

Российская Академия Наук  
Сибирское Отделение  
Институт систем информатики  
имени А.П.Ершова

На правах рукописи

ПОНОМАРЁВ Денис Константинович

**Синтаксическая однозначность  
при представлении знаний  
в логике первого порядка**

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

Научный руководитель –  
профессор А.Г.Марчук

Автореферат диссертации на соискание ученой степени  
кандидата физико–математических наук

Новосибирск 2006

Работа выполнена в Институте систем информатики  
имени А.П. Ершова СО РАН

**Научный руководитель:** Марчук Александр Гурьевич,  
доктор физико-  
математических наук

**Официальные оппоненты:** Морозов Андрей Сергеевич,  
доктор физико-  
математических наук

Витяев Евгений Евгеньевич,  
кандидат физико-  
математических наук

**Ведущая организация:** Институт автоматки и  
процессов управления ДВО  
РАН

Защита состоится 26 декабря 2006 года в 16 часов 00 минут на заседа-  
нии диссертационного совета К003.032.01 в Институте систем информатики  
имени А.П. Ершова Сибирского отделения РАН по адресу:

630090, г. Новосибирск, пр. Акад. Лаврентьева, 6.

С диссертацией можно ознакомиться в читальном зале библиотеки ИСИ  
СО РАН (пр. ак. Лаврентьева, 6).

Автореферат разослан                      ноября 2006 г.

Ученый секретарь  
Диссертационного совета,  
к.ф.-м.н.

Мурзин Ф.А.

# ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

## Актуальность проблемы

В настоящее время существует значительный интерес к методам и средствам декларативного представления знаний, который, в частности, связан с недавно озвученной концепцией "Semantic Web" и широко распространенным понятием формальной онтологии. Результатом этого является разработка и применение новых компьютерных языков для формального представления знаний, а также систем поддержки логического вывода. Каждый из новых языков соответствует некоторому подмножеству логики первого порядка, однако исходя из практики специалисты начали осознавать необходимость использования данной логики целиком для работы с теми задачами, которые стали актуальны в последнее время. Среди них можно выделить поиск в объемных слабо структурированных хранилищах информации и интеграцию гетерогенных источников данных.

В рамках данных задач декларативные описания применяются как терминологические базы знаний, в которых описаны термины вместе с заданной для них аксиоматической семантикой. Декларативные терминологические базы знаний обычно принято рассматривать с двух точек зрения: как графы связности терминов и как наборы логических утверждений. При этом широко используется понятие связи терминов по вхождению в одно (или близкие в определенной метрике) высказывания.

Связи между терминами в задаче поиска дают возможность переформулировать (усилить/ослабить) запрос в зависимости от результатов его исполнения. Аналогично этому в задаче интеграции связи влияют на точность построенного отображения между двумя описаниями данных. При этом аксиомы, заданные в декларативном описании, непосредственно определяют тип связей между терминами. Для решения конкретной задачи могут применяться различные стратегии, использующие разные типы связей в зависимости от ситуации.

В данной работе декларативные описания знаний рассматриваются как (элементарные) теории в логике первого порядка. Естественно, данное рассмотрение не является оригинальным и оно не может быть таковым в принципе, начиная с того времени, как был введен формальный аппарат логики.

С точки зрения связей терминов по вхождению для заданной теории можно рассматривать различные компоненты связности на множестве сигнатурных символов, используемых в записи предложений теории. Естественным образом возникает вопрос: являются ли компоненты связности одинаковыми для двух логически эквивалентных теорий в одной и той же сигнатуре? В целом, ответ является отрицательным и причина этому - синтаксическая суть нашего подхода. На практике это означает, что могут существовать два множества предложений, которые семантически эквивалентны, но синтаксически представлены по-разному. Возможно ли привести теорию (аксиомы теории) к такому виду, по которому однозначно определяются компоненты

связности сигнатуры теории? Данный вопрос представляет расширенную постановку проблемы разложимости, которая была сформулирована в 2003 году в связи с изучением формальных онтологий. Суть проблемы заключается в том, как определить, представима ли произвольная заданная теория в логике первого порядка в виде объединения двух (или более) теорий, имеющих непересекающиеся сигнатуры.

Вопрос разложимости имеет важное значение для формализации знаний, поскольку разложимость означает возможность разбить формальное описание интересующей предметной области на части, каждая из которых использует свой отдельный алфавит. При построении формального описания некоторой предметной области часто оказывается, что данные, полученные от экспертов (или, например, извлеченные автоматически из текстов), представляют собой некоторый набор фактов, которые необходимо структурировать, чтобы получить из них адекватную формальную модель. В частности, может быть интересно, существуют ли части знания, которые независимы друг от друга. Это в точности соответствует вопросу разложимости, если рассматривать формальное описание предметной области как логическую теорию (скажем, в некотором подмножестве логики первого порядка).

В данной работе решена проблема разложимости для произвольных элементарных теорий. Кроме того, доказана однозначность разложения сигнатуры и самой теории (с точностью до формул чистого равенства) в случае его существования. Таким образом, настоящая работа устанавливает связь между подходом к рассмотрению синтаксической связанности терминов с точки зрения графов и с точки зрения формул логики.

### **Цель работы**

- исследование возможности декомпозиции (разложения) формальных описаний в логике первого порядка на компоненты с непересекающимися алфавитами;
- изучение вопроса синтаксической однозначности при использовании элементарных теорий в логике первого порядка как средства формального представления знаний.

В результате работы автором был сформулирован критерий разложимости, доказана однозначность разложения теории на неразложимые компоненты с точностью до формул чистого равенства, показано приведение теории к такому виду, по которому однозначно определяется ее разложение.

### **Методы исследования:**

аппарат математической логики и теории моделей.

## **Научная новизна**

Результаты, полученные в работе, являются новыми и соответствуют тенденции времени в части связанных с ними информационных задач. В частности, до настоящей работы исследования по решению проблемы разложимости в теории с дизъюнктивными сигнатурами не встречались в публикациях. Не было показано ранее, что теория в логике первого порядка определяет компоненты связности сигнатуры однозначным образом. Также не была исследована взаимосвязь этих вопросов с современными постановками задач информационного поиска и интеграции источников данных.

## **Практическая ценность**

Полученные результаты имеют ценность для задач поиска и интеграции данных с использованием декларативных баз знаний, декомпозиции логических программ, распределенного исполнения логических операций, таких как проверка на непротиворечивость, над декларативными базами знаний большого объема. Кроме того, факты, показанные в работе, представляют интерес в связи с исследованием *проблемы релевантности* в логике и философии.

## **Апробация работы**

Результаты работы докладывались в рамках следующих научных встреч: Informatik-2006 - немецкой конференции по информатике (г. Дрезден, Германия), русско-немецком симпозиуме по биоинформатике в 2005 году (г. Билефельд, Германия), семинаре СОМО (г. Дармштадт, Германия, 2005 г.), международной конференции по биоинформатике BGRS'2006 (Новосибирск, 2006 г.), конференции "Технологии Майкрософт в информатике и программировании" (Новосибирск, 2006 г.), Международной Научной Студенческой Конференции в 2003 году (Новосибирск), а также на семинарах в Институте систем информатики СО РАН, Институте математики СО РАН, Институте цитологии и генетики СО РАН.

По теме диссертации опубликовано 10 печатных работ.

