

Построение интеллектуальных систем для исследований энергетики на основе алгебраических сетей и онтологий: подход и реализация*

Л. В. МАССЕЛЬ, А. Н. КОПАЙГОРОДСКИЙ, В. Л. АРШИНСКИЙ
Институт систем энергетики им. Л.А. Мелентьева СО РАН, Иркутск, Россия
e-mail: massel@isem.sei.irk.ru, digger@istu.edu, pochemzria@mail.ru

This article describes a project based on algebraic nets tool and ontology and oriented at the development of interacting intellectual systems for solution of the complex power engineering problems.

Введение

В 1990-х гг. энергетическая наука понесла значительные потери знаний, опыта, фактов, накопленных за 50 лет. Практически было расформировано большинство отраслевых научных институтов, примерно 40 % научных сотрудников покинули стены академических институтов. В то же время в Институте систем энергетики им. Л.А. Мелентьева (ИСЭМ) СО РАН накоплены существенные интеллектуальные ресурсы, которые аккумулированы в виде математических моделей, алгоритмов, методик и программ для исследований отдельных систем энергетики. Для сохранения основных информационных и программных продуктов в институте ведется работа по созданию IT-инфраструктуры научных исследований, включающей информационную, вычислительную и телекоммуникационную составляющие [1, 2].

В ИСЭМ выполняются комплексные исследования, например по проблемам энергетической безопасности, проведение которых требует совместного использования базовых математических моделей, комплексов программ и баз данных, разработанных для исследования отдельных систем энергетики с целью применения их для развития топливно-энергетического комплекса (ТЭК) в целом с позиций энергетической безопасности. В ТЭК выделяются следующие основные отраслевые системы: электроэнергетические, системы газоснабжения, системы нефте- и нефтепродуктоснабжения, системы углеснабжения, системы теплоснабжения. Каждая из них предназначена для добычи (производства, получения), переработки (преобразования), передачи (транспортирования), хранения и распределения энергоресурсов и снабжения ими (потребления). Энергетическая безопасность рассматривается как часть национальной безопасности, а именно как защищенность граждан, общества, государства, экономики от угроз дефицита в обеспечении их обоснованных потребностей топливно-энергетическими ресурсами приемлемого качества. В этих исследованиях выявлены и систематизированы

*Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (гранты № 07-07-00265а, № 08-07-00172) и РГНФ (грант № 07-02-12112в).

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2008.

неблагоприятные события, получившие название “угроз” энергетической безопасности: техногенных, экономических, природных, социально-политических, а также вызванных несовершенством управления (управленческо-правовых).

Традиционно исследования проблемы энергетической безопасности проводились с использованием программных комплексов и баз данных, в которых хранились информационные модели ТЭК страны, представляющие собой матрицы условий для решения общей задачи линейного программирования. Возмущения (чрезвычайные ситуации) задаются сценариями, представляемыми в виде численных корректировок информационных моделей. Исследования носят многовариантный характер, причем “дерево” вариантов может быть достаточно разветвленным. Предлагаемый в последнее время комбинаторный подход, в котором рассматриваются все возможные варианты, приводит к увеличению потребностей в вычислительных ресурсах и возрастанию нагрузки на эксперта. Если первая проблема достаточно легко разрешается с помощью распараллеливания и использования современной вычислительной техники, то отсутствие или недостаток экспертов существенно ограничивает возможности решения задачи.

Выходом из этой ситуации представляется создание совокупности интеллектуальных систем, которые, с одной стороны, могли бы аккумулировать опыт и знания в специализированных базах знаний, а с другой — позволили бы реализовать выдвинутую еще в 1970-х гг. идею о двухуровневой схеме проведения исследований, связанных с анализом состояния и прогнозированием направлений развития ТЭК страны (Л.Д. Криворуцкий, Ю.Д. Кононов, Б.Г. Санеев и др.). Предлагалось создание “верхнего уровня” агрегированных моделей и алгоритмов, на котором исследователь в режиме экспресс-анализа мог бы проанализировать возможные варианты исследований и выбрать необходимые, которые на “нижнем уровне” можно рассмотреть более подробно, с использованием соответствующих программных комплексов и баз данных, хранящих уже не агрегированную, а детализированную информацию об объектах энергетики. Формализация этой идеи и решение проблемы традиционными математическими методами оказались сложными, а теория и практика систем искусственного интеллекта в то время были недостаточно развиты.

