

УДК 519.688+519.7

## ЛАДА - ПРОГРАММНАЯ СИСТЕМА ЛОГИЧЕСКОГО АНАЛИЗА ДАННЫХ

Е.Е.Витлев, А.А.Москвитин

### В в е д е н и е

Программная система ЛАДА, ориентированная на решение задач Логического Анализа Данных пользователями, не владеющими программированием, возникла как результат синтеза работ в следующих трех направлениях:

- решение конкретных задач анализа данных из разных областей (геология, биология, медицина, экономика, техника, сельское хозяйство и др.) [1-8, 11, 12];
- теоретические работы по эмпирическому предсказанию [8, 9, 11, 12, 23], логике эмпирических исследований [10] и логическому анализу данных [1, 5-8, 11-23, 34-41];
- разработке технологии построения проблемно-ориентированных, инструментальных высокорентабельных программных систем, рассчитанных на пользователей, не умеющих программировать [24-26].

Основное содержание решаемых задач - построение эмпирических теорий на основании результатов измерений и некоторых, обычно неформализованных, представлений (качественной теории), имеющих у специалистов из данной предметной области. Накопленный в ходе решения таких задач опыт, как собственный [1-3], так и других исследователей (см., например, [5-7]), явился основой для достаточно полного выявления свойств данного круга задач, а также для построения обобщенного портрета пользователя программных средств для решения этих задач. Этот же опыт явился отправным пунктом фундаментальных теоретических разработок.

I. Обобщенная характеристика пользователя. Как показывает опыт, пользователь на содержательном уровне достаточно хорошо разбирается в своей проблемной области, владеет ее терминами и поня-

тиями, знает конечные цели и задачи исследования. Вместе с тем его математическая подготовка, как правило, оказывается недостаточной для представления этих знаний на формальном уровне. Ему трудно дать формальное представление терминов и понятий своей проблемной области, строго сформулировать свою задачу и тем более представить процесс ее решения на алгоритмическом языке.

Нередко пользователь не может сразу точно сформулировать, что именно он хочет от ЭВМ, и в процессе решения задачи пересматривает исходные предпосылки и предположения.

Встречаются случаи, когда пользователь, при работе с системой, из-за незнания типа шкал величин, входящих в данные, допускает ошибки в их определении.

Как правило, пользователь в области анализа данных и искусственного интеллекта привык иметь дело с числовой формой представления величин и зависимостей и логической формой представления знаний, выводов и решений и предполагает получать их от ЭВМ именно в таком виде.

2. Характеристика задач. Задачи логического анализа экспериментальных данных встречаются фактически во всех областях человеческой деятельности. Поскольку в каждой из областей свои традиции, терминология и понятия, то на содержательном уровне приходится иметь дело с неизмеримо большим разнообразием терминов, понятий, постановок задач и т.п. В то же время оказалось, что экспериментальные данные, теоретические представления исследуемой предметной области, а также получаемые закономерности могут быть представлены в единообразной форме на языке логики предикатов [1,7,8,12,14,17,20-22,35-39]. Более того, такая форма представления оказалась не только универсальной, но и удобной для математически неподготовленного пользователя [1,7,8,11,12,29] и, что особенно важно, заметно упрощает программирование [27-29].

Логическая форма представления оказалась адекватной для описания объектов с разношкальными признаками, так как все признаки удается представить в единообразном виде без потери (или привнесения неинтерпретируемой) информации [1,7-9,11-14,17,20-23,35-38].

Важная особенность данного круга задач связана с конечной целью исследований пользователей – собрать воедино всю сумму знаний о изучаемой предметной области, выявить и описать структуру отношений и на этой основе построить целостную модель данной области.

Таким образом, данный круг задач предъявляет к программным средствам высокие требования как по эффективности самого решения, так и по времени настройки на особенности задачи и по простоте реконструкции этих средств.