## **Структура и объем работы**

Диссертационная работа состоит из введения, трех глав и списка литературы. Объем диссертации - 89 стр. Список литературы содержит 63 наименования. Работа включает 8 рисунков.

## **Содержание работы**

**В первой главе** описан опыт создания автором прикладной логической теории в области биоинформатики.

Показывается, как такие теории могут возникнуть на практике при разработке конкретных программных приложений. Последовательно изложены все основные шаги от теоретической работы до реализации в качестве программного продукта.

Данная глава проясняет суть подхода автора, исходя из практических соображений, и предлагает достаточно примеров для освоения последующих глав.

Неудивительно, что биоинформатика в данное время является одним из наиболее важных поставщиков задач для информационных технологий. Объемы данных, доступные в этой области, дают широкие возможности для апробации теоретических результатов на практике. А сама суть этих данных уже, как правило, гарантирует актуальность проводимых исследований.

Нами было разработано формальное представление анатомии "модельного" растения *Arabidopsis thaliana* в развитии, которое используется для навигации в базе данных экспрессии генов этого растения и анализа взаимосвязей фенотипических аномалий. Само формальное представление является набором предложений логики первого порядка, реализованных программно в виде дедуктивной базы данных в системе Protégé.

Разработки по формальному описанию анатомии в развитии, равно как и сама база данных были выполнены в Институте Цитологии и Генетики СО РАН, внедрены в этом институте и являются частью проекта "Computable Plant" (ИЦиГ СО РАН; Институт Геномики и Биоинформатики, Университет Калифорнии, Ирвайн, США; Отдел Биологии, Технологический Институт Калифорнии, Пасадена, США; <<http://computableplant.org>>).

Для биологии важность исследований модельных организмов объясняется возможностью проецировать полученные результаты на другие организмы. В данное время для определенных классов организмов выделены "модельные" представители. Как правило, они выбираются исходя из ряда генотипических особенностей и учитывая также практическое удобство их изучения в экспериментальных условиях.

Среди растений в биологии, как модель, выделен вид *Arabidopsis thaliana*. Согласно международной исследовательской инициативе, к 2010 году rozpoзнанные функции каждого из генов этого растения должны быть интегрированы в генные сети, дополняя и/или корректируя их. Информация о генах Арабидопсиса, накопленная к настоящему времени, должна помочь в систематизированном освоении новой информации о растении и использовании ее для других видов.

Описание функций генов на уровнях РНК, белков, клеток, тканей, а также на уровне органов и всего организма для разных генотипов являются сложными и емкими знаниями. Огромный объем разрозненной информации из публикаций требует объединения и систематизации для того, чтобы понять, что уже известно об экспрессии генов Арабидопсиса и что еще предстоит узнать. В связи с этой задачей в Институте Цитологии и Генетики СО РАН была разработана и постоянно развивается база данных AGNS (*Arabidopsis GeneNet Supplementary database*), которая является открытым Интернет-ресурсом (<http://www.bionet.nsc.ru/agns>) и предоставляет доступ к описанию функций известных генов Арабидопсиса на различных уровнях

мРНК, белков, клеток, тканей, а также уровне органов и всего организма в целом, как для дикого, так и для мутантных типов.

Вся информация содержится в двух модулях AGNS: базе данных по экспрессии генов и базе данных по фенотипам растения (смысл данных поясняется в разделе 2 первой главы диссертации). Кроме того, разработаны два контролируемых словаря содержимого базы данных. В дополнение к этому потребовалось разработать формальное описание анатомического строения растения на уровне клеток, тканей и органов в развитии. Следуя популярным терминам в инженерии знаний, для краткости в данной работе мы называем это формальное описание онтологией. Основой для создания онтологии послужили статьи по паттернам экспрессии генов, фенотипам и непосредственно процессам развития растения.

В то время как контролируемые словари были разработаны, в основном, для аннотаторов статей, вводящих данные в AGNS, целью построения онтологии было применить это формальное описание в алгоритмах анализа данных экспрессии генов и фенотипических данных по аномалиям развития.

В первой главе представлены результаты формализации анатомии растения в развитии на примере задачи навигации в базе данных AGNS. Мы считаем этот пример достаточно простым и, в то же время, иллюстративным для нашей работы. Опыт работы показывает, что, имея конкретную практическую задачу как цель, выбор формализации можно сделать более осознанно и обосновать лучше, чем в некотором абстрактном случае разработки формализации для предметной области.

В первой главе формулируется задача навигации в базе данных при помощи онтологии. Исходя из ее постановки и особенности процессов развития выбранного организма, обосновывается выбор сигнатуры для разрабатываемой логической теории. Сигнатура выбирается предикатной, состоящей и приведенных ниже предикатов.

*Anatomy\_Element*<sup>1</sup>

*Developmental\_Stage*<sup>1</sup>

*Has\_Developmental\_Stage*<sup>2</sup>(*Anatomy\_Element* × *Developmental\_Stage*)

*Before*<sup>2</sup>(*Developmental\_Stage* × *Developmental\_Stage*)

*Occurs\_In*<sup>2</sup>(*Developmental\_Stage* × *Developmental\_Stage*)

Верхними индексами обозначена местность предикатов. Для краткости, в том числе, при записи аксиом, полагаем, что бинарные предикаты определены на декартовых произведениях тех множеств, которые выделяются унарными предикатами и указаны неформально в круглых скобках выше. Здесь целесообразно прокомментировать только предикат *Occurs\_In*, который используется в работе для описания того, что одна стадия развития протекает в другой стадии.

Далее в работе вводится система аксиом теории в заданной выше сигнатуре. Для примера приведем некоторые из них:

$$\forall x \exists s (Anatomy\_Element(x) \rightarrow Has\_Developmental\_Stage(x, s)) \quad (1)$$

$$\forall s \neg (Occurs\_In(s, s) \vee Before(s, s)) \quad (2)$$

$$\forall x, s_1, s_2 \neg (Has\_Developmental\_Stage(x, s_1) \wedge Has\_Developmental\_Stage(x, s_2) \wedge Occurs\_In(s_1, s_2)) \quad (3)$$

$$\forall s_1, s_2 \neg (Before(s_1, s_2) \wedge Occurs\_In(s_1, s_2)) \quad (4)$$

$$\forall s_1, s_2, s_3 (Before(s_1, s_2) \wedge Before(s_2, s_3) \rightarrow Before(s_1, s_3)) \quad (5)$$

$$\forall s_1, s_2, s_3 (Occurs\_In(s_1, s_2) \wedge Occurs\_In(s_2, s_3) \rightarrow Occurs\_In(s_1, s_3)) \quad (6)$$

С точки зрения практической реализации, предложения 1-4 задают ограничения на содержание фактов об индивидуумах, вводимых в онтологию. Предложения 5, 6 определяют транзитивность отношений *Before* и *Occurs\_In* и, фактически, задают правила для вывода новых фактов из уже имеющихся.

Для задачи навигации в базе данных и упрощения практической реализации далее сигнатура расширяется дополнительными бинарными предикатами. В частности, это делается для того, чтобы обеспечить возможность раскрытия анатомической структуры в процессе развития с необходимым уровнем детализации.

