

Синтез онтологических представлений по данным астрометрических наблюдений

А.Е.Авраменко

Пушинская радиоастрономическая обсерватория (ПРАО АКЦ ФИАН)

avr@prao.psn.ru

Аннотация

Рассматриваются функциональные возможности интегрированной коллекции астрометрических пульсарных данных для извлечения онтологически значимых свойств предметной области по результатам долговременных наблюдений. Приводятся основные требования к консолидации данных, методов и правил, определяемых спецификой предметной области. Анализируются особенности компонентной архитектуры приложений и промежуточного слоя, соответствующие терминам и понятиям, которые используются в моделях и представлениях предметной области. Приведены примеры обобщенных правил извлечения характерных признаков на основе содержательного анализа наблюдательных данных. Обозначены типы ресурсов, доступных из приложений в распределенной объектной среде.

1. Введение

Работы, выполненные в целях интеграции астрометрических пульсарных данных в открытую многофункциональную коллекцию на основе компонентной архитектуры [1], открывают новые возможности для ее приложений в предметной области, связанной с выявлением и систематизацией новых знаний о физических свойствах наблюдаемых объектов, оценкой закономерностей, достоверности, визуализацией получаемых результатов.

Содержательное представление предметной области формируется в проблемно-ориентированном приложении на основе наблюдательных данных, которые упорядочены на серверах в унифицированных форматах хранения, обмена, и требуемые фрагменты могут быть извлечены по запросам. Можно выделить следующие особенности приложений:

- Использование одних и тех же

наблюдательных данных в приложениях с различным контекстом предметной области;

- Применение альтернативных моделей предметной области в приложениях;
- Повторное использование, сопоставление онтологически значимых компонентов приложений применительно к новым и уже имеющимся наблюдательным данным.

Наблюдательные данные выражаются синтаксической структурой символов, с которой связывается смысловое содержание, используемое в приложении в контексте его предметной области. Следует принять во внимание, что, во-первых, символические языки описания данных, в общем виде, не связаны напрямую с контекстом предметной области и, во-вторых, практически невозможно найти такое символическое представление данных, которое в одинаковой мере отвечало бы требованиям контекста широкого спектра приложений. По этим причинам содержательное представление предметной области, определяемой условиями приложения, не может быть сформировано только на основе символического описания структуры данных. Однако оно может быть синтезировано на основе этих описаний структур данных с помощью правил и процедур, трансформирующих эти описания структур в термины, соответствующие понятиям предметной области. Задачу такой трансформации данных и синтеза представления предметной области в общем виде, для широкого круга возможных приложений решить достаточно сложно [7], обычно приводятся только концептуальные разработки и демонстрационные примеры. Однако применительно к достаточно узким, специализированным областям приложений такие трансформации данных и синтез представлений предметной области могут быть выполнены с помощью композиционного сценария взаимодействия данных, правил и методов, консолидируемых в приложении [9].

В приложении должен поддерживаться определенный набор функциональных возможностей, отвечающих следующим требованиям консолидации и управления компонентами, связанными с синтезом и представлением свойств предметной области:

- Отбор, кластеризация релевантных исходных и редуцированных данных;
- Оценивание, параметрическая оптимизация онтологически значимых компонентов;
- Выявление, представление признаков, свойств предметной области.

Этим требованиям соответствуют следующие задачи, которые должны быть рассмотрены и решены с целью консолидации ресурсов, извлечения и представления свойств предметной области:

1. Формирование обобщенной структуры метапроцесса синтеза свойств предметной области.
2. Управление данными, сервисами в соответствии с функциональными требованиями и правилами выявления и отображение свойств предметной области.
3. Создание интегрирующей оболочки промежуточного слоя в распределенной среде.

