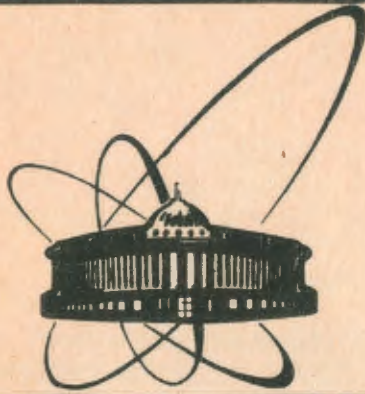


90-132



сообщения
объединенного
института
ядерных
исследований
Дубна

Φ 435

P11-90-132

Ф. Фернандес Нодарсе

**ОБ ИНТЕГРАЦИИ БАЗ ДАННЫХ
И ЭКСПЕРТНЫХ СИСТЕМ.
РАЗРАБОТКА СИСТЕМЫ EXDB**

1990

I. Введение

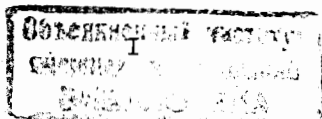
Реляционные базы данных (БД) и экспертные системы (ЭС) широко применяются в различных областях науки и техники, включая исследования в области физики высоких энергий. Это обстоятельство послужило хорошим стимулом к созданию более мощных систем, предназначенных для обработки больших баз знаний и фактов. Развитие систем обработки знаний (СОЗ) на основе СУБД стало естественным шагом дальнейшего прогресса в этой области.

Как результат объединения логического программирования и технологии баз данных, дедуктивные базы данных могут служить основой интеллектуальных прикладных программных систем. Развитие дедуктивных баз данных и экспертных систем до настоящего времени шло независимо, несмотря на то, что их общими корнями являются математическая логика и искусственный интеллект.

В настоящее время происходит развитие технологии создания систем на основе знаний и укрепления их связи с СУБД. Данную эволюцию можно разделить на три фазы: первая характеризовалась низкой совместимостью со стандартным программным обеспечением; вторая характеризуется большей совместимостью со стандартным аппаратным обеспечением и операционной системой, а также возможностью интерфейса с различными системами баз данных от dBASE до dBASE2. Наличие такого интерфейса позволяет запрашивать и обновлять внешние базы данных. Большая часть действующих в настоящее время shell (скелетных) систем находится в этой фазе. В третьей фазе, в середине 90-х годов, ожидается, что базы данных и системы на основе знания будут полностью интегрированы и совместимы со стандартным аппаратным и программным обеспечением.

Развивая систему екдв на основе знаний и используя СУБД как инфраструктуру, можно решить некоторые проблемы СОЗ. Большинство применений логических баз данных имеют серьезные ограничения при внедрении в существующих shell-системах.

Целью настоящей работы является исследование СОЗ новой архитектуры на основе создания прототипа системы и проверки его практических возможностей. Для этого была разработана система екдв, описываемая в данном сообщении.



II. Постановка задачи и принципы работы системы EXDB

Система EXDB ориентирована на решение следующей задачи: определение конфигурации микрокомпьютера в зависимости от характера его применения.

Основная идея системы формулируется следующим образом.

Исходя из ответов пользователя на вопросы, касающиеся стоимости, размеров оперативной памяти, файлов данных, требований к графическим устройствам и условий применения, требуется получить оптимальную конфигурацию микрокомпьютера.

Эта задача решается в несколько этапов.

На первом - на основе ответов пользователя на заданные ему вопросы система определяет предварительную конфигурацию микрокомпьютера и выдает ее на экран.

На втором - пользователь под руководством системы в автоматическом или полуавтоматическом режиме уточняет состав оборудования, предприятие-изготовитель, производящее микрокомпьютеры, подтверждает свои требования и, наконец, уточняет стоимость. После этого система выдает окончательную конфигурацию микрокомпьютера. Общение пользователя с системой производится в интерактивном режиме с помощью меню и других вспомогательных (HELP) возможностей. Как окончательные, так и предварительные результаты можно выдавать на печатающее устройство.

