. [20, 2, 5, 9]).

Можно предложить следующий рабочий критерий отделения тех методов АИ, которые уже сегодня могут использоваться практически для индексирования текстов второго или даже первого из выделенных выше типов, от тех более сложных и перспективных, без которых нельзя, в частности, реализовать АИ понятий и которые в настоящее время могут и должны быть предметом экспериментальных исследований.

К первым относятся методы, использующие только информацию о вхождении слов в текст или в те или иные его разделы и о взаимном расположении этих слов. Для вторых характерно использование информации о грамматической форме слова, о его принадлежности к более или менее узкому классу слов типа «часть речи», к той или иной модели изменения или управления, к семантическим классам типа «предмет», «процесс», «свойство» или более дробным — для определения синтаксической и семантической структуры связного текста (см., например, [11, с. 18—19, 42]).

При этом надо учитывать, что, во-первых, здесь нельзя провести абсолютной границы: уже сейчас вполне возможно использовать, например, такой класс слов, как «единица измерения» для

автоматизации индексирования числовых характеристик; во-вторых, что методы второго типа могут стать практически актуальными в самом близком будущем; это будет зависеть в основном от степени общности теоретических и экспериментальных проработок и от уровня используемой техники; и, в-третьих, что они уже сейчас могут быть применимы для таких специальных классов текстов третьего типа, как информационные запросы и заголовки документов.

Таким образом, задача выбора оптимального ИПЯ для ИПС с АИ есть задача на компромисс по двум линиям:

Что выгодней автоматизировать, а что оставить на человека (например, индексирование посредством КС уже сейчас вполне целесообразно автоматизировать, тогда как классификационное индексирование пока еще, как правило, выгоднее возлагать на человека).

Что выгоднее возложить на процедуру АИ (например, опознавание КС), а что — на процедуру поиска (например, опознавание понятий, выраженных сочетаниями КС).

Следует отметить, что языки индексирования текстов документов и запросов не тождественны. Только в случае простейшего унитарного индексирования можно считать, что они совпадают: в обоих случаях выражения представляют собой наборы унитарных, связанных связкой «и»; связка «или» может использоваться при поиске в виде развертывания запроса в набор подзапросов, т. е. внеязыковыми средствами. Но уже в простых дескрипторных системах «без грамматики» язык запросов, точнее, поисковых предписаний, часто использует операторы «или», «не», отсутствующие в языке ПОДов (ср., напр., [24, с. 113]). Дальнейшее усложнение языка ПОДов, например за счет введения в него грамматических средств, приводит к появлению новых специфических операторов в языке ИП и достаточно сложной и специфической грамматики для них (ср., например, [16, 17]).

Таким образом, строго говоря, всякий ИПЯ должен рассматриваться как пара подязыков: подязык запросов (ПП) и подязык документов (ПОДов).

Исходя из сказанного, на сегодняшний день и на ближайшее будущее для широкого класса информационно-поисковых ситуаций можно рекомендовать следующую концепцию ИПЯ, ориентированного на АИ текстов второго и отчасти первого типа.

Дескрипторный язык ПОДов с так называемой «слабой» (скобочной) грамматикой, учитывающей разбиение текста на фрагменты разного уровня и порядок вхождения дескрипторов в эти фрагменты, с тезаурусом дескрипторов и дескрипторосочетаний, ориентированным на использование при поиске. Процедура АИ для такого языка состоит в отождествлении слов и элементарных (неразложимых) словосочетаний, являющихся прообразами дескрипторов.

Соответствующий язык ИП должен максимально использовать возможности, предоставляемые скобочной структурой ПОТов;

в частности, помимо уже достаточно распространенных операторов «расстояние между дескрипторами» и «вхождение дескрипторов в одно предложение или параграф» (ср. операторы *N*, *SEN*, *PAR* в [16], ч. 2, с. 56)), надо использовать более общие операторы, определяющие выполнение запроса или подзапроса на выделенных и(или) динамически выделяемых комбинациях фрагментов ПОТов (ср. оператор «Релевантные фрагменты» в [17, с. 41]).

В обоснование предлагаемой концепции можно вкратце привести следующие соображения.

Принцип в основном пословного (локального) дескрипторного индексирования делает реальным и выгодным использование АИ.

Опознавание слов по словоформам и замена их дескрипторными кодами фиксированной длины в процессе АИ позволит делать это точнее, чем простым отсечением окончаний в процессе поиска (заметим, что разработчики системы PIAF DOC сочли это целесообразным даже для французского языка с его сравнительно меньшей степенью словоизменения [28]).

Принцип отождествления дескрипторосочетаний в процессе поиска делает использование связей между понятиями более гибким, чем при отождествлении словосочетаний в процессе индексирования (в частности, позволяет менять условия опознавания словосочетаний без переиндексации поискового массива).

Последовательное использование возможностей скобочной грамматики ИПЯ ПОТов вместе с тезаурусом дескрипторосочетаний обещает заметное повышение качества поиска по сравнению с действующими дескрипторными ИПС (ср. [8, 9]).

Разумеется, полное обоснование этих положений может быть дано только на базе достаточно серьезных сравнительных исследований.

Следует отметить также особую роль дескрипторных языков (в частности, со скобочной, а в перспективе и с более сильной грамматикой) как языков локального АИ в общей системе ИПЯ; по-видимому, только на их основе можно решать задачу автоматизации индексирования более глобального типа, ориентированного на выделение понятий, характеризующих документ в целом, — классификационного [21] и предметного [7].

ЛИТЕРАТУРА

1. Автоматизированная информационная система Международного центра научной и технической информации (АИС МЦНТИ-1): Опыт проектирования и эксплуатации. М.: МЦНТИ, 1979. 66 с.
2. Белоногов Г. Г., Богатырев В. И. Автоматизированные информационно-поисковые системы. М.: Сов. радио, 1973. 328 с.
3. Бернштейн Э. С. К вопросу об автоматическом индексировании. — НТИ, 1963, № 10, с. 22—25.
4. Бухалева Э. И., Космачева Н. В. Определение силы связей между дескрипторами и методика расширения поискового предписания. — НТИ. Сер. 2, 1974, № 12, с. 26—29.
5. Гиндин С. С., Леонтьева Н. Н. Задачи и общее строение системы автоматического индексирования с использованием информационного языка

- словосочетаний.— В кн.: Вопросы информационной теории и практики. М.: ВИНТИ, 1975, вып. 27, с. 88—93.
6. Гусек М. Эксперимент по автоматическому индексированию документов МСИС НИР. М.: МЦНТИ, 1978. 69 с.
 7. Добронравов И. С. Об алгоритмической предметизации книг (на примере книг по электротехнике).— В кн.: Материалы II Всесоюз. конф. «Автоматическая переработка текста методами прикладной лингвистики». Кишинев: КПИ, 1977, с. 172—174.
 8. Добронравов И. С., Лажути Д. Г., Либинзон Н. И. Структура и методика построения тезауруса дескрипторосочетаний в области электротехники.— В кн.: Материалы II Всесоюз. конф. «Автоматическая переработка текста методами прикладной лингвистики». Кишинев: КПИ, 1977, с. 143—144.
 9. Добронравов И. С., Лажути Д. Г., Малинин С. Г., Федоров Е. Б. Семантическая документальная АИПС «Скобки» с элементами грамматики и автоматическим индексированием.— В кн.: Всесоюз. конф. «Проблемы автоматизированной обработки научно-технической информации». Тез. докл. М., 1976, секция 2, ч. 1, с. 10—14.
 10. Добронравов И. С., Лажути Д. Г., Федоров Е. Б. К вопросу о стратегии поиска для АИПС с автоматизированной процедурой индексирования.— В кн.: Проблемы построения и развития алгоритмических документальных ИСП. М.: Информэлектро, 1973, вып. 1(5), с. 45—48.
 11. Иванов В. С., Пащенко Н. А. Автоматическое варьирование смысловых записей информационных запросов.— В кн.: Вопросы информационной теории и практики. М.: ВИНТИ, 1977, № 3.
 12. Китов А. И., Наумань П. Разработка алгоритма и программ индексирования научных сообщений.— В кн.: Проблемы кибернетики. М.: Наука, 1977, вып. 32, с. 119—135.
 13. Лажути Д. Г. Оценки поисковых систем: Дис. ... канд. техн. наук. М.: ВИНТИ, 1971.
 14. Лажути Д. Г., Добронравов И. С., Пархоменко В. Ф. О задаче автоматизации формирования поисковых предписаний в АИПС «Скобки».— В кн.: II Всесоюз. науч.-техн. конф. «Проблемы автоматизированной обработки научно-технической информации»: Тез. докл. М.: ВИМИ, 1976, секция 2, с. 18—19.
 15. Малинин С. Г., Лажути Д. Г. Автоматизация документальных ИПС как информационной базы современных ОАСУ: Проблемы и подход к их решению.— В кн.: Проблемы построения и развития алгоритмических документальных ИПС. М.: Информэлектро, 1973, вып. 1(5), с. 3—16.
 16. Методические материалы по применению пакета прикладных программ АСОД. М.: ВЦНТИ, 1978, ч. 1—5.
 17. Пархоменко В. Ф., Лажути Д. Г., Масковский И. Ф. Программное обеспечение документальной АИПС с грамматическим и автоматическим индексированием.— В кн.: II Всесоюз. конф. «Проблемы автоматизированной обработки научно-технической информации». Тез. докл. М.: ВИМИ, 1976, секция 3 и 4, с. 40—41.
 18. Пащенко Н. А., Иванов В. С., Кнорина Л. В., Костомарова Н. Н. Автоматический словарь в экспериментальной системе автоматического индексирования текста (ЭСАИТ).— В кн.: Вопросы информационной теории и практики. М.: ВИНТИ, 1977, № 3, с. 11—20.
 19. Певзнер Б. Р. Информационно-поисковые системы и информационно-поисковые языки. М., 1974. 50 с.
 20. Салтон Г. Автоматическая обработка, хранение и поиск информации/ Пер. с англ. М.: Сов. радио, 1973. 328 с.
 21. Тафиловская М. Я. Исследование и разработка принципов и методов автоматизации предкоординатного индексирования. Автореф. дис. ... канд. техн. наук. М.: ВИНТИ, 1976.
 22. Федоров Е. Б. Проблемы автоматизации семантических процессов обработки научно-технической литературы: Автореф. дис. ... канд. пед. наук. Л.: Ленингр. ин-т культуры им. Н. К. Крупской, 1976.
 23. Франц Ю. И., Войсунский В. Г., Франц В. И. Об оценке индексиро-

- вания и об одном способе формального составления поисковых предписаний.— НТИ. Сер. 2, 1970, № 12, с. 12—15.
24. Черный А. И. Введение в теорию информационного поиска. М.: Наука, 1975. 233 с.
 25. Meulen W. A., v. d., Janssen P. J. F. C. Automatic versus manual indexing.— Inform. Process. and Manag., 1977, vol. 13, N 1, p. 13—24.
 26. Perry J., Kent A., Berry M. Machine literature searching. N. Y., 1956.
 27. Perschke S., Fangmayer H., Fassone G., Geoffrion G. Les travaux du CETIS dans le domaine de l'informatique documentaire.— Documentaliste, 1974, Num. spéc., p. 47—57.
 28. PIAF DOC. «Bureau national de l'information scientifique et technique».— Bull. inform., 1978, N 3, p. 4—6.
 29. Salton G. The performance of interactive information retrieval. Inform. Storage and Retrieval: Rep. Text Anal., Dyn. Index., Feedback Search., Dictionary Constr. and File Organiz. Ithaca (N. Y.): 1972, p. VIII/1—VIII/18.
 30. Taube M. Storage and retrieval of information by means of the association of ideas.— Amer. Docum., 1955, vol. 6, N 1, p. 1—18.
 31. Taube M., Gull C., Wachell.— Unit terms in coordinate indexing.— Ibid., 1952, vol. 3, N 4, p. 213—218.