В результате в работе приводится список запросов наиболее актуальных для задачи навигации в базе данных и контекста самой предметной области, которые выразимы в терминах построенной онтологии:

- S - стадия развития; найти стадии раньше/позже S.
- A - анатомический элемент, S - его стадия развития; найти элементы, принадлежащие A на стадии S.
- A - анатомический элемент, S - его стадия развития; найти элементы, содержащие A на стадии S.
- A - анатомический элемент; найти анатомические элементы, из которых он развивается.
- A - анатомический элемент; найти анатомические элементы, которые развиваются из A.
- A, B - анатомические элементы; найти стадии, которые претерпевает A, будучи частью B.



- А, В - анатомические элементы; найти стадии, которые претерпевает А, в то время как он содержит В.
- А, В - анатомические элементы; является ли А частью В на некоторой стадии развития В (или А)? (по-другому: является ли А частью В когда-либо?).

Другие соображения, которые помогают оценить представленную онтологию, относятся к ее наполнению реальными данными, полученными из экспериментов по развитию растения *Arabidopsis thaliana*. Здесь мы рассматриваем онтологию в ее программной реализации - как дедуктивную базу данных, понимая наполнение данными как добавление формул от констант, не присутствовавших ранее в сигнатуре. В то же время полагаем, что формулы используют только те предикатные символы, которые введены нами изначально.

В процессе тестирования текстовая информация об анатомическом строении и развитии растения была выделена из общедоступных научных статей и введена в онтологию в форме, заданной формальными конструкциями выше. В этом процессе мы столкнулись с двумя основными проблемами, относящимися к недостаточной полноте информации, а именно:

1) стадии развития были определены не у всех анатомических элементов, которые мы хотели бы включить в онтологию;

2) развитие многих анатомических элементов описывалось в статьях в виде "*X обладает следующими свойствами, когда Y находится на стадии S*" (например, "*LDS2 leaf shows meristematic divisions throughout the mesophyll*"), но в то же время подобные описания приводились лишь для некоторых стадий Y.

Это привело к необходимости введения *искусственных* стадий развития у анатомических элементов.

Другой вопрос, который стоит здесь упомянуть, касается циклов в развитии анатомических элементов. Например, это относится к итеративному формированию листьев на поздних стадиях развития растения. В работе показано, что в сигнатуре представленной теории подобная информация может быть выражена только формулой второго порядка. В практической реализации пришлось обойти данное препятствие с помощью специального вычисления элементов второго порядка программным образом.

В выборе практической реализации мы руководствовались двумя побуждениями: применить достаточно широко распространенный и достаточно выразительный язык представления онтологий, а также использовать редактор онтологий, обладающий богатыми функциями импорта-экспорта и визуализации. В результате это привело к выбору языка OWL-DL и редактора Protégé (версия 3.1, <http://protege.stanford.edu>), а сама теория была реализована как дедуктивная база данных в этой системе. Текстовая информация о

развитии растения *Arabidopsis thaliana*, выделенная из научных статей, была представлена в формальном виде.

Проблема представления аксиом-правил не рассматривалась нами как существенная, поскольку в представленной теории число таких аксиом невелико. Таким образом, их возможно переписать вручную на язык используемой в каждом конкретном случае машины логического вывода. В нашем случае мы использовали машину вывода Algernon для редактора Protégé.

Часть предложений построенной теории была реализована как ограничения целостности данных, другая часть - как правила прямого или обратного вывода. Понятие транзитивности отношений реализовано в OWL с помощью конструкции *TransitiveProperty*, но, так как в данное время целиком язык OWL поддерживается не всеми существующими машинами вывода, было принято решение реализовать аксиомы транзитивности и в OWL, и отдельно на языке применяемой машины Algernon.

Как было сказано выше, описанный формализм был разработан с целью использования в алгоритмах анализа данных по экспрессии генов и фенотипическим аномалиям растения. В частности, с его применением автором был разработан метод нахождения взаимосвязей между аномалиями развития растения на основе анализа данных по его фенотипам. Основная проблема описания анатомии в развитии была сведена к частичному упорядочению на множестве стадий развития. Сигнатура построенной теории имеет небольшое число элементов и одновременно соответствует существенным выразительным характеристикам с точки зрения описания реальных данных.

В процессе исследований было проведено сравнение разработанной теории с широко известными информационными ресурсами "TAIR" и "Plant Ontology", которые предоставляют таксономии терминов по анатомии и развитию растений (в частности, и модельного растения *Arabidopsis thaliana*). В частности, была проанализирована связь терминологических систем и структур данных в онтологии, разработанной консорциумом "Plant Ontology", и онтологии, разработанной автором. В результате было доказано, что формальное представление, разработанное автором, превосходит разработку консорциума в возможности использования для анализа экспериментальных данных высокой детализации по времени.

**Во второй главе** изложены основные результаты работы, касающиеся проблем разложимости и синтаксической однозначности при построении элементарных теорий.

Автором сформулирован критерий разложимости для произвольной элементарной теории, доказана однозначность разложения сигнатуры в дизъюнктивное объединение, отвечающее разложению теории, и однозначность разложения самой теории с точностью до формул чистого равенства.

Поясним суть полученных результатов, введя необходимые определения.

**Определение 1** Рассмотрим сигнатуру  $\Sigma$  и теорию  $\mathcal{T}$  этой сигнатуры. **Теория  $\mathcal{T}$  – разложимая**, если можно подобрать такое дизъюнктивное разложение сигнатуры  $\Sigma = \Sigma_1 \cup \Sigma_2$ ,  $\Sigma_1 \cap \Sigma_2 = \emptyset$ ,  $\Sigma_1 \neq \emptyset \neq \Sigma_2$  и такую систему аксиом теории  $\mathcal{S}$ , которая представляется объединением  $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ , в котором в формулах  $S_i$ ,  $i = 1, 2$  используются только символы из части  $\Sigma_i$  сигнатуры  $\Sigma$ . Будем далее обозначать разложение теории как  $\mathcal{T} = \mathcal{S}_1 \otimes \mathcal{S}_2$ , а разложение сигнатуры как  $\Sigma = \Sigma_1 \amalg \Sigma_2$ .

Для произвольной теории  $\mathcal{T}$  можно было бы рассматривать тривиальное разложение, в котором  $\Sigma_1 = \emptyset$  или  $\Sigma_2 = \emptyset$ , а множество формул  $\mathcal{S}_1$  (соответственно  $\mathcal{S}_2$ ) образовано формулами чистого равенства теории  $\mathcal{T}$ , предложениями теории не использующими сигнатурных символов. Однако такое разложение малосодержательно, поэтому, согласно введенному определению, далее под *разложимой теорией* будем понимать такую теорию, которая обладает нетривиальным разложением, в котором  $\Sigma_1 \neq \emptyset \neq \Sigma_2$ . Если имеются только тривиальные разложения, то теорию будем называть **неразложимой**.

К примеру, если сигнатура  $\Sigma$  состоит только из одного предиката, то любая теория этой сигнатуры  $\Sigma$  (даже определенная пустым множеством аксиом) неразложима.

**Проблема 1** Рассмотрим теорию  $\mathcal{T}$  сигнатуры  $\Sigma$ , определенную некоторым множеством замкнутых формул  $\Phi$  сигнатуры  $\Sigma$ . Как исходя из системы аксиом  $\Phi$  определить разложимость теории  $\mathcal{T}$ ?