2. Метапроцесс синтеза свойств

Наблюдаемые объекты, их сущности воспринимаются на основе смысловых понятий как некая упорядоченность, которая подлежит познанию. Представляющие интерес и доступные наблюдению признаки, свойства отображаются в виде знаков (символов, текстов, образов), которые выступают в качестве медиумов (носителей) значений. Знаки позволяют выделять существенные качества, которые характеризуют означаемые объекты. На основе интерпретации знаков и связанных с ними значений достигается понимание, осмысление сущностных представлений, отождествляемых с означаемыми объектами [3].

Процесс познания на основе синтеза присущих объекту свойств по данным наблюдений реализуется с учетом баланса между целеполагающими действиями человека-субъекта в рамках предметной области приложений и возможностями автоматизированных информационных технологий (управление данными и сервисами, промежуточный слой, пользовательский интерфейс). Поддерживаемый информационными технологиями процесс синтеза и отображения свойств предметной области направлен на выявление новых знаний, их представление и интеграцию с целью доступа, совместного и повторного использования в контексте приложений для решения индивидуальных и коллективных задач, с возможностью их персонализации для пользователя [12]. Доступными, управляемыми из приложения оказываются не только наблюдательные данные и не только представленные результаты, но и непосредственно сам процесс синтеза свойств предметной области с использованием этих данных и результатов. Моделирование, создание сценариев процесса извлечения новых знаний и представлений

включает как машинно-ориентированные, так и субъектно-ориентированные задачи [4]. Обобщенная схема метапроцесса синтеза свойств предметной области приведена на Рис.1.

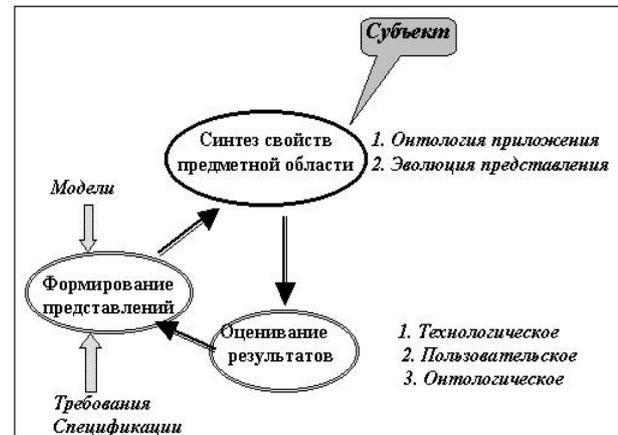


Рис.1. Обобщенная схема метапроцесса синтеза свойств

Ей присущи необходимые качества итерационной цикличности, повторяемости с учетом получаемых оценок, присоединения, вписывания, сопоставления новых представлений с известными, возможности включения в процесс новых требований и спецификаций. Оценивание результатов осуществляется по нескольким критериям:

- онтологический,
- технологический,
- пользовательский.

На основе прагматической оценки методов и результатов выявляются условия синтаксической и семантической согласованности используемых средств, интероперабельности ресурсов, формируется онтологическая модель предметной области на основе сущностных представлений свойств наблюдаемых объектов. Пользовательский аспект должен учитывать наследование функциональных, технологических решений, поддержку уже применяющихся методов и средств, возможность системной адаптации новых требований, задач проблемно-ориентированных приложений.

2. Онтологическая база предметной области

При разработке информационных средств (ИС) акцент постепенно смещается от вопросов количественного накопления данных, доступа, интеграции, обмена к содержательному использованию ресурсов применительно к задачам проблемно-ориентированных приложений (Рис.2).



Рис.2. Эволюция отношений данные-методы в ИС

Данные наблюдений, упорядоченные в виде баз или временных рядов по некоторому набору параметров, связываются в компонентной объектной среде приложения с релевантными методами, с помощью которых реализуется требуемая последовательность функциональных преобразований с целью получения результатов, отождествляемых с характеристиками предметной области. Информационные средства в компонентной среде поддерживают доступ к упорядоченным структурам данных и выборку требуемых фрагментов, интероперабельное взаимодействие данных и методов, их консолидацию и повторное использование в приложениях через промежуточный слой, связывающий данные и методы, необходимые для решения целевой задачи.