УДК 621.395:681.3

Я. П. Выставкин, Е. К. Гусева

ФОРМАЛЬНЫЕ ЯЗЫКОВЫЕ МОДЕЛИ ДИАЛОГОВЫХ ВЗАИМОДЕЙСТВИЙ

1. Одной из важных прикладных проблем кибернетической лингвистики является моделирование процессов информационных обменов в человеко-машинных системах обработки и распределения информации, получивших широкое применение в различных сферах человеческой деятельности. Диалоговый режим обработки информации с дистанционным доступом многих пользователей к ресурсам распределенной вычислительной системы — одна из наиболее прогрессивных форм использования ЭВМ, обеспечивающая эффективную обработку информации в реальном масштабе времени. Процедуры управления интерактивными взаимодействиями в диалоговых человеко-машинных системах определяются протоколами, представляющими собой наборы правил обмена данными между вычислительными устройствами (программами и др.) и соглашения по формату и семантике передаваемых данных [1]. Существует некоторая аналогия между использованием естественного языка в общении людей и информационными обменами в распределенных вычислительных системах.

Язык как системно-структурное образование в широком смысле представляет собой сложную многоуровневую систему, порождающую речь и выполняющую коммуникативные функции. В процессе общения между людьми происходит обмен сообщениями на естественном языке, построенными по правилам определенной грамматики. В процессе диалога между машинами происхо-

дит обмен сообщениями в соответствии с определенными протоколами, регулирующими коммуникативные функции и интерактивные взаимодействия в данной системе. Эта аналогия используется в настоящей работе для моделирования коммуникативных процессов в диалоговых человеко-машинных системах с привлечением математического аппарата теории формальных грамматик, которые определяются как системы правил, задающих множество упорядоченных элементов (цепочек символов). Множество цепочек символов, порождаемых данной грамматикой, составляют язык данной системы [2].

В ряде работ формальные языки с конечным алфавитом, тесно связанные с теорией конечных автоматов, были использованы для анализа цепочек коммутаций и передач сообщений в сетях связи [3—8]. В работе [9] формальный язык был использован для определения спецификации протокола, регулирующего коммуникационные процедуры на уровне логического канала. Формальная модель, использующая контекстно-свободную грамматику, была применена в работе [10] для определения спецификации и для верификации сетевого коммуникационного протокола.

В данной работе рассматривается формальная языковая модель коммуникационных процессов, происходящих в распределенных информационно-вычислительных системах, и показано, что множество цепочек коммутаций, составленных из индексов или номеров узлов коммутаций распределенной системы в той временной последовательности, в которой эти коммутации выполнялись в этой системе, в общем случае не является контекстно-свободным языком. В работе определены дополнительные условия, при удовлетворении которых множество цепочек коммутаций в распределенной вычислительной системе в определенных случаях представляет собой ограниченный контекстно-свободный язык. Эти результаты использованы в работе для вывода формальных условий логической правильности конструкции коммуникационного протокола. Полученные результаты иллюстрируются на примере сетевого протокола, регулирующего межпроцессные связи в сети ЭВМ.

2. Полулинейные множества чисел вхождения индексов в цепочки коммутаций. Рассмотрим множество цепочек коммутаций, составленных из индексов узлов коммутаций в той временной последовательности, в которой эти коммутации выполнялись при передаче данных в распределенной вычислительной системе. Обозначим через d^{**} максимальную длину сообщения, допускаемую протоколом, определяющим форматы сообщений в данной системе, через d^* — минимальную длину сообщения, через M — общее число узлов коммутации распределенной вычислительной системы. Далее, через ω ($\omega \in \Omega$) — цепочку, составленную из индексов T_k ($k = 1, 2, 3, \dots, M$) узлов коммутаций в порядке той временной последовательности, в которой эти коммутации выполнялись при передаче сообщений в процессе информационного обмена в данной системе.

При нормальной работе системы цепочки ω должны иметь ограниченную длину и все множество цепочек Ω ограничено. Таким образом, цепочка ω (длиной $s \geq 0$) над алфавитом T_k ($k = 1, 2, \dots, M$) есть конечная последовательность $T_1, T_2, \dots, T_k, \dots, T_M$ элементов T_k . Под алфавитом в данном случае подразумевается непустое конечное множество индексов T_k ($k = 1, 2, \dots, M$).

Обозначим через (T_k^*, T_k^{**}) ($0 \leq T_k^* \leq T_k^{**}$) допустимый интервал времени, в течение которого при наличии подлежащих передаче сообщений в k -м узле должен наступить момент коммутации и передачи сообщения. Таким образом, при наличии очереди сообщений, подлежащих передаче по определенному выходному каналу k -го узла распределенной вычислительной системы, момент коммутации и начало передачи данного сообщения должны произойти не раньше некоторого времени T_k^* , отсчитываемого с момента окончания передачи по этому же каналу предыдущего сообщения, и не позже момента времени T_k^{**} , равного максимальному времени передачи сообщения длиной d^{**} из k -го узла в другой соседний узел данной распределенной вычислительной системы. Если специально не будет оговорено, то везде в рассматриваемых распределенных вычислительных системах отсчеты времени коммутаций и передач сообщений выполняются в каждой системе по единому системному времени. Обозначим через $H_{T_k}(\omega)$ число вхождений T_k в цепочку ω , число вхождений различных узлов коммутаций в цепочку ω — через

$$r(\omega) = \{H_{T_1}(\omega), H_{T_2}(\omega), \dots, H_{T_M}(\omega)\}.$$

Множество различных комбинаций чисел вхождения T_k ($k = 1, 2, \dots, M$) в цепочки ω ($\omega \in \Omega$) за период эксплуатации системы в течение достаточно длительного времени обозначим через $R = \{r(\omega) (\omega \in \Omega)\}$. Следующая теорема утверждает, что в распределенных вычислительных системах множество $R = \{r(\omega) (\omega \in \Omega)\}$ не является полулинейным множеством.

Теорема 1. В распределенной вычислительной системе, состоящей из ограниченного числа узлов более двух или из двух узлов, соединенных каналами передачи данных, работающими в дуплексном режиме, если величины T_k^* , T_k^{**} и моменты времени коммутаций выражаются действительными числами, то $R = \{r(\omega) (\omega \in \Omega)\}$ не является полулинейным множеством.

Доказательство. Рассмотрим при $M > 3$ любую пару k -го и c -го ($k \neq c$; $k = 1, 2, \dots, M$; $c = 1, 2, \dots, M$) несоседних узлов системы. При $M = 3$ в случае соединения узлов в цепочку будем рассматривать вначале крайние узлы этой цепочки. При $M = 3$ в случае кольцевой схемы соединения узлов с односторонним направлением потока сообщений по кольцу и возможностью в любом узле производить одновременно прием сообщения из предыдущего узла и передачу сообщения в последующий узел по кольцу и в случае кольцевой схемы передачи сообщений в обоих направлениях по кольцу будем рассматривать любые два узла коль-

цевой системы. В случае $M = 2$ с дуплексной системой передачи данных будем рассматривать эти два узла.

Пусть для определенности приняты следующие неравенства:

$$T_c^* < T_k^*; \quad T_c^{**} < T_k^{**},$$

при $M = 2$ $k = 1$; $c = 2$. Пусть величины T_c^* и T_k^{**} выбраны или заданы так, что отношение T_c^*/T_k^{**} не является рациональным числом. Рассмотрим точечное множество B , элементами которого являются величины $H_{T_k}(\omega)/H_{T_c}(\omega)$, представляющие собой отношения чисел вхождений $H_{T_k}(\omega)$, $H_{T_c}(\omega)$ величин T_k и T_c во все цепочки ω ($\omega \in \Omega$), начиная с некоторого момента времени, принимаемого за начало работы данной системы обработки и распределения информации. Среди точек этого множества минимальная соответствует величине T_c^*/T_k^{**} . Эта величина в соответствии с сделанным выше предположением не является рациональной.

Предположим, что множество $R' = \{r'(\omega) \mid \omega \in \Omega\}$, где $r'(\omega) = \{H_{T_k}(\omega), H_{T_c}(\omega)\}$, является полулинейным множеством;

тогда R' является объединением $\bigcup_{b=1}^u A_b$ конечного числа линейных множеств A_b [2]. Причем в области неотрицательных целых чисел существуют двумерные векторы

$$(L_k^b, L_c^b), (P_{1k}^b, P_{1c}^b), \dots, (P_{m_b k}^b, P_{m_b c}^b) \in N^2$$

$$A_b = \{(L_k^b, L_c^b) + Z_1^b(P_{1k}^b, P_{1c}^b) + \dots + Z_{m_b}^b(P_{m_b k}^b, P_{m_b c}^b)\}$$

$$Z_1^b, Z_2^b, \dots, Z_{m_b}^b \in N,$$

где N — неограниченное множество неотрицательных целых чисел; N^2 — декартово произведение ($N^2 = N \times N$).

Рассмотрим точечное множество, элементами которого являются величины $H_{T_k}(\omega)/H_{T_c}(\omega)$, т. е. отношения чисел вхождений T_k и T_c во все цепочки ω ($\omega \in \Omega$). Среди точек этого множества минимальная определяется минимальной величиной P_{xk}^b/P_{xc}^b отношения компонент всех векторов (P_{xk}^b, P_{xc}^b) ($1 \leq b \leq u$), ($1 \leq x \leq m_b$). Это число является рациональным и противоречит сделанному вначале предположению, что среди точек рассматриваемого множества минимальная соответствует величине T_c^*/T_k^{**} , не являющейся рациональным числом. Следовательно, R' не является полулинейным множеством. В случае $M = 2$ когда $k = 1$, $c = 2$, доказательство теоремы закончено. В случае $M > 2$ аналогичные вышеприведенному доказательству можно привести для всевозможных других пар несоседних узлов, т. е. соединенных между собой транзитно, через другие узлы, или других пар узлов, из которых хотя бы один соединен линией связи с одним каким-либо другим узлом системы. Поскольку число узлов и число линий связи между ними предполагается ограниченным, общее число вышеуказанных пар узлов ограничено. Следовательно,

множество R может быть определено как объединение конечного числа ненулевых подмножеств R' . Из этого следует, что множество R не является полулинейным. (Конец доказательства.)

3. Использование ограниченных контекстно-свободных грамматик для описания коммуникационных процессов. Из теоремы 1 следует, что множество цепочек коммутаций сообщений, передаваемых по каналам связи в распределенных вычислительных системах, не является контекстно-свободным языком и в общем случае трудно решать задачи проектирования и верификации коммуникационных протоколов, применяя математический аппарат теории формальных грамматик непосредственно ко всему множеству цепочек коммутаций, происходящих во всех узлах распределенной вычислительной системы. В связи с этим рассмотрим коммуникационный процесс между двумя дистанционно-удаленными вычислительными системами, соединенными между собой сетью связи, обслуживающей информационные обмены между этими системами.

