

05.13.11

X 584

На правах рукописи

ХЛОПОТОВ Михаил Викторович

**МЕТОДЫ И АЛГОРИТМЫ ОБНАРУЖЕНИЯ СЕМАНТИЧЕСКИХ ОШИБОК В
ГИПЕРТЕКСТОВЫХ ИНФОРМАЦИОННЫХ РЕСУРСАХ НА ОСНОВЕ
ОНТОЛОГИЧЕСКИХ МОДЕЛЕЙ**

Специальность 05.13.11 – «Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей»

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

ЮУрГУ

Уфа 2004

СЛУЖБА ДЕЛОПРОИЗВОДСТВА		
Южно-Уральский государственный университет		
Вх. №	13-16-3653	
• 22 •	11	2004.

Работа выполнена в Южно-Уральском государственном университете

Научный руководитель – доктор технических наук, профессор
Мельников Андрей Витальевич

Официальные оппоненты – доктор технических наук, профессор
Куликов Геннадий Григорьевич

доктор технических наук, профессор
Чапцов Рудольф Петрович

Ведущая организация – Институт проблем управления им. Трапезникова
Российской академии наук

Защита состоится «17» декабря 2004 года в 10:00 часов, на заседании
диссертационного совета К-212.288.01 Уфимского государственного авиационного
технического университета по адресу: 450000, Уфа-центр, ул. К. Маркса, 12, УГАТУ

С диссертацией можно ознакомиться в библиотеке Уфимского государственного
авиационного университета.

Автореферат разослан «15» ноября 2004г.

Ученый секретарь
диссертационного совета,
кандидат физико-математических наук,
доцент



Р.А. Гараев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Благодаря развитию, распространению и доступности глобальной информационной сети Internet её ресурсы – гипертекстовые информационные ресурсы (ГИР) – стали одним из основных источников повседневной, деловой и оперативной информации. Для того, чтобы информационный ресурс был действительно полезным, представленная в нем информация не должна содержать *семантических ошибок*.

На практике большинство документов, доступных через ГИР, содержат фактические противоречия даже с документами того же ГИР.

Часто это связано с тем, что документ, однажды размещенный в рамках ГИР, хранится в нем и не претерпевает никаких изменений за время своего существования, несмотря на то, что представленная в нем информация могла устареть по сравнению с информацией, приведенной в других документах.

Чем крупнее ресурс, тем больше вероятность того, что очередной документ содержит смысловые ошибки – редакторы ресурса не знают о необходимости внести исправления в некоторые документы.

Концептуальные проблемы создания моделей и систем для интеллектуального анализа информации рассматриваются в научных трудах Д.А. Поспелова, И. Хоррокса, Э. Йордана, Н.И. Юсуповой, О.П. Кузнецова В.И. Васильева и др.

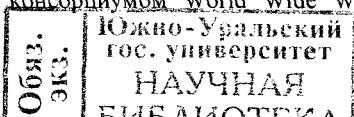
Классические подходы к поиску ошибок в текстах основаны на анализе синтаксиса и морфологии. Очевидно, что поиск взаимопротиворечащих фрагментов может быть осуществлен только при помощи анализа смысла, содержащегося в ГИР, т.е. при помощи *анализа семантики*.

Несмотря на интенсивность работ в области обработки естественных текстов и их семантики, в настоящее время отсутствует научнообоснованный подход к обнаружению семантических ошибок в текстах на естественном языке, и, в частности, в ГИР. В связи с этим разработка методов и алгоритмов для обнаружения семантических ошибок является актуальным направлением исследований.

Для обнаружения смысловых ошибок ГИР при проведении анализа семантики необходима информация о предметной области, свойственных ей понятиях и отношениях между ними, а также ограничениях, свойственным этим отношениям. Такую информацию принято называть *Онтологией*, и автор считает, что онтологическая информация может быть использована для эффективного обнаружения нарушений правил предметной области.

Исследования проблемы формирования моделей онтологий посвящены работы Ю.М. Гусева, А.С. Нариньяни, Т.А. Гавриловой, Э.В. Попова, В.Ш. Рубашкина и др.