**Проблема 2** Рассмотрим конечно аксиоматизируемую теорию  $\mathcal{T}$  сигнатуры  $\Sigma$ , определенную некоторым конечным множеством предложений  $\Phi$  сигнатуры  $\Sigma$ . Как исходя из системы аксиом  $\Phi$  определить разложимость теории  $\mathcal{T}$ ?

Для точной формулировки результатов работы расширим определение 1, сформулировав его не для систем аксиом, а на языке теорий.

**Определение 2** Рассмотрим теории  $\mathcal{S}_1, \mathcal{S}_2$  дизъюнктивных сигнатур  $\Sigma_1, \Sigma_2$ ;  $\Sigma_1 \cap \Sigma_2 = \emptyset$ ,  $\Sigma_1 \neq \emptyset \neq \Sigma_2$ .

Теория  $\mathcal{T}$  сигнатуры  $\Sigma = \Sigma_1 \amalg \Sigma_2$  **разлагается в произведение теорий  $\mathcal{S}_1, \mathcal{S}_2$**  (обозначаем  $\mathcal{T} = \mathcal{S}_1 \otimes \mathcal{S}_2$ ), если она получается дедуктивным замыканием всех предложений теории  $\mathcal{S}_1$  (сигнатуры  $\Sigma_1$ ) и всех предложений теории  $\mathcal{S}_2$  (сигнатуры  $\Sigma_2$ ) в исчислении предикатов расширенной сигнатуры  $\Sigma$ :  $\mathcal{S}_1, \mathcal{S}_2 \vdash \mathcal{T}$ .

При этом теории  $\mathcal{S}_1, \mathcal{S}_2$  называем **компонентами (разложимости) теории  $\mathcal{T}$** .

Основываясь на интерполяционной теореме Крейга, в работе доказано, что для любой (элементарной) теории  $\mathcal{T}$  сигнатуры  $\Sigma$  можно однозначно определить дизъюнктивное разложение сигнатуры  $\Sigma = \Sigma_1 \amalg \dots \amalg \Sigma_n$ , которое с точностью до перестановки сомножителей отвечает любому разложению теории  $\mathcal{T} = \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$  на неразложимые компоненты разложимости  $\mathcal{S}_i$  – теории сигнатуры  $\Sigma_i$ ,  $i = 1, \dots, n$  (см. теорему 1 ниже). Также показано, какие шаги необходимо выполнить для того, чтобы привести систему аксиом теории к такому виду, по которому однозначно определяется это разложение.

Интересным является тот факт, что, независимо от того, какую систему аксиом теории выбрать, компоненты разложимости получаются одни и те же.

Обсуждение данных проблем следует начать с такого замечания.

**Замечание 1** Если теорию  $\mathcal{T}$  сигнатуры  $\Sigma$  можно определить такой системой аксиом, в которой используются только символы части сигнатуры  $\Sigma' \subset \Sigma$ , то такая теория – разложимая. Она разлагается на теорию меньшей сигнатуры  $\Sigma'$  и теории символов из дополнения  $\Sigma \setminus \Sigma'$ , определенные множествами тождественно истинных предложений.

В условиях замечания проблема разложимости по существу сводится к проблеме разложимости теории меньшей сигнатуры. Это рассуждение приводит к таким понятиям.

**Определение 3** Рассмотрим сигнатуру  $\Sigma$  и теорию  $\mathcal{T}$  этой сигнатуры. Теория  $\mathcal{T}$  – **приводимая**, если можно подобрать такое подмножество  $\Sigma' \subset \Sigma$  сигнатуры  $\Sigma$  и такую систему аксиом теории  $\mathcal{S}$ , в формулах которой используются только символы из части  $\Sigma'$  сигнатуры  $\Sigma$ . В этом случае будем говорить точнее, что теория  $\mathcal{T}$  **приводится к теории  $\mathcal{T}'$  меньшей сигнатуры  $\Sigma'$** .

Если же такое невозможно, и любая система аксиом теории  $\mathcal{T}$  использует все сигнатурные символы, то теорию  $\mathcal{T}$  называем **неприводимой**.

Символы сигнатуры  $\Sigma$  теории  $\mathcal{T}$ , которые невозможно исключить из любой ее системы аксиом, называем **существенными** (символами теории  $\mathcal{T}$ ).

Можно считать, что несущественные символы ничего не добавляют с семантической точки зрения, но изменяют предложение синтаксически. Примером несущественного вхождения символа  $p$  в предложения теории может быть вхождение в виде подформулы  $(p \vee \neg p)$ .

Таким образом, решение проблемы разложимости должно включать предварительный этап. На нем следует определить приводимость или неприводимость теории и сигнатуру, к которой она сводится, выделить существенные символы теории и аксиомы приводимости. А уже далее устанавливать разложимость или неразложимость неприводимой теории меньшей сигнатуры и определить аксиомы разложимости.

Относительно приводимости теорий в работе доказываются следующие ключевые утверждения.

**Замечание 2** Рассмотрим сигнатуру  $\Sigma'$  и ее расширение  $\Sigma$ .

Пусть  $\phi$  и  $\psi$  – предложения меньшей сигнатуры  $\Sigma'$ , причем в исчислении предикатов большей сигнатуры  $\Sigma$  выводима секвенция  $\phi \vdash \psi$ . Тогда эта секвенция выводима и в меньшей сигнатуре  $\Sigma'$ .

**Предложение 1** Рассмотрим расширение сигнатуры  $\Sigma' \subseteq \Sigma$  и теорию  $\mathcal{P}$  сигнатуры  $\Sigma'$  соответственно. Рассмотрим предложение  $\varphi$  сигнатуры  $\Sigma$ .

Тогда если формула  $\varphi$  следует из теории  $\mathcal{P}$ , то найдется предложение  $\theta$  теории  $\mathcal{P}$  такое, что  $\theta$  включает только те символы сигнатуры  $\Sigma'$ , которые встречаются в формуле  $\varphi$  и из которого предложение  $\varphi$  следует:  $\mathcal{P} \vdash \theta$ ,  $\theta \vdash \varphi$ .

Это предложение, доказательство которого опирается на интерполяционную теорему Крейга, делает возможным следующее определение.

**Определение 4** Рассмотрим теорию  $\mathcal{T}$  и предложение теории  $\varphi \in \mathcal{T}$ .

Предложение  $\varphi$  назовем **приводимым в теории  $\mathcal{T}$** , если найдется такое предложение  $\theta$  теории  $\mathcal{T}$ ,  $\theta \in \mathcal{T}$ , которое использует строго меньшее число сигнатурных символов формулы  $\varphi$ , и выполнена секвенция  $\theta \vdash \varphi$ . При этом, предложение  $\theta$  назовем **сужением формулы  $\varphi$  в теории  $\mathcal{T}$** .

Если таких предложений  $\theta$  найти нельзя, то формулу  $\varphi$  будем называть **неприводимой** в теории  $\mathcal{T}$ .

**Замечание 3** Любая неприводимая формула  $\varphi$  теории  $\mathcal{T}$  включает только существенные символы теории  $\mathcal{T}$ .

Из этого замечания следует, что существенные символы теории определяются по любой системе аксиом теории, образованной неприводимыми предложениями. Отметим к тому же, что любая система аксиом теории  $(Ax)$  может быть переработана в неприводимую систему аксиом  $(Ax)'$  если ее предложения заменять их неприводимыми сужениями в данной теории.