Наборы программных модулей, выполняющих в соответствующих вычислительных системах операции управления и обработки информации при осуществлении коммуникативных функций приема и передачи сообщений между вычислительными системами, будем называть подпроцессами \hat{P}_{kc} , где k — индекс вычислительной системы ($k = 1, 2$); c — индекс иерархического уровня данного подпроцесса в иерархической структуре всех подпроцессов, выполняющих функции управления связью, обработки информации и передачи сообщений ($c = 1, 2, \dots, C$). Иерархическая структура организации подпроцессов \hat{P}_{kc} объясняется тем, что в больших человеко-машинных системах, состоящих из множества узловых вычислительных систем и терминальных комплексов, которые могут быть неоднородными по программному обеспечению и по используемым наборам кодовых знаков с различными значениями буквенных символов, в качестве базы организации коммуникационных процессов и диалоговых взаимодействий должен служить общий для всей человеко-машинной системы код, определенный на самом верхнем уровне иерархии подпроцессов \hat{P}_{k0} ($k = 1, 2$). Прием и передача сообщений на следующем уровне иерархии подпроцессов \hat{P}_{k1} ($k = 1, 2$) осуществляется кодовой системой, определенной в результате взаимодействия подпроцессов \hat{P}_{k0} ($k = 1, 2$). В результате взаимодействия подпроцессов \hat{P}_{11} и \hat{P}_{21} определяется кодовая система и коммуникативные функции, которые должны быть выполнены на следующем уровне иерархии подпроцессов \hat{P}_{12} и \hat{P}_{22} . Аналогично взаимодействие между подпроцессами на $(c - 1)$ -м уровне иерархии подпроцессов определяет семантику и форму представления информационных данных, а также все детали коммуникативных функций, выполняемых на следующем, c -м уровне иерархии подпроцессов. Например, при

пересылке файла данных между вычислительными системами, расположенными в дистанционно-удаленных узлах распределенной вычислительной системы, на самом верхнем уровне иерархии подпроцессов инициируются процедуры выбора из множества прикладных протоколов, используемых в данной системе, необходимого протокола, обслуживающего передачу файла данных.

На следующем уровне иерархии подпроцессов определяется кодовая система и структура передаваемого файла, а также те коммуникативные функции, которые необходимы для его передачи. Конкретные операции пересылки файла данных выполняются на следующем уровне иерархии подпроцессов в соответствии с теми определениями, которые были сделаны на предыдущем уровне. Таким образом, при осуществлении коммуникационного процесса информационные обмены сообщениями могут происходить только между подпроцессами одного и того же иерархического уровня. В соответствии с этим процесс взаимодействия двух дистанционно-удаленных вычислительных процессов в двух территориально удаленных друг от друга узлах распределенной вычислительной системы реализуется на каждом уровне иерархии подпроцессов \hat{P}_{kc} множеством цепочек ω_c , составленных из индексов T_{kc} ($k = 1, 2$) в той временной последовательности, в которой эти коммутации выполнялись при передаче соответствующих сообщений между подпроцессами \hat{P}_{kc} ($k = 1, 2$). При проектировании и верификации протоколов, регулирующих вышеуказанные коммуникационные процедуры и межпроцессные связи, для эффективного решения этих задач необходимо знать, к какому семейству формальных языков относятся указанные множества цепочек ω_c на каждом c -м уровне иерархии взаимодействующих подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 1, 2, \dots, C$). Ответ на этот вопрос дает следующая теорема.

Теорема 2. Если коммуникационный процесс на каждом c -м уровне иерархии подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 1, 2, \dots, C$) реализуется на ограниченном множестве цепочек $\omega_c = \{T_{kc}^j\}$, причем число j и число $H_{T_{kc}}$ вхождений символов T_{kc} ($k = 1, 2$) в цепочки ω_c ограничены сверху и выражаются целыми неотрицательными числами, и если удовлетворяются условия

$$T_{1c}^* \leq T_{1c} \leq T_{1c}^{**}, \quad (I)$$

$$T_{2c}^* \leq T_{2c} \leq T_{2c}^{**}, \quad (II)$$

где T_{kc}^* и T_{kc}^{**} ($k = 1, 2; c = 1, 2, \dots, C$) — неотрицательные рациональные числа, то множество $R_c = \{r_c(\omega_c) (\omega_c \in \Omega)\}$ является полулинейным.

Доказательство. Предположим, что множество $R_c = \{r_c(\omega_c) (\omega_c \in \Omega)\}$ не является полулинейным. Тогда $\bigcup_{b=1}^U A_{bc}$

$$\bigcup_{b=1}^U A_{bc} = \bigcup_{b=1}^U (X_{0c}^b + Z_{1c}^c X_{1c}^b + \dots + Z_{m_c c}^b X_{m_c c}^b + \dots),$$

где X_{ic}^b ($i = 0, 1, 2, \dots, m_{bc}, \dots$) — двумерные векторы с компонентами из различных комбинаций чисел T_{kc}^* , T_{kc}^{**} ; Z_{mc}^b ($m = 1, 2, \dots, m_b, \dots$) — неотрицательные целые числа, представляет собой объединение бесконечного числа множеств A_{mc}^b , которые не являются линейными. Это противоречит тому, что множество различных возможных комбинаций $r_c(\omega_c)$ чисел $H_{T_{kc}}$ вхождений T_{kc} в различные цепочки ω_c является конечным, поскольку числа j и $H_{T_{kc}}$ ограничены сверху и являются неотрицательными целыми числами и соблюдаются условия (I), (II). Следовательно, множество $R_c = \{r_c(\omega_c) (\omega_c \in \Omega)\}$ является полулинейным. (Конец доказательства.) Из теоремы 2 следует, что множество цепочек ω_c является ограниченным контекстно-свободным языком, поскольку $R_c = \{r_c(\omega_c) (\omega_c \in \Omega)\}$ является полулинейным множеством и каждой цепочке ω_c сопоставлено в R_c упорядоченное выражение, соответствующий элемент которого представляет собой число вхождений T_{kc} в цепочку ω_c [2].

4. Верификация логической правильности коммуникационного протокола. Теория ограниченных контекстно-свободных языков хорошо развита, и все известные и представляющие практический интерес алгоритмические проблемы, связанные с этими языками, разрешимы. Это может быть использовано для решения практических задач проектирования и верификации протоколов. С увеличением числа различного рода диалоговых взаимодействий и информационных обменов в распределенных вычислительных системах становится все более сложным предусмотреть и определить в протоколе все возможные варианты и случаи передач и приемов различных сообщений в различных ситуациях, возникающих при выполнении коммуникационного процесса. Например, протокол X 21 МТТККТ, несмотря на большие усилия, затраченные на его проектирование, содержит логические ошибки [11]. Конструкция протокола является логически полной, если в протоколе однозначно определены все необходимые приемы и передачи сообщений так, что не допускается никаких безусловных протоколом сообщений при работе системы в нормальных условиях. Под нормальными подразумеваются условия работы системы без потерь, изменений и безусловных протоколом дублирования сообщений при обработке их в порядке поступления.

Некоторые задачи верификации коммуникационных протоколов, включая задачи определения достижимости соответствующих состояний коммуникационного процесса, начиная с исходного или с некоторого промежуточного состояния этого процесса, и задачи определения правильности логической структуры коммуникационного протокола, являются алгоритмически разрешимыми на основе построения моделей коммуникационных процессов с использованием теории ограниченных контекстно-свободных языков и с учетом вероятностных свойств моделируемых коммуникационных процессов. Дополним контекстно-свободную грам-

матику $G_c(V_T, V_H, P, \sigma)$, где $V_T = \{T_{1c}, T_{2c}\}$, V_H — множество нетерминальных символов, соответствующих возможным вариантам цепочек ω_c ; P — множество правил левостороннего вывода цепочек ω_c ; σ — начальный нетерминальный символ, отображением f из P на отрезок действительной оси $(0, 1)$. Сумма всех $f(P_i)$, $P_i \in P$, имеющих одну и ту же левую часть, равна 1, т. е. f задает распределение вероятностей на множестве всех выводов, имеющих одну и ту же левую часть. Обозначим через B всю совокупность процессов левосторонних выводов с помощью правил грамматики G_c , начиная с нетерминального символа σ . Всю совокупность цепочек ω_c , выведенных с помощью правил грамматики G_c с использованием левосторонних выводов, входящих в множество B , обозначим через E .

Рассмотрим марковскую цепь, имеющую стационарные во времени вероятности переходов, пространством состояний которой является счетное множество D , матрицей вероятностей переходов — матрица Y , удовлетворяющая указанным ниже условиям (III), (IV), и начальное распределение которой выражается последовательностью действительных чисел $(a_i)_{i=1}^{\infty}$, где через a_i обозначена вероятность того, что некоторая начальная случайная величина равна σ_i . Если $G_c(V_T, V_H, P, \sigma)$ — контекстно-свободная грамматика, не имеющая правил вывода типа $A \rightarrow \sigma$, $A \in V_H$, и если f — отображение из P на отрезок $(0, 1)$ действительной оси, задающее распределение вероятностей на множестве всех выводов, имеющих одну и ту же левую часть, причем сумма всех $f(P_i)$, $P_i \in P$, имеющих одну и ту же левую часть, равна 1, то марковская цепь, соответствующая грамматике G_c и отображению f , выражается тройкой $(D, Y, (a_i)_{i=1}^{\infty})$, удовлетворяющих условиям:

1) $Y_{ij} \geq 0$ ($i, j \in N$, N — все множество целых чисел); (III)

Y_{ij} — элемент матрицы переходных вероятностей Y , означающий условную вероятность перехода из состояния σ_i в состояние σ_j , $\sigma_i, \sigma_j \in D$, i, j — неотрицательные целые числа;

2) $\sum_{j=1}^{\infty} Y_{ij} = 1$ ($i \in N$); (IV)

3) $\hat{D}U\{\sigma\}$; (V)

4) элементы Y_{ij} матрицы переходных вероятностей Y имеют следующие свойства для любых состояний $\sigma_i, \sigma_j \in D$:

а) при $\sigma_i = \sigma$, если σ_j соответствует такому правилу вывода $P^* \in P$, которое содержит в левой части σ , то $Y_{ij} = f(P^*)$, в остальных случаях

$Y_{ij} = 0$; (VI)

б) если $\sigma_i \in \hat{B} - E$, то если существует применимое к σ_i правило вывода $P^* \in P$ и если $\sigma_j = \sigma_i P^*$, то $Y_{ij} = f(P^*)$, если $\sigma_j = P^*$, то $Y_{ij} = f(P^*) = 1$, в остальных случаях

$Y_{ij} = 0$; (VII)

в) если σ_i ($\sigma_i \in E$) содержит на правом конце цепочки, соответствующей этому состоянию, правило вывода $P^* \in P$, которое имеет в левой части σ , то, если $\sigma_j = \sigma$, $Y_{ij} = 1$; в остальных случаях

$Y_{ij} = 0$; (VIII)

г) если $\sigma_i \in \hat{B} - E$ и $\sigma_i = P^* P^{(1)} P^{(2)} \dots P^{(n)}$ ($P^*, P^{(1)}, \dots, P^{(n)} \in P$) и если $\sigma_j = P^*$, то $Y_{ij} = 1$, в остальных случаях

$Y_{ij} = 0$; (IX)

5) $(a_i)_{i=1}^{\infty}$ — последовательность действительных чисел, удовлетворяющих условию

$\sum_{i=1}^{\infty} a_i = 1$. (X)

Кроме того, предполагается, что марковская цепь, соответствующая грамматике G_c и отображению f , удовлетворяет следующим условиям, отражающим древовидную структуру процесса порождения цепочек ω_c с помощью формальной грамматики $G_c(V_T, V_H, P, \sigma)$ и отображения f :

число состояний, достижимых из некоторого состояния за один шаг марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ ограничено; (XI)

состояния, достижимые за один шаг марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$, определяются однозначно характеристиками коммуникационного процесса и правилами вывода цепочек ω_c ; (XII)

число траекторий достижения исхода из начального состояния σ некоторого другого состояния марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ через последовательность промежуточных состояний ограничено, и каждая траектория определяется однозначно характеристиками коммуникационного процесса и правилами выбора цепочек ω_c . (XIII)

Если $\sigma_i, \sigma_j \in D$, то достижимость состояния σ_j из состояния σ_i и соответственно с этим логическая полнота и правильность структуры коммуникационного протокола означают, что либо $\sigma_i = \sigma_j$, либо существуют такие состояния коммуникационного процесса и соответствующей марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ $\sigma_{j_1}, \sigma_{j_2}, \dots, \sigma_{j_n} \in D$ и такие вероятности перехода $Y_{i,j_1}, Y_{j_1,j_2}, \dots, Y_{j_n,j} > 0$, что при удовлетворении условий (XI), (XII), (XIII) и нормальной работы системы без потерь, изменений и не обусловленных протоколом дублирования сообщений при их обслуживании в порядке поступления обеспечивается переход $\sigma_i \rightarrow \sigma_j$.