В данной статье рассматривается возможность построения совокупности интеллектуальных систем для исследований энергетики, в первую очередь, по проблемам энергетической безопасности.

1. Подход к созданию интеллектуальных систем, основанный на применении алгебраических сетей

Учитывая разнородность и сложность разработанных программ, моделей и баз данных, трудно реализовать их совместное использование, или “склеивание” (joining) вручную, тем более что проблема усугубляется разным уровнем агрегирования общих параметров на разных уровнях исследования. В настоящем проекте предлагается разработать и использовать для этого новую технологию, основанную на применении алгебраических сетевых моделей процессов и моделей данных. Предлагается создать совокупность интеллектуальных систем по типу пятиконечной звезды, или “пентаграммы”. Аналогичная технология весьма активно развивается в американских и европейских интернациональных энергетических корпорациях, где пять основных проблем увязаны в “пятиконечную звезду”: физика (генерация, передача, потребление), география, экономика, финансы, события.

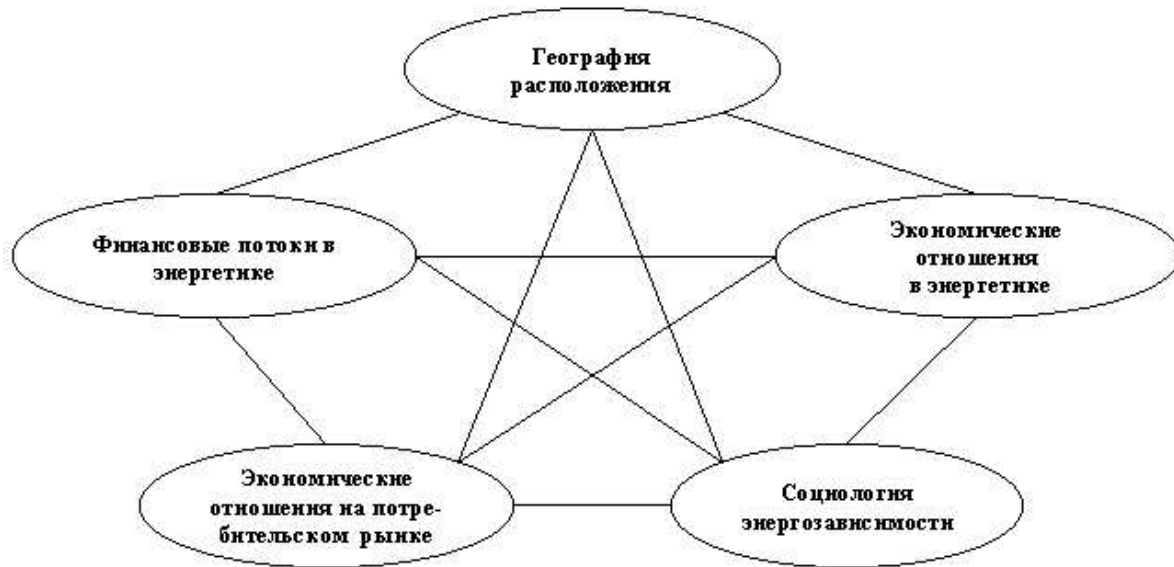


Рис. 1. Предметные области для построения интеллектуальных систем

Выделены следующие предметные области исследований проблем энергетики (рис. 1):

- 1) география расположения генерирующих и потребительских мощностей;
- 2) финансовые потоки в энергетике;
- 3) экономические отношения в энергетике;
- 4) экономические отношения на потребительском энергетическом рынке;
- 5) социология энергозависимости.