Тем самым это доказывает следующее предложение.

**Предложение 2** Пусть  $\mathcal{T}$  – теория сигнатуры  $\Sigma$ . Рассмотрим множество существенных символов  $\Sigma'$  сигнатуры  $\Sigma$ :  $\Sigma' \subseteq \Sigma$ .

Тогда теорию  $\mathcal{T}$  можно определить некоторой системой аксиом сигнатуры  $\Sigma'$ . При этом, данная система аксиом в сигнатуре  $\Sigma'$  определяет неприводимую теорию  $\mathcal{T}'$ .

Далее в работе рассматриваются уже неприводимые теории и неприводимые системы аксиом.

Решающим шагом к определению компонент разложимости служит следующий вариант интерполяционной теоремы Крейга.

**Предложение 3** Рассмотрим разложение сигнатуры  $\Sigma = \Sigma_1 \amalg \Sigma_2$  и теории  $\mathcal{P}, \mathcal{Q}$  сигнатур  $\Sigma_1, \Sigma_2$  соответственно. Рассмотрим предложение  $\varphi$  сигнатуры  $\Sigma$ .

Тогда если  $\varphi$  следует из объединения теорий  $\mathcal{P}, \mathcal{Q} \vdash \varphi$ , то найдутся такие предложения  $\theta$  и  $\phi$  теорий  $\mathcal{P}$  и  $\mathcal{Q}$  соответственно, из которых следует предложение  $\varphi$ :  $\mathcal{P} \vdash \theta$ ,  $\mathcal{Q} \vdash \phi$ ,  $\theta, \phi \vdash \varphi$ , причем, предложение  $\theta$  включает только такие символы сигнатуры  $\Sigma_1$ , которые встречаются в формуле  $\varphi$ , а предложение  $\phi$  содержит только те символы сигнатуры  $\Sigma_2$ , которые используются в формуле  $\varphi$ .

Из предложения 3 выводится способ установить разложимость теории  $\mathcal{T}$ . Для его описания определим понятие **разложимого предложения теории**.

**Определение 5** Рассмотрим теорию  $\mathcal{T}$  и предложение теории  $\varphi \in \mathcal{T}$ .

Предложение  $\varphi$  назовем **разложимым в теории  $\mathcal{T}$** , если найдутся такие предложения  $\theta$  и  $\psi$  теории  $\mathcal{T}$ ,  $\theta \in \mathcal{T}, \psi \in \mathcal{T}$ , которые включают символы

только из  $\varphi$ , не содержат общих сигнатурных символов, не являются формулами чистого равенства, и для которых выполнена секвенция  $\theta, \psi \vdash \varphi$ . При этом  $\theta$  и  $\psi$  будем называть **фрагментами разложения** предложения  $\varphi$ .

Если таких  $\theta$  и  $\psi$  найти нельзя, то предложение  $\varphi$  будем называть **неразложимым в теории  $\mathcal{T}$** .

**Замечание 4** Если предложение  $\varphi$  теории  $\mathcal{T}$  неприводимо, то фрагменты его разложения  $\theta$  и  $\psi$  также являются неприводимыми предложениями теории.

Для теории  $\mathcal{T}$  через  $\mathcal{T}^\#$  обозначим теорию чистого равенства, содержащуюся в  $\mathcal{T}$ . Она определяется замкнутыми формулами чистого равенства, принадлежащими теории  $\mathcal{T}$ . Для произвольной теории  $\mathcal{S}$  через  $\langle \mathcal{S}, \mathcal{T}^\# \rangle$  обозначим теорию  $\mathcal{S}$  вместе с предложениями чистого равенства теории  $\mathcal{T}$ .

Основным утверждением на пути к доказательству однозначности разложения является следующая лемма и ее обобщение.

**Лемма 1** При любом разложении теории  $\mathcal{T} = \mathcal{S}_1 \otimes \mathcal{S}_2$  в произведение теорий сигнатур  $\Sigma_1, \Sigma_2$  соответственно ( $\Sigma_1 \neq \emptyset \neq \Sigma_2$ ) каждое неразложимое предложение  $\varphi$  теории  $\mathcal{T}$ , включающее сигнатурные символы, выводится либо только из теории  $\langle \mathcal{S}_1, \mathcal{T}^\# \rangle$ , либо только из теории  $\langle \mathcal{S}_2, \mathcal{T}^\# \rangle$ .

Если формула  $\varphi$ , к тому же, неприводима, то она содержится либо в теории  $\langle \mathcal{S}_1, \mathcal{T}^\# \rangle$  сигнатуры  $\Sigma_1$ , либо в теории  $\langle \mathcal{S}_2, \mathcal{T}^\# \rangle$  сигнатуры  $\Sigma_2$ . В частности, она содержит символы либо только из сигнатуры  $\Sigma_1$ , либо только из сигнатуры  $\Sigma_2$ .

**Следствие 1** Рассмотрим произвольное нетривиальное разложение теории  $\mathcal{T} = \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$ . Тогда любое неприводимое и неразложимое предложение теории  $\varphi \in \mathcal{T}$  принадлежит некоторой теории  $\langle \mathcal{S}_i, \mathcal{T}^\# \rangle$ .

Данное следствие показывает, что для любого разложения теории любое неприводимое и неразложимое предложение теории содержит сигнатурные символы только из какой-то одной компоненты разложимости теории. Это позволяет определить разложение сигнатуры и компоненты разложимости теории  $\mathcal{T}$ .

Следует формализовать полученные результаты с помощью следующего ниже понятия. Приведем его для произвольной системы аксиом  $\Phi$  некоторой теории сигнатуры  $\Sigma$ .

**Определение 6** Два сигнатурных символа  $P, Q \in \Sigma$  называются **непосредственно связанными** (системой аксиом  $\Phi$  теории  $\mathcal{T}$ ), если  $P$  и  $Q$  входят в одну аксиому  $\psi \in \Phi$ .

Символы  $P, Q$  называются **связанными** (системой аксиом  $\Phi$ ), если найдется такая последовательность символов  $P = T_1, \dots, T_q = Q$ , в которой любая соседняя пара  $T_i, T_{i+1}$  связана непосредственно.

Отношение связности является отношением эквивалентности для неприводимой теории  $\mathcal{T}$ , так что множество  $\Sigma$  разбивается на компоненты связности – эквивалентности. По сути, получаем неориентированный помеченный граф (с петлями в вершинах), вершинами которого служат символы сигнатуры, а аксиомы определяют связывающие их ребра.

Условие неприводимости теории гарантирует отсутствие в таком графе изолированных вершин без петель – синглетонов. Каждый сигнатурный символ включается в некоторую аксиому неприводимой теории  $\mathcal{T}$ , так что в этом случае каждая вершина графа инцидентна некоторому ребру. При этом условии отношение связности можно определить двойственным образом через отношения не вершин, а ребер графа инцидентности.

**Определение 7** Две аксиомы  $\varphi, \psi \in \Phi$  назовем **непосредственно связанными**, если они содержат некоторый общий символ сигнатуры  $\Sigma$ .

Аксиомы  $\varphi, \psi$  назовем **связанными** (системой аксиом  $\Phi$ ) если найдется такая последовательность аксиом из  $\Phi$ :  $\varphi = \phi_1, \dots, \phi_q = \psi$ , в которой любая соседняя пара  $\phi_i, \phi_{i+1}$  связана непосредственно.

