

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Факультет математики и информатики
Кафедра информационных и управляющих систем
Направление подготовки 09.03.02 – Информационные системы и технологии
Направленность (профиль) образовательной программы: Безопасность информационных систем

ДОПУСТИТЬ К ЗАЩИТЕ
Зав. кафедрой

_____ А.В. Бушманов
« _____ » _____ 201_ г.

БАКАЛАВРСКАЯ РАБОТА

на тему: Разработка подсистемы синтаксического анализа текстовой информации

Исполнитель

студент группы 355об

(подпись, дата)

Ю.Д. Шестаков

Руководитель

доцент, канд. техн. наук

(подпись, дата)

А.В. Бушманов

Консультант

по безопасности и экологичности

доцент, канд. техн. наук

(подпись, дата)

А.Б. Булгаков

Нормоконтроль

инженер кафедры

(подпись, дата)

В.В. Романико

Благовещенск 2017

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Факультет математики и информатики
Кафедра информационных и управляющих систем

УТВЕРЖДАЮ
Зав.кафедрой
_____ А.В.Бушманов
« _____ » _____ 2017 г.

З А Д А Н И Е

К бакалаврской работе студента Шестакова Юрия Дмитриевича.

1. Тема бакалаврской работы: Разработка подсистемы синтаксического анализа текстовой информации

(утверждено приказом от 25.04.2017 № 929-уч)

2. Срок сдачи студентом законченной работы 21.06.2017 г.

3. Исходные данные к бакалаврской работе: отчет по преддипломной практике.

4. Содержание бакалаврской работы: анализ предметной области, проектирование подсистемы, разработка программного обеспечения, исследование вопросов информационной безопасности, рассмотрение аспектов безопасности жизнедеятельности.

5. Перечень материалов приложения: организационная линейная структура, диаграммы IDEF0, изображения экранных форм, схемы базы данных.

6. Консультант по бакалаврской работе консультант по безопасности и экологичности доцент, канд. техн. наук Булгаков А.Б.

7. Дата выдачи задания 09.05.2017 г.

Руководитель бакалаврской работы Бушманов Александр Вениаминович, доцент, канд. техн. наук.

Задание принял к исполнению (дата): _____ Ю.Д. Шестаков

РЕФЕРАТ

Бакалаврская работа содержит 67 с., 26 рисунков, 20 таблиц, 6 приложений, 20 источников.

ПРОЕКТИРОВАНИЕ, РАЗРАБОТКА, БАЗА ДАННЫХ, ПРИЛОЖЕНИЕ, ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ, ПОДСИСТЕМА, АНАЛИЗ, ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА, ЭКСПЕРТНАЯ СИСТЕМА, ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ, БЕЗОПАСНОСТЬ ЖИЗНЕДЕЯТЕЛЬНОСТИ

Для данной бакалаврской работы объектом исследования была выбрана предметная область экспертных систем, в частности – автоматизированная обработка естественного языка, использующаяся для извлечения знаний из текстовых источников и последующей их записи в базы знаний.

Целью работы является разработка компонента экспертных систем, основанного на автоматической начальной обработке естественного языка – подсистемы синтаксического анализа текстовой информации.

Работа выполнялась последовательно в соответствии со следующими этапами: анализ предметной области, проектирование подсистемы, разработка программного обеспечения, рассмотрение вопросов информационной безопасности, а также исследование требований безопасности жизнедеятельности.

Разработанная подсистема позволит автоматизировать процесс построения начальной системы связей – синтаксических бинарных деревьев, на основе имеющейся в ней базы знаний в области синтаксиса русского языка. Данная подсистема является компонентом экспертных систем, обеспечивающим автоматизированную начальную обработку естественного языка.

					<i>ВКР.135194.09.03.02.ПЗ</i>			
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подпись</i>	<i>Дата</i>				
<i>Разраб.</i>		<i>Шестаков Ю.Д.</i>			Разработка подсистемы синтаксического анализа текстовой информации	<i>Лит.</i>	<i>Лист</i>	<i>Листов</i>
<i>Пров.</i>		<i>Бушманов А.В.</i>				<i>У</i>	<i>З</i>	<i>77</i>
<i>Консульт.</i>		<i>Булгаков А.Б.</i>				АМГУ кафедра ИУС		
<i>Н. контр.</i>		<i>Романико В.В.</i>						
<i>Зав.каф.</i>		<i>Бушманов А.В.</i>						

СОДЕРЖАНИЕ

Введение	7
1 Анализ предметной области	9
1.1 Интеллектуальные системы	9
1.1.1 Экспертные системы	9
1.1.2 Анализ текстов на естественном языке	12
1.2 Постановка задачи	13
2 Проектирование подсистемы	15
2.1 Анализ используемых программно-технических средств	15
2.2 Обоснование выбора средств разработки	15
2.3 Характеристика функциональных подсистем	16
2.3.1 Характеристика подсистемы регистрации, идентификации, аутентификации и авторизации	16
2.3.2 Характеристика подсистемы обмена данными БД с синтаксическим анализатором	18
2.3.3 Интерактивный пользовательский интерфейс	18
2.4 Характеристика обеспечивающих подсистем	19
2.5 Проектирование БД	19
2.5.1 Инфологическое проектирование	19
2.5.2 Логическое проектирование	27
2.5.3 Физическое проектирование	32
2.6 Проектирование синтаксического анализатора	35
2.7 Проектирование клиентского приложения	35
3 Разработка программного обеспечения	37
3.1 Разработка морфологической и синтаксической БД	37
3.2 Разработка синтаксического анализатора	42
3.3 Разработка подсистемы обмена данными БД с синтаксическим анализатором	46

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		4

3.4 Руководство пользователя	47
4 Описание сервисов обеспечения ИБ	51
4.1 Угрозы информационной безопасности	51
4.2 Ролевая модель разграничения доступа в системе 1С:Предприятие 8	54
5 Безопасность и экологичность	58
5.1 Требования безопасности, предъявляемые к работе с ЭВМ	58
5.1.1 Эргономические требования	58
5.1.2 Требования электробезопасности	60
5.2 Обеспечение экологичности	61
5.3 Требования к обеспечению пожарной безопасности при работе с ЭВМ	62
Заключение	64
Библиографический список	65
Приложение А Технологическая сеть проектирования	68
Приложение Б Логическая модель БД	70
Приложение В Физическая модель БД	71
Приложение Г Ролевое разграничение доступа	72
Приложение Д Представление синтаксического дерева в БД	73
Приложение Е Диаграмма процессов в подсистеме синтаксического анализа текстовой информации	77

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

ВКР – выпускная квалификационная работа;

НСД – несанкционированный доступ;

ОС – операционная система;

ПК – персональный компьютер;

ПО – программное обеспечение;

РФ – Российская Федерация;

ЧС – чрезвычайная ситуация;

ЭС – экспертная система;

ЕЯ – естественный язык;

БД – база данных;

БЗ – база знаний;

ИС – информационная система;

ИБ – информационная безопасность;

БЖД – безопасность жизнедеятельности;

СУБД – система управления базами данных.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		6

ВВЕДЕНИЕ

В настоящее время, в связи с быстрым развитием информационных технологий, растет объем получаемой человеком информации, на обработку которого уже не хватает ни времени, ни интеллектуальных ресурсов. Подавляющее большинство баз данных позволяют облегчить процесс запоминания и структурирования информации, однако обрабатывать ее все же приходится человеку.