Условия достижимости любого состояния коммуникационного процесса из другого состояния, в частности из начального, и соответственно с этим логическая правильность структуры протокола, регулирующего этот коммуникационный процесс, выражаются следующей теоремой.

Теорема 3. При работе системы в нормальных условиях и удовлетворении условий (XI), (XII), (XIII) любое состояние комму-

никационного процесса между подпроцессами \hat{P}_{c1} и \hat{P}_{c2} на c -м уровне иерархии ($c = 1, 2, \dots, C$) подпроцессов, выполняющих функции управления связью и передачи сообщений, достижимо из любого другого состояния коммуникационного процесса на том же уровне иерархии подпроцессов \hat{P}_{c1} и \hat{P}_{c2} , и логическая структура коммуникационного протокола, регулирующего функции связи и передачи сообщений на том же c -м ($c = 1, 2, \dots, C$) уровне иерархии подпроцессов \hat{P}_{c1} и \hat{P}_{c2} сконструирована правильно тогда и только тогда, когда марковская цепь $(D, Y, (a_i)_{i=1}^{\infty})$, соответствующая формальной грамматике $G_c(V_T, V_H, P, \sigma)$ и отображению f , порождающих цепочки ω_c данного коммуникационного процесса, является неприводимой.

Доказательство. Для доказательства теоремы вначале покажем, что любое состояние, например начальное состояние σ , достижимо из любого другого состояния σ_i марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$. Если $\sigma_i = \sigma$, то из этого сразу следует $\sigma_i \rightarrow \sigma$. Если $\sigma_i \in E$, то в соответствии со свойствами (VI)–(IX) матрицы Y и в соответствии с условиями (XI), (XII), (XIII) σ_i содержит на правом конце цепочки, соответствующей этому состоянию, правило вывода $P^* \in P$, имеющее в левой части нетерминальный символ σ , поэтому $Y_{ij} = 1$ ($\sigma_j = \sigma$) и, следовательно, $\sigma_i \rightarrow \sigma$. Если $\sigma_i \in \hat{B} - E$, то в соответствии со свойствами матрицы Y и при удовлетворении условий (XI)–(XIII) существует процесс вывода $\sigma_j = \sigma_i P_1 P_2 \dots P_n$ ($P_1 P_2 \dots P_n \in P$). Если $\sigma_{j1} = \sigma_i P_1$, $\sigma_{j2} = \sigma_i P_1 P_2$, \dots , $\sigma_{j(n-1)} = \sigma_i P_1 P_2 \dots P_{n-1}$, то $Y_{i,j1} = f(P_1)$, $Y_{j1,j2} = f(P_2)$, \dots , $Y_{j(n-1),j} = f(P_n)$. По определению f , $Y_{i,j1}$, $Y_{j1,j2}$, \dots , $Y_{j(n-1),j} > 0$; следовательно, $\sigma_i \rightarrow \sigma_j$. σ_j соответствует некоторый процесс вывода, как было показано выше, в силу свойств матрицы Y и условий (XI)–(XIII) $\sigma_j \rightarrow \sigma$. Учитывая это и используя свойство транзитивности двуместного отношения \rightarrow , получим $\sigma_i \rightarrow \sigma$.

Теперь покажем достижимость любого состояния σ_j из начального состояния σ . Если $\sigma_j \in E$ или $\sigma_j \in \hat{B} - E$, то в соответствии со свойствами матрицы Y и условиями (XI)–(XIII) существует процесс вывода $\sigma_j = \sigma P_1 P_2 \dots P_n$. Если $\sigma_{j1} = \sigma P_1$, $\sigma_{j2} = \sigma P_1 P_2$, \dots , $\sigma_{j(n-1)} = \sigma P_1 P_2 \dots P_{n-1}$, то $Y_{\sigma,j1} = f(P_1)$, $Y_{j1,j2} = f(P_2)$, \dots , $Y_{j(n-1),j} = f(P_n)$. По определению f , $Y_{\sigma,j1}$, $Y_{j1,j2}$, \dots , $Y_{j(n-1),j} > 0$; откуда $\sigma \rightarrow \sigma_j$. Если $\sigma_j = \sigma$, то тогда $\sigma \rightarrow \sigma_j$.

Из вышеизложенного следует, что начальное состояние σ достижимо из любого другого состояния марковской цепи и любое состояние цепи достижимо из начального состояния σ . Поскольку каждое состояние марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ достижимо из другого любого состояния цепи, марковская цепь $(D, Y, (a_i)_{i=1}^{\infty})$ является неприводимой. Из соответствия марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ формальной грамматике $G_c(V_T, V_H, P, \sigma)$ и отображению f , порождающих цепочки ω_c данного коммуникационного процесса, и из свойства неприводимости марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$

следует, что каждое состояние данного коммуникационного процесса на соответствующем уровне иерархии подпроцессов \hat{P}_{c1} , \hat{P}_{c2} достижимо из любого другого состояния, и в частности из начального на том же уровне иерархии подпроцессов \hat{P}_{c1} , \hat{P}_{c2} , и в соответствии с этим логическая структура протокола, выполняющего функции управления связью и передачей сообщений на c -м ($c = 1, 2, \dots, C$) уровне иерархии подпроцессов \hat{P}_{c1} , \hat{P}_{c2} , сконструирована правильно. (Конец доказательства.)

Следствие 1. Логическая правильность конструкции коммуникационного протокола на каждом c -м уровне иерархии подпроцессов \hat{P}_{c1} , \hat{P}_{c2} ($c = 1, 2, \dots, C$) должна удовлетворять следующим формальным условиям.

1. В формальной грамматике $G_c(V_T, V_H, P, \sigma)$, порождающей цепочки ω_c коммуникационного процесса, начальное состояние σ обоих взаимодействующих подпроцессов \hat{P}_{c1} , \hat{P}_{c2} должно быть одно и то же. Это означает, что при логически правильной конструкции протокола в начальный момент взаимодействия на каждом c -м уровне иерархии подпроцессов оба подпроцесса \hat{P}_{c1} и \hat{P}_{c2} должны находиться в одном и том же состоянии, т. е. после окончания взаимодействия между подпроцессами на $(c - 1)$ -м уровне иерархии должен быть обеспечен переход в начальное состояние σ . Это обеспечивается тем, что в соответствии со свойствами (VIII) матрицы Y переходных вероятностей конечное состояние на данном уровне иерархии подпроцессов содержит на правом конце цепочки, соответствующей этому состоянию, правило вывода, имеющее в левой части σ , поэтому с вероятностью, равной 1, происходит переход из конечного состояния на данном уровне взаимодействия между соответствующими подпроцессами в начальное состояние σ и обеспечивается синхронизация начального состояния взаимодействия на следующем уровне иерархии подпроцессов \hat{P}_{c1} , \hat{P}_{c2} .

2. В правилах вывода P формальной грамматики $G_c(V_T, V_H, P, \sigma)$ в правой части каждого правила $P^* \in P$ может быть только один терминальный символ, принадлежащий множеству V_T , соответствующий одному принимаемому или передаваемому сообщению, т. е. взаимодействующие подпроцессы \hat{P}_{c1} и \hat{P}_{c2} не могут быть одновременно в двух состояниях: передачи собственного сообщения и приема сообщения от другого подпроцесса. В противном случае при наложении этих состояний возникает коллизия событий и нарушается логика управления связью и информационными обменами.

3. Логика построения правил вывода P формальной грамматики $G_c(V_T, V_H, P, \sigma)$ и отображения f на отрезок $(0, 1)$ действительной оси должны соответствовать заданной системе управления потоками в межпроцессных обменах таким образом, чтобы обеспечить требуемую определенную последовательность приемов и передач сообщений между взаимодействующими подпроцессами

и определенную последовательность передач двух или большего числа сообщений от одного подпроцесса другому на данном уровне иерархии взаимодействия подпроцессов.

Для иллюстрации применения полученных результатов рассмотрим следующий пример процедур установления межпроцессных связей между двумя дистанционно-удаленными вычислительными процессами, выполняемыми в двух главных ЭВМ, расположенных в узлах распределенной вычислительной сети. Пусть вначале из первого узла, инициирующего соответствующую межпроцессную связь с вычислительным процессом в главной ЭВМ в другом узле, высылается во второй в момент времени T_1 управляющее сообщение $RTS^{(1)}$, указывающее номер транспортного порта, номер канала, буферные емкости и другие идентификаторы инициируемого процесса. Если в период времени T_2 будет получено из узла 2 уведомление об ошибке $NAC^{(2)}$, то первое сообщение из узла 1 будет повторено; если уведомление о приеме будет положительное $AC^{(2)}$, то в соответствии с установленной процедурой из узла 2 в узел 1 высылается в период времени T_2 ответное управляющее сообщение $STR^{(2)}$ с идентификаторами инициированного процесса в главной ЭВМ узла 2. При конструировании формальной грамматики $G_c(V_T, V_H, P, \sigma)$, порождающей цепочки ω , составленные из индексов T_1 и T_2 в той временной последовательности, в которой соответствующие сообщения передавались из первого и второго узлов, $V_T = \{T_1, T_2\}$, правила вывода цепочек ω , соответствующих событиям передачи из узла 1 сообщений $RTS^{(1)}, NAC^{(1)}, AC^{(1)}$ в узел 2 и сообщений $STR^{(2)}, NAC^{(2)}, AC^{(2)}$ из узла 2 в узел 1, можно представить следующим образом:

- $P_1: \sigma \rightarrow T_1 K_2$ (передача сообщения $RTS^{(1)}$);
- $P_2: K_2 \rightarrow T_2 K_3$ (передача уведомления $AC^{(2)}$);
- $P_3: K_2 \rightarrow T_2 \sigma$ (передача уведомления $NAC^{(2)}$);
- $P_4: K_3 \rightarrow T_2 K_9$ (передача сообщения $STR^{(2)}$);
- $P_5: K_9 \rightarrow T_1 O_2$ (передача уведомления $AC^{(1)}$);
- $P_6: K_9 \rightarrow T_1 K_3$ (передача уведомления $NAC^{(1)}$).