Последние годы наибольшие усилия ученых сосредоточены вокруг направления SemanticWeb, поддерживаемого консорциумом World Wide Web



Consortium, и основанного на таких технологиях, как XML и RDF. Это обеспечивает эффективное взаимодействие ученых со всего мира и позволяет воплотить их труды в виде технических рекомендаций и стандартов.

Однако все эти модели носят декларативный характер, и не обладают продукционными возможностями, необходимыми для семантического поиска ошибок.

На основании сказанного выше сформулирована основная цель диссертационной работы и задачи исследования.

Цель работы и задачи исследования. Основной целью диссертационной работы является разработка методов и алгоритмов обнаружения семантических ошибок в ГИР на основе семантических моделей выявления парадоксов в информационных ресурсах.

В процессе достижения данной цели были сформулированы и решены следующие задачи:

- разработка модели представления семантики ГИР, способной отражать понятия и структуры, свойственные текстам естественного языка;
- разработка алгоритмов обработки семантики ГИР и проверки её соответствия ограничениям, характерным для конкретной предметной области;
- разработка формальных методов записи и алгоритмов обработки *контекста отношений*, с учетом необходимости наложения ограничений на них;
- предложены методы и алгоритмы проверки семантической целостности ГИР и, как следствие, обнаружения семантических ошибок;
- разработка базовой онтологии «ГИР организации» для использования при проверке информационных ресурсов;
- разработка информационной системы «TergereLibrum», осуществляющей обнаружение семантических ошибок в ГИР.

Методы исследования. При выполнении исследования был использован математический аппарат теории множеств, теории графов, теории формальных семантик, логики двухместных и N-местных предикатов.

Результаты, выносимые на защиту:

1. Результаты анализа ошибок, характерных для гипертекстовых информационных ресурсов.
2. Теоретическое описание модели онтологий, её особенностей и элементов, обеспечивающих синтез парадоксов в семантической сети.
3. Типовые алгоритмы поиска семантических ошибок в ГИР.
4. Архитектура информационной системы «TergereLibrum».

Научная новизна работы заключается в следующем:

- разработан метод обнаружения смысловых ошибок в ГИР, основанный на исследовании семантической целостности семантической сети, моделирующей

ГИР, относительно *онтологии*, описывающей структуру и свойства предметной области ГИР;

- предложен формализм «семантические правила», позволяющие описывать алгоритмы семантической обработки в терминах семантических сетей. Эти правила используются для создания особых объектов-парадоксов, позволяющих описывать нарушения семантической целостности сети;

- разработана модель представления *контекстной зависимости* свойств онтологии предметной области, соответствующей N-местным предикатам, а также новый алгоритм обработки N-местных предикатов в виде набора 2 местных предикатов.

Практическая значимость и внедрение результатов. Основными практическими результатами исследования является следующее:

1. Разработана информационная система «TergereLibrum», предназначенная для обнаружения семантических ошибок в ГИР и управления его содержимым.
2. Разработана базовая онтология предметной области «ГИР организации».
3. Разработана методика и алгоритмы обнаружения семантических ошибок в ГИР путем проверки семантической целостности семантической сети.

Разработанные в ходе диссертационного исследования модели, методы и алгоритмы используются в программных системах управления Web-сайтами ряда промышленных предприятий и организаций, в том числе ОАО «Автомобильный завод «Урал», ЗАО «Интерсвязь», ГОУ ВПО «Челябинский государственный университет».

Апробация работы. Положения диссертации и результаты исследования докладывались на 6 научных и научно-практических конференциях, в том числе:

- Международная научно-практическая конференция «Автоматизированные системы управления промышленным предприятием АСУ-ПП 2003», г. Москва, ИПУ РАН 2003 г,

- Международная научно-практическая конференция «Computer Science & Information Technology, CSIT 2003» г. Уфа, УГАТУ, 2003г.

- Международная научно-практическая конференция «Computer Science & Information Technology, CSIT 2004», 2004 г. Будапешт, Венгрия

- Российская научно-практическая конференция «Реинжиниринг бизнес-процессов на основе современных информационных технологий. Системы управления знаниями. РБП-СУЗ 2004», 2004 г. Москва, МЭСИ и РАИИ, 2004.

