

Ф-435

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

На правах рукописи

10-92-432

Франсиско Фернандес Нодарсе

УДК 681. 3: 539. 12

**СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ
РЕШЕНИЙ В МОДЕЛИ
КОМПЬЮТЕРНОЙ ОБРАБОТКИ ЭКСПЕРИМЕНТА**

Специальность: 05.13.16 - Применение вычислитель-
ной техники, математического моделирования и ма-
тематических методов в научных исследованиях

Автореферат диссертации на соискание ученой степени
кандидата физико-математических наук

Дубна 1992

Работа выполнена в Лаборатории вычислительной техники и автоматизации Объединенного института ядерных исследований (ОИЯИ, г. Дубна).

Научный руководитель

Кандидат физико-математических наук
Старший научный сотрудник

В. Г. Иванов

Официальные оппоненты:

Доктор физико-математических наук

Глаголев
Виктор Викторович

Кандидат технических наук

Исмаилова
Лариса Юсифовна

Ведущая организация:

Институт проблем кибернетики, РАН, Москва

Защита состоится "____" _____ 1992 г. в "____" часов на заседании специализированного совета Д047.01.04 при Лаборатории вычислительной техники и автоматизации Объединенного института ядерных исследований, 141980 г. Дубна, Московской области.

С диссертацией можно ознакомиться в библиотеке ОИЯИ.

Автореферат разослан "____" _____ 1992 г.

Ученый секретарь

Специализированного совета

Кандидат физико-математических наук

Иванченко
Иванченко
Зинаида Мироновна

1. Общая характеристика работы

Опыт подготовки и проведения электронных экспериментов на больших коллайдерах показывает, что одним из важных этапов процесса подготовки и планирования эксперимента является моделирование процесса офф-лайн обработки данных с целью оценки требующихся вычислительных ресурсов и выбора на этой основе оптимального варианта компьютерной среды эксперимента.

Современные системы обработки экспериментальных данных в физике высоких энергий (ФВЭ) являются большими и сложными распределенными вычислительными комплексами, предназначенными для сбора и накопления данных, обработки результатов регистрации событий, интерактивного анализа результатов, получаемых сотнями специалистов из десятков институтов различных стран. В качестве примера на Рис. 1 приведена схема типичного компьютерного комплекса эксперимента на электрон-позитронном коллайдере ЦЕРНа.¹

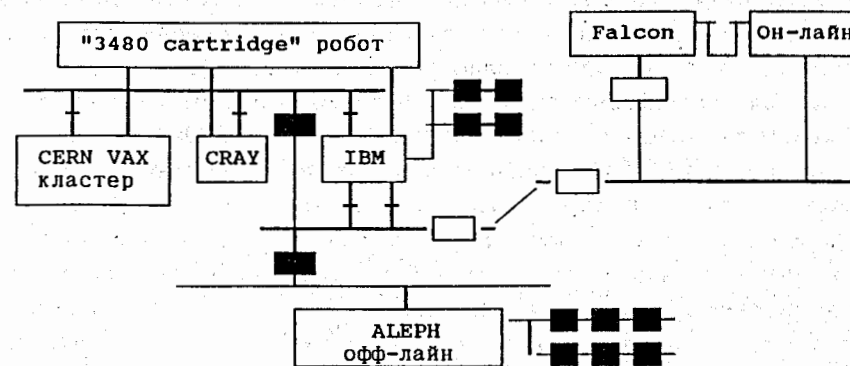


Рис. 1. Схема компьютерного комплекса эксперимента ALEPH

¹M. Delfino et al.-In: Proc. of Int. Conf. on Computing in HEP'91. Universal Academy Press, Inc., Japan, Tsukuba, 1991, p.115-117

В состав комплекса входят следующие подсистемы, объединенные на основе сети: а) VAX-кластер для сбора данных, б) локальная сеть VAX-кластера для реконструкции событий (Falcon), в состав которой входят две системы типа микроVAX, файловые серверы и 14 рабочих станций, в) сеть рабочих станций (VAX-кластер) для физического анализа в интерактивном режиме (ALEPH офф-лайн), которая включает 90 узлов, 50 Гигабайт памяти на диске и файловые серверы, г) IBM 3090/600 и CRAY XMP/48, обеспечивающие хранение данных на машинных носителях, использование "cartridge" робота и другие задачи.