Онтологическая база в логическом смысле включает обобщенные объекты представления предметной области, набор типов которых ориентирован на извлечение и отображение ее существенных свойств [6]. Типы обобщенных объектов:

1. Термины, факты, их роль в отображении понятий предметной области.
2. Правила, условия, ограничения, определяющие отношения сущностей.
3. Функции, методы выявления и представления сущностей предметной области.

Первые два типа объектов составляют концептуальную основу онтологической базы (репозиторий знаний), с помощью которой достигается семантическая независимость информационных решений от особенностей их системной реализации. Хотя функции и методы не относятся непосредственно к онтологической базе – они в значительной степени определяются особенностями выбранных системных средств, моделей предметной области, синтаксических средств описания данных, – их включение в состав рассматриваемых типов объектов оправдано стремлением к поддержанию объектной целостности проблемно-ориентированного приложения, функциональной полноты средств

извлечения и представления свойств предметной области. Кроме того, этот тип объектов, наряду с двумя первыми, также оказывается включенным в состав интероперабельных ресурсов, доступных приложению в распределенной объектной среде, обеспечивая тем самым компонентную целостность приложения.

Онтологическая база коллекции астрометрических пульсарных данных основывается на четырехмерной пространственно-временной метрике, с помощью которой любое наблюдаемое событие может быть описано в едином четырехмерном континууме, образуемом пространством и временем. Детальное описание типов объектов онтологической базы определяется используемыми моделями предметной области, особенностями проблемно-ориентированных приложений.

Извлечение (mining) характерных признаков осуществляется на основе содержательного анализа наблюдательных данных в контексте предметной области с использованием определенных правил и процедур. Здесь приведены применительно к пульсарным наблюдениям лишь некоторые, наиболее общие, используемые в различных приложениях, примеры таких правил и процедур, связываемых непосредственно с онтологически значимыми параметрами:

а) Нахождение подобия, совпадений (Sim) временных рядов $X(t)$ и $Y(t)$:

$\text{Sim} \{ X(t), Y(t) \}$, где:

$$X(t) = x_1(t_1), x_2(t_2), \dots, x_n(t_n)$$

$$Y(t) = y_1(t_1), y_2(t_2), \dots, y_n(t_n)$$

например, для поиска одинаковых последовательностей или выявления характерных отличий наблюдательных данных.

б) Выявление трендов, периодичностей, характерных спектров временных рядов для адекватного выбора моделей предметной области, которые учитывали бы эти признаки, их влияние на свойства объектов.

в) Подбор групп для синхронного суммирования сигналов в K выбранных каналах регистрации пульсарных профилей и M дискретных отсчетов в периоде по критерию наилучшего приближения (минимальная величина случайных погрешностей):

$$x(t) = \sum_{j=1, \dots, K} x_{j1}(t_1), \dots, x_{ij}(t_i), \dots, x_{MK}(t_M).$$

г) Выявление пульсарных событий $x(T)$ – моментов, соответствующих положению t_n суммарных наблюдаемых профилей по критерию минимальной статистической ошибки:

$$x_1(t_1), x_2(t_2), \dots, x_n(t_n) \rightarrow x(T).$$

д) Регуляризация присоединяемых временных рядов (преобразование случайным образом расположенных на оси времени событий в

упорядоченные последовательности с шагом, кратным постоянной величине Δt с последующим временным или спектральным анализом долговременных рядов:

$$x_1(t_1), x_2(t_2), \dots, x_n(t_n) \rightarrow \\ \rightarrow x_1(t_1), x_2(t_1 + \Delta t), \dots, x_n(t_1 + (n-1)\Delta t)$$

Рассмотренные правила извлечения и связанные с ними признаки соответствуют специфике предметной области и используются в проблемно-ориентированных приложениях.