Связь с государственными программами. Исследование выполнено в рамках тематики пункта № 5 плана дополнительных мероприятий федеральной целевой программы «Электронная Россия»: «Разработка концепции и реализация практических мер по обеспечению общедоступности государственных информационных ресурсов с учетом мероприятий по обеспечению их

целостности и достоверности» и федеральной целевой программы «Информатизация образования».

Публикации. Основные положения исследования нашли отражение в 6 публикациях, в том числе в сборниках научных трудов и материалах научно-технических конференций.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы и приложения. Общий объем работы составляет 125 страниц. В работу входит 23 рисунка, 17 таблиц. Список литературы содержит 109 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы диссертационной работы, формулируются цель, научная новизна и практическая значимость полученных результатов.

В первой главе рассматриваются понятие ошибки, особенности ошибок, свойственных гипертекстовым информационным ресурсам, основные методы и модели обнаружения смысловых ошибок.

Для обнаружения смысловых ошибок ГИР при проведении анализа семантики необходима информация о предметной области, свойственных ей понятиях и отношениях между ними, а также ограничениях, свойственным этим отношениям. Такую информацию принято называть *Онтологией*. Онтологическая информация может быть использована для эффективного обнаружения нарушений правил предметной области.

Определим понятие «семантической целостности» как отсутствие противоречий между знаниями о предметной области, описанными в виде онтологии, и знаниями, приведенными в текстах ГИР.

На сегодняшний день разработано и разрабатывается множество моделей онтологий. Среди них IDEF5, DAML+OIL, OIL, KIF, OWL, CYCL. Все они так или иначе базируются на модели семантических сетей или сводимы к ней.

Среди моделей наблюдается разнообразие используемых терминов. Так, модель OIL использует понятия фрейма и слота, модели OWL и DAML используют в том же смысле понятия классов и свойств (отношений), а модель CYCL – коллекция и предикат.

Рассмотренные модели имеют три основных цели их создания: автоматизированное понимание текстов, построение логического вывода на основе знаний о предметной области, и интероперабельное описание предметной области. Ни одна из рассмотренных моделей не проектировалась с целью

обнаружения семантических ошибок, кроме того, большинство моделей не поддерживаются стандартизирующими организациями.

Модель OWL, разработанная консорциумом W3 в рамках проекта SemanticWeb, учитывает опыт разработчиков DAML, OIL и многих других моделей. В настоящее время она принята как стандарт. В работе используется модель OWL как основа для синтеза модели онтологий, удовлетворяющей задачам исследования.

Во второй главе разрабатывается онтологическая модель, совместимая со стандартом OWL и дополняющая его поддержкой контекстно-зависимых свойств, семантическими правилами и возможностью интерпретации парадоксальной семантики.

Для анализа смысла, содержащегося в семантической сети, составим её *интерпретацию*.

Под *интерпретацией* будем понимать теоретико-множественное представление знаний, записанных в семантической сети, в котором идентификаторам URL сопоставлены абстрактные сущности, соответствующие понятиям реального мира. Будем считать, что семантическая сеть описывает «мир», вернее – минимальный набор условий, которым «мир» должен соответствовать, чтобы каждому утверждению сети соответствовал конкретный аспект «мира». Такой подход называется *концепцией открытого мира*, т.е. мира, в котором есть всё, что описано семантической сетью, а всё остальное возможно, но не обязательно.

Определение 1

Интерпретация I семантической сети – это отображение из множества сетей во множество логических значений {истина, ложь}, определяемое кортежем C_1 и рекурсивными правилами. Кортеж C_1 имеет вид

$$C_1 = \langle V_1, R_1, P_1, EXT_1, S_1, L_1, LV_1 \rangle \quad (1)$$

где V_1 – множество URL, словарь ПрО,

R_1 – множество *ресурсов*,

S_1 – отображение $V \Rightarrow R_1$, т.е. отображение множества URL во множество ресурсов,

$P_1 \subset R_1$ – множество *свойств*,

EXT_1 : отображение $P_1 \Rightarrow R_1 \times R_1$ – из множества свойств во множество множеств пар вида $\langle x \in R_1, y \in R_1 \rangle$, расширение свойства,