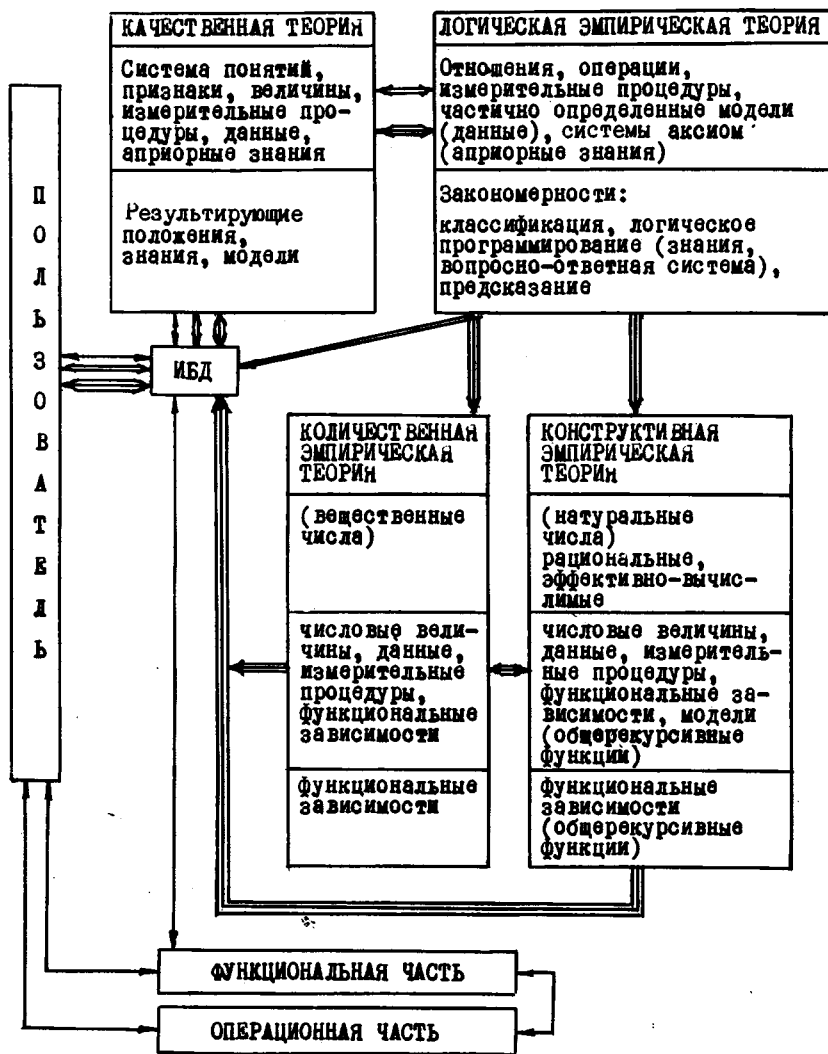
В согласии со спецификой данного круга задач и особенностей профессиональной и математической подготовки пользователей можно выделить три уровня решения задачи (см.схему).

На первом уровне пользователь проверяет правильность своих представлений о типах шкал величин данных, систематизирует имеющиеся у него сведения о данной задаче, по которым система должна синтезировать необходимые программные средства. Другими словами, на этом уровне он подготавливает свои данные для трансляции их в логическую форму.

Второй уровень является основным в системе ЛАДА. Начинается он с трансляции всех исходных данных (величин, понятий, измерительных процедур, априорных знаний и т.п.) в логические отношения, операции, измерительные процедуры, частично определенные модели, системы аксиом и др. Затем решается одна из задач логической эмпирической теории (обнаружение закономерностей, классификация или предсказание и т.д.), и выдается результат в логической форме. Математически подготовленный пользователь при решении своих задач уже может ограничиться этим уровнем.

На третьем уровне результаты, представленные в логической форме, преобразуются в числовую и выражаются либо на языке функциональных зависимостей, либо общерекурсивными функциями. Включение этого уровня в общую схему вызвано пожеланиями пользователей представлять результаты вычислений в традиционном для них виде.

В §2 дается подробное изложение схемы построения эмпирических теорий (см.схему), а в §3 и 4 в согласии с [24] дается описание программной системы ЛАДА.



### Схема решения задачи.

ИБД - интеллектуальная база данных;

— — — I уровень решения задачи,

— — — II уровень решения задачи,

— — — III уровень решения задачи.

## §1. Схема построения эмпирических теорий<sup>\*)</sup>

1. Первым шагом построения количественной теории является фиксация качественной теории (существующих понятий, знаний и представлений) в базе данных. Как следует из результатов теории измерений [20–22, 38], теории принятия решений [35–37], логики эмпирических исследований и эмпирического предсказания, дальнейшее построение количественной теории должно происходить на основе логического анализа предметной области. Такой анализ может быть осуществлен методами логического анализа данных. Результатом его является логическая эмпирическая теория – некоторая теория первого порядка. Количественная эмпирическая теория строится на основании логической при использовании результатов теории измерений и теории принятия решений о числовых представлениях величин и закономерных связей. Конструктивная эмпирическая теория строится на основании логической при использовании результатов теории конструктивных моделей [30]. Она дает представление величин и закономерных связей натуральными или рациональными числами (в общем случае кодами). Все построенные эмпирические теории интерпретируются в понятиях качественной теории и составляют совместно с ней результирующую количественную теорию предметной области.

2. Качественная теория является исходным пунктом дальнейшего построения эмпирических теорий. Носителем ее является пользователь. Допускаются мало разработанные и чисто интуитивные теории. Минимальные требования к качественной теории состоят в том, чтобы, во-первых, существовала система понятий, в которой она формулируется и интерпретируется, и, во-вторых, существовали свойства, признаки, величины и соответствующие измерительные процедуры, интерпретируемые в этих понятиях. Имеющиеся знания, сведения, данные и модели должны храниться в интеллектуальной базе данных. Построе-

<sup>\*)</sup> Термин "эмпирические теории" имеет здесь несколько иной смысл, чем, например, в [10]. Результаты измерений и закономерности являются эмпирическими, но эмпирические системы и системы аксиом теории измерений являются уже идеализациями процессов измерения и не являются, строго говоря, эмпирическими. Термин "эмпирические" сохранен за ними (как и в теории измерений) потому, что рассматриваются идеализации только реальных измерительных процедур. Статус этих идеализаций тот же, что и у аналогичных идеализаций в физике. В рассматриваемой системе переход от эмпирического уровня к идеализированному и обратно осуществляется пользователем.