4. Синтез и отображение свойств предметной области

Высокая точность измерения параметров, достижимая при наблюдениях объектов астрометрическими методами, а также развитые математические модели исследуемой предметной области позволили с помощью многопараметрического фиттинга согласовывать вновь получаемые результаты наблюдений, приводя выявляемые свойства объектов в соответствие с уже известными по каталогам. Так, применительно к рассматриваемой области, в ПРАО ФИАН были разработаны алгоритм и программа TIMAPR высокоточного фазового анализа наблюдений пульсаров (О.В.Дорошенко, С.М.Копейкин, 1990) для обработки моментов прихода импульсов (МПИ) пульсаров с точностью порядка 10 нс с учетом нескольких десятков влияющих физических факторов. Поскольку формально выходным результатом фиттинга является единственная переменная (МПИ пульсара), а эффекты влияющих факторов учтены в нем интегрально, всех вместе и сразу, то возникает необходимость выявления небольшого числа (1-3) наиболее значимых для данного приложения параметров, которые существенным образом связаны с исследуемыми свойствами предметной области и позволяют целенаправленно интерпретировать наблюдаемые результаты. Таким образом, на основе декомпозиции предметной области, ограничивающей число переменных, удастся представить в явном виде существенные зависимости, которые определяют выявляемые свойства и поведение объекта.

Переход к оперированию понятиями и терминами предметной области с целью синтеза и отображения ее свойств по наблюдательным данным требует приведения в соответствие структур, форматов данных тем значениям, терминам и понятиям, которые будут использоваться в моделях и представлениях предметной области [5]. Предметная специфика, следовательно, требует создания среды с новыми компонентами, поддерживающей не только наблюдательные данные и связываемые с ними методы управления, но также компоненты, означающие сущности предметной области, на основе которых синтезируются ее свойства.

Существует два подхода к решению проблемы интеграции гетерогенных компонентов, отображающих онтологическое разнообразие предметной области приложений [12]:

1. Создание интероперабельных ресурсно-ориентированных систем методом «снизу-вверх» (bottom-up) путем последовательного обобщения спецификаций ресурсов от нижнего уровня структур исходных данных до описания в терминах предметной области.
2. Использование метода «сверху-вниз» (top-down), основанного на исходном определении общих концепций онтологии предметной области с последующей адаптацией ресурсов нижнего уровня.

Второй подход, основанный на определении общих концепций предметной области, ориентированных на достаточно широкий круг поддерживаемых приложений, более эффективен, в то время как методом «снизу-вверх», когда для каждого нового приложения требуется обновление спецификаций, сложно получить удовлетворительное обобщение для всех случаев.

Полным решением проблемы можно было бы считать создание интероперабельных (shared) онтологий, которые включают общую концептуальную терминологию всех приложений, а выполнение проблемно-ориентированных правил и процедур поддерживается средствами интегрирующих распределенных систем. Однако, ввиду большого разнообразия тех и других, удовлетворительным можно считать и некоторые ограниченные решения, например, создание персональных онтологий, которые содержат согласованные терминологии для группы пользователей и в определенных границах предметной области (например, для специалистов, изучающих свойства пульсаров методами хронометрирования).

Модель, как некоторое сущностное обобщение, выступает в качестве содержательного ядра предметной области, которое, будучи относительно независимым от деталей конкретных приложений, задач, аккумулирует базовые знания и представления о предметной области как объекте изучения [10]. Моделирование исходно не предполагает разделять (share) концептуальные спецификации всех возможных приложений, но в то же время должно удовлетворять представляющим интерес для пользователя правилам всех тех приложений, которые предполагается поддерживать в предметной области. Должен соблюдаться баланс между достаточно широким набором правил, содержащихся в онтологической модели, с тем чтобы достигалась интероперабельность в пределах возможно более широкого круга приложений, и концептуальной целостностью модели.

Таким образом, метод интеграции компонентов «снизу-вверх» наполняет модель онтологическим разнообразием приложений, а метод «сверху-вниз» -

сохраняет ее фундаментальные свойства, общие для всех приложений. И, следовательно, онтологический подход реализует две взаимосвязанные функции: моделирование предметной области и создание интероперабельных онтологий.