Особо остро данный вопрос стоит перед проблематикой процессов пополнения баз знаний экспертных систем. Автоматизация этих процессов позволила бы сократить как финансовые расходы на специалистов в этой области, так и временные затраты.

Известно, что большая часть информации представлена на естественном языке, и поэтому одним из решений для автоматизации процесса пополнения баз знаний экспертных систем может являться автоматическая обработка слов и предложений естественного языка, начиная от морфологического, и заканчивая семантическим анализами.

В соответствии с вышесказанным, объектом исследования и разработки в данной работе был выбран механизм автоматической обработки информации на естественном языке и занесения его результатов в базу знаний экспертных систем, в частности начальный этап – синтаксический анализатор.

Перед процессом разработки подсистемы синтаксического анализа текстовой информации были поставлены следующие задачи:

- проанализировать предметную область экспертных систем;
- описать проблематику и обосновать актуальность разрабатываемой подсистемы;
- выбрать и обосновать выбор программных средств проектирования и разработки;
- разработать проект БД;
- разработать проект функциональных компонентов подсистемы;

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		7

- создать спроектированную подсистему и описать этапы ее реализации;
- написать руководство для пользователей подсистемы;
- выявить основные угрозы подсистеме и описать пути их нейтрализации;
- рассмотреть основные аспекты БЖД пользователей подсистемы.

Итогом выполнения ВКР будет являться разработанная подсистема синтаксического анализа текстовой информации, выполняющая функции начального этапа обработки естественного языка, а так же имеющая простой механизм самообучения, на основе записи результатов анализа в синтаксическую БД подсистемы.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		8

1 АНАЛИЗ ДЕЯТЕЛЬНОСТИ ПРЕДПРИЯТИЯ

Тема данной ВКР относится к области экспертных систем. Для полного понимания темы и целей разработки, проведем анализ этой предметной области, берущей свое начало еще с середины прошлого века, однако не менее актуальной и по сей день.

1.1 Интеллектуальные системы

Первым человеком, который ввел определение термина интеллектуальная система, был британский математик Алан Тьюринг, опубликовав свою работу «Вычислительная машина и интеллект» в журнале «Mind» в 1950 году. В ней Тьюринг предложил называть интеллектуальным такое поведение программы, которое будет моделировать разумное поведение человека.

Интеллектуальные системы разрабатываются для решения задач, алгоритм решения которых заранее не известен. Следствием этого является важная особенность интеллектуальных систем – предполагается запрограммированное наличие выбора в условиях неопределенности, обычно опирающееся на имеющуюся базу знаний системы. Другими словами интеллектуальная система должна иметь свободу действий в принятии решений, подобно человеку.

Одним из наиболее распространенных и значительных достижений в области интеллектуальных систем стали экспертные системы, призванные решать не формализуемые и трудно формализуемые задачи.

1.1.1 Экспертные системы

Под экспертной системой понимается программа, использующая знания специалиста (эксперта) в некоторой конкретной узко специализируемой предметной области и способна принимать решения в пределах данной области, имитируя мышление и действия этого эксперта.

Важнейшей составляющей экспертных систем является наличие базы знаний, содержащей не данные, а правила анализа информации и вывода результата.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		9

Как правило, база знаний включает в себя факты (достоверные сведения о предметной области) и правила – набор инструкций, применяя которые к известным фактам можно получать новые факты.

Основными категориями решаемых экспертными системами задач являются: диагностика, управление (в том числе технологическими процессами), интерпретация, прогнозирование, проектирование, отладка и ремонт, планирование, наблюдение (мониторинг), обучение.

Рассмотрим различные способы классификации экспертных систем.

По назначению экспертные системы делятся на:

- ЭС общего назначения;
- проблемно-ориентированные для задач диагностики, проектирования и прогнозирования;
- предметно-ориентированные для специфических задач, например, контроля маршрутов авиарейсов.

В степени зависимости от внешней среды выделяют:

- статистические экспертные системы, не зависящие от внешней среды;
- динамические, учитывающие динамику внешней среды и предназначенные для решения задач в реальном времени. Время реакции в таких системах может задаваться в миллисекундах, и поэтому для их реализации предпочтительно использовать такие языки, как, например, C++.

По типу использования различают:

- изолированные ЭС;
- ЭС на входе или выходе других систем;
- гибридные ЭС, т.е. интегрированные с базами данных и другими программами продуктами.

По сложности решаемых задач различают:

- простые ЭС – до 1000 простых правил;
- средние ЭС – от 1000 до 10000 структурированных правил;
- сложные ЭС – более 10000 структурированных правил.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		10

При разработке экспертных систем возникает острая проблема: извлечение знаний эксперта. Эту функцию выполняет определенный специалист – инженер знаний. Он должен:

- во-первых, иметь, так называемую, харизму, т.к. ему необходимо в разговоре с экспертом убедить его поделиться своими знаниями и опытом;
- во-вторых, быть достаточно подкованным, для понимания и выявления знаний в процессе общения;
- в-третьих, иметь навыки пользования интерфейсом ЭС, для занесения знаний в базу знаний.

Частое нежелание человека-эксперта передавать свой опыт вызвано тем, что для него его знания являются наиболее ценным ресурсом, благодаря которому эксперт обеспечивает себя достойным заработком. Тем более программа, оперирующая знаниями даже одного эксперта, способна заменить сразу несколько экспертов в этой предметной области. Данный аспект и делает особо актуальной вышеупомянутую проблему извлечения знаний.

Альтернативой может являться получение инженером экспертных знаний из литературных источников. Однако это требует большого времени на вникание инженера знаний в суть предметной области, особенно если ранее он был с ней слабо знаком. Слабая окупаемость интеллектуальной работы человека по извлечению знаний в сравнении с временными затратами и послужила толчком к обращению разработчиков ЭС к отдельной области интеллектуальных систем в общем, и экспертных систем в частности – автоматизированной обработки и анализа естественного языка, или, другими словами, естественно-языковых систем.

В качестве примера подобной реализованной интеллектуальной системы можно привести разработанный компанией IBM суперкомпьютер Watson, оснащенный вопросно-ответной системой, отвечающей на вопросы, представленные на естественном языке.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		11

Проектируемая подсистема является компонентом отдельно разрабатываемой экспертной системы направленной на решение вопросов в различных практических областях.

1.1.2 Анализ текстов на естественном языке

Компьютерная лингвистика является одной из популярных тем исследований искусственного интеллекта, начиная с 50-х годов прошлого века. Проблемы понимания естественного языка возникают в такой области, как выявление сути или контекста разговора, целей разговора, учет недосказанности и т.д. Для ее решения требуется наличие знаний о предметной области, в которой происходит разговор и определенный опыт общения на данном языке и с данным человеком.

Рассмотрим основные понятия естественного языка. Слово – одна из единиц языка, служащая для именованя предметов, лиц, процессов, свойств и т.д. Оно имеет различные лингвистические характеристики, такие как морфологические признаки, отдельные части образующие состав слова (приставка, корень, суффикс, окончание), наличие синонимов, антонимов и т.д. Все это в совокупности определяет правила написания и, например, склонения данного слова, а так же, что самое главное, правила его связи с другими словами, что в совокупности образует словосочетания и предложения. Словосочетание - простейшая единица речи, которая образуется на основе подчинительной связи (согласования, управления, примыкания) двух и более слов. Словосочетания составляют главные и второстепенные части предложения и уже являются единицами, несущими информацию и определенный смысл. Предложение – любое высказывание, являющееся сообщением о чем-либо. Можно сказать, что предложение - это единица естественного языка, описывающая какое-либо событие, явление или поведение объекта.