Возможность управления сверхсложными комплексами моделей процессов и моделей данных западные IT-специалисты видят в применении алгебраических сетей (АН), которые обладают “интеллектуальными” процедурами, позволяющими решать проблемы риска, когда неблагоприятные события распространяются по сетям “звезды”, порождая их элементы. Аналогично решаются проблемы распространения благоприятных событий (“фарта”), которые улучшают характеристики процессов и свойств объектов.

Теория и практика алгебраических сетей разрабатывалась в трудах российских ученых еще в 1970-х — 1980-х гг. (А.И. Мальцев, А.П. Ершов, Э.Х. Тыгу и др.) [3–5]. По сути, на идеях алгебраических сетей Э.Х. Тыгу в 1970–1973 гг. создана система ПРИЗ, которая только из-за реализации на громоздких ЭВМ типа IBM/360 не получила широкого распространения. На основе идей Э.Х. Тыгу в ИСЭМ СО РАН (ранее — Сибирском энергетическом институте) в свое время разрабатывалась система автоматизированного синтеза программ (в первой версии для БЭСМ-6 — система машинного построения программ СМПП, автор — В. Эпельштейн), версия СМПП-ПК для ПЭВМ реализована в середине 1990-х гг. С.К. Скрипкиным и Т.Н. Ворожцовой, и до настоящего времени она успешно используется для моделирования теплоэнергетических установок непрерывного действия.

Основной особенностью алгебраических сетей является возможность эквивалентных преобразований элементов сети, изменяющих направление управляющих воздействий. Ценность построения информационных систем с учетом алгебраических свойств сети заключается в следующем: всегда можно проверить противоречивость заново включенных в систему моделей процессов и данных; провести тестирование на возможность

(разрешимость), вычислимость при различных конфигурациях входных данных; построить заведомо выполнимые сценарии до начала процесса вычислений.

2. Построение алгебраических сетей на основе Joiner-сетей

В данной работе развивается теория Joiner-сетей (JN), предложенная в диссертационной работе К.В. Новика под руководством Л.Н. Столярова, которые можно рассматривать как расширение сетей Петри [6, 7] для учета изменяемых связей.

Основная цель теории Joiner-сетей — построение математической и информационной технологии для моделирования систем, состоящих из множества взаимодействующих процессов. Joiner-сеть является обобщением известных моделей асинхронных сетей и представляет собой граф с двумя видами вершин — позициями и переходами. Сеть строится из базовых элементов, соответствующих каждому процессу. Правила порождения сети основаны на отождествлении вершин — позиций этих элементов. Графическому представлению JN ставится в соответствие система логических уравнений из пусковых и флаговых функций.

Базовые элементы сети определяются следующим образом: позиции представляют собой распределенную память о событиях сети, а переходы интерпретируются элементарными автоматами, которые запускаются входными событиями, описываемыми пусковыми функциями, и генерируют выходные события, описываемые флаговыми функциями. Каждый такой переход представляет собой отдельный процесс или его состояние (подпроцесс).

Всякую алгебраическую сеть (AN) можно в общем виде представить как некоторую JN-сеть, имеющую два “сорта” вершин (функций и подпроцессов), и связи между ними. Сеть необязательно должна быть связанной, возможно представление сети в виде множества элементов “пусковые функции — подпроцесс — флаговые функции” (рис. 2). Конечной целью создания JN-сети в рассматриваемом контексте является автоматизация процесса выполнения вычислительных экспериментов на описанных программных комплексах, поэтому построение связанной сети не столь критично, как описание отдельных вычислительных ядер в виде элементов “пусковые функции — подпроцесс — флаговые функции”.