Синтез свойств наблюдаемого объекта заключается в выявлении существенных признаков по совокупности наблюдательных данных и трансформацию их в онтологически значимые характеристики объекта. Признаки, извлекаемые непосредственно из наблюдательных данных, определяются типом и характером моделей, структурой наблюдаемых данных, отображаемых синтаксическими конструкциями и атрибутами отношений. В выявлении свойств основополагающую роль играет интерес субъекта к содержанию предметной области, реализуемый через приложение. Схема отношений в процессе синтеза свойств приведена на Рис.3.

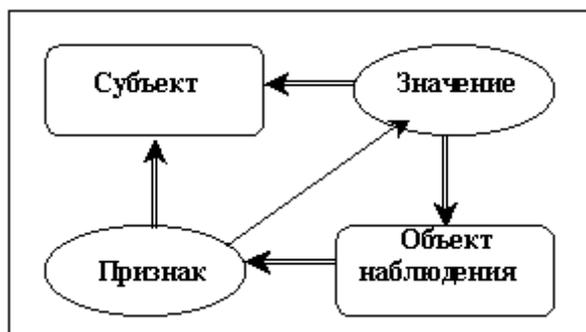


Рис.3. Схема отношений в процессе синтеза свойств

Оценка наблюдательных данных и связанных с ними признаков предметной области может носить как объективный, так и субъективный характер. Первый связан с анализом компонентов в терминах их структуры, качества и достоверности исходных данных. Так, включение $X \rightarrow Y$, как функция вероятностей $P(X)$, $P(Y)$ и $P(X \wedge Y)$, используется при выборе критерия отбора релевантных наблюдательных данных. Например, текущие наблюдательные данные хронометрирования пульсаров присоединяются к долговременным рядам только в том случае, если соответствующие им моменты прихода импульсов (МПИ) различаются не более чем на некоторое предельное (по условиям выбранной модели) значение. Однако интерес не исчерпывается только формальными критериями и простыми правилами. Наиболее значимые характеристики, отражающие существенные свойства, особенности наблюдаемых объектов, могут быть выявлены только на основе содержательных представлений субъекта, в том числе его прошлого опыта, знаний, всестороннего восприятия доступных атрибутов [8].

5. Компонентная интеграция

С расширением типов информационных ресурсов, включением в их состав, наряду с данными и методами, семантически значимых компонентов, существенно расширяется круг объектов, которые должны поддерживаться в среде проблемно-ориентированного приложения. Поскольку создание интероперабельных онтологий, отражающих все многообразие свойств предметной области, представляется пока задачей трудно разрешимой, приходится выбирать ограниченные, но практически реализуемые варианты, удовлетворяющие требованиям проблемно-ориентированных приложений.

При этом, если выбрана концептуальная модель интеграции по методу «сверху-вниз», потребуется создание некоторой промежуточной среды, удовлетворяющей спецификациям объектов предметной области [12]. Изменение профиля предметной области и соответственно спецификаций потребует внесения изменений в организацию промежуточной среды. Определенное упрощение этой задачи достигается применением логики управления базами данных (DB-style) к организации онтологических информационных систем [6]. Эта методология предусматривает разделение фактов и правил, причем факты представляются в виде лексических выражений – терминов, сопровождаемых идентификатором, и вместе с ним специфицируемых как уникальные концепты (понятия), характеризующиеся некоторой смысловой ролью. Релевные признаки определяются по принципу атрибутов БД, а правила, входящие в онтологическую базу, описываются в виде отношений, что и позволяет использовать модель БД для реализации специализированного промежуточного слоя (software agent services), поддерживающего все многообразие приложений предметной области. Естественно, спецификации приложений не должны выходить за пределы принятых понятий, атрибутов и отношений. В противном случае специализированные программные средства промежуточного слоя должны быть приведены в соответствие с новым профилем предметной области.