Для извлечения знаний из текста на естественном языке, необходимо, в первую очередь, эти знания выявить, правильно их интерпретировать и в конце преобразовать, т.е. привести в соответствие языку представления знаний, используемому данной экспертной системой. Одним из лучших решений дан-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		12

ной задачи является последовательный анализ естественно-языкового текста от выявления морфологических характеристик до поиска синтаксических конструкций и их семантического интерпретирования. Опираясь на все вышеизложенное, в качестве темы выпускной квалификационной работы были выбраны два первых подпункта, а именно: морфологический и синтаксический автоматизированный анализ, в виде реализации их в качестве полноценной системы синтаксического анализа текстовой информации на русском языке.

1.2 Постановка задачи

На рисунке 1.1 приведена приблизительная схема реализации вышеописанного анализа ЕЯ текстов. Как видно из схемы, вначале производится морфологический анализ всех слов, содержащихся в тексте. Для этого необходимо разработать морфологическую базу данных, содержащую аналоги слов русского языка с описанием их морфологических характеристик.



Рисунок 1.1 – Схема последовательного анализа ЕЯ текстов

Изм.	Лист	№ докум.	Подп.	Дата

Входными данными являются любые слова, словосочетания и предложения на русском языке, вводимые пользователем вручную через соответствующий интерфейс, либо считываемые из текстовых файлов.

Неизвестные слова записываются в БД в качестве морфологических конструкций вручную. Этим занимается пользователь – эксперт по лингвистике, либо любой другой пользователь, опирающийся, при записи, на сторонний морфологический словарь. Для данной функции необходимо разработать соответствующий удобный пользовательский интерфейс.

Так же необходимо спроектировать и разработать синтаксическую базу данных, содержащую результаты синтаксического анализа, следующего сразу после морфологического. Таким образом будет реализован небольшой цикл самообучения подсистемы, т.к. при последующей ее встрече с таким же, ранее проанализированным, фрагментом не потребуется затрат ресурсов на его повторный анализ. Возможность ее администрирования так же должна быть доступна пользователю через интерфейс.

Для выполнения механизма синтаксического разбора текстов, отдельно должен быть разработан синтаксический анализатор с базой знаний, содержащей синтаксические правила русского языка, в дальнейшем активно работающий с обеими базами данных.

Выходными данными в результате обработки текста через анализатор должны являться древовидные структуры – бинарные синтаксические деревья, отображающие иерархию синтаксических связей между словами и фрагментами в тексте. Выходные данные анализатора, затем, должны записываться в синтаксическую БД.

Подсистему следует разрабатывать на платформе 1С:Предприятие, т.к. это требуется для последующего ее объединения с отдельно разрабатываемой экспертной системой на той же технологической платформе.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		14

2 ПРОЕКТИРОВАНИЕ ПОДСИСТЕМЫ

На данном этапе производится анализ и выбор программных средств, необходимых для реализации подсистемы, а так же описывается процесс проектирования конечного программного продукта.

В приложении А представлена технологическая сеть проектирования, описывающая основные этапы проектирования и разработки подсистемы.

2.1 Анализ используемых программно-технических средств

Аппаратной платформой для разрабатываемой ИС является персональный компьютер под управлением ОС Windows.

В качестве СУБД и инструментального программного средства для создания пользовательского интерфейса был сделан выбор в пользу технологической платформы 1С:Предприятие 8.3. Разработанное на ней ПО должно включать в себя следующие компоненты:

- СУБД для создания и администрирования баз данных;
- подсистема регистрации, идентификации, аутентификации и авторизации пользователя;
- подсистема преобразования естественно-языковых структур в базу фактов и обмена ими с синтаксическим анализатором;
- интерактивный пользовательский интерфейс.

В качестве инструментального программного средства для создания базы знаний и синтаксического анализатора была выбрана среда разработки Visual Prolog 5.2 и используемый в ней язык логического программирования Prolog.

2.2 Обоснование выбора средств разработки

Для описания правил синтаксического анализа слов русского языка лучше всего подходит язык, основанный на логике предикатов первого порядка. Одним из них является Prolog. Среда разработки Visual Prolog 5.2, хотя и является далеко не современной, однако, предоставляет весь необходимый набор инструментальных средств для проектирования и создания базы знаний пред-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		15

ставленной в виде синтаксических правил русского языка.

В качестве основных достоинств среды можно выделить:

- удобная и гибкая среда программирования с подсветкой синтаксиса;
- совместимость с большинством версий операционной системы Windows: 2000/XP/Vista/7/8/8.1/10;
- язык программирования Prolog основан на логике предикатов первого порядка и использует метод поиска с возвратом, что позволяет уделить наибольшее внимание не процедурным составляющим, а логическому описанию синтаксических правил русского языка.

Кроме того, данный язык, в сравнении с процедурными ЯП, имеет преимущества при осуществлении поиска в глубину в бинарных деревьях, которые являются входными и выходными данными для синтаксического анализатора.

Выбор, в качестве СУБД для разрабатываемой подсистемы, платформы 1С:Предприятие 8.3 был сделан по ряду причин:

- подсистема будет являться частью отдельно разрабатываемой экспертной системы на платформе 1С:Предприятие 8.3;
- простота создания и администрирования БД, а так же пользовательского интерфейса для работы с ними;
- наличие всех необходимых инструментов для создания форм представления пользовательского интерфейса;
- предварительный опыт работы с платформой 1С:Предприятие 8.3.

2.3 Характеристика функциональных подсистем

Приведем подробное описание задач, решаемых каждой из подсистем, для полного их понимания и выбора наиболее верного способа реализации данных подсистем.

2.3.1 Характеристика подсистемы регистрации, идентификации, аутентификации и авторизации

Данная подсистема, в первую очередь, необходима для обеспечения безопасности разрабатываемой подсистемы в целом, и содержащейся в ней ин-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		16

формации в частности. Она состоит из следующих четырех компонентов: регистрации, идентификации, аутентификации и авторизации. Необходимый функционал данных компонентов уже реализован в платформе 1С:Предприятие 8.3.

Идентификация – это процедура распознавания пользователя по его идентификатору, присвоенному данному пользователю ранее и занесенному в базу данных в момент его регистрации в качестве легального пользователя системы.

Аутентификация – процедура проверки подлинности входящего в систему объекта (пользователя, процесса или устройства), предъявившего свой идентификатор.

Авторизация – процедура предоставления пользователю (процессу или устройству) определенных прав доступа к ресурсам системы после успешного прохождения им процедуры аутентификации, иными словами, авторизация устанавливает сферу действия пользователя и доступные ему ресурсы.

Компонент регистрации пользователя реализуется в виде занесения в базу его индивидуальной пары логин-пароль, а так же предоставления ему роли, либо ролей, и производится администратором системы средствами конфигулятора платформы.

Компонент идентификации производит ее путем ввода или выбора из списка пользователем своего логина, а так же ввода пароля.

Компонент аутентификации удостоверяет подлинность пользователя, пытающегося войти в систему. Она производится по введенной им паре логин-пароль и осуществляется средствами платформы.