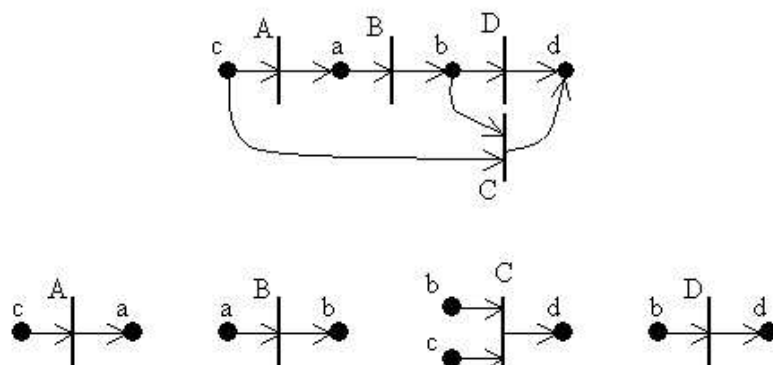


Рис. 2. Пример JN-сети в связанном виде и поэлементно

3. Общая архитектура интеллектуальной системы “Пентаграмма”

Предлагается следующая архитектура интеллектуальной системы “Пентаграмма” (ИС “Пентаграмма”), которая позволит проводить исследования в режиме экспресс-анализа с последующим более детальным рассмотрением возможных вариантов исследований (рис. 3). Описания объектов и процессов энергетики и их взаимодействий представляются онтологиями, хранимыми в специально разработанной базе знаний на основе онтологий, реализуемой средствами информационной инфраструктуры исследований энергетики [10, 11]. Ядро базы знаний ИС “Пентаграмма” поддерживается центральным компонентом информационной инфраструктуры — репозитарием [12]. Там же могут храниться разработанные JN-представления решенных задач энергетики.

Пользователь моделирует интересующие его задачи в среде создания сценариев на основе JN, используя знания, описанные онтологиями. Можно сказать, что, моделируя задачу JN, элементами которой могут быть объекты и процессы энергетики, описанные онтологиями из базы знаний ИС “Пентаграмма”, пользователь описывает онтологию конкретной задачи. Построенная таким образом модель задачи может быть проанализирована и оптимизирована с помощью аппарата JN. Однажды созданная модель задачи может быть вызвана повторно из базы знаний ИС “Пентаграмма” для доработки или исследования.

Использование аппарата JN позволяет не только моделировать и исследовать задачи, но и фактически создавать сценарии, управляющие процессом выполнения. Сервер ИС “Пентаграмма”, используя готовый сценарий, организует управление разнородными ресурсами. Комплексы программ и базы данных, хранящие детализированную информацию об объектах энергетики, взаимодействуют через протокол SOAP.

Таким образом, в рамках ИС “Пентаграмма” предполагается использовать Joiner-сети как для манипулирования вычислительными ресурсами, так и для построения онтологий конкретных задач и преобразования информационных ресурсов.

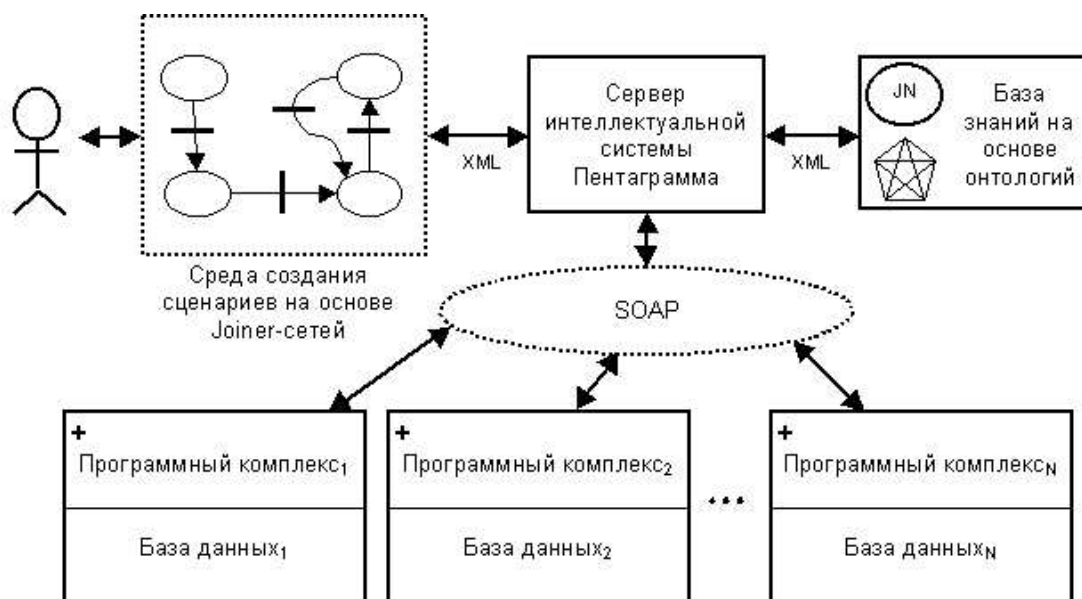


Рис. 3. Общая организация интеллектуальной системы “Пентаграмма”