Появление на компьютерном рынке недорогих рабочих станций типа RISC, создание на их основе мощных вычислительных систем, стоимость которых, по крайней мере, на порядок меньше стоимости больших ЭВМ такой же производительности, коренным образом изменило характер вычислительной среды эксперимента. Основным источником вычислительной мощности становятся распределенные системы на базе рабочих станций, а большие ЭВМ все чаще начинают играть роль серверов данных. Кроме того, рынок компьютеров развивается весьма интенсивно. Поэтому для определения состава и структуры вычислительной среды эксперимента с учетом имеющихся компьютерных средств общего назначения необходимо хорошее знание не только методов обработки и организации работы больших коллабораций, но и компьютерного рынка, включая перспективы его развития. Кроме того, имеются неопределенности в оценках вычислительных ресурсов физических экспериментов в ФВЭ, основывающиеся на следующих факторах: а) невозможно детально планировать вычислительные ресурсы физических экспериментов, т.к. многое может меняться в процессе их подготовки и проведения; б) скорость набора событий может значительно изменяться по сравнению с планируемой; в) места, где должны выполняться различные вычислительные задачи, могут меняться; г) некоторые из оценок могут оказаться настолько большими, что исследователи постараются их уменьшить; д) системы обработки — это системы с обратной связью, так как исследователи начинают лучше понимать некоторые проблемы работы с данными и на этой основе корректировать процесс анализа. Так как задача планирования связана со многими динамично изменяющимися факторами,

то для ее решения требуется программная поддержка.

Поскольку в ОИЯИ, как и в других физических центрах, планируется сооружение крупных экспериментальных установок, а также участие сотрудников института в экспериментах на больших коллайдерах, то необходимо иметь программные средства для моделирования процессов офф-лайн обработки данных с больших электронных установок и для выбора необходимых компьютерных комплексов. Поэтому разработка соответствующей программной системы поддержки принятия решений (СППР) является в настоящее время актуальной задачей.

Целью разработанной СППР COMEX является решение следующих задач, связанных с планированием (определением) компьютерных ресурсов для проведения реальных экспериментов:

1. ознакомление пользователей с методами офф-лайн обработки экспериментальных данных и вычислительными системами (обучающее и информационное обеспечение);
2. оценка требующихся вычислительных ресурсов и объемов информации;
3. моделирование процессов офф-лайн обработки данных с больших электронных установок;
4. выбор оптимальной конфигурации компьютерного комплекса для обеспечения экспериментов в режиме диалога под руководством системы или пользовательский выбор.

С точки зрения методов решения, задача была поставлена как проблема, характеризующаяся неточно и неполно определенными данными с использованием эвристических знаний для определения значений части переменных модели и её последующего решения. Структура нашей задачи допускает разбиение исходной задачи на подзадачи, каждую из которых можно решать относительно независимо от других, но подзадачи могут взаимодействовать, в результате чего решение одной подзадачи будет зависеть от решений остальных. Для ее реализации потребовалось исследовать и создать интегрированную экспертную СППР, базирующуюся на знаниях и взаимосвязи с реляционной СУБД, и моделировать предметную

область с использованием алгоритмических и неалгоритмических знаний. Для создания "компьютерной модели"² автором рассмотрена проблема соединения математической модели с экспертной СППР путем единого представления модели объекта и создана на её основе СППР СОМЕХ. Кроме того, необходимо было разработать информационное обеспечение.