ние такой системы в диалоге с пользователем составляет задачу I.1 (см. табл. I<sup>ж</sup>). Построенные далее логическая, количественная и конструктивная эмпирические теории будут интерпретироваться в системе понятий качественной теории. Полученные результирующие положения, знания и модели также должны заноситься в базу данных.

3. Начальное состояние логической эмпирической теории получается из качественной теории формализацией ее логико-операциональной составляющей – формальным заданием качественных сравнений, отношений, суждений, действий, операций и т.д. В соответствии с методологическим принципом теории измерений – свойства определяются отношениями – из всех свойств, величин и признаков качественной теории пользователь в диалоге выделяет интерпретируемые в системе понятий отношения и операции. Они образуют словарь  $v = \{P_1, \dots, P_n, \rho_1, \dots, \rho_k\}$ . Выделение словаря проводится по методике [I] в диалоге с пользователем (задача П.1). Результаты измерительных процедур и данные качественной теории представляются частично-определенными моделями  $\mathcal{M} = \langle A; v \rangle$ ,  $A = \{a_1, \dots, a_n\}$  – множество объектов. Это может быть сделано простыми программами преобразования данных (задача П.2). Априорные знания качественной теории представляются системой аксиом  $S^W$  в словаре  $v$  (задача П.3). Можно показать [I, 2I, 3I], что из всех эмпирически осмысленных аксиом можно удалить кванторы существования, вводя в них интерпретируемые (в системе понятий качественной теории) операции над объектами (скулемовские функции). В результате получим систему аксиом  $S^V$  как множество универсальных формул [3I]. В работе [I5] показано, что множество универсальных формул логически эквивалентно множеству формул вида

$$\forall x_1, \dots, x_n (A_1^{\epsilon_1} \& \dots \& A_1^{\epsilon_1} \rightarrow A_0^{\epsilon_0}), \quad (1)$$

где  $A_0, A_1, \dots, A_1$  – атомарные формулы,

$$\epsilon = \begin{cases} 1, & \text{если атомарная формула берется без отрицания,} \\ 0, & \text{если атомарная формула берется с отрицанием.} \end{cases}$$

Алгоритм перевода универсальных формул в формулы (I) описан в [I5, 32]. Такой путь преобразования системы аксиом  $S^W$  в совокупность формул (I) составляет задачу П.3,б. В последнее время фор –

---

ж)

Все таблицы даны в приложении.

мулы (I) часто используются при задании базы знаний в экспертных системах [29,33]. Поэтому некоторые априорные знания качественной теории могут быть с самого начала заданы формулами (I) (задача П.3,а). Дальнейшее построение логической эмпирической теории происходит путем усиления системы аксиом  $S^V$  методом обнаружения закономерностей [15]. Применяя к имеющимся данным  $M$  этот метод с некоторым уровнем доверия  $\alpha$ , получим множество закономерностей  $F_\alpha$  - совокупности формул (I) (задача П.4). В [16] доказано, что этим методом можно обнаружить любую закономерность, выраженную универсальной формулой. Доказательство проведено для предикатной сигнатуры, но его легко распространить на общий случай. Ставя различные эксперименты по анализу зависимостей между различными величинами и обрабатывая получающиеся данные, можно дополнить систему аксиом  $S^V$  различными множествами  $F_\alpha$  и получить в результате теорию  $T$  предметной области как совокупность формул (I). Эту теорию назовем логической эмпирической теорией.

Если независимо определить атомарные формулы и их отрицания и включить формулы  $A \vee \neg A$  в теорию  $T$ , то формулы (I) можно представить хорновыми дизъюнктами, а теорию  $T$  - программой в логическом программировании, например, в языке ПРОЛОГ [28]. Это обеспечивает эффективное использование теории  $T$  и получение из нее различных следствий (задача П.5). К теории  $T$ , как к программе в языке ПРОЛОГ, можно обращаться с любыми вопросами вида:  $? - A_1 \& \dots \& A_n$  и автоматически получать ответ. Использование логического программирования или  $\Sigma$ -программирования [27] реализует вопросно-ответную систему.

Теория  $T$  может использоваться для классификации объектов методом [34] (задача П.6), а также для получения предсказаний неизвестных значений признаков методом [15] (задача П.7).

Все результаты логической эмпирической теории нужно интерпретировать в системе понятий качественной теории. Интерпретируемость результатов следует из интерпретируемости отношений и операций, интерпретируемости формул (I) и следствий из них, а также из интерпретируемости результатов классификации и предсказания [15,34].