Компоненты связности однозначно определяются либо перечислением входящих в них сигнатурных символов (вершин графа), либо перечислением аксиом (ребер графа), включающих эти символы.

Применим введенные выше определения 6, 7 к максимальной системе неприводимых и неразложимых аксиом  $\Psi(\mathcal{T})$  неприводимой теории  $\mathcal{T}$ . Определим *стандартные компоненты связности сигнатуры*  $\Sigma = \Sigma_1 \amalg \dots \amalg \Sigma_n$  относительно этой системы аксиом. Построим *стандартное разложение теории* на компоненты разложимости, отвечающее этому разложению сигнатуры.

По построению каждая аксиома из системы  $\Psi(\mathcal{T})$  содержит символы только из какой-то одной дизъюнктивной части сигнатуры  $\Sigma_i$ . Обозначим для каждого  $i = 1, \dots, n$  через  $S_i$  все аксиомы, содержащие символы из  $\Sigma_i$ . Поскольку теория неприводимая, каждый символ сигнатуры присутствует в некоторой аксиоме из  $\Psi$ . Значит,  $S_i$  – множество аксиом соответствующей компоненты связности системы аксиом. Вместе эти множества аксиом определяют *стандартное разложение теории*  $\mathcal{T} = S_1 \otimes \dots \otimes S_n$ .

Заметим, что аксиомы  $S_i, i = 1, \dots, n$  определяют неразложимые теории. Действительно, если бы некоторая система  $S_i$  определяла разложимую теорию, то по лемме 1 сигнатура  $\Sigma_i$  и система аксиом  $S_i$  была бы несвязна. А по построению это компонента связности.

Приведем заключительные утверждения перед формулировкой основного результата - Теоремы 1.

**Лемма 2** Любые системы неприводимых и неразложимых аксиом  $\Phi$  неприводимой теории  $\mathcal{T}$  определяют одно и то же отношение связности на символах сигнатуры  $\Sigma$  теории  $\mathcal{T}$ . В частности, они определяют одни и те же компоненты связности  $\Sigma_i, i = 1, \dots, n$  сигнатуры  $\Sigma$ .

**Лемма 3** Пусть  $\mathcal{T}$  - неприводимая теория сигнатуры  $\Sigma$ , а  $\Sigma = \Sigma_1 \amalg \dots \amalg \Sigma_n$  - есть стандартное разложение сигнатуры на компоненты связности.

Рассмотрим произвольное разложение теории на компоненты  $\mathcal{T} = \mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_m$ . Тогда разложение сигнатуры, отвечающее этому разложению теории, получается из разложения  $\Sigma = \Sigma_1 \amalg \dots \amalg \Sigma_n$  объединением некоторых компонент  $\Sigma_i$ .

В частности, из связности некоторой системы неприводимых и неразложимых аксиом теории  $\mathcal{T}$  следует неразложимость теории  $\mathcal{T}$ .

**Лемма 4** Пусть теория  $\mathcal{T}$  сигнатуры  $\Sigma = \Sigma_1 \cup \Sigma_2$  разлагается в произведение теорий  $\mathcal{S}_1$  и  $\mathcal{S}_2$  дизъюнктивных сигнатур  $\Sigma_1, \Sigma_2$ ,  $\Sigma_1 \cap \Sigma_2 = \emptyset$ ,  $\Sigma_1 \neq \emptyset \neq \Sigma_2$ . Предположим, что эти теории содержат все формулы чистого равенства теории  $\mathcal{T}$ :  $\mathcal{S}_i \supseteq \mathcal{T}^\#$ ,  $i = 1, 2$ .

Тогда каждая теория  $\mathcal{S}_i$  состоит из всех предложений теории  $\mathcal{T}$ , которые включают только символы из части сигнатуры  $\Sigma_i$ ,  $\mathcal{S}_i = \{\varphi \in L(\Sigma_i) \mid \varphi \in \mathcal{T}\}$ ,  $i = 1, 2$ . В частности, такое разложение определяется по разложению сигнатуры однозначно.

**Замечание 5** Утверждение леммы 4 легко обобщается на случай произвольного числа сомножителей в разложении теории.

**Теорема 1** Для любой неприводимой теории  $\mathcal{T}$  сигнатуры  $\Sigma$  однозначно определяется дизъюнктивное разложение сигнатуры  $\Sigma = \Sigma_1 \amalg \dots \amalg \Sigma_n$ , которое с точностью до перестановки сомножителей отвечает любому разложению теории  $\mathcal{T} = \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$  на неразложимые компоненты разложимости  $\mathcal{S}_i$  – теории сигнатуры  $\Sigma_i$ ,  $i = 1, \dots, n$ . Зафиксируем некоторое такое разложение.

Для произвольного разложения теории на неразложимые компоненты  $\mathcal{T} = \mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_m$  имеем  $n = m$  и после подходящей перенумерации компонент каждая теория  $\mathcal{S}_i$  отличается от теории  $\mathcal{T}_i$  некоторыми предложениями чистого равенства из исходной теории  $\mathcal{T}$ :  $\langle \mathcal{S}_i, \mathcal{T}^\# \rangle = \langle \mathcal{T}_i, \mathcal{T}^\# \rangle$ .

Можно заметить, что перечисленные выше утверждения дают основу для доказательства этой теоремы. Здесь приведем лишь итоговые рассуждения.

Для произвольного  $i = 1, \dots, n$  обозначим через  $\mathcal{S}_i$  теорию сигнатуры  $\Sigma_i$ , порожденную системой аксиом  $S_i$ , образующей компоненту связности системы  $\Psi$ . Также обозначим через  $\mathcal{S}'_i = \langle \mathcal{S}_i, \mathcal{T}^\# \rangle$  – теорию  $\mathcal{S}_i$  вместе с предложениями чистого равенства теории  $\mathcal{T}$ . Имеем разложение теории на неразложимые компоненты  $\mathcal{T} = \mathcal{S}'_1 \otimes \dots \otimes \mathcal{S}'_n$ .

Рассмотрим произвольное разложение теории на неразложимые компоненты  $\mathcal{T} = \mathcal{T}_1 \otimes \dots \otimes \mathcal{T}_m$ . По лемме 3 сигнатура каждой компоненты  $\mathcal{T}_j$  получается объединением некоторых компонент  $\Sigma_{i_j}$ ,  $i_j = 1, \dots, n$  стандартного разложения общей сигнатуры.

Дополним каждую теорию формулами чистого равенства:  $\mathcal{T}'_i = \langle \mathcal{T}_i, \mathcal{T}^\# \rangle$ . Тогда также имеем  $\mathcal{T} = \mathcal{T}'_1 \otimes \dots \otimes \mathcal{T}'_m$ .

В силу леммы 4 и замечания 5, заключаем, что каждая теория  $\mathcal{T}'_j$  образована всеми формулами теории  $\mathcal{T}$ , которые включают символы из соответствующей ей сигнатуры  $\Delta_j$ . Однако по лемме 3  $\Delta_j$  получается объединением некоторых  $\Sigma_i$ ,  $i = 1, 2, \dots, n$ .