В состав информационного обеспечения СППР СОМЕХ входят:

- методика анализа информации, получаемой с помощью современных экспериментальных установок, используемых в физике высоких энергий;
- описание моделей обработки экспериментальных данных с электронных установок;
- характеристики компьютерных средств и систем, применяемых для обработки экспериментальных данных в физике высоких энергий;
- описания методов анализа экспериментальных данных с различных электронных установок;
- методика и примеры оценки вычислительных ресурсов, требующихся для проведения экспериментальных исследований;
- прогнозы развития различных компьютерных средств и систем, включая средства связи;
- стоимость и методы оценки компьютерных средств с учетом изменения цен в течение нескольких последующих лет.

СППР СОМЕХ обеспечивает автоматизацию работы пользователя на всех этапах процесса подготовки и принятия решения, таких как:

1. подготовка входных данных;

²В. А. Геловани, В. В. Юрченко. Компьютерное моделирование.- Математическое моделирование, М.: Академия Наук СССР, 1989, Т.1, N.1, с.3-12

2. накопление и хранение информации, требующейся в процессе моделирования и принятия решения;
3. моделирование разных вариантов;
4. выбор оптимального варианта;
5. проверка решения.

В процессе проектирования распределенных неоднородных вычислительных комплексов, представляющих собой совокупность взаимосвязанных различных технических и программных средств, необходим компромисс между такими факторами, как улучшение технических и программных средств, сокращение их стоимости и одновременно повышение требований к компьютерным средствам, необходимым в экспериментах особенно в физике высоких энергий. Основным принципом создания компьютерного комплекса, который используется с СППР СОМЕХ, является принцип модульной структуры компьютерного комплекса на базе серийно выпускаемых унифицированных технических и программных средств вычислительной техники с применением стандартных аппаратно-программных интерфейсов.

Модель включает в основном две подмодели: одна из них характеризует типичную задачу для решений, т.е. физический эксперимент, а вторая - стратегию (принципы) проектирования компьютерных комплексов. Полученные решения, т.е. компьютерные комплексы, отражают принципы проектирования и зависят от баз данных процессоров и внешних устройств, на основе которых СППР моделирует решения.

2. Научная новизна

Разработка интегрированной системы на основе знаний для решения задач оптимального выбора компьютерного комплекса и создание компьютерной модели области применения являются оригинальным и эффективным решением проблемы планирования компьютерного комплекса, обеспечивающего научные эксперименты в области физики высоких энергий. Компьютерная модель области

применения представлена в виде модели классификации. Она соединяет математическую модель с экспертной СППР путем единого представления модели объекта. Эта интегрированная СППР использует различные средства представления данных и знаний, различные средства интерфейса и различные методы решения подзадач в многоуровневой схеме принятия решений.

3. Практическая значимость

СППР COMEX обеспечивает следующие функциональные возможности:

1. выбор модели процесса офф-лайн обработки,
2. оценку вычислительных ресурсов и объемов информации,
3. выбор схемы компьютерного комплекса эксперимента,
4. выбор оптимальной конфигурации компьютерного комплекса в режиме диалога под руководством системы или пользовательский выбор,
5. информационное и обучающее обеспечение.

С помощью СППР COMEX смоделирован процесс офф-лайн обработки данных в распределенных системах с учетом имеющихся ресурсов и выбраны различные варианты компьютерных комплексов для экспериментов ALERH, L3, DELPHI, ГИПЕРОН, ЭКСЧАРМ и МАРС-МЧС. С помощью СППР COMEX также был разработан проект развития вычислительного центра Ядерного Исследовательского Центра в Гаване. Она также используется для решения задач, связанных с оптимальным выбором компьютерного комплекса Исполнительного секретариата по ядерным вопросам на Кубе. Применение СППР COMEX позволяет получать оптимальные и обоснованные результаты. Система позволяет пользователю после знакомства и небольшой тренировки на имеющихся примерах решать задачу моделирования варианта комплекса за 1-2 часа работы на ПЭВМ. Чем лучше пользователь освоит область применения, тем более эффективно будет её моделирование и использование СППР COMEX.

4. Структура работы

Работа состоит из введения, трех глав, заключения, списка цитируемой литературы. Работа содержит 161 страницу машинописного текста и включает 64 рисунка и 3 таблицы. Список литературы включает 71 наименование.