LV_1 – множество строковых значений литералов,

L_1 – отображение из LV в множество ресурсов,

а правила заданы таблично:

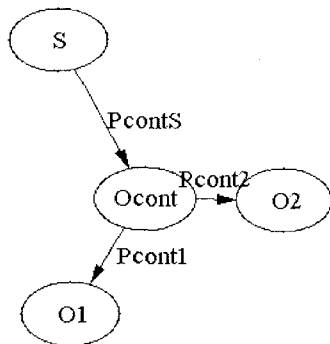


Рис. 2. Реификационная запись контекстного отношения

В такой форма записи ограничения контекстного свойства средствами OWL выразимы не полностью. Для возможности описания ограничений на контекстных свойствах и дополнительных ограничений, которые могут потребоваться при составлении онтологии предметной области, включим в состав модели онтологий *семантические правила*.

Запишем правило в виде трех RDF-сетей: C, N и A, где C – сеть условий, N – сеть отрицательных условий, и A – сеть расширения.

Для обеспечения возможности записи гибких условий введем понятие *пустого узла*. Пустой узел может участвовать в отношениях наравне с URL и литералами (в качестве субъекта, объекта или предиката), и подразумевает собой существование ресурса, о природе и идентификации которого ничего не известно.

Определим правило кортежем

$$R = \langle C, N, A \rangle, \quad (3)$$

где C – RDF-сеть условий,

N – RDF-сеть отрицательных условий,

A – RDF-сеть расширения.

Если существует преобразование Z из анонимных узлов во множество ресурсов, такое что любая интерпретация I сети G истинна для сети условий и ложна для всех утверждений сети отрицательных условий, то эта интерпретация истинна и для сети расширения.

$$\exists Z: Q \Rightarrow R, \forall I(G) = \text{true}: I(Z(C)) = \text{true} \ \& \ \forall S \in Z(N), I(S) = \text{false} \Rightarrow I(Z(A)) = \text{true}$$

В контексте задачи данного диссертационного исследования будем различать правила, определяющие интерпретацию семантически целостных сетей, и правила, определяющие интерпретацию сетей с нарушениями семантической целостности.

Это определяет существенное отличие разрабатываемой модели онтологий от модели OWL: в модели OWL, сеть, имеющая нарушения семантической

целостности, не может быть интерпретирована, тогда как в данной модели для нарушений семантики существует особая интерпретация.

Важнейшей особенностью разработанной модели онтологий является её способность выявлять и локализовать семантические ошибки. Для обеспечения такой возможности в модели определено понятие *парадокса*:

Парадокс – явление существования в семантической сети утверждений, не удовлетворяющих условиям ассоциированной онтологии.

С другой стороны, Парадокс – особый объект семантической сети ГИР, описывающий нарушение семантической целостности, имеющееся в этой сети.

Парадоксы определяются при помощи правил (ограничений), формирующих онтологию.

Парадокс может быть выявлен относительно ограничений классов – для некоторого объекта заявлена принадлежность классу, но как минимум одно из условий принадлежности не выполнено.

В третьей главе разрабатываются алгоритмы обнаружения семантических ошибок с использованием онтологических моделей. Определяются основные структуры и форматы данных.

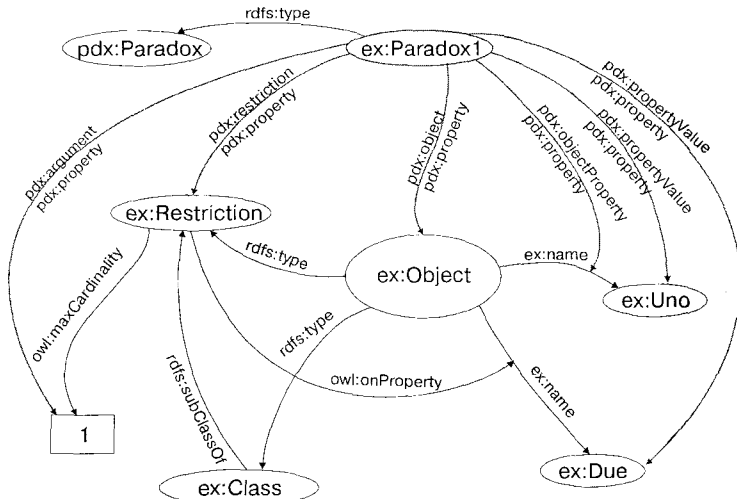


Рис. 3. Объект-парадокс

В терминах семантической сети парадокс локализуется следующим образом:

Р – объект класса `pdx:Paradox`, может быть наделен следующими свойствами:

Р `pdx:property` PP, где PP – свойства парадокса, которые могут быть:

Р `pdx:restriction` R, где R – нарушенное ограничение,

Р `pdx:argument` A, где A – аргументы нарушенного ограничения,

$R \text{ pdx:object } O$ – объект, нарушающий ограничение R ,

$R \text{ pdx:objectProperty } OP$, где OP – свойства объекта, нарушающие ограничение, и

$R \text{ pdx:value } V$, где V – значения свойств, нарушающие ограничение.

Разрабатывая общий алгоритм обнаружения семантических ошибок в ГИР, рассмотрим наиболее общий случай изменения ресурса – когда к исходному ГИР добавляется новый документ, и требуется произвести поиск ошибок в обновленном ГИР.

Обобщенный алгоритм проверки семантической целостности сети:

1. Построить семантическую сеть G добавляемого документа D .
2. Добавить семантическую сеть G нового документа в сеть V . ($V+G=V'$)
3. Выполнить расширительные правила ($V' \Rightarrow V^{\text{расш}}$).
4. Выполнить проверочные правила ($V^{\text{расш}} \Rightarrow V^{\text{пров}}$).
5. Найти множество P объектов-парадоксов и извлечь их свойства.
6. Сформировать отчет R о найденных парадоксах P и их причинах.

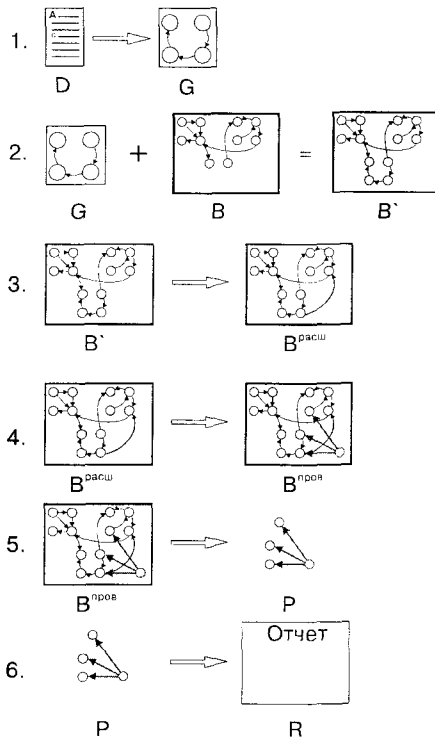


Рис. 4. Алгоритм поиска семантических ошибок в ГИР

Будем выполнять правила последовательно и в цикле, до тех пор, пока очередной проход цикла вносит изменения в семантическую сеть. Формализуем этот процесс в виде следующего алгоритма:

1. Для каждого правила составить все возможные отображения Z из множества пустых узлов во множество URL.
2. Добавить исполнительную часть правила в исходную сеть с использованием операции слияния семантических сетей.
3. Если хотя бы одно правило добавило хотя бы одно утверждение в исходную сеть, то процесс повторяется с пункта 1.

Рассмотрим алгоритм поиска отображений Z .

Составим список утверждений, формирующих оба графа условий, при этом для каждого утверждения определим *признак существования*. Если утверждение получено из сети положительных утверждений, то признак существования будет иметь значение «ИСТИНА», а если из сети отсутствующих утверждений – то «ЛОЖЬ».

Тогда каждое из утверждений (сейчас они уже рассматриваются как условия) может содержать до трех пустых узлов. Если идентификатор пустого узла встречается впервые за время выполнения данного правила, то будем искать подходящее для него значение URL, а если повторно – то заменим его на найденным ранее значением. Для представления соответствия между идентификаторами пустых узлов и URL будем использовать ассоциативный массив, где в качестве ключа будет выступать имя пустого узла, а в качестве значения – найденный для него URL, например:

```
Z={
    a => "http://comp.susu.ac.ru/a",
    b => "http://comp.susu.ac.ru/b"
},
```

где a и b – имена пустых узлов, а «http://comp.susu.ac.ru/a» и «http://comp.susu.ac.ru/b» – найденные для них значения.