4. Количественная эмпирическая теория строится на основании результатов теории измерений и теории принятия решений о числовых представлениях величин, законов и функций полезности. Если в теории  $T$  содержится какая-либо система аксиом теории измерений или

теории принятия решений, то, используя соответствующие процедуры шкалирования (задача Ш.2), можно получить числовые представления величин, функциональных зависимостей и функций полезности. Включение в Т систем аксиом может быть установлено логическим выводом (задача П.5) либо методом обнаружения закономерностей на реальных данных (задача Ш.1). Шкалирование может быть осуществлено либо по данным методами решения линейных неравенств [22] (задача Ш.2,а), либо в диалоге с экспертом, как это делается, например, в [36] (задача Ш.2,б), либо планированием эксперимента (задача Ш.2,б), в случае величин, измеряемых приборами. Функциональные зависимости и функции полезности получаются из систем аксиом с одновременным построением числовых представлений входящих в них величин [22]. Это позволяет простыми функциональными зависимостями описывать целые классы функций [22,18]. В теории измерений [22] найдены системы аксиом для многих физических величин и фундаментальных физических законов. Результаты функциональной теории измерений [39] показывают, что подобную систему величин, связанных между собой фундаментальными законами, можно получить и во многих других областях. Многие системы аксиом [21,22,35-38], описывающие функциональные зависимости и функции полезности, содержат только отношение линейного порядка для одной из взаимосвязанных величин. Это позволяет применять их в различных предметных областях. Числовые представления величин, функциональных зависимостей и функций полезности, получаемые из систем аксиом, адекватны [21] и интерпретируемы в системе понятий качественной теории.

5. В теории измерений и теории принятия решений нельзя получить числовые представления некоторых величин и закономерных связей в силу ограниченности используемого в них понятия числового представления. Величины и закономерные связи, описываемые частичными порядками, толерантностями, решетками и т.д. не могут быть сильным гомоморфизмом вложены в поле вещественных чисел. Для числового представления таких величин и закономерных связей можно использовать конструктивные числовые представления [13,19,40,41]. Значениями величин в этом случае являются натуральные, рациональные или другие эффективно вычислимые числа (например, какие-либо коды). Ближе всего к понятию числового представления теории измерений понятие конструктивного числового представления, основанное на конструктивизации [30] эмпирических систем. Требование конструктивности процесса шкалирования приводит к другой формализации



конструктивного числового представления. Конструктивные числовые представления интерпретируются в системе понятий качественной теории.

6. Качественная теория и логическая, количественная и конструктивная эмпирические теории строятся последовательно. Циклы их формирования обозначены соответственно одинарной, двойной и тройной пунктирными линиями. Совместное их размещение в базе данных, при котором они бы дополняли друг друга и составляли совместно результирующую количественную теорию предметной области, можно осуществить на основе многосортных алгебр [33], реализуемых реляционной базой данных.

Результирующая количественная теория характеризуется тем, что в ней есть реляционная база данных, есть база знаний, включающая теорию  $T$ , функциональные и конструктивные зависимости и есть вопросно-ответная система, основанная на логическом программировании. Все данные, знания и результаты интерпретируемы в понятиях предметной области. Интерпретация доступна конечному пользователю. В количественной теории можно получить сколь угодно точное описание величин и зависимостей.

## §2. Основные черты программной системы ЛАДА

Как указывалось во введении, к программным средствам системы ЛАДА предъявляются самые высокие требования.

1. Организация непосредственного общения с ЭВМ пользователя, не обладающего достаточно высокой математической подготовкой и не владеющего программированием. Традиционно такой пользователь общается с ЭВМ через посредников: математика и программиста (нередко представленных в одном лице). На математика обычно ложится формализация знаний пользователя, строгая постановка задачи, выбор алгоритма решения, а также обсуждение результатов и внесение изменений в постановку задачи или алгоритм ее решения. На программиста — запись алгоритма на формальном языке, отладка программы и решение задачи на ЭВМ.

Теперь эти посреднические функции должны быть переложены на программные средства. Наиболее трудной является задача формализации функций математика. Математик, для того чтобы он мог правильно формализовать задачу пользователя, должен владеть не только аппаратом математики, но и достаточно хорошо разбираться в предметной области пользователя. В процессе общения с пользователем

он ведет диалог: задает вопросы, переспрашивает, уточняет, требует подтверждения того или иного тезиса и т.п. Он, образно говоря, "извлекает" из пользователя необходимую информацию по определенному сценарию, который у него выработался в ходе подобных общений. Формализация такого сценария весьма сложная задача, поэтому более естествен другой путь, а именно, заранее на начальной стадии разработки системы очертить круг задач пользователя. "Извлечь" из него всю необходимую информацию по каждой из задач, формализовать эти задачи, указать алгоритмы их решения.

Программист анализирует эту совокупность алгоритмов, выделяет типовые программные модули и ищет для каждого из алгоритмов главную программу, определяющую порядок работы модулей.

После того, как такая работа проделана, роль сценария диалога с пользователем становится вполне конкретной. Она заключается в последовательном выяснении следующих вопросов: 1) не сводится ли предлагаемая пользователем задача к одной из ранее решенных (для которой уже имеется головная программа)? 2) может ли данная задача быть решена с помощью уже имеющихся совокупностей модулей?

При положительном ответе на второй вопрос выдается задание на автоматический синтез головной программы. При отрицательном - пользователь отсылается к математику (проблемному программисту), который определяет, какие именно программные модули нужно разработать для решения данной задачи, и достаточно ли этого, т.е. решает, справится ли с этой работой он сам либо нужна помощь системного программиста, либо разработчика системы.

Таким образом, система ЛАДА должна обладать диалоговой системой, способной вести в терминах конкретной предметной области диалог с пользователем на естественном языке.