Компонент авторизации определяет заданную пользователю роль в системе и предоставляет ему доступ к соответствующим функциям.

Необходимо настроить список ролей для распределения прав на выполнение определенных действий пользователей в системе, например, ограничить, либо запретить доступ определенной группе пользователей к чтению и записи в базу данных, во избежание искажения или потери информации. Создаваемые группы пользователей могут одновременно иметь несколько ролей.

2.3.2 Характеристика подсистемы обмена данными БД с синтаксическим

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		17

анализатором

Эта подсистема должна быть своеобразным посредником между базой данных и синтаксическим анализатором.

Язык Prolog оперирует не данными, в традиционном их понимании, а фактами и правилами, которые являются отношением между объектами (в виде данных или структур данных) и заключением, для которого известно, что оно истинно, если одно из нескольких найденных заключений или фактов являются истинными, соответственно. По этой причине входные данные при записи в базу данных, необходимо программным образом преобразовывать в факты, соответствуя синтаксису их представления на языке Prolog.

В основные функции подсистемы входят:

- вышеупомянутое преобразование данных при записи в БД;
- выявление соответствия ЕЯ конструкций из анализируемого текста содержащимся в БД данным и составление списка фактов с последующей передачей его в анализатор;
- прием выходных данных из анализатора и запись их в БД.

Все вышеупомянутые функции должны частично реализовываться на языке запросов к БД.

2.3.3 Интерактивный пользовательский интерфейс

Пользовательский интерфейс должен обеспечивать все возможности интерактивного взаимодействия пользователя с подсистемой. В его задачи входят:

- предоставление форм для интерактивного просмотра, записи, добавления, изменения и удаления данных из БД;
- предоставление форм для ввода пользователем текста на русском языке или выбора текстового файла для анализа, а так же вывода результатов анализа в наглядном для пользователя виде.

Разработку интерфейса предполагается производить по окончании создания баз данных и синтаксического анализатора, по причине заранее неизвестной структуры выходных данных.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		18

2.4 Характеристика обеспечивающих подсистем

Рассмотрим некоторые обеспечивающие подсистемы, имеющие наибольшую важность для разработки.

Основной обеспечивающей подсистемой является техническое обеспечение. В качестве аппаратного средства используется пользовательский ПК, либо, что наиболее предпочтительно, ПК-сервер с 64 Гб оперативной памяти.

Так же, не менее важной обеспечивающей подсистемой для разрабатываемой ИС является информационное обеспечение. Все данные должны быть представлены в соответствующем принципам и методам создания баз данных виде и формате. Преобразование данных должно осуществляться без потерь. Для всех форм необходимо написать краткую справочную документацию.

2.5 Проектирование БД

Процесс проектирования базы данных состоит из трех последовательных этапов: инфологического, логического и физического проектирования. Результатом должна стать наглядная схема БД для последующего ее переноса в СУБД.

2.5.1 Инфологическое проектирование

В первую очередь, выделим сущности для будущей БД.

Источником данных является множество русского языка, а так же морфологический словарь. Следовательно, опираясь на тему ВКР, из исследуемой области можно выявить следующие сущности:

- морфологическая конструкция;
- синтаксическая конструкция;
- часть речи;
- морфема;
- составная часть слова;
- ветви синтаксической конструкции;
- слово в синтаксической конструкции;
- связанное слово;

– тип связи.

Подробное описание данных сущностей представлено в таблице 1. Значения в колонке количество экземпляров могут быть пустыми, вследствие невозможности определить его заранее, т.к. данные таблицы будут заполняться в процессе эксплуатации подсистемы.

Таблица 1 – Определение и описание сущностей

Сущность	Описание	Количество экземпляров
Морфологическая конструкция	содержит морфологические конструкции русского языка	> 4 млн.
Синтаксическая конструкция	содержит синтаксические конструкции русского языка	
Часть речи	содержит ограниченный список наименований частей речи	12
Морфема	содержит список приставок, корней, суффиксов и окончаний, входящих в состав некоторых слов	
Составная часть слова	содержит список морфем входящих в состав определенного слова	
Ветви синтаксической конструкции	содержит список синтаксических конструкций, входящих в состав определенного экземпляра из этих конструкций	
Слово в синтаксической конструкции	содержит список морфологических конструкций, входящих в состав определенного экземпляра синтаксических	
Связанное слово	содержит список синонимов, омонимов, антонимов и паронимов определенного слова	
Тип связи	содержит ограниченный список лингвистических связей: синоним, омоним, антоним, пароним	4

Теперь, когда определены все сущности, приведем подробное описание их атрибутов в таблицах 2-10.

Некоторые сущности содержат мигрирующие атрибуты, которые так же могут мигрировать из сущности, являющейся родителем для данных сущностей. Это объясняется тем фактом, что, например, синтаксическая конструкция может включать в себя синтаксические конструкции уровнем ниже.

Таблица 2 – Атрибуты сущности Морфологическая конструкция

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код слова</u>	уникальный идентификатор морфологической конструкции	>1	Число	83
Слово	слово на естественном языке	-	Строка	лодка
Свойства	конструкция, содержащая список морфологических свойств слова, в синтаксисе языка Prolog	-	Строка	см. пункт 3.2
Код части речи	уникальный идентификатор части речи, относящий данное слово к определенной части речи	>1	Число	2

Таблица 3 – Атрибуты сущности Синтаксическая конструкция

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код конструкции</u>	уникальный идентификатор морфологической конструкции	>1	Число	906
Синтаксическая конструкция	синтаксическая конструкция, в синтаксисе языка Prolog	-	Строка	см. пункт 3.2

Таблица 4 – Атрибуты сущности Часть речи

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код части речи</u>	уникальный идентификатор части речи	>1	Число	10
Наименование	наименование части речи	-	Строка	глагол

Таблица 5 – Атрибуты сущности Морфема

Атрибут	Описание	Диапазон значений	Тип значения	Пример
1	2	3	4	5
<u>Код морфемы</u>	уникальный идентификатор морфемы	>1	Число	33

Продолжение таблицы 5

1	2	3	4	5
Наименование	наименование морфемы	-	Строка	пре
Тип морфемы	наименование типа морфемы	-	Строка	приставка

Таблица 6 – Атрибуты сущности Составная часть слова

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код слова</u>	уникальный идентификатор слова	>1	Число	3391
Наименование	наименование морфемы	-	Строка	ова
Тип морфемы	имя типа морфемы	-	Строка	суффикс
Код морфемы	уникальный идентификатор морфемы	>1	Число	73

Таблица 7 – Атрибуты сущности Ветви синтаксической конструкции

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код конструкции</u>	уникальный идентификатор конструкции верхнего подуровня	>1	Число	384
Синтаксическая конструкция левая	наименование левой конструкции в бинарном дереве	-	Строка	см. пункт 3.2
Код конструкции 1	уникальный идентификатор левой конструкции	>1	Число	732
Синтаксическая конструкция правая	наименование правой конструкции в бинарном дереве	-	Строка	см. пункт 3.2
Код конструкции 2	уникальный идентификатор правой конструкции	>1	Число	9813

Таблица 8 – Атрибуты сущности Связанное слово

Атрибут	Описание	Диапазон значений	Тип значения	Пример
1	2	3	4	5
<u>Код слова</u>	уникальный идентификатор слова	>1	Число	783943