Наличие в Z к моменту проверки условия одной из указанных в нем переменных уменьшает количество неизвестных в условии.

Алгоритм поиска отображения Z для расширительных правил:

1. Z – пустое.
2. В очередном условии заменим пустые узлы в соответствии с Z .
3. Если в условии остались пустые узлы, то найдем для них все возможные значения URL.

4. Если признак существования для рассматриваемого условия «ЛОЖЬ», и утверждение существует в сети, то обработка правила прекращается с признаком «не выполнено», иначе переход к следующему условию в цепочке и возврат значения, полученного от выполнения этого условия.

5. Если признак существования «ИСТИНА», и в текущем шаге не найдено ни одного нового элемента Z , то обработка правила прекращается с признаком «не выполнено», иначе выполняется следующее условие со всеми найденными в текущем шаге вариантами Z . При этом если для одного из вариантов Z будет возвращен статус «выполнено», то и текущее правило возвращает «выполнено».

6. Если далее по цепочке правил нет, то перейти к выполнению исполнимой части. Выполнение исполнимой части состоит в слиянии исходной сети с исполнимой частью условия, преобразованной с учетом Z . Операция слияния возвращает количество измененных утверждений, и если это количество оказалось равным 0, то возврат из правила осуществляется с признаком «не выполнено».

В четвертой главе разрабатывается архитектура информационной системы, предназначенной для обнаружения семантических ошибок в гипертекстовых информационных ресурсах, а также описывается процесс разработки онтологии для конкретной организации и её гипертекстового ресурса.

Наиболее эффективным способом внедрения программного обеспечения, предназначенного для обнаружения семантических ошибок в ГИР, является разработка информационной системы, основанной на *системе управления содержимым* информационного ресурса.

Для каждого документа, поступающего в ГИР, создается собственная семантическая сеть, которая вливается в общую семантическую сеть ГИР. Специальный программный модуль – онтопроцессор – на основании онтологического описания предметной области ГИР производит анализ и выявление парадоксов, указывая на возможный источник ошибки. Сведения об ошибках поступают редактору ГИР (оператору ИС), который имеет возможность внести изменения в документы ИС и исправить парадокс, действуя по той же цепочке, что и при первоначальном вводе документа.