5. Содержание работы

Во введении обосновывается актуальность исследований, формулируются основные положения, выносимые на защиту, их научная новизна и практическая значимость.

В первой главе обсуждаются вопросы компьютерного моделирования с помощью СППР. В частности, обосновывается положение о необходимости разработки экспертных СППР и даются их общие характеристики. Обсуждается проблема представления знаний, вопрос о поиске решения задачи, выборе средства для построения экспертных СППР и ее архитектуре.

Для СППР COMEX необходим персональный компьютер IBM PC/AT или совместимый с ним, с оперативной памятью не менее 640 Кб, с жестким диском и с операционной системой DOS 3.0 или выше.

СППР COMEX в основном была написана на основе языка реляционной СУБД FOXBASE +, а также используются возможности, которые обеспечивают LOTUS 123, редакторы текстов (например SHIWRITER, MS -MicroStart), Pascal и пустая (shell) экспертная система, у которой основным средством представления знаний являются продукционные правила с коэффициентом вероятности / уверенности, с прямой и обратной цепочкой рассуждений, и которая имеет возможности обращения к внешним подпрограммам.

Во второй главе подробно описывается компьютерная модель области применения, представленная в виде модели классификации. Далее описываются основные характеристики СППР COMEX и стратегии решения задачи, а также объясняется архитектура системы. Общая методика работы СППР COMEX показана на Рис. 2.

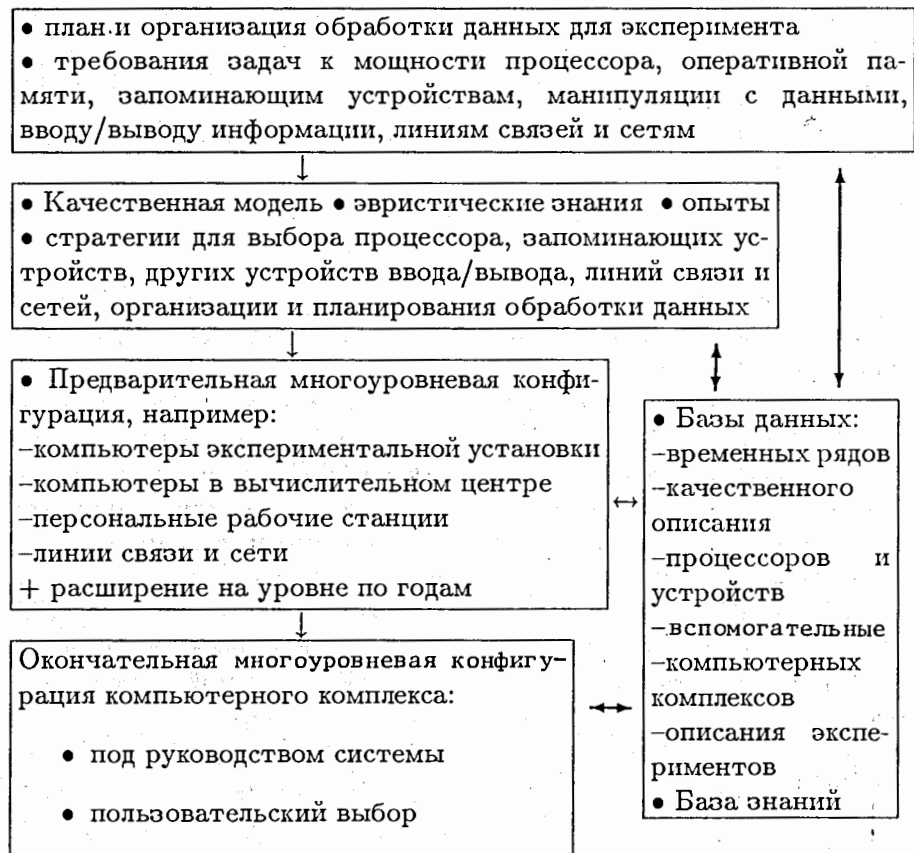


Рис. 2. Общая методика работы СППР COMEX

СППР COMEX состоит из следующих основных компонентов:

1. средств для создания, доступа, обновления и обработки баз данных (БД);
2. средств для работы с базой знаний (БЗ): редактора и отладчика;

3. механизма получения решения задачи (МР) или вывода, который содержит интерпретатор правил, диспетчер, средства обращения к ассоциированным подпрограммам, предназначенным для выполнения подзадач принятия частичных или глобальных решений, связанных с приобретением данных и с выбором процессоров, внешних устройств, характеристик линий связи и сетей и т.д.;
4. средств обработки текстов и гистограмм. Во время постановки СППР пользователь может выбирать программу обработки текстов. Кроме того, пользователь имеет возможность использовать графическую систему;
5. средств для интерфейса с пользователем, использующих в основном интерактивные меню и электронные таблицы для ввода / вывода данных, модуль объяснения и другие вспомогательные средства;
6. баз данных, в том числе БД оборудования (процессоры, внешние запоминающие устройства, устройства ввода / вывода), БД описания экспериментов, БД текстов статей, связанных с экспериментом, БД конфигураций компьютерных комплексов разных экспериментов, БД временных рядов (исходных данных) и временных и вспомогательных баз данных, необходимых в процессе решения задачи;
7. базы знания, предназначенной для решения проблем оптимального или близкого к оптимальному выбора компьютерного комплекса.

Следует также отметить, что между этими частями нет абсолютной границы. Принципы работы МР тесно связаны со способами представления знаний в БЗ. В любом случае МР содержит алгоритм получения решения, т.е. специальное алгоритмическое знание. Так как в МР содержится, по крайней мере, часть семантики БЗ, определенная в процедурной форме, это подтверждает относительность границы между МР и БЗ. Для описания модели прикладной области используются уравнения, неравенства и другие математические конструкции с соответствующими методами получения

оптимального или близкого к оптимальному решения, а также применяются средства представления знаний в виде правил и объяснений получаемых с их помощью решений. Правила включают такие элементы, как область определения, область действия, условия применения, условия выбора, действие и др.

В процессе решения задачи появляются частичные результаты (решения), содержащиеся во вспомогательной базе данных СППР СОМЕХ. Одним из них является предварительная конфигурация компьютерного комплекса. С использованием этих частичных результатов как входных данных, под управлением модуля контроля (диспетчера) выполняются прикладные программы для дополнительного накопления необходимых данных и для принятия частичных или глобальных решений, связанных с выбором окончательной (оптимальной) конфигурации компьютерного комплекса.

СППР СОМЕХ имеет возможность хранения требований к компьютерному комплексу и различных вариантов решения, доступа к базам данных процессоров и внешних устройств, установления заранее имплицитных (экспертных) значений параметров и их модификации.

Третья глава посвящена вопросам применения СППР СОМЕХ для моделирования процесса "офф-лайн обработки" в ФВЭ, определения оптимальной конфигурации вычислительного комплекса, выявления узких мест и принятия мер по их устранению. Описываются схема решения задачи, общие принципы проектирования компьютерного комплекса для обработки экспериментальных данных, модель области применения и применение системы поддержки принятия решений (СППР) СОМЕХ для выбора оптимальной конфигурации компьютерного комплекса на примере совместного ЦЕРН-ОИЯИ эксперимента DELPHI. Блок-схема процесса оценки вычислительных ресурсов для электронного эксперимента, разработанная с учетом возможности её программной реализации, приведена на Рис. 3.

Практика показывает, что необходимо моделировать стратегию решения задачи, т.е. оптимизировать рабочие методы (организацию обработки) и соответствующий выбор оптимальной конфигурации компьютерного комплекса с учетом имеющихся ресурсов и других особенностей. Данные используются вместе с пользователь-

ской обратной связью для обеспечения моделирования и внесения изменений в решение. Используется многоуровневая схема принятия решений.