Более гибкие возможности предоставляют объектные информационные технологии в сочетании с компонентной архитектурой ресурсов для управления процессом моделирования сложных объектов на основе декомпозиции предметной области - выделение ряда простых объектов в виде подзадач, решаемых соответствующими методами, и последующей сборки результатов в рамках общей модели [2]. В рассматриваемой системе средства промежуточного слоя реализованы в компонентной архитектуре CORBA, обеспечивающей интероперабельность ресурсов в приложении и программирование объектов представления предметной области и управления процессами (объектами) в распределенной среде.

Таблица 1

Тип, назначение ресурсов	Функции промежуточного слоя
Данные наблюдений, каталоги Приложения Сервисы обработки, редуцирования данных Сервисы извлечения признаков Сервисы консолидации данных	Форматы, синтаксическая совместимость Связывание данных и методов Методы трансформации, извлечения данных Сценарии извлечения, представления Динамическая интеграция, доступ, выборка

Архитектура DCOM с развитыми средствами промежуточного слоя, интегрированными в программируемые оболочки многофункциональных системных приложений, позволяет, в дополнение к возможностям CORBA, консолидировать разнотипные компоненты многоцелевого объектно-ориентированного репозитория с широким набором сервисов, отвечающих требованиям не только извлечения и представления свойств, но и создания сценариев моделирования объектов предметной области средствами визуального программирования [1]. Следовательно, сервисы в среде приложения поддерживают не только онтологическую базу предметной области, но и ее фактологическую основу: трансформацию, редактирование, форматирование наблюдательных данных, а также представление свойств объектов по результатам моделирования. Такая многофункциональность позволяет интегрировать в среде приложения широкий спектр проблемно-ориентированных задач обработки и интерпретации результатов наблюдений, используя системные возможности компонентного взаимодействия через пользовательский интерфейс промежуточного слоя.

Типы ресурсов, доступных пользователю из приложения:

- Полученные по данным наблюдений результаты моделирования в виде онтологически значимых характеристик (таблицы, графики, диаграммы и т.п.)
- Исходные данные, правила и методы моделирования предметной области, которые могут быть частично или полностью использованы в приложениях.

Приложение может интегрировать функциональные компоненты разного уровня смысловой значимости и завершенности, обеспечивая извлечение, и представление свойств предметной области с помощью существующих технологических возможностей:

- Повторное использование данных, методов управления, правил извлечения признаков предметной

области в проблемно-ориентированных приложениях.

- Создание и выполнение сценариев моделирования, извлечения свойств, экспертных оценок с использованием необходимых данных и сервисов.

Функции промежуточного слоя, определяемые типом, назначением используемых ресурсов, на которых реализуются задачи приложений, приведены в Таблице 1.

Несмотря на значительное разнообразие функций, привлекаемых в процессе создания и выполнения сценариев синтеза и представления свойств, технологические возможности компонентной объектно-ориентированной архитектуры позволяют сформировать интегрирующую среду промежуточного слоя в соответствии с требованиями приложений [11]. Функции промежуточного слоя консолидируются в клиентской оболочке пользователя и доступны приложению через его интерфейс.

Заключение

Интеграция астрометрических пульсарных данных и связанных с ними методов управления, редуцирования и обработки открывает путь к синтезу онтологически значимых свойств предметной области путем включения в компонентную среду коллекции терминов, правил и понятий предметной области, консолидируемых в приложении. Поскольку набор этих терминов, правил и понятий, используемых в специализированных приложениях, представляется достаточно компактным и в значительной степени устойчивым, эта задача может быть сведена к определенному расширению функциональности, выраженному в терминах понятий предметной области, которое реализуется в компонентной среде промежуточного слоя и приложения. В процессе синтеза свойств на основе консолидации семантически значимых компонентов выявляются и систематизируются новые сведения, отождествляемые с характеристиками предметной области, формируется содержательное представление сущностей наблюдаемых объектов.