III. Архитектура EXDB

Общая архитектура системы EXDB показана на рис. I. Система состоит из следующих элементов:

- программы применения базы данных (ППБД);
- входа-диалога конечного пользователя;
- transaction базы данных, экранов и таблицы;
- экспертной подсистемы (ЭПС), состоящей из следующих компонент:
 - модуль извлечения;
 - машина логического вывода (МВ);
 - база знаний (БЗ);
 - экспликационный модуль;
 - модуль приобретения знания;
 - рабочее место;
- подсистемы базы данных (ПБД), в которую входят:
 - база данных;
 - система управления базой данных (СУБД);
 - модуль извлечения.

Архитектура логической базы данных отличается от обыкновенной применением машины вывода и базы знаний.

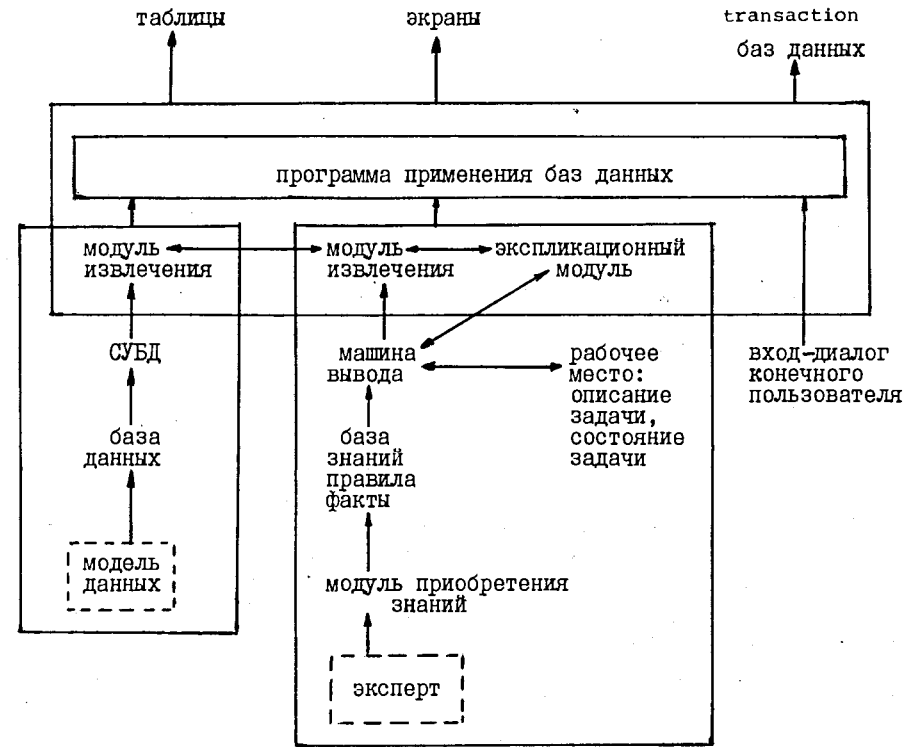


Рис. I. Общая архитектура системы EXDB.

Исходя из того, что СОЗ и СУБД различные, комбинируя SHELLE-системы продукции и СУБД, можно использовать слабую (сообщение есть, но каждая из частей поддерживает свое тождество) или сильную связь (построение интегрированной системы, используя компоненты обеих частей). В EXDB в основном используется слабая связь. Над реляционной СУБД *BASE III PLUS* имеется ППБД. В зависимости от характера применения (рис. 2), стоимости, емкости оперативной памяти, размеров файлов и требований к графике, ЭПС определяет предварительную конфигурацию микрокомпьютера в диалоговом режиме с возможностью хранения и сравнения промежуточных результатов. База знания содержит совокупность сведений о данной предметной области, то есть технические правила определения конфигурации микрокомпьютера в зависимости не только от требования совместимости каждой из компонент, но также и общей эффективности системы. База знаний была построена с помощью правил типа IF-THEN на основе дерева решений с коэффициентами уверенности, соответствующими некоторым узлам.