Кроме того, система ЛАДА должна обладать и достаточно развитой синтезирующей системой, позволяющей по заданию диалоговой системы синтезировать головную программу.

2. Под требованием к эффективности системы ЛАДА в согласии с § I и п. I настоящего параграфа следует понимать:

- требование к быстрой настройке на параметры задачи пользователя, т.е. на эффективный синтез программных средств. Эта настройка включает в себя: работы I уровня (см. схему); синтез программных средств и контрольный счет, по результатам которого проверяется правильность выбора параметров и постановки задачи. При необходимости этот процесс повторяется многократно;

- простоту реконструкции, т.е. разработку и включение новых программных модулей;

- эффективность счета, т.е. высокое качество реализации самих модулей и головной программы (см. [24]).

Все указанные качества обеспечиваются  $\pi$ -технологией и подробно обсуждены в предыдущей работе [24].

3. Система ЛАДА, как указывалось в §1, в перспективе должна стать основой для построения экспертных систем для конкретных предметных областей, т.е. в ней должна быть предусмотрена весьма совершенная организация баз данных и знаний.

Эту большую работу целесообразно вести постепенно, определив уже на данной стадии основные принципы ее создания.

Этот вопрос так же достаточно подробно обсужден в [24], и, чтобы не повторяться, укажем лишь, что в первой версии системы ЛАДА используется файловая организация базы данных, обеспечиваемая средствами ОС РВ СМ ЭВМ. Сейчас ведутся работы по созданию базы данных на основе использования вынесенных списков и выделенных структур файлов [24].

Все указанные выше требования полностью совпадают с теми, которые положены в основу создания  $\pi$ -технологии [24]. Поэтому естественно было создавать систему ЛАДА как  $\pi$ -систему в рамках указанной  $\pi$ -технологии.

### §3. Первая версия $\pi$ -системы ЛАДА

Построение системы ЛАДА как  $\pi$ -системы позволяет использовать большой арсенал средств, которыми располагает  $\pi$ -технология. Задача в этом случае сводилась к адаптации  $\pi$ -системы-прародительницы на выполнение функций логического анализа данных и к последующему выделению системы ЛАДА в качестве самостоятельной (дочерней)  $\pi$ -системы.

В согласии с идеологией  $\pi$ -технологии система ЛАДА состоит из пяти систем: анализа и обработки данных, диалоговой, синтеза, операционной и баз данных. Для первой версии системы ЛАДА (табл. I) можно было взять в готовом виде простейшие варианты двух последних (операционной и баз данных).

Таким образом, адаптация коснулась двух систем: анализа и обработки данных и диалоговой. Для системы анализа и обработки данных имелся задел в виде 64 программ на языке ФОРТРАН. Требовалось

придать им вид SP-модулей, т.е. вынести из тела каждой из программ переменные величины. Это вынесение коснулось сбора основных переменных величин в COMMON-блоках и не потребовало больших усилий, так как при написании программ учитывалась возможная локализация таких параметров. Связь системы анализа и обработки данных с диалоговой и синтеза выражалась в добавлении в SP-модули COMMON-блоков и вставки в тело модулей операторов вызова модулей связи.

Для построения диалоговой системы у  $\pi$ -системы-прародительницы имелись готовые средства и подходящие конструкции фраз естественного языка. Требовалось создать языковые средства (табл.2-9)  $\pi$ -системы ЛАДА и построить таблицу соответствий между конструкциями входного языка и последовательностями программных модулей,необходимую для системы синтеза, а также установить регламентацию работ конечного пользователя с  $\pi$ -системой ЛАДА (табл.10).

Операционная система и базы данных не потребовали каких-либо доработок, а для системы синтеза потребовалось разработать средства синтаксического анализа и генерации (синтеза) модулей по формулам (I) входного языка. Для операционной системы и баз данных использовались средства ОС РВ СМ ЭВМ.

Все это (в первую очередь,возможности  $\pi$ -технологии, а также наличие готовых программных модулей на языке ФОРТРАН) позволило создать  $\pi$ -систему ЛАДА за весьма короткий срок - всего за два человеко-месяца.

Уже в таком виде  $\pi$ -система ЛАДА оказалась вполне жизнеспособной и применяется в ряде организаций страны. При этом была продемонстрирована эффективность настройки на задачу пользователя.Если ранее на составление и отладку соответствующей головной программы уходили недели, то теперь  $\pi$ -системе ЛАДА требуется на это несколько (5-10) минут.

В ходе внедрения потребовалось также расширить круг решаемых задач, а это потребовало расширить языковые средства (табл.8) и дополнительно написать и отладить 2I модуль. Такая реконструкция заняла всего три недели, что также подтверждает указанные [24] возможности  $\pi$ -технологии.

В настоящее время ведутся работы по созданию  $\pi$ -системы для принятия решений [41] и реализации идей  $\Sigma$ -программирования [27], которые существенно обогатят  $\pi$ -систему ЛАДА.