Модель, выступающая в качестве содержательного ядра предметной области, должна удовлетворять основным правилам ее контекста. Поддержка онтологического разнообразия приложений достигается interoperability семантически значимых компонентов на основе согласованных спецификаций, их использования в приложениях наряду с традиционными типами ресурсов (данные наблюдений и методы редуцирования). Как следствие, пользователю доступны не только данные и готовые результаты моделирования, но также методы моделирования и правила извлечения свойств предметной области, с помощью которых можно создавать и выполнять собственные сценарии моделирования с повторным использованием необходимых данных и результатов¹.

Литература

- [1] А.Е.Авраменко. Управление коллекцией астрометрических пульсарных данных в объектной среде. //Труды Четвертой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». Дубна, 2002, т.1, с.311-321.
- [2] Копысов С.П., Красноперов И.В., Рычков В.Н. Реализация объектно-ориентированной модели метода декомпозиции на основе параллельных распределенных компонентов CORBA. //Вычислительные методы и программирование, т. 4, №1, 2003, с.194-206. <http://num-meth.srcc.msu.su>
- [3] Б.В.Марков. Знаки бытия.// Санкт-Петербург, “Наука”, 2001, с.9-14, 447-448.
- [4] C.Bussler, D.Fencel, A.Maedche. A Conceptual Architecture for Semantic Web Enabled Web Services. //SIGMOD Record, v.31, No.4, Dec.2002. <http://www.acm.org/sigmod/record/issues/0212/SPECIAL/4.Bussler1.pdf>
- [5] Z.Cui, D.Jones and P.O’Brien. Semantic B2B Integration: Issues in Ontology-based Approaches. // SIGMOD Record, v.31, No.1, March 2002. <http://www.acm.org/sigmod/record/issues/0203/SPECIAL/7.cui.pdf>.
- [6] M.Jarrar, R.Meersman. Formal Ontology Engineering in the DOGMA Approach. //CoopIS/DOA/ODBASE 2002, LNCS 2519, pp. 1238-1254. Springer-Verlag Berlin Heidelberg 2002.
- [7] L.A.Kalinichenko, N.A.Skvortsov. Extensible Ontological Modeling Framework for Subject Mediation. // Труды RCDL 2002, Дубна, 2002, т.1, с.99-119.
- [8] T.Palpanas. Knowledge Discovery in Data Warehouses. SIGMOD Record, v.29, No.3, Sept.2000, <http://www.acm.org/sigmod/record/issues/0009/index.html>.
- [9] M.P.Singh. The Pragmatic Web: Preliminary Thoughts. <http://lsdis.cs.uga.edu/SemNFS/Sig-Position.pdf>.
- [10] P.Spyns, R.Meersman, M.Jarrar. Data modeling versus Ontology engineering. // SIGMOD Record, v.31, No.4, Dec.2002. <http://www.acm.org/sigmod/record/issues/0212/SPECIAL/2.Meersman.pdf>
- [11] M.Stonebraker. Too Much Middleware. // SIGMOD Record, v.31, No.1, March 2002. <http://www.acm.org/sigmod/record/issues/0203/industry-ms.pdf>.
- [12] Y.Sure, S.Staab, R.Studer. Methodology for Development and Employment of Ontology based Knowledge Management Applications. //SIGMOD Record, v.31, No.4, Dec.2002. <http://www.acm.org/sigmod/record/issues/0212/SPECIAL/3.Sure.pdf>.

Ontology Engineering in the Astrometry Applications

A.E.Avrmenko

There are considered the functional possibilities of the pulsar astrometry collection for ontology significant features of application domain extraction, are summarized the main requirements in consolidation of data, methods and rules, determined by specificity of application domain, are analyzed the particularities of the intermediate layer, based on component architecture, which correspondent to the terms and concepts of application models. There are generalized rules of typical properties extraction, based on meaningful analysis of observed astrometry data sets, are defined the resources types, which are available from problem oriented applications through object oriented intermediate layer.

¹ Работа поддерживается грантом РФФИ № 03-07-90219