После получения узлом 2 уведомления $AC^{(1)}$ из узла 1 о приеме сообщения $STR^{(2)}$, из узла 2 передается в узел 1 сообщение R_1^* , указывающее номер транспортного принимающего порта и другие идентификаторы, определяющие спецификацию приема сообщений из узла 1 в узел 2 на следующем уровне иерархии взаимодействующих коммуникационных подпроцессов в узлах 1 и 2. При получении отрицательного уведомления $NAC^{(1)}$ узел 2 повторяет передачу сообщения R_1^* . Для упрощения рассматриваемого иллюстративного примера сократим возможные варианты информационных обменов управляющими сообщениями между узлами 1 и 2, связанными с различными случаями системных отказов и с соответствующей обработкой такого рода аварийных ситуаций в

этих узлах в сети связи, а также связанными с различными логическими ошибками в сообщениях, с искажениями принятых сообщений в узле-адресате, с обработкой ошибок, совершаемых коммуникационными подпроцессами и т. д. Учет такого рода различных случаев и факторов при конструировании формальной грамматики $G_c(V_T, V_H, P, \sigma)$ приводит к увеличению числа элементарных правил вывода $P^* \in P$, к усложнению логики правил вывода цепочек ω , и к увеличению числа нетерминальных и терминальных символов формальной грамматики G_c . При этом существо данного метода моделирования коммуникационного процесса и верификации коммуникационного протокола меняться не будет, поэтому в рассматриваемом примере будем считать, что после нормального приема в узле 1 сообщения R_1^* из узла 2 и послышки соответствующего уведомления $AC^{(1)}$ из узла 1 в узел 2, из узла 1 высылается в узел 2 управляющее сообщение $R-CLS^{(1)}$ о выполнении процедур по разъединению межпроцессной связи на данном уровне иерархии взаимодействующих коммуникационных подпроцессов. После послышки в узел 1 из узла 2 соответствующего уведомления $AC^{(2)}$ или $NAC^{(2)}$ о корректном или ошибочном приеме сообщения $R-CLS^{(1)}$ в первом случае узел 2 высылает в узел 1 сообщение $R-CLS^{(2)}$ о выполнении процедур по разъединению связи на данном уровне иерархии взаимодействующих подпроцессов; во втором случае — при ошибочном приеме сообщения $R-CLS^{(1)}$ — узел 1 повторяет передачу сообщения $R-CLS^{(1)}$. После получения узлом 2 уведомления $AC^{(1)}$ о корректном приеме сообщения $R-CLS^{(2)}$ взаимодействие между коммуникационными подпроцессами на данном уровне иерархии подпроцессов прекращается, коммуникационный процесс возвращается в начальное состояние и начинается взаимодействие между подпроцессами на следующем, более низком уровне иерархии подпроцессов. Логика этих взаимодействий можно выразить следующим образом, используя аппарат формальной грамматики $G_c(V_T, V_H, P, \sigma)$ и соответствующей марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$:

- $P_7: O_2 \rightarrow T_2 K_{11}$ (передача сообщения R_1^*);
- $P_8: K_{11} \rightarrow T_1 O_2$ (передача уведомления $NAC^{(1)}$);
- $P_9: K_{11} \rightarrow T_1 O_4$ (передача уведомления $AC^{(1)}$);
- $P_{10}: O_4 \rightarrow T_1 C_1$ (передача сообщения $R-CLS^{(1)}$);
- $P_{11}: C_1 \rightarrow T_2 O_4$ (передача уведомления $NAC^{(2)}$);
- $P_{12}: C_1 \rightarrow T_2 C_2$ (передача уведомления $AC^{(2)}$);
- $P_{13}: C_2 \rightarrow T_2 K_{13}$ (передача сообщения $R-CLS^{(2)}$);
- $P_{14}: K_{13} \rightarrow T_1 C_2$ (передача уведомления $NAC^{(1)}$);
- $P_{15}: K_{13} \rightarrow T_1 \sigma$ (передача уведомления $AS^{(1)}$).

Диаграмма перехода состояний марковской цепи $(D, Y, a_i)_{i=1}^{\infty}$, построенная в соответствии с вышеописанной логикой порожде-

ния цепочек ω_c формальной грамматикой $G_c(V_T, V_N, P, \sigma)$, представлена на рис. 1.

Соответствующая матрица вероятностей перехода имеет следующий вид:

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P_1	0	0	$f(P_2)$	$f(P_3)$	0	0	0	0	0	0	0	0	0	0	0
P_2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P_3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P_4	0	0	0	0	0	$f(P_6)$	$f(P_6)$	0	0	0	0	0	0	0	0
P_5	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
P_6	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P_7	0	0	0	0	0	0	0	$f(P_8)$	$f(P_9)$	0	0	0	0	0	0
P_8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
P_9	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
P_{10}	0	0	0	0	0	0	0	0	0	0	$f(P_{11})$	$f(P_{12})$	0	0	0
P_{11}	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
P_{12}	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P_{13}	0	0	0	0	0	0	0	0	0	0	0	0	$f(P_{14})$	$f(P_{15})$	0
P_{14}	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P_{15}	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

В матрице переходных вероятностей Y условные вероятности $f(P_i)$ определяются исходя из измерений статистики потоков или аналитическими методами на основе имитационных моделей, численного решения уравнений, описывающих рассматриваемый процесс. Как видно из диаграммы перехода состояний (рис. 1), при ненулевых значениях переходных вероятностей $f(P_i)$ в матрице Y построенная марковская цепь является неприводимой, поэтому в соответствии с теоремой 3 логическая структура процедур управления связью и информационными обмена на данном уровне иерархии взаимодействующих подпроцессов удовлетворяет условиям (1), (2), (3) логически правильного конструирования коммуникационного протокола.

При детальном учете всевозможных ситуаций, возникающих в процессе взаимодействия между соответствующими подпроцессами, увеличивается число состояний марковской цепи $(D, Y, (a_i)_{i=1}^{\infty})$ и усложняется структура формальной грамматики $G_c(V_T, V_N, P, \sigma)$ и логика правил вывода P цепочек ω_c , поэтому в этих случаях, а также при большом числе уровней в иерархической структуре взаимодействующих коммуникационных подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 1, 2, \dots, C$) применение данного метода верификации коммуникационного протокола необходимо осуществлять с использованием машинных методов проектирования распределенных вычислительных систем.

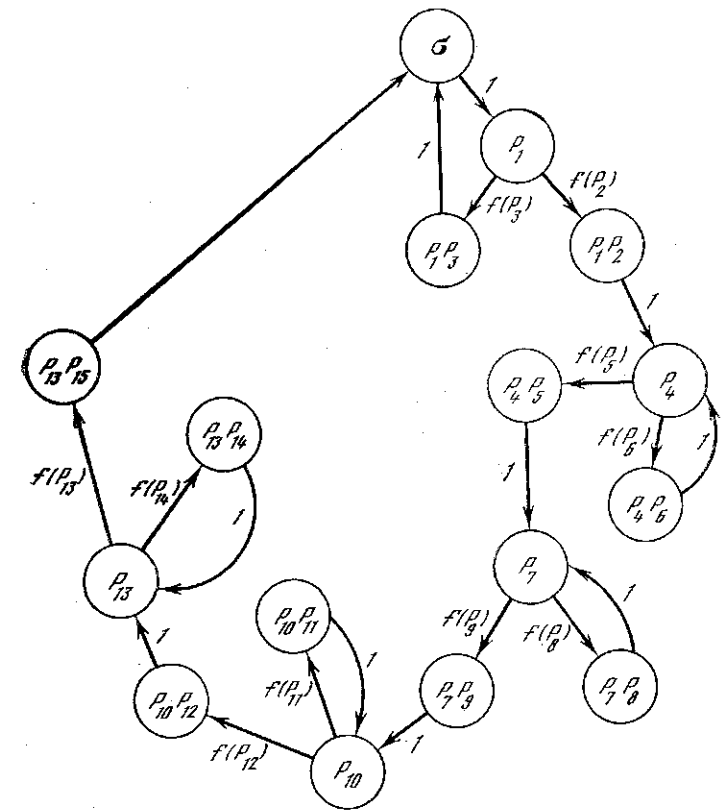


РИС. 1. Диаграмма перехода состояний марковской цепи

5. Интерактивный язык межуровневых взаимодействий коммуникационного процесса. В ходе выполнения коммуникационного процесса между дистанционно-удаленными вычислительными процессами или процессами обработки информации в территориально распределенных вычислительных системах, соединенных между собой каналами сети связи, вначале производится иницирование процессов в соответствующих вычислительных системах, затем наступает фаза установления соединений и выполнения информационных обменов на различных уровнях иерархии подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 1, 2, \dots, C$); после этого следует фаза выполнения операции разъединения связи на соответствующих уровнях подпроцессов \hat{P}_{kc} . Обозначим через G_{dc} ($c = 0, 1, 2, \dots, C$) формальную грамматику, порождающую во время диалога между подпроцессами P_{kc} ($k = 1, 2; c = 1, 2, \dots, C$) на c -м уровне иерархии подпроцессов цепочки слов, определяющих коммуникативные функции, форматы и семантику данных, которые

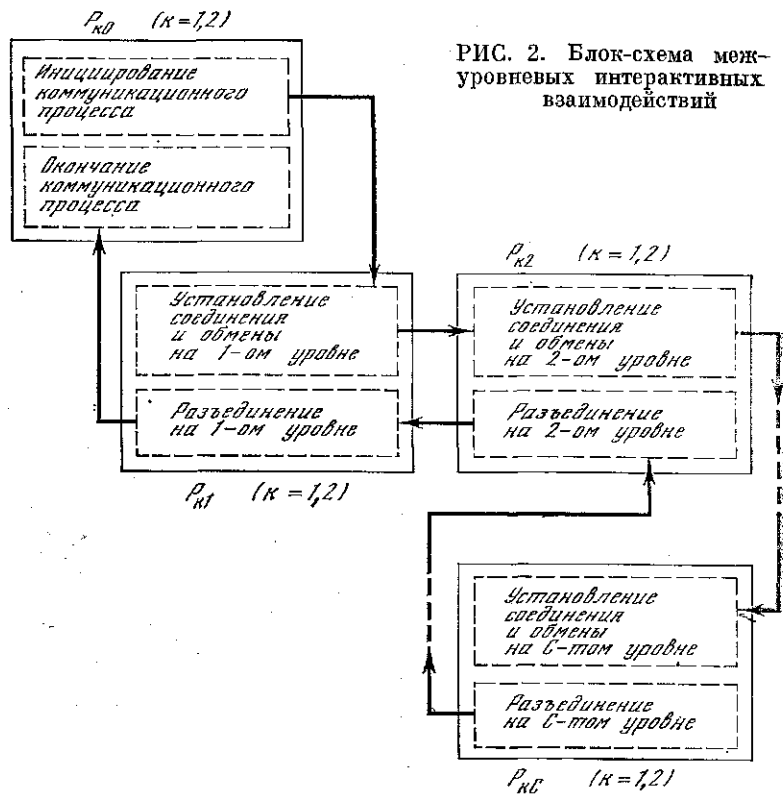


РИС. 2. Блок-схема межуровневых интерактивных взаимодействий

будут передаваться на следующем, $(c + 1)$ -м уровне взаимодействия подпроцессов. Вышеуказанной последовательности выполняемых фаз коммуникационного процесса соответствует определенная последовательность формальных грамматик G_{dc} ($c = 0, 1, 2, \dots, C$), в которой каждая предыдущая формальная грамматика порождает цепочки управляющих слов, передаваемых последующей формальной грамматике и используемых для порождения цепочек управляющих слов на следующем уровне иерархии взаимодействующих коммуникационных подпроцессов. Форма и характер связей в цепочке взаимодействующих формальных грамматик G_{dc} определяется структурой и содержанием процедур управления, предусмотренных коммуникационным протоколом, регулирующим межпроцессные связи в данной системе. Например, если порядок следования фаз коммуникационного процесса, предусмотренный протоколом, такой, как показано на рис. 2, то цепочка интерактивных последовательных взаимодействий формальных грамматик G_{dc} имеет закрученную форму, как показано на рис. 3.