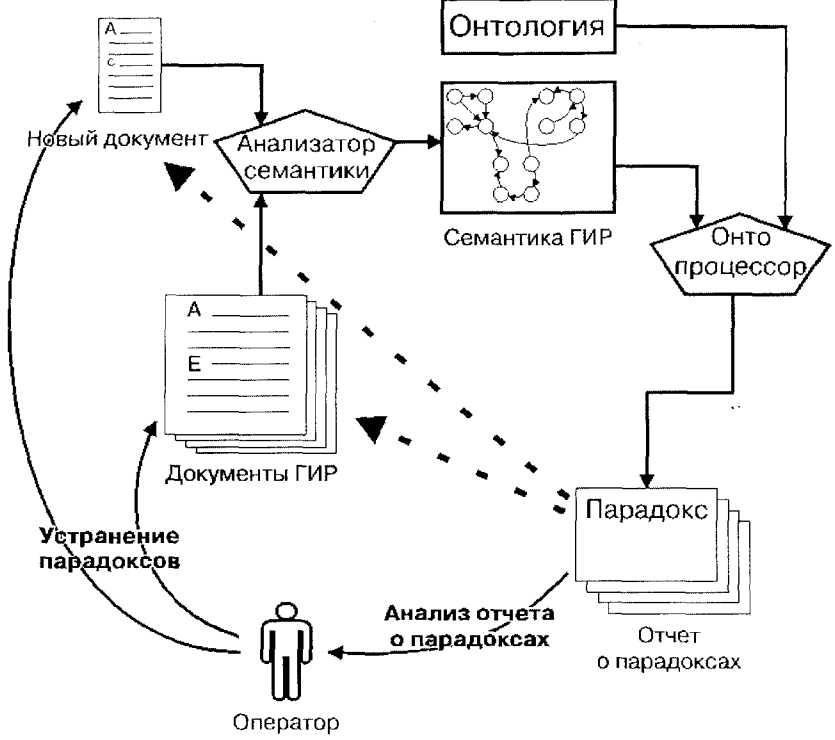


Рис. 5. Общая технологическая схема поиска семантических ошибок

Основная функциональность по обнаружению парадоксов лежит на онтологическом процессоре. Онтологический процессор – это программный модуль, входными данными для которого являются *естественная семантическая сеть* (см 2.3.3.) и *онтология*. На выходе семантического процессора – расширенная семантическая сеть, содержащая весь объем логических выводов, сделанных при помощи онтологии, и объекты-парадоксы, возникшие из-за нарушения условий онтологии. Онтопроцессор выполняет все правила, содержащиеся в онтологии согласно алгоритмам, описанным в главе 3.

Поскольку хранилище семантических сетей построено на базе Redland RDF Application Framework, то онтопроцессор строится на базе той же программной среды. Процессор онтологий реализуется как надстройка к REDLAND, и осуществляет все манипуляции над сетями при помощи Redland API.

XSLT – универсальный язык преобразования документов XML, разработанный W3C. Процессор XSLT является ядром системы управления

документами. Он используется для преобразования XML-форматов и диалектов. XSLT-процессор выполняет следующие основные функции:

- преобразование документов, поступающих от интерфейсной части ИС в формате HTML в формат, используемый при хранении документов (XML);
- преобразование из формата хранения (XML) в формат XHTML для выдачи посетителю ресурса (применение стилевого оформления);
- синтез информации о структуре ресурса для клиентской части в формате RDF;
- обновление индексных файлов при добавлении, перемещении или удалении документов;
- изменение, удаление или добавление атрибутов к существующим документам и разделам.

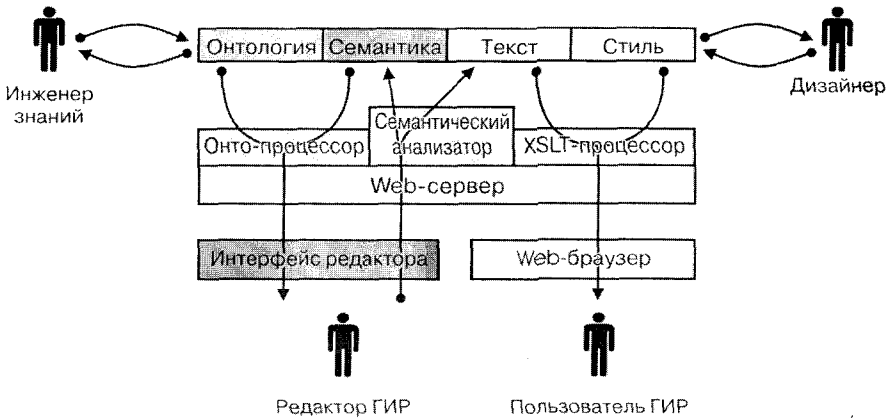


Рис. 6. Архитектура информационной системы «TergereLibrum»

Использование архитектуры, построенной на открытых стандартах линейки XML, позволяет получить гибкую и открытую к изменению систему, которая может быть дополнена модулями или включена в информационную среду предприятия.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

Проведенные в рамках данной работы исследования образуют теоретическую и практическую основу для решения задачи обнаружения семантических ошибок в ГИР. К основным результатам диссертационного исследования следует отнести следующее:

1. Проведен анализ категорий ГИР и ошибок ГИР, на основе которого выделена предметная область «ГИР организации» как основной объект применения разрабатываемых методов и алгоритмов.

2. Определено понятие *семантической целостности*, которое определяется как отсутствие противоречий между знаниями о предметной области и знаниями, представленными в текстах ГИР, которое позволяет делать выводы о наличии семантических ошибок в текстах информационного ресурса.