Для произвольного  $j$  рассмотрим произведение теорий  $\mathcal{S}'_i$  сигнатур  $\Sigma_i \subseteq \Delta_j$ . Оно содержится в теории  $\mathcal{T}$ , является непротиворечивым и определенным. Применяя замечание 5, получаем, что теория  $\mathcal{T}'_j$  совпадает с этим произведением. В силу неразложимости теории  $\mathcal{T}'_j$ , она совпадает с некоторой теорией  $\mathcal{S}'_i$ , использующей те же символы сигнатуры.

Этим построено отображение  $j \rightarrow i$ . Компоненты разложения  $\mathcal{T}'_j = \mathcal{S}'_i$  по замечанию 5 однозначно определяются используемыми символами сигнатуры – компонентами стандартного разложения  $\Sigma_i$ . Поэтому это отображение неоднозначно. Сюръективность отображения очевидна, так как, в силу разложимости, любой символ сигнатуры должен использоваться некоторой теорией  $\mathcal{T}'_j$ .

Тем самым теорема 1 доказана.

Из леммы 2 следует, что для определения компонент связности сигнатуры достаточно использовать любую систему  $(Ax)''$  аксиом теории, образованную неприводимыми и неразложимыми предложениями. Ее нетрудно получить исходя из неприводимых аксиом  $(Ax)'$  теории  $\mathcal{T}$ .

Действительно, согласно предложению 2 для теории  $\mathcal{T}$  существует система аксиом  $(Ax)'$ , состоящая из неприводимых предложений. Применим к этой системе следующие преобразования.

Каждую разложимую аксиому  $\varphi \in (Ax)'$  заменим на ее фрагменты  $\theta, \psi$ , каждый из которых выводится из  $(Ax)'$  и для которых  $\theta, \psi \vdash \varphi$ . Неразложимые аксиомы не изменяем. Обозначим полученную после преобразования систему аксиом как  $T(Ax)'$ .

По построению система аксиом  $T(Ax)'$  также определяет теорию  $\mathcal{T}$ . Будем повторять это преобразование и построим последовательность систем аксиом  $T(Ax)', T^2(Ax)', \dots$

Каждая формула содержит лишь конечное число сигнатурных символов, поэтому каждая аксиома  $\varphi$  из  $(Ax)'$  может быть последовательно разложена только конечное число раз (оговоримся, что рассматриваемые сигнатуры – конечные). Таким образом, начиная с некоторого  $n$  в системе аксиом  $T^m(Ax)'$ ,  $m \geq n$ , будут только неразложимые фрагменты аксиомы  $\varphi$ .

Обозначим через  $(Ax)''$  такие предложения  $\varphi$ , каждое из которых встречается во всех членах построенной последовательности кроме, может быть, конечного их числа. Это неразложимые фрагменты аксиом из  $(Ax)'$ . По замечанию 4 система  $(Ax)''$  состоит из неразложимых и неприводимых предложений теории  $\mathcal{T}$  и представляет систему аксиом теории  $\mathcal{T}$ .

С учетом леммы 2 система неприводимых и неразложимых аксиом  $(Ax)''$  теории  $\mathcal{T}$  определяет разложение сигнатуры, а с учетом леммы 3 и самой теории. Вместе с заключительным утверждением леммы 3 это приводит нас к критерию разложимости элементарной теории  $\mathcal{T}$ .

**Критерий разложимости** Теория  $\mathcal{T}$  сигнатуры  $\Sigma$  является разложимой

тогда и только тогда, когда соответствующая ей система неприводимых и неразложимых аксиом  $(Ax)''$  индуцирует более, чем одну компоненту связности на сигнатуре  $\Sigma$ .

**В третьей главе** рассматриваются практически вопросы разложимости.

Рассмотрен случай конечно аксиоматизируемых теорий. изучены стандартные преобразования, такие как сведение к предикатной сигнатуре и скулемизация. Установлено, что первое преобразование сохраняет свойство разложимости. Для второго преобразования показано, что, хотя оно не сохраняет разложимость, все равно свойство разложимости исходной теории может быть распознано после его применения. Оба преобразования определяются формулами исчисления предикатов и приводят к более простому виду предложений теории.

В случае конечно аксиоматизируемой теории  $\mathcal{T}$  проблема разложимости сводится к задаче минимизации системы аксиом.

Каждой формуле  $\psi$  сигнатуры  $\Sigma$  можно сопоставить множество символов  $\text{supp}(\psi)$ , использованных в записи этой формулы. Обозначим их количество через  $np(\psi) = \#(\text{supp}(\psi))$ .

Определим *вес* формулы  $\psi$  как натуральное число

$$w(\psi) = 3^{np(\psi)} = 3^{\#(\text{supp}(\psi))}.$$

**Замечание 6** Основание степени 3 выбрано так, чтобы для любых натуральных чисел  $n, m$  выполнялось неравенство  $3^{n+m} > 3^n + 3^m$  (в частности, чтобы суммарный вес фрагментов разложения формулы был всегда меньше, чем вес самой формулы).

**Определение 8** Весом системы аксиом  $\Psi$  теории  $\mathcal{T}$  назовем сумму весов составляющих ее аксиом:

$$w(\Psi) = \sum_{\psi \in \Psi} w(\psi).$$

*Минимальной системой аксиом теории  $\mathcal{T}$  будем называть такую систему аксиом  $\Psi$ , вес которой минимален.*

Поскольку веса образуют множество натуральных чисел, всегда можно выбрать систему аксиом  $\Psi$  минимального веса (из известного множества систем аксиом). Далее будем предполагать, что система аксиом  $\Psi$  имеет минимальный вес. Возможность использования такой системы для решения проблемы разложимости объясняет следующее утверждение.

**Предложение 4** Пусть  $\Psi$  – есть минимальная система аксиом теории  $\mathcal{T}$ . Тогда она образована такими предложениями теории  $\mathcal{T}$ , которые являются одновременно неразложимыми и неприводимыми.

В решении проблемы разложимости сигнатуру  $\Sigma$  можно предполагать предикатной, включающей только символы предикатов. Для этого в работе используется стандартный способ представления функций  $x = f(x_1, \dots, x_n)$  их графиками  $F(x_1, \dots, x_n, x)(= 0)$  и показывается, что свойство разложимости при этом сохраняется.

Более точно, рассматривается преобразование (обозначено в работе как  $'$ ) теории, соответствующее этому представлению, и для него доказывается следующее предложение.

**Предложение 5** *Рассмотрим теорию  $\mathcal{T}$  сигнатуры  $\Sigma$  и преобразование  $'$  формул сигнатуры  $\Sigma$  в формулы предикатной сигнатуры  $\Sigma'$ , описанное выше.*

*Тогда  $\mathcal{T}$  разложима тогда и только тогда, когда разложима производная теория  $\mathcal{T}'$ .*

Другим естественным преобразованием, которое приводит к более простому виду формул, является скулемизация. Для него замечается следующее.