Процесс интерактивного взаимодействия для вывода цепочек с использованием управляющих слов был определен в работе [12].

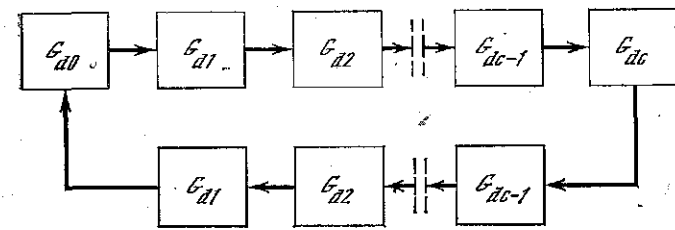


РИС. 3. Цепочка интерактивных взаимодействий формальных грамматик

Интерактивные взаимодействия в сложных системах рассмотрены в работе [13]. В работе [14] интерактивные языки, порождаемые в процессе интерактивного вывода, были использованы для моделирования межпроцессных связей в вычислительных сетях.

Рассмотрим две формальные грамматики: $G_{d1}(V_{dT1}, V_{dN1}, P_d^{(1)}, \sigma_d^{(1)})$, $G_{d2}(V_{dT2}, V_{dN2}, P_d^{(2)}, \sigma_d^{(2)})$. Обозначим через $P_d^{(1)} = (P_1^{(1)}, P_2^{(1)}, \dots, P_r^{(1)})$, $P_d^{(2)} = (P_1^{(2)}, P_2^{(2)}, \dots, P_n^{(2)})$ множества правил вывода цепочек в словарях соответственно V_{dT1} и V_{dT2} . Для того чтобы между формальными грамматиками G_{d1} и G_{d2} было возможно интерактивное взаимодействие, необходимо выполнение следующего условия [14]: $V_{dT1} \subseteq P_d^{(2)}$; $V_{dT2} \subseteq P_d^{(1)}$.

Цепочки левостороннего вывода, существующие в грамматике G_{dc} ,

$$S_{dc} = S_1 \xrightarrow{P_{j(1)}}_{G_{dc}} S_2 \xrightarrow{P_{j(2)}}_{G_{dc}} \dots \xrightarrow{P_{j(n)}}_{G_{dc}} S_{n+1} \quad (n \geq 1)$$

с использованием управляющих слов $P_{j(1)}, P_{j(2)}, \dots, P_{j(n)}$ сокращенно будем обозначать в виде $G_{G_{dc}}^{(Y)} = S_{n+1}$, где Y — множество управляющих слов [15].

Определим отношение между словами в языке $\kappa(G_{d1})$

$$\varphi_1 \xrightarrow{G_{d1}G_{d2}} \varphi_2$$

таким образом, чтобы удовлетворялись следующие условия:

$$\varphi_1, \varphi_2 \in \kappa(G_{d1}) \subseteq V_{dT1}^* \subseteq P_d^{(2)*},$$

$$\exists r \in \kappa(G_{d2}) \subseteq V_{dT2}^* \subseteq P_d^{(1)*},$$

$$r = G_{G_{d2}}^{(\varphi_1)}; \quad \varphi_2 = G_{G_{d1}}^{(r)},$$

где r является ассоциативным языком [16]. Отношение $\xrightarrow{G_{d1}G_{d2}}$ называется интерактивным выводом между грамматиками G_{d1} и G_{d2} .

Интерактивным языком $\kappa(G_{d1}, G_{d2}; \sigma_{d0})$, порождаемым двумя формальными грамматиками G_{d1} и G_{d2} , называется множество слов, получаемых с помощью интерактивного вывода, если начальное слово $\sigma_{d0} \in \kappa(G_{d1})$. Множество ассоциативных слов, соответ-

вующих словам языка $\kappa(G_{d1}, G_{d2}; \sigma_{d0})$, называется ассоциативным языком

$$\theta(G_{d1}, G_{d2}; \sigma_{d0}).$$

В случае последовательных взаимодействий нескольких формальных грамматик, например в случае цепочки взаимодействующих

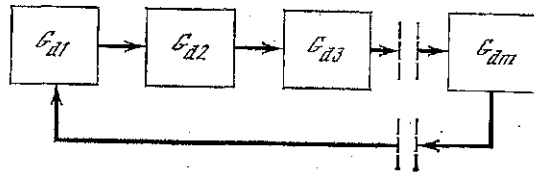


Рис. 4. Цепочка взаимодействующих формальных грамматик

грамматик, показанной на рис. 3, определим отношение между словами

$$\varphi_j, \varphi_{j+1} \in \kappa(G_{d0})$$

$$\varphi_j \xrightarrow{G_{d0}, G_{d1}, \dots, G_{dc}} \varphi_{j+1}$$

таким образом, чтобы удовлетворялись следующие условия:

$$\varphi_j = l_{j0} \in \kappa(G_{d0}), \quad G_{G_{d1}}^{(l_{j0})} = l_{j1} \in \kappa(G_{d1}),$$

$$G_{G_{d2}}^{l_{j1}} = l_{j2} \in \kappa(G_{d2})$$

$$G_{G_{dc}}^{(l_{jc(c-1)})} = l_{jc} \in \kappa(G_{dc}),$$

$$G_{G_{d(c-1)}}^{(l_{jc})} = l'_{j(c-1)} \in \kappa(G_{d(c-1)}),$$

$$G_{G_{d(c-2)}}^{(l'_{j(c-1)})} = l'_{j(c-2)} \in \kappa(G_{d(c-2)}),$$

$$G_{G_{d1}}^{(l'_{j2})} = l'_{j1} \in \kappa(G_{d1}),$$

.....

$$G_{G_{d0}}^{(l'_{j1})} = \varphi_{j+1} \in \kappa(G_{d0}).$$

Слова $l_{j0}, l_{j1}, l_{j2}, \dots, l_{jc}, l'_{j(c-1)}, l'_{j(c-2)}, \dots, l'_{j1}$ передаются на соответствующие уровни подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 0, 1, 2, \dots, C$) для определения формата, семантики и формы представления данных, передаваемых в диалоговых взаимодействиях на соответствующих уровнях иерархии подпроцессов \hat{P}_{kc} .

Обозначим через $\kappa(G_{d0}, G_{d1}, G_{d2}, \dots, G_{dc}; \sigma_{d0})$ множество цепочек интерактивного языка, порождаемого формальными грамматиками $G_{d0}, G_{d1}, \dots, G_{dc}$ при начальном слове σ_{d0} . Пусть в общем случае цепочка взаимодействующих формальных грамматик $G_{d1}, G_{d2}, \dots, G_{dm}$ имеет такой вид, как показано на рис. 4.

Определим отношение между словами

$$U_j, U_{j+1} \in \kappa(G_{d1}) \quad U_j \xrightarrow{G_{d1}, G_{d2}, \dots, G_{dm}} U_{j+1}$$

таким образом, чтобы удовлетворялись следующие условия:

$$U_j = l_{j1} \in \kappa(G_{d1}),$$

$$G_{G_{d2}}^{l_{j1}} = l_{j2} \in \kappa(G_{d2}),$$

.....

$$G_{G_{dm}}^{(l_{j(m-1)})} = l_{jm} \in \kappa(G_{dm}),$$

$$G_{G_{d1}}^{(l_{jm})} = U_{j+1} \in \kappa(G_{d1}).$$

Если в интерактивном языке $\kappa(G_{d1}, G_{d2}, \dots, G_{dm}; \sigma_{d0})$, порождаемом интерактивными взаимодействиями формальных грамматик G_{d1}, \dots, G_{dm} , определяющимися порядком следования уровней подпроцессов \hat{P}_{kc} ($k = 1, 2; c = 1, 2, \dots, C$) и фаз коммуникационного процесса, установленных данных протоколом, для любого слова $U \in \kappa(G_{d1}, G_{d2}, \dots, G_{dm}, \sigma_{d0})$ существует такое слово U' , что $U \xrightarrow{G_{d1}, G_{d2}, \dots, G_{dm}} U'$, то это означает, что система процедур межуровневых взаимодействий между уровнями подпроцессов \hat{P}_{kc} является логически полной и обеспечивает при нормальных условиях работы системы достижимость любого уровня взаимодействующих подпроцессов, реализующих данный коммуникационный процесс.

Одной из самых сложных проблем, возникающих при проектировании распределенных вычислительных систем с многослойной иерархической структурой, является создание удобных, кратких и точных моделей для определения спецификации протоколов и архитектуры системы, а также для управления процессами в этих системах. Предлагаемая в данной работе формальная языковая модель основана на использовании теории формальных грамматик и цепей Маркова.

Коммуникационный протокол сконструирован логически правильно при удовлетворении следующих условий: во-первых, если на каждом уровне иерархии подпроцессов, реализующих данный коммуникационный процесс, является неприводимой цепью Маркова, соответствующая формальной грамматике, порождающей коммуникационные цепочки данного коммуникационного процесса, и, во-вторых, если для всех уровней иерархии взаимодействующих подпроцессов существует интерактивный язык межуровневых взаимодействий, определяющий коммуникативные функции, форматы, семантику и форму представления данных, передаваемых на соответствующих уровнях взаимодействующих подпроцессов.

КОММУНИКАЦИОННЫЕ ПРОЦЕДУРЫ В СИСТЕМАХ
КОНВЕЙЕРНОЙ ОБРАБОТКИ ДАННЫХ

Во многих областях науки и техники (метеорология, геофизические исследования и др.) требуется значительно увеличить производительность ЭВМ и более эффективно обрабатывать информацию. Дополнительные возможности увеличения производительности распределенных вычислительных систем могут быть получены путем совмещения процессорных и коммуникационных функций в единой структуре канала передачи данных между узловыми ЭВМ распределенной системы. Организация вычислительного процесса на множестве микропроцессорных элементов, распределенных в последовательную цепочку по каналу передачи данных между узловыми ЭВМ, характеризуется последовательной обработкой с точки зрения технических средств, на которых выполняется эта обработка. В такой системе может быть осуществлено последовательное управление традиционной архитектуры ЭВМ фон Неймана; конвейерная обработка данных с помощью микропроцессорных элементов, распределенных по каналу передачи данных, может быть организована программным путем с помощью специальных методов управления. Построение такой системы, совмещающей коммуникационные и процессорные функции, позволит создавать распределенные многопроцессорные вычислительные системы с высоким коэффициентом производительности/стоимость.