## Л и т е р а т у р а

1. ВИТЯЕВ Е.Е. Анализ данных с применением языка эмпирических систем. -Автореф. дис.... канд.техн.наук. - Новосибирск, 1982. - 16 с.
2. ВИТЯЕВ Е.Е., КОЛЧАНОВ А.А., СОЛОВЬЕВ В.В. Исследование закономерностей формирования вторичных структур в глобальных белках. -В кн.: Использование вычислительных машин в спектроскопии молекул и химических исследованиях: Тез.докл. Всесоюз. конф. Новосибирск, 1980, с.31-32.
3. ВИТЯЕВ Е.Е. Возможности использования теории измерений для обнаружения закономерностей и формирования понятий в психологии. -В кн.: Психологические аспекты повышения эффективности трудовой и учебно-воспитательной деятельности, Новосибирск, 1981, с.11-12.
4. КИОНОВАЛОВА Н.А. Прогнозирование швартообразования и тракционной отслойки сетчатки после диасклерального удаления внутриглазных осколков. -Автореф. дисс.... мед.наук. -М., 1981. - 25 с.
5. ЗАГОРУЙКО Н.Г., ЛЬОВ Г.С., МАШАРОВ Ю.П. Пакет прикладных программ для обработки таблиц экспериментальных данных ОТЭС-1. -В кн.: Вопросы обработки информации при проектировании систем (Вычислительные системы, вып. 69). Новосибирск, 1977, с.93-101.
6. ЛЬОВ Г.С., КОТЮКОВ В.И., МАНОХИН А.Н. Об одном алгоритме распознавания в пространстве разнотипных признаков. -В кн.: Вычислительные системы. Вып. 55. Новосибирск, 1973, с.108-110.
7. ЛЬОВ Г.С. Методы обработки разнотипных экспериментальных данных. -Новосибирск: Наука, 1981. - 160 с.
8. ЗАГОРУЙКО Н.Г. Эмпирическое предсказание. -Новосибирск: Наука, 1979. - 123 с.
9. КАРНАП Р. Философские основания физики. -М.: Прогресс, 1971. - 387 с.
10. ЗАГОРУЙКО Н.Г., САМОХВАЛОВ К.Ф., СВИРИДЕНКО Д.И. Логика эмпирических исследований. -Новосибирск, 1978. - 66 с.
11. ЗАГОРУЙКО Н.Г. Искусственный интеллект и эмпирическое предсказание. -Новосибирск, 1975. - 82 с. (Новосиб.Госуниверситет).
12. ЗАГОРУЙКО Н.Г. Методы обнаружения закономерностей. -М.: Знание, 1981. - 62 с.
13. ВИТЯЕВ Е.Е. Числовое, алгебраическое и конструктивное представления одной физической структуры. -В кн.: Логико-математические основы МОЗ (Вычислительные системы, вып. 107). Новосибирск, 1985, с. 40-51.
14. ВИТЯЕВ Е.Е. Закономерности в языках эмпирических систем и законы классической физики. -В кн.: Эмпирическое предсказание и распознавание образов (Вычислительные системы, вып.79). Новосибирск, 1979, с.45-56.
15. ВИТЯЕВ Е.Е. Метод обнаружения закономерностей и метод предсказания. -В кн.: Эмпирическое предсказание и распознавание образов (Вычислительные системы, вып.67). Новосибирск, 1976, с.54-68.

16. ВИТЯЕВ Е.Е. Обнаружение закономерностей, выраженных универсальными формулами. - В кн.: Эмпирическое предсказание и распознавание образов (Вычислительные системы, вып.79). Новосибирск, 1979, с.57-59.

17. ВИТЯЕВ Е.Е. Закономерности в языках эмпирических систем. - В кн.: Эмпирическое предсказание и распознавание образов (Вычислительные системы, вып. 76). Новосибирск, 1978, с.3-14.

18. ВИТЯЕВ Е.Е. Упрощение функциональных зависимостей за счет перешкалирования величин. - В кн.: II Всесоюз. школа-семинар по "Программно-алгоритмическому обеспечению прикладного многомерного статистического анализа", М., 1983, с.260-262.

19. ВИТЯЕВ Е.Е. Конструктивное числовое представление величин. - Настоящий сборник, с. 23-32.

20. Психологические измерения /Под ред. Л.Д.Мешалкина. -М.: Мир, 1967. - 120 с.

21. ФАНЦАГЛЬ И. Теория измерений. -М.: Мир, 1976. - 248 с.

22. Foundations of measurement. Vol.1. / Krantz D.H., Zuse R.D., Suppes P., Tversky A. - New-York and London: Academic press, 1971.- 577 p.

23. ЗАГОРУЙКО Н.Г., ЕЛИНА В.Н., ЛБОВ Г.С. Алгоритмы обнаружения эмпирических закономерностей. -Новосибирск: Наука, 1985. - 110 с.

24. КОСАРЕВ Ю.Г., МОСКВИТИН А.А. Проблемно-инструментальная технология построения программных систем. -Настоящий сборник, с. 59-76.

25. ДЕНИНГ В., ЭССИГ Г., МААС С. Диалоговые системы "Человек-ЭВМ": адаптация к требованиям пользователя. -М.: Мир, 1984.-112 с.

26. МОСТЕЛЛЕР Ф., ТЬЮКИ Дж. Анализ данных и регрессия. Вып. I. -М.: Финансы и статистика, 1982. - 319 с.