THE INTENDED APPLICATION(S) IS (ARE):

1. PROFESSIONAL ACCOUNTING OR SPREADSHEETS
2. HIGH COMPLEX DATABASE MANAGEMENT
3. LOW COMPLEX DATABASE MANAGEMENT
4. PROFESSIONAL WORD PROCESSOR
5. WORD PROCESSOR
6. HIGH COMPLEX GRAPHIC
7. LOW COMPLEX GRAPHIC
8. PROFESSIONAL CAD/CAM OR CAE
9. COMMUNICATIONS
10. SOFTWARE DEVELOPMENT
11. IMAGE PROCESSING
12. LOW RESOLUTION IMAGE PROCESSING
13. EXPERT SYSTEM
14. EDUCATION
15. ENTERTAINMENT
16. OFF LINE EXPERIMENTAL DATA PROCESSING
17. HIGH COMPLEX OFF LINE EXPERIMENTAL DATA PROCESSING
18. ON LINE EXPERIMENTAL DATA PROCESSING
19. HIGH COMPLEX ON LINE EXPERIMENTAL DATA PROCESSING
20. HIGH COMPLEX SCIENTIFIC CALCULATION
21. SCIENTIFIC CALCULATION
22. TRAINING & SIMULATION

TYPE YOUR CHOICE = _____

WILL YOU TYPE ANOTHER CHOICE (Y/N)? _____

Рис.2. Экран типов применений ПЭВМ.

Для того, чтобы поднять эффективность выполнения EXDB, под-
дерево, обеспечивающее обработку выбранной комбинации нескольких ти-
пичных предметных областей, обработали на ШБД через детерминирован-
ный конечный автомат, который был запрограммирован на СУБД dBASE III
PLUS. Полученные автоматом результаты поступают в ЭПС, которая была
построена на основе SHELL - системы EXSYS /6/. ЭПС, в свою очередь,
возвращает в ШБД предварительную конфигурацию микрокомпьютера. После
этого, с помощью ПБД, пользователь подтверждает свои требования, уточ-
няя состав оборудования и предприятие-изготовитель под руководством
системы в автоматическом или неавтоматическом режиме, и, наконец,
указывает стоимость. Для этого в ПБД имеются базы данных микрокомпью-
терного оборудования, в которых описываются и классифицируются каждый

элемент микрокомпьютера. В автоматическом режиме пользователь уточня-
ет конфигурацию оборудования в соответствии со списком компонент
предварительной конфигурации, неавтоматическом режиме - с помощью
системы экранов ПБД и списка компонент предварительной конфигурации.
Окончательные и предварительные результаты можно также посылатъ на
печатающее устройство.

Система EXDB выполняется в интерактивном режиме с помощью меню
и других вспомогательных возможностей (HELP), которые дают системы
dBASE III, ШБД и EXSYS. В системе продукции используется об-
ратная цепочка рассуждений. Комбинация вероятности конечной цели в
машине логического вывода вычисляется как среднее значение вероятнос-
тей истинных правил.

В программу применения логических баз данных имеется три входа
(рис.1):

- диалог с пользователем,
- непосредственное обращение к базам данных,
- непосредственное обращение к машине вывода.

Выходы ШБД - экраны, таблицы или TRANSACTION базы данных. Машина
логического вывода берет информацию из базы данных и базы правил.
Одно из различий между классическими СУБД и логическими базами дан-
ных в том, что несколько экранов входа и значений на экране были
сгенерированы с помощью машины логического вывода. База знаний являет-
ся сердцем интеллектуальной части логической базы данных системы

EXDB. В ней применяется явное обращение к машине вывода и базе зна-
ний из ПБД и ШБД, используя интерфейс для вызова соответствующей
процедуры. В дальнейшем предполагается исследовать возможности неяв-
ного подхода через активный словарь данных. В системе EXDB пользова-
телям не обязательно знать, какие значения хранятся, а какие получают-
ся с помощью правил. Обновление БД производится с помощью нескольких
специальных программ (EXDB update utilities). В настоящее
время есть тенденция к интеграции между модулем извлечения и програм-
мой применения базы данных. Для функционирования системы EXDB требу-
ется следующее оборудование: микрокомпьютер, совместимый с IBM PC,
с памятью 640 Кбайт RAM и диском типа "Винчестер".

IV. Некоторые объяснения пользователям EXDB

Как уже отмечалось, основная идея системы EXDB заключается в
следующем.

Исходя из ответов пользователя на пять вопросов (стоимость, размер опе-
ративной памяти, размер файлов, графические требования и область при-
менения), требуется получить оптимальную конфигурацию микрокомпьютера.

Предположим, для примера, что пользователь из предложенных ему
вариантов выбрал следующие ответы:

- THE INTENDED APPLICATION(S) IS(ARE) PROFESSIONAL CAD/CAM OR CAE
- COST IS A MAJOR FACTOR
- THE FILE STORAGE YOU EXPECT TO NEED IS 60 -> 80 MB
- GRAPHICS DATA WILL BE INPUT AND OUTPUT FROM THE COMPUTER
- THE LARGER FILE YOU EXPECT TO USE CAN BE DIVIDED INTO SECTION <= 720 KB

Рис.3. Пример возможных ответов пользователя на вопросы EXDB.