Продолжение таблицы 8

1	2	3	4	5
Слово	наименование слова	-	Строка	Холодно
Код слова 1	уникальный идентификатор связанного слова	>1	Число	506204
Связанное слово	имя связанного слова	-	Строка	Жарко
Код типа связи	уникальный идентификатор типа связи	>1	Число	3

Таблица 9 – Атрибуты сущности Тип связи

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код типа связи</u>	уникальный идентификатор типа связи	>1	Число	1
Наименование	наименование типа связи	-	Строка	синоним

Таблица 10 – Атрибуты сущности Слова в синтаксической конструкции

Атрибут	Описание	Диапазон значений	Тип значения	Пример
<u>Код конструкции</u>	уникальный идентификатор конструкции	>1	Число	384
Слово левое	наименование левого слова в дереве конструкции	-	Строка	см. пункт 3.2
Код слова 1	уникальный идентификатор левого слова	>1	Число	39621
Слово правое	наименование правого слова в дереве конструкции	-	Строка	см. пункт 3.2
Код слова 2	уникальный идентификатор правого слова	>1	Число	7382

В таблице 11 приведено описание связей между сущностями. Как видно из вышеописанного, в БД были включены дополнительные таблицы, позволяющие избежать возникновения связей многие-ко-многим, посредством связывания записей (ключей) определенной сущности с одной или несколькими записями другой сущности. Так же были введены две отдельные таблицы, свя-

занные с одной и той же синтаксической конструкцией. Первая содержит только синтаксические, вторая только морфологические конструкции (слова).

Таблица 11 – Описание связей между сущностями

Название первой сущности, участвующей в связи	Название второй сущности, участвующей в связи	Название связи	Тип связи	Обоснование выбора типа связи
1	2	3	4	5
Морфологическая конструкция	Часть речи	имеет тип	один-ко-многим	Одной записи сущности Морфологическая конструкция соответствует одна запись сущности Часть речи, каждой записи сущности Часть речи соответствует множество записей сущности Морфологическая конструкция. Морфологическая конструкция соответствует только одной части речи, в то время как каждая Часть речи может определять несколько морфологических конструкций.
Ветви синтаксической конструкции	Синтаксическая конструкция	содержит	один-ко-многим	Одной записи сущности Ветви синтаксической конструкции соответствуют множество записей сущности Синтаксическая конструкция. Ветви синтаксической конструкции могут содержать только одну левую и одну правую синтаксическую конструкцию, в то время как каждая Синтаксическая конструкция может быть ветвью нескольких таких ветвей верхнего подуровня.

1	2	3	4	5
Слова в синтаксической конструкции	Морфологическая конструкция	содержит	один-ко-многим	Одной записи сущности Слова в синтаксической конструкции соответствуют множество записей сущности Морфологическая конструкция, каждой записи сущности Морфологическая конструкция соответствует множество записей сущности Слова в синтаксической конструкции. Слова в синтаксической конструкции могут являться только одной левой и одной правой морфологической конструкцией, в то время как каждая Морфологическая конструкция может быть ветвью нескольких синтаксических конструкций.
Составная часть слова	Морфема	содержит	один-ко-многим	Одной записи сущности Составная часть слова соответствуют множество записей сущности Морфема, каждой записи сущности Морфема соответствует множество записей сущности Составная часть слова. Составная часть слова может включать в себя только одну морфему, в то время как каждая Морфема может быть включена в состав множества слов.

1	2	3	4	5
Связанное слово	Морфологическая конструкция	связано с	один-ко-многим	Одной записи сущности Связанное слово соответствуют несколько записей сущности Морфологическая конструкция, каждой записи сущности Морфологическая конструкция соответствует множество записей сущности Связанное слово. Одно слово (морфологическая конструкция) может иметь несколько связанных с ним синонимов, антонимов и т.д., в то время как каждая морфологическая конструкция может быть как синонимом для одного слова (морфологической конструкции), так и, например, омонимом для другого.
Связанное слово	Тип связи	определяется	один-ко-многим	Одной записи сущности Связанное слово соответствуют несколько записей сущности Тип связи, каждой записи сущности Тип связи соответствует множество записей сущности Связанное слово. Один тип связи может определять несколько связей слов между собой, в то время как каждой связи слов (морфологических конструкций) может соответствовать только один тип связи.

1	2	3	4	5
Связанное слово	Тип связи	определяется	один-ко-многим	Одной записи сущности Связанное слово соответствуют несколько записей сущности Тип связи, каждой записи сущности Тип связи соответствует множество записей сущности Связанное слово. Один тип связи (синоним, пароним и т.д.) может определять несколько связей слов между собой, в то время как каждой связи слов (морфологических конструкций) может соответствовать только один тип связи.

2.5.2 Логическое проектирование

В программе ERWin была составлена схема логической модели БД, представленная в приложении Б.

Целью данного этапа является построение реляционной логической модели. Реляционная логическая модель представляет собой совокупность нормализованных отношений, в которых реализованы связи между объектами предметной области и выполнены все преобразования, необходимые для ее эффективной реализации в среде конкретной СУБД.

Выполним отображение сущностей инфологической модели на отношения реляционной модели.

Связь «Морфологическая конструкция – Часть речи» является связью типа один–ко–многим. При отображении ключ порожденной сущности добавляется в исходную сущность. Исходной сущностью является сущность Морфологическая конструкция, порожденной – Часть речи. Связь показана на рисунке 2.1, на рисунке 2.2 приведены итоговые отношения.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		27

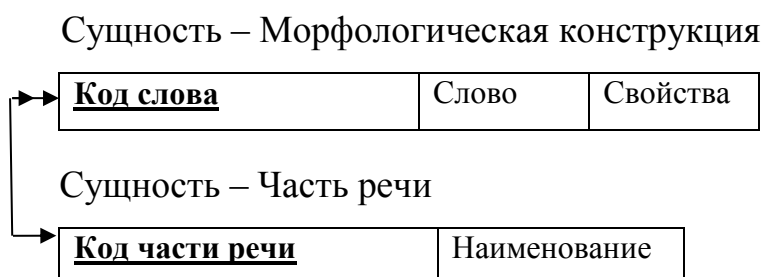


Рисунок 2.1 – Связь «Морфологическая конструкция – Часть речи»

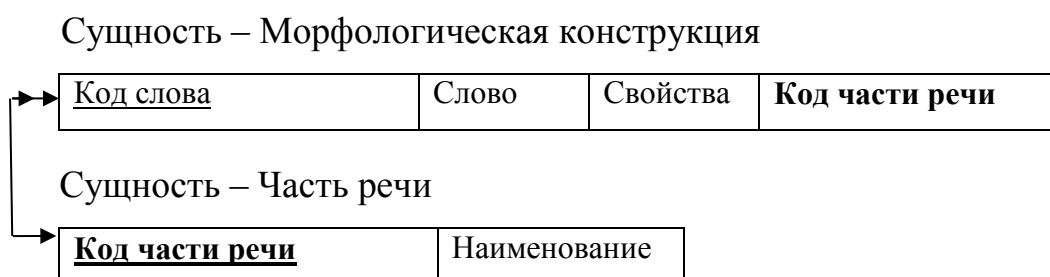


Рисунок 2.2 – Отображение связи «Морфологическая конструкция – Часть речи»

Связь «Ветви синтаксической конструкции – Синтаксическая конструкция» является связью типа один–ко–многим. При отображении ключ порожденной сущности добавляется в исходную сущность, причем сущность имеет два атрибута, содержащие отличные друг от друга ключи сущностей одного типа. Причиной этого является тот факт, что, например, абзац включает в себя набор взаимосвязанных предложений, предложения – набор словосочетаний и т.п. Все перечисленное, соответственно, является синтаксическими конструкциями. Исходной сущностью является сущность Ветви синтаксической конструкции, порожденной – Синтаксическая конструкция. Связь показана на рисунке 2.3, на рисунке 2.4 приведены итоговые отношения.