27. СВИРИДЕНКО Д.И. Проектирование  $\Sigma$ -программ. Постановка проблемы. -Настоящий сборник, с. 108-127.

28. CLOCKSIN W.F., MELLISH C.S. Programming in Prolog.- New York: Berlin Heidelberg, 1981. - 280 p.

29. CLARK K.L., MCCABE F.G. PROLOG: A Language for Implementing Expert Systems. - Machine Intelligence, 1980, v.10.

30. ЕРШОВ Ю.Л. Проблемы разрешимости и конструктивные модели. -М.: Наука, 1980. - 415 с.

31. МАЛЫЦЕВ А.И. Алгебраические системы. -М.: Наука, 1970. - 387 с.

32. Дискретная математика и математические вопросы кибернетики. Т. I /Под ред. С.В.Яблонского, О.Б.Лупанова. -М.: Наука, 1974. - 311 с.

33. Данные в языках программирования. Абстракция и типология. -М.: Мир, 1982. - 327 с.

34. ВИТЯЕВ Е.Е. Классификация как выделение группы объектов, удовлетворяющих разным множествам согласованных закономерностей. - В кн.: Анализ разнотипных данных (Вычислительные системы, вып. 99). Новосибирск, 1983, с.44-50.

35. ФИШБЕРН П.С. Теория полезности для принятия решений. -М.: Наука, 1978. -352 с.
36. КИНИ Р.Л., РАЙБА Х. Принятие решений при многих критериях: предпочтения и замещения. -М.: Радио и связь, 1981. -560 с.
37. КОЗЕЛЕЦКИЙ Ю. Психологическая теория решений. -М.: Прогресс, 1979. - 503 с.
38. KRANTZ D.H., TVERSKY A. Conjoint measurement analysis of composition rules in psychology. - Psychol.Rev., 1971,v.78,p.151-169.
39. ANDERSON M.H. Algebraic Rules in Psychological measurement.-Amer.Scientist, 1979,v.67,p.555-563.
40. ВИТЯЕВ Е.Е. О конструктивном представлении величин и законов.-В кн.: Конструктивные процессы: Тез. конф. Ижевск, 1984, с. 31-32.
41. НОВИКОВ В.Ф. Анализ экспертных предпочтений. -Настоящий сборник, с. 33-37.

Поступила в ред.-изд.отд.  
12 сентября 1985 года

ПРИЛОЖЕНИЕ  
Т а б л и ц а I

| Уро-<br>вень | Классы решаемых задач   | Реализация |           |
|--------------|---|------------|-----------|
|              |   | Алгоритмы  | Программы |
| I            | 2   | 3          | 4         |
| I            | <u>Формирование качественной теории</u>   |            |           |
|              | 1. Внесение в базу данных исходных данных, систем, понятий, признаков, величин, априорных знаний.                       | +          | +         |
|              | <u>II Формирование логической эмпирической теории</u>   |            |           |
|              | 1. Выделение отношений и операций (формирование словаря $v$ ).  | +          | +         |
|              | 2. Перевод данных на логический уровень.  | +          | +         |
|              | 3. Перевод априорных знаний в логическую форму:   |            |           |
|              | а) задание их формулами (I);  | +          | +         |
|              | б) задание их универсальными формулами, введением операций для кванторов существования и перевод их в формулы вида (I), | +          | -         |
|              | 4. Обнаружение закономерностей и внесение их в интеллектуальную базу данных:  |            |           |
|              | а) для фиксированной формулы (I) и заданных признаков;  | +          | +         |
|              | б) для формулы (I) в шкалах наименований, порядка и отношений.  | +          | +         |
|              | 5. Логический вывод.  | +          | -         |
|              | 6. Классификация объектов методом [34]  | +          | +         |
|              | 7. Предсказание в шкалах наименований, порядка и отношений методом [15].  | +          | +         |



| I | 2   | 3 | 4 |
|---|---|---|---|
| Ш | <u>Формирование количественной и конструктивной эмпирических теорий</u>   |   |   |
|   | 1. Определение шкал величин и функциональных зависимостей проверкой выполнимости одной из систем аксиом теории измерений или теории принятия решений: |   |   |
|   | - удаление из аксиом кванторов существования введением операций;  | + | - |
|   | - приведение универсальных формул к виду (I);   | + | - |
|   | - проверка формулы на реальных данных решением задачи П. 4, а.  | + | - |
|   | 2. Построение шкал и функциональных зависимостей:   |   |   |
|   | а) по данным;   | + | - |
|   | б) в диалоге с экспертом или планированием эксперимента.  |   |   |
|   | 3. Определение конструктивных числовых представлений величин и закономерных связей проверкой выполнимости системы аксиом (см. задачу Ш. I)            | + | - |
|   | 4. Построение конструктивных числовых представлений и закономерных связей:  |   |   |
|   | а) по данным;   | - | - |
|   | б) в диалоге с экспертом или планированием эксперимента.  | + | - |

Т а б л и ц а 2

| Решаемый класс задач                       | №          |
|--|------------|
| Анализ шкалы величин данных                | 1          |
| Классификация                              | 2          |
| Обнаружение закономерностей и предсказание | 3          |
| Конец работы                               | < RETURN > |