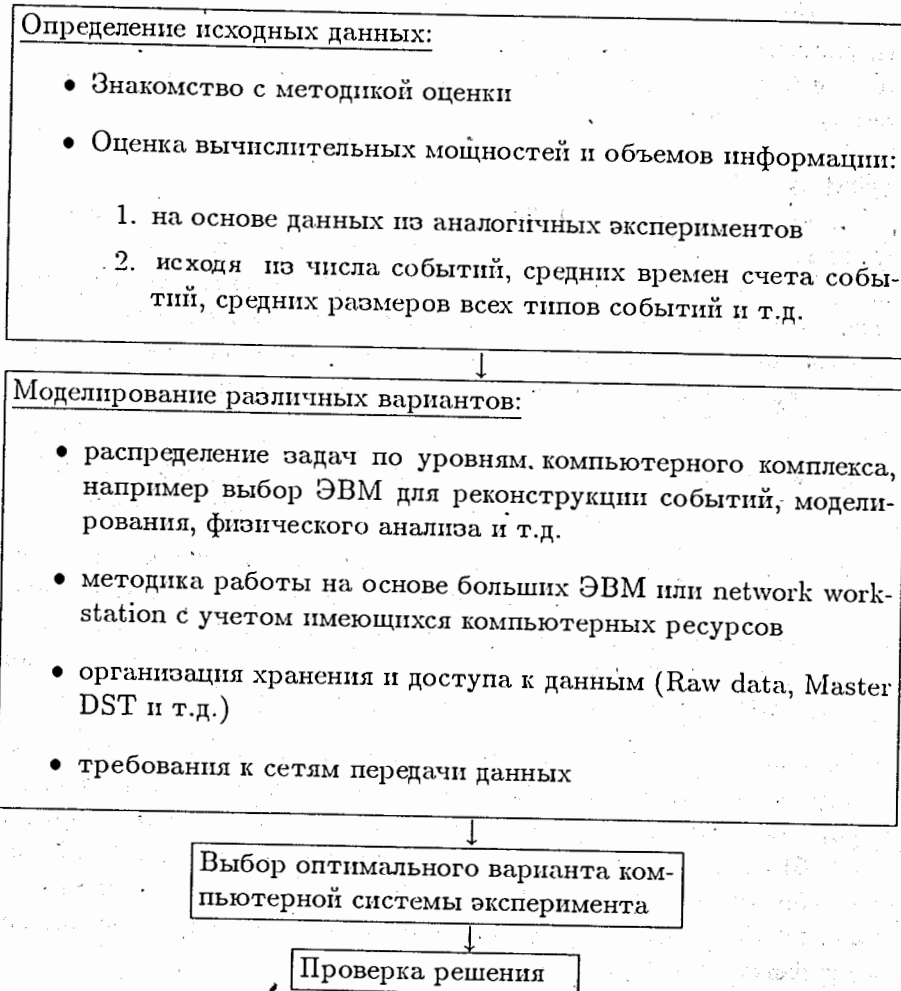


Рис. 3. Блок-схема процесса оценки вычислительных ресурсов

В процессе проектирования компьютерного комплекса СППР СОМЕХ работает в два этапа: выбор предварительного общего со-

става с использованием экспертных оценок и определение окончательного состава компьютерного комплекса. Сравнение вариантов выполняется с помощью математической модели принятия решения по критерию эффективности, в нашем случае это соотношение характеристика/стоимости. СППР COMEX также учитывает коэффициент изменения годовой стоимости каждого типа устройства и процент ошибок во входных данных, который влияет на уровень определенности в результатах.

В конце главы анализируются результаты применения СППР COMEX.