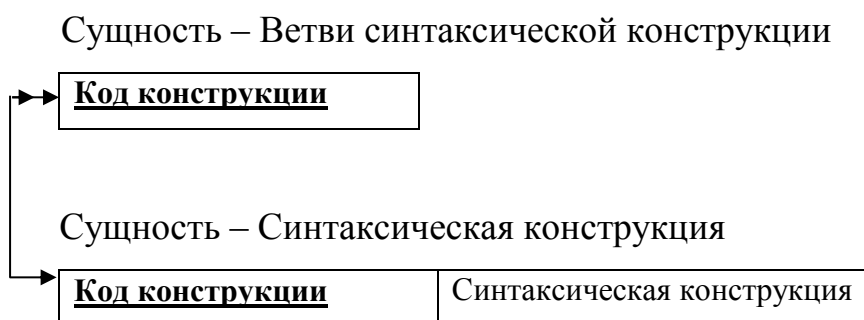


Рисунок 2.3 – Связь «Ветви синтаксической конструкции – Синтаксическая конструкция»

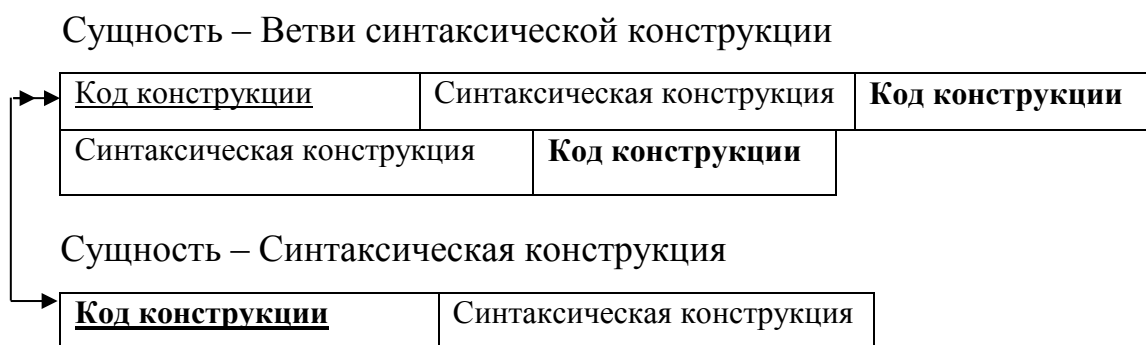


Рисунок 2.4 – Отображение связи «Ветви синтаксической конструкции – Синтаксическая конструкция»

Аналогичным образом строится связь «Слова в синтаксической конструкции – Морфологическая конструкция». При отображении ключ порожденной сущности добавляется в исходную сущность. Данная связь является связью типа один–ко–многим. Исходной сущностью является сущность Морфологическая конструкция, порожденной – Часть речи. Связь показана на рисунке 2.5, на рисунке 2.6 приведены итоговые отношения.

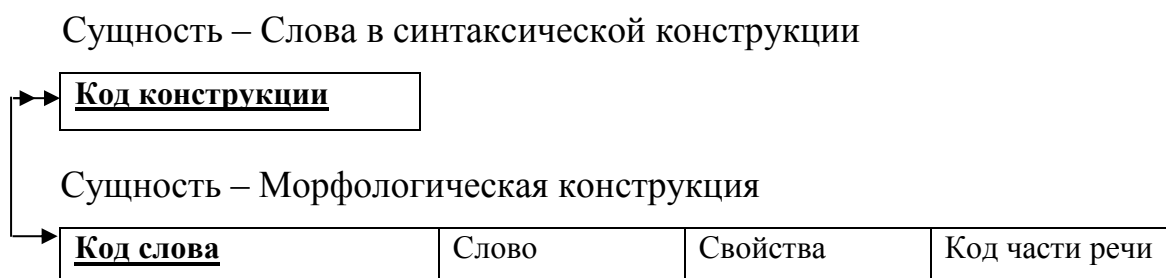


Рисунок 2.5 – Связь «Слова в синтаксической конструкции – Морфологическая конструкция»

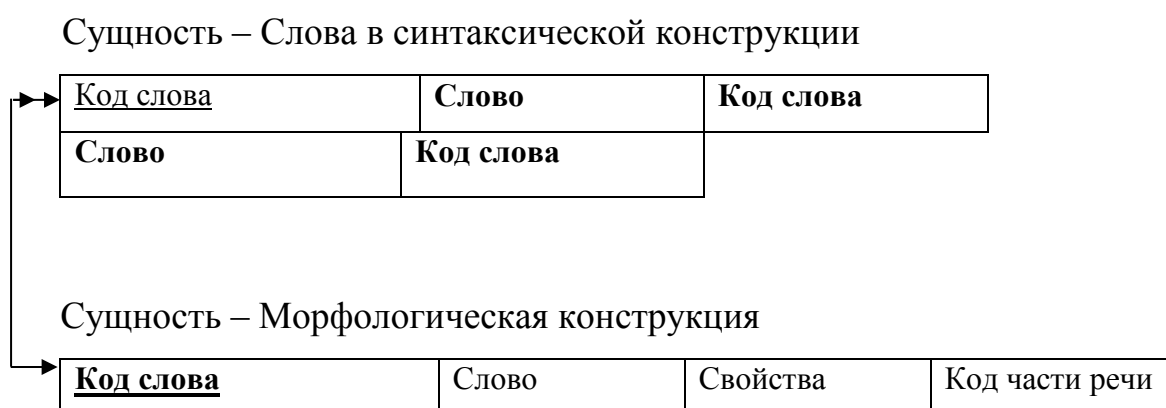


Рисунок 2.6 – Отображение связи «Слова в синтаксической конструкции – Морфологическая конструкция»

Связь «Составная часть слова – Морфема» является связью типа один–ко–многим. При отображении ключ порожденной сущности добавляется в исходную сущность. Исходной сущностью является сущность Составная часть слова, порожденной – Морфема. Связь показана на рисунке 2.7, на рисунке 2.8 приведены итоговые отношения.