Для выбора микрокомпьютера система использует процессы логического вывода, эвристики и знания эксперта (рис.4).

Система EXDB

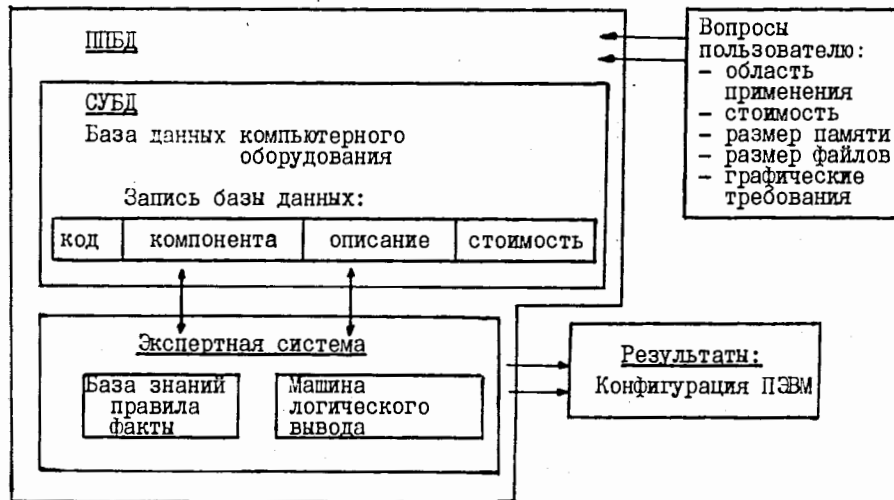


Рис.4. Принципиальная схема системы EXDB.

В системе EXDB имеется реляционная база данных микрокомпьютерного оборудования. В каждой записи имеется имя компонента, ее код, общее описание и стоимость. Если мы организуем такую базу данных с помощью традиционной СУБД, то один запрос к базе данных не позволяет извлечь оптимальную конфигурацию микрокомпьютера. Для этого необходимо провести дедуктивное рассуждение с фактами из базы данных (см. рис.3).

Возможность проведения дедуктивного рассуждения, которое позволяет установить отношения между записями файла вывода для создания файла конфигураций микрокомпьютера, обеспечивают база знаний и машина вывода экспертной подсистемы. Решить эту задачу, используя только ЭИС, очень сложно и не оптимально. Поэтому ЭИС строит только предва-

рительную конфигурацию на уровне типа его элементов. Пример такой конфигурации, соответствующей запросам пользователя, приведенным на рис.3, показан на рис.5.

THOSE ITEMS WITH A SCORE OF 10 ARE ESSENTIAL TO THE INTENDED APPLICATION. ITEMS WITH A SCORE OF LESS THAN 10 WOULD THE APPLICATION EASIER OR FASTER, OR MEANS ANOTHER SOLUTION, BUT ARE NOT ESSENTIAL.

Press any key to display results:

Values based on 0 - 10 system	VALUE
1 IBM AT COMPATIBLE (80286) 8 -> 12 MHZ	10
2 1 MB RAM	10
3 60 -> 80 MB HARD DISK + ADAPTER	10
4 COPROCESSOR 80287	10
5 FLOPPY DISK DRIVE 5" 2SHD(1,2MB) + ADAPTER	10
6 FLOPPY DISK DRIVE 3" (720KB) + ADAPTER	10
7 HIGH RESOL.COLOR GRAPHICS DISPLAY + ADAPTER	10
8 MEDIUM RESOL.COLOR GRAPHICS DISPLAY + ADAPTER	10
9 PARALLEL GRAPHICS PRINTER 120 COLUMN	10
10 PLOTTER,FORMAT A3 - A2	10
11 GRAPHICS TABLET,FORMAT A2 - A3	10
12 40 MB TAPE BACK UP + ADAPTER	10
13 SERIAL RS 232 INTERFACE	10
14 PARALLEL CENTRONIC INTERFACE	10
15 IBM S/2 M 70/80 (80386) 16 -> 25 MHZ	9
16 2 -> 4 MB RAM	9
17 COPROCESSOR 80387	9
18 MOUSE	9
19 FLOPPY DISK DRIVE 5" 2S2D(360KB) + ADAPTER	8
20 LOW RESOL MONOC GRAPHICS DISPLAY + ADAPTER	7

Press any key for more:

Рис.5. Предварительная конфигурация ПЭВМ, выданная системой EXDB.