Т а б л и ц а 3

| Параметры настройки системы  |
|--|
| Число объектов обучения  |
| Число признаков в шкале порядка и отношений                                  |
| Число значений целевого признака   |
| Мах. размер массива единичных закономерностей                                |
| Мах. размер массива двойных закономерностей                                  |
| Мах. размер массива тройных закономерностей                                  |
| Мах. число коэффициентов при отношении порядка                               |
| Число контрольных точек  |
| Количество строк матрицы A   |
| Количество столбцов матрицы A  |
| Укажите формат исходных данных   |
| Сколько чисел в строке указанного формата                                    |
| Число градаций кодирующей матрицы  |
| Число коэффициентов целевого признака  |
| Введите число раскодированных признаков, вычисляемых по формуле              |
| $N2K = (IND(9) * (N2-1) * (IND(9) + 1)) / 2 + (IND(12) * (IND(12) + 1)) / 2$ |

Т а б л и ц а 4

| Параметры организации массивов данных                        |
|--|
| Будете(Y), нет(N) вводить основной массив данных             |
| Будете(Y), нет(N) формировать массив градаций                |
| Автоматически или вручную                                    |
| Будете(Y), нет(N) вводить массив коэффициентов при отношении |
| Будете(Y), нет(N) массив количества коэф. в признаках        |

Т а б л и ц а 5

## Определение конфигурации системы

| Обнаружение закономерностей |              |                 |                           |                        |                   |
|-----------------------------|--------------|-----------------|---------------------------|------------------------|-------------------|
| Шкала наименований          |              |                 | Шкала отношений и порядка |                        |                   |
| IND1                        | IND2         | IND3            | IND4                      | IND5                   | IND6              |
| Одинарные                   | Двойные      | Тройные         | Одинарные                 | Двойные                | Тройные           |
| Печать                      |              |                 |                           | Раскодировка           |                   |
| Шкала отношений и порядка   |              |                 |                           |                        |                   |
| Закономерностей             | Предсказание | Закономерностей | Предсказание              | Матрицы исходн. данных | Целевого признака |
| IND7                        | IND8         | IND10           | IND11                     | IND9                   | IND12             |

Т а б л и ц а 6

## Доверительные уровни для отбора закономерностей

|   |                                |  |
|---|--------------------------------|--|
| UP1-одинарных<br>UP2-двойных<br>UP3-тройных   | } шкала<br>наименований        | UP7- порог для условн.вероятн.<br>UP8-вероятн. для предсказания<br>в шкале порядка и отношений<br>UP9-доверительный уровень предсказания в шкале порядка и отношений |
| UP4-одинарных<br>UP5-двойных<br>UP6-тройных   | } шкала<br>порядка и отношений | UP10-доверительный уровень предсказания для условных вероятностей в шкале наименований<br>UP11-точность вычисления условной вероятности в шкале наименований         |
| UP1 UP2 UP3 UP4 UP5 UP6 UP7 UP8 UP9 UP10 UP11 |                                |  |

Т а б л и ц а 7

Вспомогательные параметры

Введите номер целевого признака  
 Вещественное число, кодирующее пробел  
 Значение целевого признака (целое)  
 Значение целевого признака (вещественное)  
 Будете(У), нет(Н) работать с прерываниями

Т а б л и ц а 8

Настроечные и управляющие параметры задачи №1 (табл.2)

Введите общее число признаков в закономерности  
 Введите число переменных, входящих в закономерность  
 Введите число строк в матрице А  
 Введите число столбцов (признаков) матрицы А  
 Введите число объектов обучения  
 Введите число, кодирующее пробел  
 Введите уровень критериев

Т а б л и ц а 9

Организация прерываний

| Вопрос                                   | Ответ |
|--|-------|
| Продолжить счет с прежними параметрами   | У/Н   |
| Переопределить параметры в данной задаче | У/Н   |
| Переформулировать задачу                 | У/Н   |
| Закончить работу                         | У/Н   |

## Инструкция для работы первой очереди системы ЛАДА

- Шаг 1: Запуск системы ЛАДА командой.
- Шаг 2: Выделение класса решаемых задач (табл.2).
- Шаг 3: Поиск необходимой конфигурации системы для данной задачи.
- Шаг 4: Если требуемая конфигурация найдена, то переход на шаг 7. В противном случае на шаг 5.
- Шаг 5: Синтез управляющей программы по данным табл.3 или 8.
- Шаг 6: Привязка управляющей программы к ОС РВ СМ ЭВМ (производится автоматически с выдачей пользователю диагностических сообщений и возможностью вмешиваться в процесс).
- Шаг 7: Ввод исходных данных и задание вспомогательных параметров (табл. 5-7).
- Шаг 8: Решение задачи одним из выбранных методов.
- Шаг 9: Анализ полученных результатов вычислений и принятие пользователем решения по ним.
- Шаг 10: Переход на шаг 2, 7 или 8 согласно ответу на один из вопросов табл. 9.
- Шаг 10: Интерпретация полученных результатов вычислений (окончательных) и запись их в базу данных и знаний.
- Шаг 11: Выход из системы с сохранением или без результатов вычислений и конфигурации системы.