3. Разработаны методы поиска семантических ошибок в текстах ГИР, использующие модель *онтологий*, дополненную механизмом расширения на основе *семантических правил*, который позволяет интерпретировать нарушения семантической целостности при помощи объектов-парадоксов.

4. Предложены типовые алгоритмы проверки семантической целостности семантической сети, поиска парадоксов, локализации текстовых фрагментов, повлекших за собой возникновение парадоксов и, как следствие, предложен алгоритм обнаружения семантических ошибок в ГИР.

5. Разработана типовая онтология предметной области «ГИР организации», включающая в себя понятия, характерные для корпоративных Internet-сайтов: общие сведения о предприятии, продукция, сотрудники, взаимодействие с другими предприятиями.

6. Предложена методика раздельного хранения семантики ГИР и его естественного текста, позволяющая установить связь между фрагментом семантики и фрагментом текста, его породившим. Эта связь является необходимым звеном для перехода от нарушений семантической целостности семантической сети к смысловым ошибкам в текстах на естественном языке.

7. Разработана и внедрена в нескольких организациях информационная система «TergereLibrum», реализующая рассмотренные в исследовании принципы и алгоритмы. Использование в ИС индустриальных стандартов XML, RDF и IEEE 1600.1 позволило построить гибкую модульную систему, способную к расширению и совместимую с большим количеством программного обеспечения, обработки текстов, семантического анализа и анализа с учетом онтологий.

0399607

По теме диссертационной работы опубликованы следующие работы:

1. Мельников А.В., Хлопотов М.В. Парадоксы в семантике гипертекстового информационного ресурса (ГИР), достоверность ГИР. Материалы седьмой научно-практической конференции «Реинжиниринг бизнес-процессов на основе современных информационных технологий. Системы управления знаниями – 2004»//М.: МЭСИ, 2004с. 189-192

2. Хлопотов М.В. Применение RDF для описания структуры информационных ресурсов. // Интеллект, логистика, системология. Сборник научных трудов. Вып. 7 \ Под ред. Р.П. Чапцова. / Челябинск: Издание ЧНЦ РАЕН, РУО МАИ, ЧРО МАНПО, ЧРО МААНОИ, 2002. – 207 с.

3. Хлопотов М.В. Обеспечение достоверности Internet-ресурсов // Современные проблемы атомной энергетики. Сборник трудов международной научно-практической конференции/Снежинск, 2003, стр. 208-212

4. Мельников А.В., Хлопотов М.В. Поддержание достоверности гипертекстового информационного ресурса (на англ.)/Материалы 5-й Международной конференции по проблемам информатики и информационных технологий CSIT'2003.- Уфа, УГАТУ, 2003.- Т.-1.- С.197-198.

5. Мельников А.В., Хлопотов М.В. Онтологии: Семантическое программирование (на англ.)/ Материалы 6-й Международной конференции по проблемам информатики и информационных технологий CSIT'2004.- Будапешт, Венгрия, 2004.- Т.-2.- С.11-12.

6. Мельников А.В., Хлопотов М.В. Обеспечение достоверности информационного ресурса промышленного предприятия// АСУ-ПП-2003. Сборник трудов международной научно-практической конференции /Москва, 2003.



ХЛОПОТОВ Михаил Викторович

МЕТОДЫ И АЛГОРИТМЫ ОБНАРУЖЕНИЯ СЕМАНТИЧЕСКИХ
ОШИБОК В ГИПЕРТЕКСТОВЫХ ИНФОРМАЦИОННЫХ РЕСУРСАХ НА
ОСНОВЕ ОНТОЛОГИЧЕСКИХ МОДЕЛЕЙ

Специальность 05.13.11 – «Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей»

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать 12.11.2004
Формат 60x84 1/16. Бумага офсетная. Печать офсетная.
Объем 1.0 – изд. л. Тираж 100 экз.
Заказ 02 - 22.
Бесплатно

Издательство «Фрегат»
Центральный офис:
Г. Челябинск, ул. Васенко, 96.
тел. (3512) 63-34-59, 98-29-22
Лицензия ЛР № 063457 Комитета Российской Федерации по печати
Правительства Российской Федерации
Отпечатано в типографии издательства «Фрегат»
(г. Челябинск, ул. Чайковского, 9, 17-18).