Пользователь с помощью ПБД уточняет предложенную системой предварительную конфигурацию микрокомпьютера, после чего создается файл вывода с окончательной конфигурацией (рис.6).

05/09/89

<u>CONFIGURATION SELECTED</u>		
<u>EQUIPMENT</u>	<u>BRAND</u>	<u>PRICE</u>
- AT - IBM	IBM	3560.0
- 1 MB RAM	INTEL	380.0
- 6085 70MB 28 mSEC HARD DISK + ADAPTER	MINISCRIBE	819.0
- INTEL COPROCESSOR 80287 8 MHZ	INTEL	495.0
- NDO8DE- G FLOPPY DRIVE 5" (1,2MB)+. ADAPTER	TOSHIBA	119.0
- FDD4403 FLOPPY DISK DRIVE 3" (720KB) + ADAPTER	TOSHIBA	129.0
- HIGH RESOL.MONOC.GRAPHICS DISPLAY + ADAPTER	NEC	999.0
- TAXAN 440 COLOR GRAPHICS DISPLAY + ADAPTER	TAXAN	390.0
- EX 1000 PARALLEL GRAPHICS PRINTER 120 COLUMNS	EPSON	725.0
- PLOTTER A3 DMP 41	HOUSTON INSTRUMENT	2549.0
- GRAPHICS TABLET,FORMAT A2 - A3	CALCOMP	2009.0
- BERNOULLI 40 MB TAPE BACK UP + ADAPTER	IOMEGA	1648.0
- ACA 2 SERIAL PORTS	ACA	60.0
- PARALLEL/SERIAL CARD	ADVANCED	69.0
*** Total ***		13951.0

Рис.6. Уточненная конфигурация ПЭВМ, выданная системой EXDB.

В автоматическом режиме предварительная конфигурация используется как скелет для создания окончательной. Из каждого элемента предварительной конфигурации, например, 60 --> 80 MB HARD DISK + ADAPTER (см. рис.5), используя систему меню (рис.7), пользователь выбирает 6085 70MB 28mSec HARD DISK + ADAPTER BRAND: MINISCRIBE (см.рис.6).

SELECT YOUR CONFIGURATION

<u>CODE</u>	<u>EQUIPMENT</u>	<u>PRICE</u>
1401	85 MB 28 mSec HARD DISK+ADAPTER BRAND:MICROPOLIS	848.00
1402	ST4096 80 MB FULL HT HARD DISK+ADAPTER BRAND:MINISCRIBE	799.00
1403	6085 70 MB 28 mSec HARD DISK+ADAPTER BRAND:MINISCRIBE	819.00

SELECT CODE:1403

EQUIPMENT SELECTED:6085 70 MB 28 mSec HARD DISK+ADAPTER

WILL YOU SELECT ANOTHER CHOICE (Y/N)? : _____

Рис.7. Пример экрана, полученного в автоматическом режиме.

В неавтоматическом режиме с помощью иерархической системы меню (рис.8 и 9), пользователь имеет возможность получить исходный экран рис.7. В этом случае предварительная конфигурация имеет только вспомогательный характер.

<u>MAIN MENU</u>	
1.- MICROCOMPUTERS	7.- PRINTERS
2.- MONITORS AND TERMINALS	8.- MODEMS
3.- COPROCESSORS	9.- DIGITIZERS
4.- MEMORIES	10.- PLOTTERS
5.- BOARDS	11.- MISCELANEUS
6.- HARD, FLOPPY AND TAPE DEVICES	
-----> TYPE OPTION: 6	

Рис.8. Общее меню неавтоматического режима.

В обоих режимах результаты выбора автоматически переходят в состав файла окончательной конфигурации.

У. Методология разработки системы EXDB

Методология разработки системы EXDB подобна той, которая используется в процессе создания классических ЭС и СУБД, и весь процесс с методологической точки зрения можно разделить на следующие пять этапов:

1 - "Идентификация": определение типа задачи, широты постановки, участников процесса разработки (см. рис.10), необходимых ресурсов и целей системы.