4. Программа редактирования Joiner-сетей

В рассматриваемых Joiner-сетях для описания процессов и их взаимодействия можно выделить следующие элементы:

- пусковые функции — описывают входные параметры программных комплексов, которые будут использоваться для выполнения расчетов;
- подпроцессы — описывают программные комплексы, используемые для выполнения расчетов;
- флаговые функции — выходные параметры программных комплексов, которые являются результатами выполнения расчетов.

Выделенные классы информационной инфраструктуры дополняются необходимыми атрибутами: протоколом доступа, указателем (URL), функцией инициирования и др. Некоторые результаты расчетов одних программных комплексов могут являться входными параметрами других программных комплексов, поэтому параметры рассматриваются как единая логическая группа, однако для удобства создания JN производится их “виртуальное разделение”.

Для простоты ввода информации о программных комплексах, а также описания входных и выходных параметров разработана специальная программа редактирования JN, интерфейс которой представлен на рис. 4. С ее помощью в легкой форме можно описать сеть в виде подпроцессов, пусковых и флаговых функций. Программа редактирования Joiner-сетей построена с использованием компонентов информационной инфраструктуры исследований энергетики и может работать как с СУБД Firebird, так и с Microsoft SQL Server.

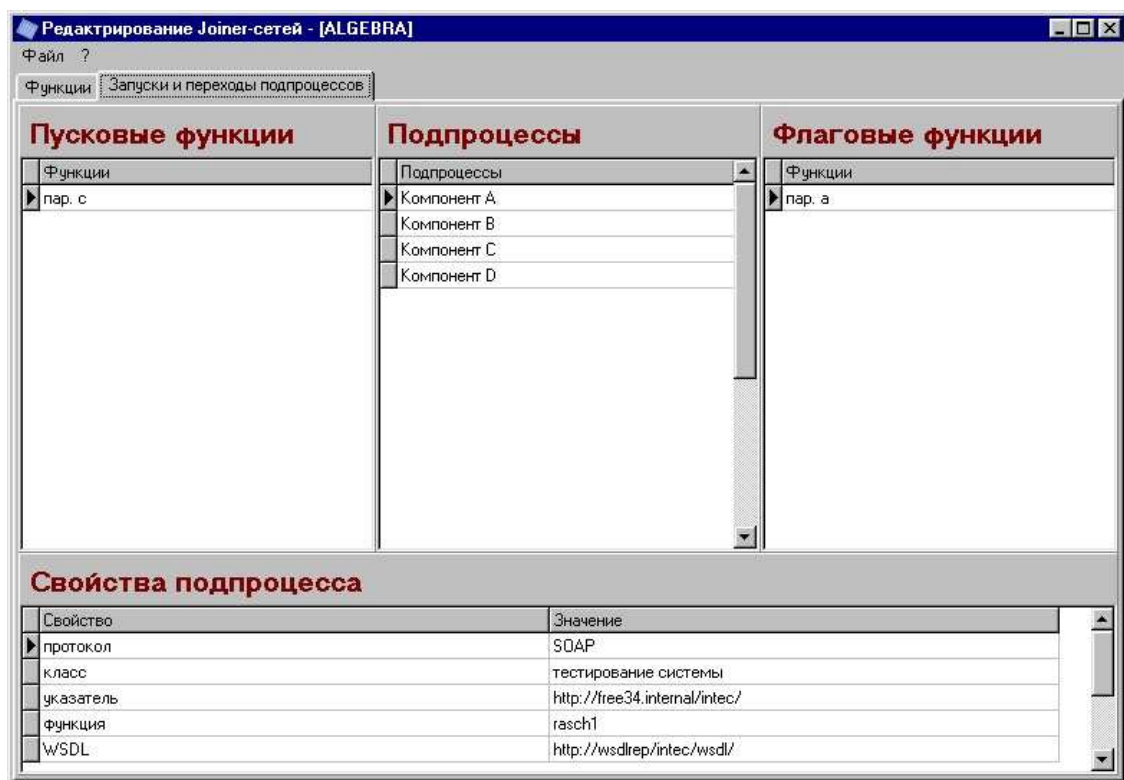


Рис. 4. Интерфейс программы редактирования Joiner-сетей

После внесения информации о программных компонентах в JN ее можно извлекать и представлять в различном виде, удобном для дальнейшего анализа и использования в интеллектуальной системе “Пентаграмма”.

5. Представление Joiner-сети в графическом виде и в виде XML-документа

Извлечение данных в различных форматах из репозитория выполняется с помощью хранимых процедур. Разработано несколько процедур для выгрузки JN в инструментальную систему SmartTools [13] (рис. 5). На концепт-карте изображены пусковые/флаговые функции и подпроцессы, а также связи (“входной параметр” и “выходной пара-”

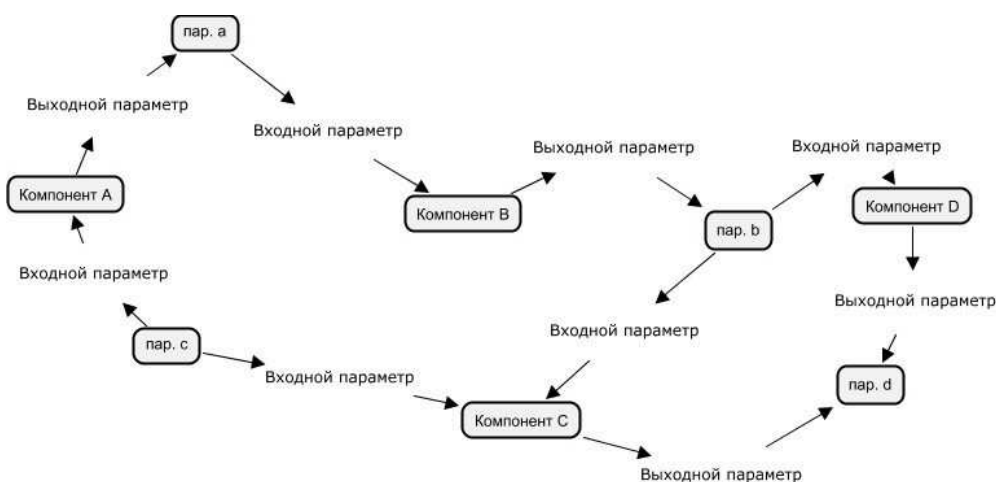


Рис. 5. Joiner-сеть в виде концепт-карты SmartTools

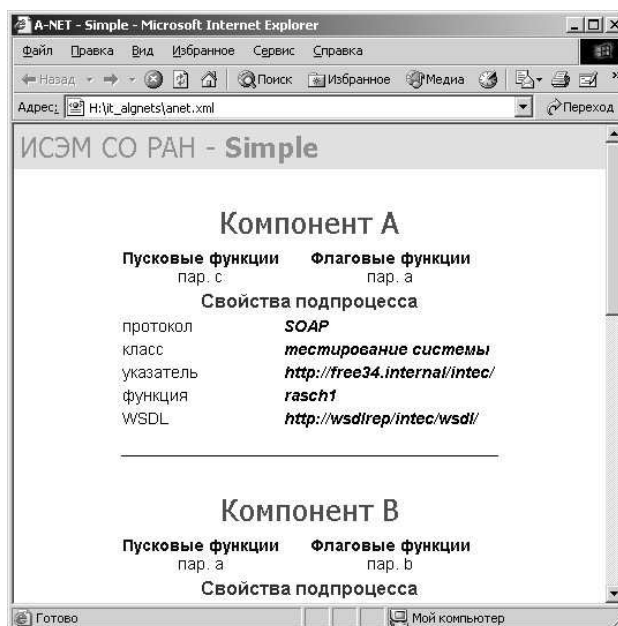


Рис. 6. Просмотр выгруженного XML-документа

метр”). После построения Joiner-сети в графическом виде можно провести ее “ручной” анализ, выявить и исправить аномалии.