**Замечание 7** *Если теория  $\mathcal{T}$  конечно аксиоматизируема, то она определяется одной формулой  $\psi$  сигнатуры  $\Sigma$ . Тогда её скулемовское обогащение в расширенной сигнатуре  $\Sigma^*$  определяется универсальной формулой  $\psi^*$  и скулемовскими предложениями  $\mathcal{S}$ . Во всяком случае,  $\psi^* \vdash \mathcal{T}$ .*

*В общем случае теория определяется некоторой системой аксиом  $\Psi$ . Её скулемовское обогащение  $\mathcal{T}^*$  определяется соответствующим набором универсальных формул  $\Psi^*$  сигнатуры  $\Sigma^*$  и теорией скулемовского обогащения  $\mathcal{S}$ . При этом  $\Psi^* \vdash \mathcal{T}$ .*

Конечно, добавление к теории формул новой сигнатуры  $\Sigma^*$  нарушает разложимость – появляются неразложимые формулы  $\varphi^*$  с символами из обеих сигнатур (исходной и  $\Delta = \{F_\psi^n \mid \psi = (\exists x)\varphi(x, x_1, \dots, x_n), \varphi\text{- формула сигнатуры } \Sigma\}$ ). Поэтому требуется расширить определение разложимой теории. Используем предикатность сигнатуры  $\Sigma$ , считая, что функциональные символы появляются только при скулемизации.

**Определение 9** *Рассмотрим теорию  $\mathcal{T}$  предикатной сигнатуры  $\Sigma$ . Теория  $\mathcal{T}$  называется **относительно разложимой** (разложимой относительно скулемовского обогащения), если можно подобрать дизъюнктное разложение обогащённой сигнатуры  $\Sigma^* = \Sigma_1^* \amalg \Sigma_2^*$ ,  $\Sigma \cap \Sigma_1^* \neq \emptyset \neq \Sigma \cap \Sigma_2^*$  и такие множества формул  $S'_1, S'_2$  соответствующих сигнатур  $\Sigma_1^*, \Sigma_2^*$ , которые содержатся в скулемовском обогащении теории  $\mathcal{T}$ :  $S'_1, S'_2 \subseteq \mathcal{T}^*$  и порождают расширение теории  $\mathcal{T}$ :  $S'_1, S'_2 \vdash \mathcal{T}$ . Обозначение:  $\mathcal{T} = \underset{\mathcal{S}}{S'_1} \otimes S'_2$ .*

Поясним данное определение. Наша цель - определить разложимость исходной теории  $\mathcal{T}$  после скулемовского обогащения. Так как  $\mathcal{T} \subseteq \mathcal{T}^*$ , мы требуем существования в  $\mathcal{T}^*$  двух множеств формул  $S'_1, S'_2$  дизъюнктных сигнатур, из которых следует сама теория  $\mathcal{T}$  (другими словами, она содержится в дедуктивном замыкании этих множеств формул).

Заметим также, что определением предусматриваются только *нетривиальные разложения* теории, приводящие к действительному разложению

исходной сигнатуры  $\Sigma : \Sigma_1 = \Sigma \cap \Sigma_1^* \neq \emptyset \neq \Sigma \cap \Sigma_2^* = \Sigma_2$ . В частности, этим исключается тот случай, когда одна из компонент  $\Sigma_i$  может совпадать с  $\Delta$ . То есть, случай, когда, может быть, одно из множеств  $S'_i$  содержит формулы только с символами из  $\Delta$ , а другое – из  $\Sigma$ .

Отметим, что если исходная теория  $\mathcal{T}$  не содержит  $\exists$ -формул, т.е.  $\Delta = \emptyset$  и  $\mathcal{T} = \mathcal{T}^*$ , мы приходим к изначальному определению разложимой теории, данному в главе 2 диссертации.

Заключительным утверждением главы 3 является предложение, показывающее возможность использования скулемизации при определении свойства разложимости.

**Предложение 6** *Разложимость теории  $\mathcal{T}$  равносильна её относительной разложимости.*

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ

1. Исследовано применение декларативных описаний в логике первого порядка для задач поиска информации и интеграции источников данных. Разработана и реализована программно логическая теория в области биоинформатики для анализа экспериментальных данных по экспрессии генов и фенотипическим аномалиям модельного организма *Arabidopsis thaliana* (L.).
2. В связи с использованием связей терминов в декларативных описаниях исследована возможность декомпозиции (разложения) формальных описаний в логике первого порядка на компоненты с непересекающимися алфавитами. Сформулирован критерий разложимости.
3. Исследован вопрос синтаксической однозначности при описании знаний как теорий в логике первого порядка. Доказана однозначность разложения теории на неразложимые компоненты с точностью до формул чистого равенства. Установлено приведение теории к такому виду, по которому однозначно определяется ее разложение.

## ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

- [1] **Романюгов Д.** Semantic Web basics in logical consideration. // Lecture Notes in Informatics: Proc. / Informatik-2006, Dresden, 2006 - Bonn, 2006. - Volume 2 - P.337-344.
- [2] **Пономарев Д.К.** Проблема разложимости при формальном описании знаний. – Новосибирск, 2006 – 21 стр. – (Препр. / СО РАН. Ин-т систем информатики; N135).

- [3] **Ponomaryov D.** On decomposability of elementary theories // Algebra and Model Theory – Novosibirsk State Technical University, 2005 – N 5 – P. 162-169.
- [4] **Ponomaryov D., Omelianchuk N, Kolchanov N., Mjolsness E., Meyerowitz E.** Semantically rich ontology of anatomical structure and development for Arabidopsis thaliana (L.). // Proc. Bioinformatics of Genome Regulation and Structure (BGRS'2006), Novosibirsk – 2006 – P. 227-230.
- [5] **Mironova V.V., Poplavsky A.S., Ponomaryov D.K., Omelianchuk N.A.** Ontology of Arabidopsis Genenet Supplementary Database(AGNS): Cross references to TAIR ontology. // Proc. Bioinformatics of Genome Regulation and Structure (BGRS'2006), Novosibirsk – 2006 – P. 209-212.
- [6] **Ponomaryov D., Omelianchuk N, Kolchanov N., Mjolsness E., Meyerowitz E.** A program method for inferring relationships between phenotypic abnormalities of Arabidopsis. // Proc. Bioinformatics of Genome Regulation and Structure (BGRS'2006), Novosibirsk – 2006 – P. 231-234.
- [7] **Пonomaryov D.** Lattice semantics for incremental data extraction from declarative knowledge bases. – Новосибирск, 2006 – 13 стр. – (Препр. / СО РАН. Ин-т систем информатики; N134).
- [8] **Пономарев Д.К.** Задача разложимости элементарных теорий и проблема минимизации из аксиом. // Тез. конференции-конкурса "Технологии Microsoft в информатике и программировании", Новосибирск, 22-24 февраля, 2006. – С. 213-215.
- [9] **Пономарев Д.К.** Применение языков описания онтологий для построения Web-ориентированных информационных систем. // Вестник Новосибирского Гос. Ун-та / Информационные технологии – Новосибирск, 2004 – Т.1, Вып.2 – С. 5-20.
- [10] **Пономарев Д.К.** Применение Web-стандартов описания метаданных для представления предметных областей. // Тез. XLI Международной Научной Студенческой Конференции (МНСК), Новосибирск, 14-17 апреля 2003. – С. 39-40.

Пономарев Д.К.

СИНТАКСИЧЕСКАЯ ОДНОЗНАЧНОСТЬ ПРИ ПРЕДСТАВЛЕНИИ  
ЗНАНИЙ В ЛОГИКЕ ПЕРВОГО ПОРЯДКА

Автореферат

---

Подписано в печать  
Формат бумаги 60x90 1/16

Объем 1,1 уч.-изд.л.  
Тираж 100 экз.

---

Отпечатано на ризографе "AL Group"  
630090, г. Новосибирск, пр. акад. Лаврентьева, 6