2 - "Концептуализация" : выбор понятий, отношений и механизмов управления, необходимых для описания решения задачи в выбранной области. Исследование подзадач, стратегии, ограничений и требований к решению.

SUBMENU

CHOICE	DESCRIPTION
12	200 --> 400 MB HARD DISK + ADAPTER
13	100 --> 200 MB HARD DISK + ADAPTER
14	60 --> 80 MB HARD DISK + ADAPTER
15	10 --> 40 MB HARD DISK + ADAPTER
16	FLOPPY DISK DRIVE 5" 2S2D(360 KB) + ADAPTER
17	FLOPPY DISK DRIVE 5" 2SHD(1,2MB) + ADAPTER
18	FLOPPY DISK DRIVE 3" (720 KB) + ADAPTER
19	60 MB TAPE BACK UP + ADAPTER

SELECT ONE CHOICE: 14

Рис.9. Подменю неавтоматического режима, полученное из рис.9.

3 - "Формализация" : формальное выражение ключевых понятий и отношений в соответствии с организацией языков ЭС EXSYS и СУБД DBASE III. Определение и нормализация структуры баз данных ПБД.

4 - "Реализация" : сформулированное знание превращается в работающую компьютерную программу. Написание программы потребовало содержание, форму и согласования. В содержание включается структура данных, правила логического вывода и стратегия управления, необходимые для решения задачи. Форма определяется инструментом, то есть SHELL системой и СУБД, которые мы выбрали. Согласование означает комбинирование и реорганизацию различных порций знаний БЗ с целью устранения глобальных неувязок между спецификациями структур данных, БД, правил и схем управления в последовательном уточнении. Было проверено решение начальных этапов.

5 - "Тестирование" : оценка EXDB, которую проверяли вместе инженер знаний и эксперт. На этой фазе еще раз уточнялись правила ПЭС, и, когда это было необходимо, переопределялась структура, переформулировались понятия или/и определялись новые требования, повторяющие процесс с начала.

VI. Заключение

Весьма вероятно, что главные логические применения баз данных будут результатом развития технологии реляционных баз данных. Методология, описанная в этой статье, была полезна для разработки систем EXDB и мы рекомендуем ее для такого типа применений.

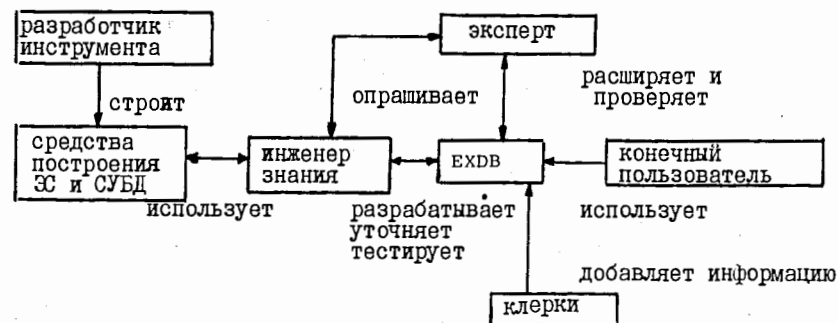


Рис.10. Участники создания EXDB.

Несмотря на существующие ограничения ЭС SHELL и СУБД в области их совместимости и интерфейса, та или иная форма интеграции при применении дает новые интеллектуальные возможности для обработки баз данных. Использование машины вывода и базы знаний совместно с базой данных дает возможность логического вывода новых данных из первоначальных или исходных.

Литература

1. Tzvielli A., "Emulating OPS5 with a data base", AI EXPERT, 1989, vol. 4, p.32-41.
2. Cohen B., "Merging Expert Systems and Databases", AI EXPERT, 1989, vol 4, p.22-31.
3. Schur S., "The Intelligent Database", AI EXPERT, jan 1988, p.26-34.
4. Sciore E. and Warren D.S., "Integrating Database & Prolog", AI EXPERT, jan 1988, p.38-44.
5. Hayes-Roth F., Waterman D.A., Lenat D.B., "Building expert systems", Addison-Wesley, 1983.
6. EXSYS V. 3 Expert System, Development Packages.
7. Ashton Tate, DBASE III PLUS, User Guide.

Рукопись поступила в издательский отдел
23 февраля 1990 года.