В заключении кратко сформулированы основные результаты работы. Разработка интегрированной системы на основе знаний для решения задач оптимального выбора компьютерного комплекса и создание компьютерной модели являются эффективным решением проблемы планирования компьютерного комплекса, обеспечивающего научные эксперименты в области физики высоких энергий. В итоге:

1. разработаны принципы создания СППР для решения такого типа задач с применением средств представления знаний;
2. создана компьютерная модель области применения, позволяющая описывать не только количественные, но и качественные параметры задачи. Она представлена в виде модели классификации и соединяет математическую модель с экспертной СППР путем единого представления модели объекта;
3. разработана и создана интегрированная развивающаяся СППР, состоящая из следующих основных компонентов: а) средств для создания, доступа, обновления и обработки баз данных; б) средств для работы с базой знаний; в) средств для обработки текстов и графики; г) баз данных; д) базы знаний; е) механизма получения решения задач; ж) средств для интерфейса с пользователем;
4. создано необходимое информационное обеспечение;
5. СППР COMEX применена для решения ряда задач, связанных с планированием (оценкой) компьютерных ресурсов, моделированием процесса офф-лайн обработки экспериментов

в ФВЭ и оптимальным выбором компьютерных средств для проведения вычислительных экспериментов в ОИЯИ и на Кубе. Применение СППР COMEX позволяет получать оптимальные и обоснованные результаты.

Апробация и публикации

Материалы диссертации опубликованы в работах /1-10/, выполненных автором в 1986-1992 годах.

Основные результаты доложены и обсуждены на III Международной конференции "Программное обеспечение ЭВМ" /Тверь, 1991/; IV научном семинаре с международным участием "Автоматизированные библиотечно-информационные системы" /Академия наук СССР, Государственная публичная научно-техническая библиотека, Новосибирск, 1991/, и "Second International Workshop on software engineering, Artificial Intelligence and Expert System for High Energy and Nuclear Physics" under sponsorship of the European Physical Society /L'Agelonde, France, 1992/ (Proc. of Conf., "World Scientific", in press).

Основные результаты опубликованы в следующих работах:

1. F. Fernandez Nodarse, V. G. Ivanov. Blackboard architecture and qualitative model in a computer aided assistant designed to define computers for HEP computing.-PREPRINT JINR E11-91-560, Dubna, 1991.
2. Ф. Фернандес Нодарсе. Об интеграции баз данных и экспертных систем. Разработка системы EXDB.-Сообщение ОИЯИ P11-90-132, Дубна, 1990.
3. В.Г.Иванов, Ф. Фернандес Нодарсе, А.Г. Белов, Х. Зуньига. Автоматизированное рабочее место оператора микротрона МТ-25. Концепция и основные функции.- Сообщение ОИЯИ P10-91-302, Дубна, 1991.
4. В.Г.Иванов, Ф. Фернандес Нодарсе, Т. Эрдэнэдэлгер. Назначение и структура программного обеспечения автомати-

- зированного рабочего места в распределенной системе обработки экспериментальных данных.- Сообщение ОИЯИ P10-91-105, Дубна, 1991.
5. F. Fernandez Nodarse. NETSYS, an AI system building tool.- JINR Communication E10-90-474, Dubna, 1990.
 6. Ф. Фернандес Нодарсе, В. Г. Иванов, Г. Балунова, С. Лима. Примеры разработки интеллектуальных программных систем.- Краткие сообщения ОИЯИ No 7 (46)-90, Дубна, 1990.
 7. F. Fernandez Nodarse. Automated Systems and Computer Techniques in cuban's nuclear development.- Nucleus (Cuba), 1986, N.1, p. 49 - 52.
 8. Ф. Фернандес Нодарсе. Модель компьютерной обработки экспериментов в Системе Поддержки Принятия Решений COMEX.- Сообщение ОИЯИ P11-92-335, Дубна, 1992.
 9. В.Г. Иванов, Ф. Фернандес Нодарсе. Система Поддержки Принятия Решений COMEX.- Сообщение ОИЯИ P10-92-336, Дубна, 1992.
 10. В.Г. Иванов, Ф. Фернандес Нодарсе. Моделирование процесса "офф-лайн" обработки данных в физике высоких энергий с помощью СППР COMEX.- Сообщение ОИЯИ P10-92-386, Дубна, 1992.

Рукопись поступила в издательский отдел
23 октября 1992 года.