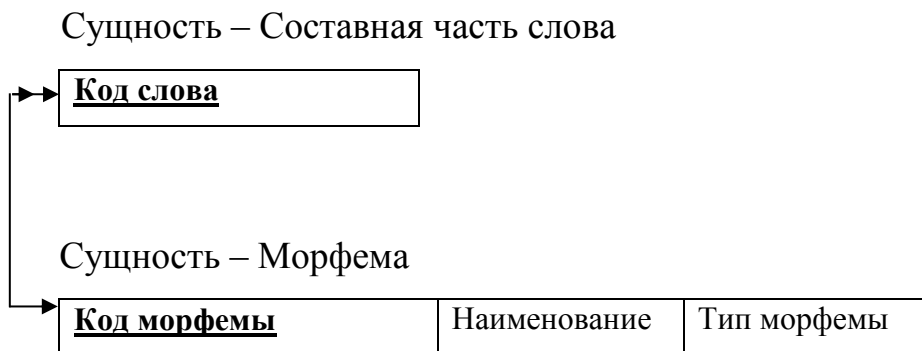


Рисунок 2.7 – Связь «Составная часть слова – Морфема»

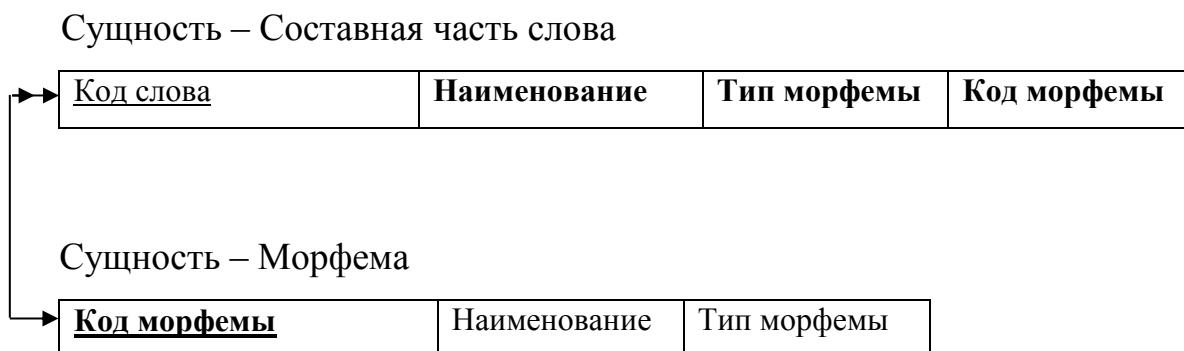


Рисунок 2.8 – Отображение связи «Составная часть слова – Морфема»

Связь «Связанное слово – Морфема» является связью типа один–ко–многим. При отображении ключ порожденной сущности добавляется в исходную сущность. Исходной сущностью является сущность Связанное слово, порожденной – Морфологическая конструкция. Связь показана на рисунке 2.9, на рисунке 2.10 приведены итоговые отношения.

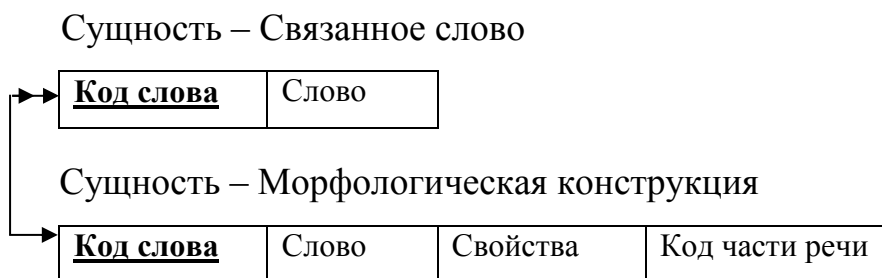


Рисунок 2.9 – Связь «Связанное слово – Морфологическая конструкция»

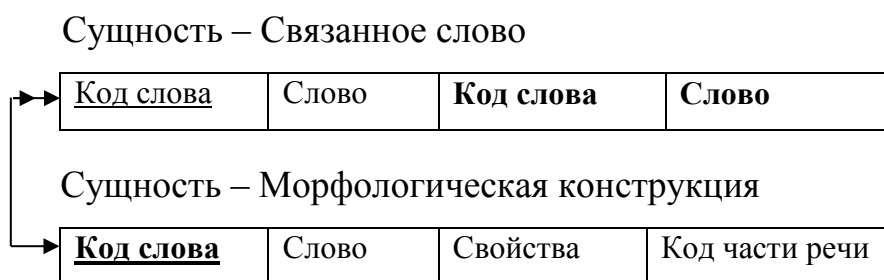


Рисунок 2.10 – Отображение связи «Связанное слово – Морфологическая конструкция»

Связь «Связанное слово – Тип связи» является связью типа один–ко–многим. При отображении ключ порожденной сущности добавляется в исходную сущность. Исходной сущностью является сущность Связанное слово, порожденной – Тип связи. Связь показана на рисунке 2.11, на рисунке 2.12 приведены итоговые отношения.



Рисунок 2.11 – Связь «Связанное слово – Тип связи»

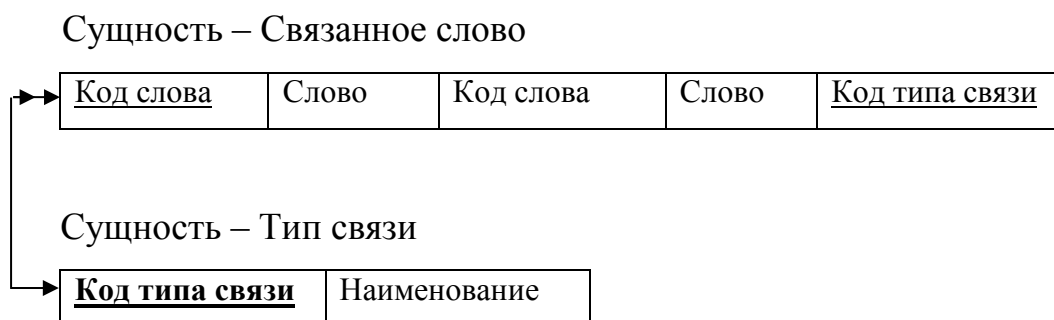


Рисунок 2.12 – Отображение связи «Связанное слово – Тип связи»

Теперь переходим к этапу нормализации отношений между сущностями.

Этап нормализации состоит в следующем: готовую реляционную модель БД необходимо привести к трем нормальным формам.

Так как в разрабатываемой подсистеме, большинство данных содержатся

в БД в формате синтаксиса языка Prolog, и используются исключительно как часть базы знаний, то нормализация отношений является нецелесообразным.

2.5.3 Физическое проектирование

Завершив этап логического проектирования, мы можем перейти к проектированию физическому. Данный этап заключается в приведении спроектированной реляционной модели к форме таблиц, которые затем будут реализованы в СУБД.

Согласно пункту 2.1, в подсистеме используется встроенная в платформу 1С:Предприятие 8.3 СУБД. В таблицах 12-20 представлены данные в соответствии с формой их представления в ней.

Для представления в физической модели сущность Морфологические конструкции была переименована в Словарь.

Таблица 12 – Словарь

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Код слова	Число	12	-	-	нет	да (совпадения не допускаются)
Слово	Строка	100	-	-	нет	да
Свойства	Строка	200	-	-	нет	нет
Часть речи	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)

Таблица 13 – Синтаксические конструкции

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Код конструкции	Число	12	-	-	нет	да (совпадения не допускаются)
Наименование	Строка	неогр.	-	-	нет	нет

Таблица 14 – Часть речи

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Часть речи	Строка	фикс.	(наименование частей речи русского языка)	-	нет	да (совпадения не допускаются)

Таблица 15 – Морфемы

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Наименование	Строка	25	-	-	нет	да (совпадения не допускаются)
Тип морфемы	Строка	25	-	-	нет	да (совпадения не допускаются)

Таблица 16 – Составные части слова

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Родитель	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Слово	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Тип морфемы	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)

Таблица 17 – Типы связи

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Наименование	Строка	25	(синоним, антоним, омоним, пароним)	-	нет	да (совпадения не допускаются)

Таблица 18 – Связанные слова

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Родитель	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Связанное слово	Ссылка	фикс.	Данные по ссылке	-	нет	да
Тип связи	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)

Таблица 19 – Ветви синтаксической конструкции

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Родитель	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Левая ветвь	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Правая ветвь	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)

Таблица 20 – Слова в синтаксической конструкции

Название поля	Тип данных	Длина	Ограничения	Значения по умолчанию	NULL	Индексация
Родитель	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Левая ветвь	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)
Правая ветвь	Ссылка	фикс.	Данные по ссылке	-	нет	да (совпадения не допускаются)

Спроектированная физическая модель БД представлена в приложении В.