Применяя обозначения циклических интерактивных грамматик, использованные в предыдущем разделе, процесс конвейерной обработки j -й большой задачи φ_j микропроцессорными элементами M_i ($i = 1, 2, \dots, n$), каждый из которых выполняет i -ю подзадачу с помощью программы $P_{p_i^j}$, обрабатывающей

данные D_i^j , выражается следующим образом:

$$G_{G_1}^{l_{j_0}} = l_{j_1} = (P_{p_1^j} \cup D_1^j); G_{G_2}^{l_{j_1}} = l_{j_2} = (P_{p_2^j} \cup D_2^j); \dots$$

$$\dots; G_{G_n}^{l_{j(n-1)}} = l_{j_n} = (P_{p_n^j} \cup D_n^j)$$

Здесь l_{j_0} — начальные данные решения задачи φ_j , соответствующие корневному узлу дерева решения задачи φ_j , которое можно отобразить в форме таблицы стеков номеров программ $P_{p_i^j}$ и обрабатываемых программами данных, соответствующих определенным узлам дерева решения задачи. G_i — формальная грамматика, порождающая программу $P_{p_{i+1}}$ и данные D_{i+1} , необходимые для обработки информации в следующем микропроцессорном элементе M_{i+1} . В этом случае каждая программа $P_{p_i^j}$ автоматически генерируется в предыдущем микропроцессорном элементе M_{i-1} на основе заданной цели решения задачи φ_j . Другой подход к генерации программ $P_{p_i^j}$ базируется на том, что в узловой ЭВМ определяется дерево решения задачи φ_j и строится соответствующая таблица стеков подзадач решения задачи φ_j . Эта таблица вместе с необходимыми данными для решения задачи φ_j распределяется по микропроцессорным элементам, которые, имея набор всех необходимых программ $P_{p_i^j}$, определяют по таблице стеков подзадач необходимую программу и производят соответствующую обработку данных. Если по логике решения задачи φ_j , помимо последовательной структуры управления вычислительным процессом, в дереве решения задачи φ_j имеется также альтернативная структура управления, предполагающая разветвление вычислительного процесса, то соответствующие предикатные логические функции могут быть добавлены в конце предыдущего узла дерева решения задачи φ_j или выделены в самостоятельный предикатный узел, чтобы выравнивать время обработки данных, выполняемой каждым микропроцессорным элементом M_i . В случае итеративной

структуры вычислительного процесса логические функции принятия решения могут быть выделены в самостоятельный предикатный узел.

При выполнении конвейерной обработки все процедуры обмена блоками данных — фреймов между микропроцессорными элементами M_i ($i = 1, 2, \dots, n$) регулируются коммуникационным протоколом, алгебраическая спецификация которого приведена ниже.

Алгебраическая спецификация фреймов. Независимые переменные: булевы (bool), неотрицательные целые числа (NNI), бит (BIT). Наборы независимых переменных: фрейм (F); последовательный номер (SQN); данные (D); биты направления конвейерной обработки (PF); биты направления передач между узловыми ЭВМ (RL); начальные биты управления передачей фрейма: конвейерная обработка или режим сквозной передачи без конвейерной обработки (G) (передачи данных между микропроцессорными элементами M_i ($i = 1, 2, \dots, n$) производится через интерфейсные блоки микропроцессоров I_{bi} ($i = 1, 2, \dots, n$), выполняющие коммуникационные функции в соответствии с установленным в данной системе коммуникационным протоколом); биты начала блока данных, передаваемых из I_{bi} в M_i для обработки (HA); биты конца блока данных, передаваемых из I_{bi} в M_i для обработки (KO). SQN=NN(0,1); D=BIT; PF=BIT=B; RL=BIT=B; HA=BIT; KO=BIT. Обозначение «=» определяется как базисный предикат со свойствами равенства, т. е. это выражение эквивалентности левой и правой части формулы.

Структура функций и их определение: информационный фрейм IF(G,SQ,N,HA,D,KO,PF,RL)=F; управляющие фреймы: RRF(G,SQ,N,PF,RL)=F; RNR(F,G,SQ,N,PF,RL)=F; функции, определяющие тип фрейма: TI(F)=BOOL; TRR(F)=BOOL; TRNR(F)=BOOL. Функции, определяющие SQN фрейма: SSQN(F)=S (номер передачи); RSQN(F)=R (номер приема); функции, определяющие бит набора PF: PFB(F)=BIT; функция, определяющая бит набора RL: RLB(F)=BIT; функция индикации режима обработки фрейма (режим передачи с конвейерной обработкой или режим сквозной передачи без конвейерной обработки): GB(F)=BIT; функции, определяющие информационный блок данных, начало и конец этого блока: INF(F)=D; HAB(F)=BIT; KOV(F)=BIT; функции контроля ошибки фрейма: ERF(F)=TRUE; ERCH(F)=BOOL.

Алгебраические выражения, определяющие значения функций: TI(IF(G,S,HA,D,KO,PF,RL))=TRUE; TI(RRF(G,R,PF,RL))=FALSE; TI(RNR(F,G,R,PF,RL))=FALSE; TRR(IF(G,S,HA,D,KO,PF,RL))=FALSE; TRR(RRF(G,R,PF,RL))=TRUE; TRR(RNR(F,G,R,PF,RL))=FALSE; TRNR(RNR(G,R,PF,RL))=TRUE; TRNR(IF(G,S,HA,D,KO,PF,RL))=FALSE; TRNR(RR(G,R,PF,RL))=FALSE; ERCH(ERF)=TRUE; ERCH(F)=FALSE;

Алгебраическая спецификация процедур управления приемом/передачей фреймов (R/SC). Спецификация R/SC включает в себя все независимые переменные, наборы независимых переменных и функции алгебраической спецификации фреймов. Множество состояний C блока I_{bi} , передающего фреймы блоку $I_{b_{i+1}}$, определяется сменой операций приема/передачи фреймов и таймаутом, вызванным ошибками во фрейме.

Структура функций и их определение. Три типа основных операций выполняют интерфейсные блоки I_{bi} : передачу и прием информационных, управляющих и служебных фреймов и операции, связанные с таймаутом, возникающим при ошибочных передачах и повторных передачах фреймов. В соответствии со сменой этих работ происходит изменение состояний блоков I_{bi} . В функциях, соответствующих указанным работам, выполняемым блоками I_{bi} , аргументами являются состояние интерфейсных блоков непосредственно перед выполнением работы и необходимые параметры; значением функции является состояние интерфейсного блока сразу после выполнения соответствующей работы.

Функции, определяющие передачу фреймов: SDIF(G,C,B,B)=C; SDRRF(G,C,B,B)=C; SDRNR(G,C,B,B)=C. Функция, определяющая прием фрейма: RCE(C,F)=C. Функция, определяющая последовательный номер передачи фрейма информационного формата: SQ(C)=SQN. Функция, определяющая величину, получаемую прибавлением 1 по модулю 1 к последователь-

ному номеру очередного передаваемого фрейма информационного формата: $HH(C) = NNI(0,1)$. Функция, определяющая последовательный номер информационного фрейма, ожидаемого в принимающем блоке I_{bi} , при условии, что информационный фрейм с последовательным номером, равным $(LW - 1)$, принят без ошибок: $LW(C) = NNI(0,1)$.

Значение $B=1$ в формате фрейма, переданного из блока I_{bi} в блок $I_{b(i+1)}$ в направлении конвейерной обработки, определяет окончание передачи из блока I_{bi} в блок $I_{b(i+1)}$ и ожидание ответной передачи из блока $I_{b(i+1)}$. Значение $B=0$ в формате фрейма, переданного из блока $I_{b(i+1)}$ в блок I_{bi} в направлении, противоположном направлению конвейерной обработки, установленному в данный момент времени в канале между узловыми ЭВМ, определяет окончание передачи из блока $I_{b(i+1)}$ в блок I_{bi} и ожидание следующей передачи из блока I_{bi} .

Функция, определяющая последовательный номер фрейма, переданного из блока I_{bi} с PF битом $B=1$: $CHP(C) = NNI(0,1)$. Функция, определяющая состояние, когда фрейм с соответствующим значением B набора битов PF передан, но в ответ на это фрейм с противоположным значением величины B еще не принят: $RY(C) = \text{BOOL}$. Функция, определяющая состояние, когда, после управляющего фрейма RNRFF еще не был получен фрейм, отменяющий это состояние: $BY(C) = \text{BOOL}$. Функция, определяющая, являются ли операции, определенные аргументами этой функции, недопустимыми: $IVD(C) = \text{BOOL}$. Обозначим через C_0 начальное состояние работы системы; тогда справедливы следующие выражения: $SQ(C_0) = 0$; $HH(C_0) = 0$; $LW(C_0) = 0$; $CHP(C_0) = 0$; $BY(C_0) = \text{FALSE}$; $IND(C_0) = \text{FALSE}$

Алгебраические выражения эквивалентностей, определяющие приемы / передачи фреймов в канале передачи данных с конвейерной обработкой.

$SQ(SDIF(G,C,B,\dot{B})) = SQ(C) + 1 \text{ MOD } 1$;
 $HH(SDIF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDIF(G,C,B,\dot{B}))) \text{ THEN } \text{IF } SQ(C) = HH(C) \text{ THEN } HH(C) + 1 \text{ MOD } 1 \text{ ELSE } HH(C)$;
 $LW(SDIF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDIF(G,C,B,\dot{B}))) \text{ THEN } LW(C)$;
 $CHP(SDIF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDIF(G,C,B,\dot{B}))) \text{ THEN } \text{IF } B=1 \text{ THEN } SQ(C) \text{ ELSE } CHP(C)$;
 $BY(SDIF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDIF(G,C,B,\dot{B}))) \text{ THEN } BY(C)$;
 $RY(SDIF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDIF(G,C,B,\dot{B}))) \text{ THEN } (B=1) \text{ OR } RY(C)$;
 $SQ(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } SQ(C)$;
 $HH(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } HH(C)$;
 $LW(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } LW(C)$;
 $CHP(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } \text{IF } B=1 \text{ THEN } SQ(C) \text{ ELSE } CHP(C)$;
 $BY(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } BY(C)$;
 $RY(SDRRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRRF(G,C,B,\dot{B}))) \text{ THEN } B=1 \text{ OR } RY(C)$;
 $SQ(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } SQ(C)$;
 $HH(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } HH(C)$;
 $LW(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } LW(C)$;
 $CHP(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } \text{IF } B=1 \text{ THEN } SQ(C) \text{ ELSE } CHP(C)$;
 $BY(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } BY(C)$;
 $RY(SDRNRF(G,C,B,\dot{B})) = \text{IF NOT}(IND(SDRNRF(G,C,B,\dot{B}))) \text{ THEN } B=1 \text{ OR } RY(C)$;
 $SQ(RCE(C,F)) = \text{IF NOT}(IND(RCE(C,F))) \text{ THEN } \text{IF } ERCH(F) \text{ THEN } SQ(C) \text{ ELSE } \text{IF } PFB(F)=1 \text{ AND } (CHP(C) - LW(C)) \text{ MOD } 1 > (RSQN(F) - LW(C)) \text{ MOD } 1 \text{ THEN } RSQN(F) \text{ ELSE } SQ(C)$;
 $HH(RCE(C,F)) = \text{IF NOT}(IND(RCE(C,F))) \text{ THEN } HH(C)$;

$CHP(RCE(C,F)) = \text{IF NOT}(IND(RCE(C,F))) \text{ THEN } \text{IF } ERCH(F) \text{ THEN } CHP(C) \text{ ELSE } \text{IF } (CHP(C) - LW(C)) \text{ MOD } 1 \leq (RSQN(F) - LW(C)) \text{ MOD } 1 \text{ THEN } RSQN(F) \text{ ELSE } CHP(C)$;
 $BY(RCE(C,F)) = \text{IF NOT}(IND(RCE(C,F))) \text{ THEN } \text{IF } ERCH(F) \text{ THEN } BY(C) \text{ ELSE } \text{IF } TRR(F) \text{ OR } (TI(F) \text{ AND } PFB(F)=1) \text{ THEN } \text{FALSE} \text{ ELSE } TRNR(F) \text{ OR } BY(C)$;
 $RY(RCE(C,F)) = \text{IF NOT}(IND(RCE(C,F))) \text{ THEN } \text{IF } ERCH(F) \text{ THEN } RY(C) \text{ ELSE } \text{IF } PFB(F)=1 \text{ THEN } \text{FALSE} \text{ ELSE } RY(C)$.