Продобным образом создана процедура для извлечения JN в виде XML-документа (рис. 6). Для более наглядного просмотра XML-документов применяется технология XSL Transformations (XSLT).

Заключение

При выполнении проекта предусмотрено три этапа. На первом этапе (в 2007 г.) разработаны:

- 1) новые элементы теории алгебраических сетей для вычислительного анализа, событийного анализа и онтологического анализа;
- 2) архитектура интеллектуальной системы “Пентаграмма”;
- 3) модель онтологической базы знаний для всех фрагментов интеллектуальной системы “Пентаграмма”;
- 4) технология решения содержательных задач с использованием элементов теории алгебраических сетей.

Кроме того, выполнена постановка содержательных задач, использующих общую сеть программ, данных и знаний, на примере исследования проблемы энергетической безопасности.

При проектировании и реализации разрабатываемой системы используются современные CASE-средства для моделирования программ и данных (Rational Rose, CA ErWin), Java-технологии, интеграционные технологии веб-сервисов и SOAP, языки XML, XTM, OWL и инструментальная система Smart Tools для графического построения онтологий.

Список литературы

- [1] ИНТЕГРАЦИЯ информационных технологий в системных исследованиях энергетики / Л.В. Массель, Е.А. Болдырев, А.Ю. Горнов и др. / Под ред. Н.И. Воропая. Новосибирск: Наука, 2003. 320 с.
- [2] МАССЕЛЬ Л.В. Методы и технологии создания ИТ-инфраструктуры научных исследований // Тр. X Байкальской Всерос. конф. “Информационные и математические технологии в науке, технике и образовании”. Иркутск: ИСЭМ СО РАН, 2005. С. 57–65.
- [3] МАЛЬЦЕВ А.И. Алгебраические системы. М.: Наука, 1970. 392 с.
- [4] ЕРШОВ А.П. Введение в теоретическое программирование. М.: Наука, 1977. 288 с.
- [5] ТЫГУ Э.Х. Концептуальное программирование. М.: Наука, 1984. 256 с.
- [6] ПИТЕРСОН Дж. Теория сетей Петри и моделирование систем. М.: Мир, 1984. 264 с.
- [7] КОТОВ В.Е. Сети Петри. М.: Наука, 1984. 160 с.
- [8] НОВИК К.В. Сеть автоматов для моделирования асинхронного взаимодействия процессов / Автореф. дис. ... канд. физ.-мат. наук. М.: МФТИ, 2006. 22 с.
- [9] СТОЛЯРОВ Л.Н., НОВИК К.В. Реализация параллельных процессов с помощью сетей Joiner-net // Информационные и математические технологии: Тр. Байкальской Всерос. конф. Иркутск: ИСЭМ СО РАН, 2004. С. 11–14.

- [10] Копайгородский А.Н., Массель Л.В. Разработка и интеграция основных компонентов информационной инфраструктуры научных исследований // Вест. ИрГТУ. 2006. № 2 (26). С. 20–24.
- [11] Копайгородский А.Н., Макагонова Н.Н., Трипутина В.В. Информационная инфраструктура научных исследований // Тр. X Байкальской Всерос. конф. “Информационные и математические технологии в науке, технике и образовании”. Иркутск: ИСЭМ СО РАН, 2005. С. 72–80.
- [12] Копайгородский А.Н. Репозитарий как ядро информационной инфраструктуры системных исследований в энергетике // Сб. тр. молодых ученых ИСЭМ СО РАН. Иркутск: ИСЭМ СО РАН, 2006. С. 274–281.
- [13] IHMC SMARTTOOLS. <http://smar.ihmc.us/>

Поступила в редакцию 25 января 2008 г.