При передаче фреймов информационного формата, подвергаемых конвейерной обработке, из интерфейсного блока I_{bi} в следующий по ходу конвейерной обработки, интерфейсный блок $I_{b(i+1)}$ контроль времени передачи фреймов осуществляется интерфейсным блоком I_{bi} .

Алгебраическая спецификация процедур, связанных с контролем времени при передаче фреймов и обработкой ошибок. Обозначим через $TM(C)$ состояние таймаута интерфейсного блока I_{bi} , передающего информационные фреймы по направлению конвейерной обработки в следующий интерфейсный блок $I_{b(i+1)}$. Логическую конструкцию: « $\text{IF NOT}(IND(\dots)) \text{ THEN}$ », в которой в скобках указывается аргумент функции IND , обозначим $INT(\dots)$.

Функции, определяющие состояние ошибки и возникновение таймаута: $TER(C_0) = \text{FALSE}$; $TER(SDIF(G,C,B,\dot{B})) = INT(SDIF(G,C,B,\dot{B})) \text{ TER}(C)$;
 $IND(SDIF(G,C,B,\dot{B})) = \text{IF } IND(C) \text{ THEN } \text{TRUE}, \text{ ELSE } LW(C) = (SQ(C) + 1) \text{ MOD } 1 \text{ OR } (RY(C) \text{ AND } B=1) \text{ OR } TER(C) \text{ OR } BY(C)$;
 $TER(SDRRF(G,C,B,\dot{B})) = INT(SDRRF(G,C,B,\dot{B})) \text{ NOT } (B=1) \text{ AND } TER(C)$;
 $IND(SDRRF(G,C,B,\dot{B})) = \text{IF } IND(C) \text{ THEN } \text{TRUE} \text{ ELSE } (RY(C) \text{ AND } B=1)$;
 $TER(SDRNRF(G,C,B,\dot{B})) = INT(SDRNRF(G,C,B,\dot{B})) \text{ TER}(C)$;
 $IND(SDRNRF(G,C,B,\dot{B})) = \text{IF } IND(C) \text{ THEN } \text{TRUE} \text{ ELSE } (RY(C) \text{ AND } B=1) \text{ OR } TER(C)$;
 $TER(RCE(C,F)) = INT(RCE(C,F)) \text{ TER}(C)$;
 $SQ(TM(C)) = INT(TM(C)) \text{ SQ}(C)$; $HH(TM(C)) = INT(TM(C)) \text{ HH}(C)$;
 $LW(TM(C)) = INT(TM(C)) \text{ LW}(C)$; $CHP(TM(C)) = INT(TM(C)) \text{ CHP}(C)$;
 $TER(TM(C)) = \text{TRUE}$; $RY(TM(C)) = \text{FALSE}$; $BY(TM(C)) = INT(TM(C)) \text{ BY}(C)$;
 $IND(TM(C)) = IND(C)$.

Функции, определяющие состояние интерфейсного блока $I_{b(i+1)}$ при приеме им информационных фреймов, подвергаемых конвейерной обработке, из интерфейсного блока I_{bi} : $RY(C_0) = \text{TRUE}$; $IND(SDIF(G,C,B,\dot{B})) = \text{TRUE}$;
 $IND(SDRRF(G,C,B,\dot{B})) = \text{IF } IND(C) \text{ THEN } \text{TRUE} \text{ ELSE } RY(C)$;
 $IND(SDRNRF(G,C,B,\dot{B})) = \text{IF } IND(C) \text{ THEN } \text{TRUE} \text{ ELSE } RY(C)$.

Обозначим через $PT(BF, b, D)$ функцию, определяющую процедуры помещения в буфер интерфейсного блока данных D с последовательным номером передачи b . Функцию, определяющую процедуры выдачи из буфера интерфейсного блока данных D , имеющих последовательный номер передачи b , обозначим через $ACC(BF, b)$. $b = SQN(0,1)$. Множество постоянных, обозначающих включение таймера, отсчитывающего время с момента передачи фрейма, выключение таймера и безразличное состояние обозначим соответственно YCT, VCT и HT . Функция, определяющая процедуры передачи фрейма информационного формата с учетом включения таймера, отсчитывающего время с момента передачи этого фрейма, может быть представлена следующим образом: $SDIF(G,C,B,\dot{B}) = \text{IF } B=1 \text{ THEN } YCT \text{ ELSE } HT$. Аналогичным образом может быть определено состояние включенного таймера при определении структуры функций $SDRRF(G,C,B,\dot{B})$ и $SDRNRF(G,C,B,\dot{B})$, а также функций $RTR(F)$, определяющей повторную передачу фрейма F .

Обозначим состояние, когда информационный фрейм принят с ошибками и информационные данные этого фрейма не могут быть переданы на обработку микропроцессорному элементу, через $HD(F)$. Состояние, когда информационный фрейм принят правильно и информационные данные фрейма переданы в микропроцессорный элемент для обработки, обозначим через $INF(F)$.

При этом, функция, определяющая прием фрейма, RCE(C,F) может быть представлена следующим образом: RCE(C,F)=IF ERCH(F) THEN HD AND HT ELSE IF NOT(TI(F)) THEN HD AND HT ELSE IF SSQN(F)= =ACK(F)=b THEN PT(BI,b,D) AND INF(F) AND BCT ELSE HD ANDHT.

Функция IYD(C), определяющая недопустимость операций, выражается IYD(SDIF(G,C,B,B))=IF IYD(C) THEN TRUE ELSE NOT(SQ(C)=HH(C)); IYD(RTR(F))=IF IYD(C) THEN TRUE ELSE SQ(C)=HH(C).

Множество функций IF, RRF, RNR, C₀, SDIF, SDRRF, SDRNR, RCE, TM обозначим через K-функции. Множество остальных функций обозначим через Y-функции. Тогда левые части вышеприведенных формул можно представить в виде Y(K(X₁, X₂, . . . , X_m), X_{m+1}, . . . , X_n).

X₁, X₂, . . . , X_n — взаимно различные параметры, среди которых могут быть переменные, соответствующие K-функциям. Пары функций Y и K в левых частях вышеуказанных формул представлены в различных формулах различными комбинациями и отличаются друг от друга. Следовательно, для системы, выражаемой этими формулами, справедливы свойства CHURCH—ROSSER [17] и можно утверждать, что две системы алгебраических базисных выражений различные и не являются эквивалентными в этих формулах, определяющих вышеуказанные спецификации; то есть если значения Y-функций в вышеуказанных формулах определены, то эта система определяется однозначно. Это определяет непротиворечивость вышеуказанных алгебраических спецификаций, выражаемых этими формулами.

ЛИТЕРАТУРА

1. Якубайтис Э. А. Архитектура вычислительных сетей. М.: Статистика 1980.
2. Гинзбург С. Математическая теория контекстно-свободных языков. М.: Мир, 1970.
3. Hoffman H. On linguistic aspects of communication line control procedures. IBM Res., Calif. Rep. RZ345, 1970.
4. Hopgood F. R. A. Compiling techniques. N. Y.: Amer. Elsevier, 1969.
5. Bartlet K., Scantlebury R., Vilkinson P. A note on reliable full duplex transmission over half duplex lines.— Commun. Assoc. Comput. Mech., 1969, May, vol. 12.
6. Linch W. Commentary on the foregoing note.— Ibid., 1969, May, vol. 12.
7. Gray J. Line control procedures.— Proc. IEEE, 1972, vol. 60, N 11.
8. Linch W. Reliable full duplex file transmission over half duplex telephone lines.— Commun. Ass. Comput. Mech., 1968, June, vol. 11.
9. Harangozo J. An approach to describing a link level protocol with a formal language.— In: Proc. 5th Data Common Symp. Utah., 1977, p. 4.37.
10. Teng A., Liu M. T. A formal model for automatic implementation and Logical validation of network communication protocols.— In: Proc. Comput. Networking Symp. Nat. Bur. Stand., 1978, p. 114—123.
11. C. H. West and, Zafiropulo P. Automated validation of a communication protocol: the X.21 recommendation.— IBM J. Res. Develop., 1978, vol. 21, N 1.
12. Ginsburg S., Spanier E. H. Control sets on grammars.— Math. Syst. Theory, 1968, vol. 2, p. 169.
13. Gabrielian A. The theory of interacting local automata.— Inform. and Control., 1970, vol. 16, p. 360.
14. Выставкин Я. П. Сети обмена информацией между ЭВМ. М.: Наука, 1975.
15. Hopcroft J. E., Ulman J. D. Formal languages and their relation to automata. Addison Wesley, 1969.
16. Moriya E. Associate languages and deridational Complexity of formal grammars and languages.— Inform. and Control, 1973, vol. 22 p. 139.
17. Huet G. Confluent reductions abstract properties and applications to term rewriting sistem.— Proc. IEEE, 18 Amer. Aymp. on found of Comp. Science, pp. 30—45 (1977).

Предисловие	3
I. ПСИХОФИЗИОЛОГИЧЕСКИЕ И ТЕХНИЧЕСКИЕ АСПЕКТЫ ПОРОЖДЕНИЯ И РАСПОЗНАВАНИЯ РЕЧИ ЧЕЛОВЕКОМ И МАШИНОЙ	
Влч. Вс. Иванов Естественный язык — мозг, искусственный язык — машина . .	5
Ю. М. Лотман Культура и текст как генераторы смысла	23
Л. А. Чистович, <u>В. А. Кожевников</u> , Ж. А. Першин Проблемы моделирования слуховой обработки речевого сигнала	30
Е. К. Гусева К проблеме понимания ЭВМ графематического уровня языка	40
II. ОБЩЕНИЕ С ЭВМ И МОДЕЛИРОВАНИЕ ЕСТЕСТВЕННОГО ЯЗЫКА	
Б. В. Суголин Моделирование грамматических категорий средствами тензорной алгебры	55
Ю. С. Мартельянов О связанном задании синтаксиса	72
Д. Г. Лагути, Е. В. Федоров, И. С. Добронравов, В. Ф. Пархоменко Автоматическое индексирование текстов в документальных ИПС	84
Я. П. Выставкин, Е. Г. Гусева Формальные языковые модели диалоговых взаимодействий . .	95

**КИБЕРНЕТИЧЕСКАЯ
ЛИНГВИСТИКА**

Утверждено к печати
Научным советом
по комплексной проблеме «Кибернетика»
АН СССР

Редакторы издательства *Р. Х. Дорионов, С. С. Матвеев*
Художник *А. Г. Кобрин*. Художественный редактор *Н. Н. Власин*
Технический редактор *В. Д. Прилепская*
Корректоры *Т. С. Козлова, И. А. Талалай*

ИБ № 27568

Сдано в набор 04.01.83
Подписано к печати 21.03.83
Т-02757. Формат 60×90¹/₄.
Бумага книжно-журнальная
Гарнитура обыкновенная новая
Печать высокая
Усл. печ. л. 7,5. Усл. кр. отт. 7,88
Уч.-изд. л. 8,4. Тираж 2550 экз. Тип. зак. 2395
Цена 85 коп.

Издательство «Наука»
117864 ГСП-7, Москва, В-485, Профсоюзная ул., 90
2-я типография издательства «Наука»
121099, Москва, Г-99, Шубинский пер., 10