В приведенных таблицах 12-20 многие поля имеют тип Ссылка по той причине, что в системе «1С:Предприятие» реализован механизм ссылок. Общий смысл данного механизма заключается в том, что в режиме

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		34

1С:Предприятие можно получать текстовые ссылки на данные, хранящиеся в информационной базе. Условно такие ссылки разделяются на интерактивные и неинтерактивные. В разрабатываемой подсистеме данная особенность представляет интерес в плане программного обращения к ссылкам в коде или запросе, для более быстрого получения необходимой информации.

2.6 Проектирование синтаксического анализатора

Вторым важным компонентом разрабатываемой подсистемы является синтаксический анализатор, который берет на себя основной процесс преобразования входных данных в выходные. Как следует из пункта 2.1, в качестве платформы для его разработки должна быть использована среда разработки Visual Prolog 5.2.

Основная задача, при проектировании и реализации анализатора, состоит в создании простой базы знаний, содержащей синтаксические правила построения словосочетаний, предложений и прочих синтаксических конструкций русского языка, а так же правил вывода результата. Опираясь на синтаксис языка Prolog, для их описания используются предложения Хорна, являющиеся подмножеством логики предикатов.

За основу построения алгоритмов анализатора следует взять один из наиболее эффективных и легко реализуемых – алгоритм восходящего парсера.

Сам анализатор может быть представлен в качестве дополнительного ПО, активно взаимодействующего с приложением, разработанным, в соответствии с предыдущими пунктами, на платформе 1С:Предприятие 8.3.

В качестве вспомогательного методического материала следует использовать учебную и научную литературу по синтаксису и морфологии русского языка.

2.7 Проектирование клиентского приложения

В своей окончательной структуре приложение должно иметь пользовательский интерфейс, организующий взаимодействие пользователя с подсистемой и программные модули, организующие исполнение программного кода при совершении пользователем определенных действий на формах.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		35

Для модуля, организующего взаимосвязь информационной базы с приложением-анализатором необходимо написать оптимизированный код на языке 1С, использующий язык запросов к БД, а так же преобразование выходных данных из конфигурации во входные данные для анализатора и обратно.

Построение форм легко реализуется встроенным конструктором платформы 1С, с описанием обработки событий для объектов форм. Они должны содержать весь необходимый функционал и информационные окна для полноценного взаимодействия пользователя с подсистемой.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		36

3 РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

Завершив этапы анализа предметной области и проектирования, можно приступать к реализации подсистемы. В первую очередь разрабатывается база данных слов русского языка и морфологических признаков, а так же синтаксическая БД.

Параллельно с этим идет разработка синтаксического анализатора на Prolog и последним этапом станет разработка пользовательского интерфейса, на основании представления входных и выходных данных подсистемы.

Опишем поочередно каждый из вышеприведенных этапов.

3.1 Разработка морфологической и синтаксической БД

СУБД для создания и администрирования БД, согласно пункту 2.1, является платформа 1С:Предприятие 8.3. Следовательно, основные таблицы физически реализуются в виде Справочников, а промежуточные таблицы, связывающие данные одного и того же, либо двух разных справочников между собой – в виде Табличных частей Справочников. Таблицы, содержащие ограниченный набор записей, использующихся как основные свойства данных из справочников, реализуются в виде объекта конфигурации – Перечисление. В общем виде данная структура выглядит следующим образом:

– справочник «Словарь» содержит табличные части: «Составные части слова», содержащая список морфем, входящих в состав данного слова; «Связанные слова», включающая слова из этого же справочника, связанные с данным словом как синонимы, антонимы и т.д.;

– справочник «Синтаксические конструкции» содержит табличные части: «Ветви синтаксической конструкции», содержащая список конструкций-поддеревьев, входящих в состав данной конструкции-дерева; «Слова в синтаксической конструкции», включающая слова – морфологические конструкции-поддеревья, входящие в состав данной конструкции-дерева;

– справочник «Морфемы», данные которого пополняются при мор-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		37

фемном разборе слов и используются в табличной части «Состав слова»;

– перечисления «Тип связи» и «Часть речи».

Для заполнения описанных выше таблиц был разработан общий программный модуль конфигурации на языке 1С – «Заполнение Словаря», циклично считывающий данные из морфологического словаря, представленного в виде текстового файла, и посредством языка запросов записывающий их в БД. В итоге, на своем начальном этапе, подсистема содержит базу почти 4,5 млн. слов русского языка в различных формах и состоящих в строгой иерархии, согласно их морфологическим признакам.

Справочники, табличные части и перечисления, данные, которых, используются справочником «Словарь», и составляют в совокупности морфологическую базу данных подсистемы. Относящиеся же к справочнику «Синтаксические конструкции» – составляют, в свою очередь, синтаксическую БД.

На рисунке 3.1 представлена таблица «Словарь», которая, согласно пункту 2.5.1, содержит морфологические конструкции. Колонка «Наименование» содержит конструкцию, представленную на естественном языке в виде слова, к значению которой подсистема обращается посредством запроса, при поиске соответствия в БД слову из анализируемого текста. В колонке «Свойства» содержится структура из морфологических свойств слова, представленная в соответствии синтаксису языка Prolog. При выполнении анализа, каждые найденные соответствия словам из текста записываются в специально созданный для этого массив, из которого, затем, в модуле обмена фактами с синтаксическим анализатором происходит преобразование данных из всей строки данной таблицы в факты на языке Prolog. Далее из всех найденных и преобразованных слов составляется список и записываются в текстовый документ, являющийся внешней базой фактов для синтаксического анализатора.

Форма записи слова в БД представлена на рисунке 3.2. При записи слово преобразуется в морфологическую структуру таблицы «Словарь», описанную в колонке «Свойства» на вышеприведенном рисунке. На рисунке 3.3 приведены склонения по падежам того же слова, имеющие атрибут «Родитель», содержа-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		38

щий, в данном случае, ссылку на именительный падеж данного слова из рисунка 3.2.

Для глаголов, например, данная иерархия выражена в ссылке на инфинитив, от которого наследуются основные признаки для различных, его, представлений в тексте.

Наименование	Код	Свойства
информатики	33 685	свойства_сущ(0000033685,"Неодушевленное",0,"", "Множественное", "Имени...
информатор	33 686	свойства_сущ(0000033686,"Одушевленное",0,"Мужской", "Единственное", "И...
информаторы	33 687	свойства_сущ(0000033687,"Одушевленное",0,"", "Множественное", "Имени...
информационность	33 688	свойства_сущ(0000033688,"Неодушевленное",0,"Женский", "Единственное", "...
информационности	33 689	свойства_сущ(0000033689,"Неодушевленное",0,"", "Множественное", "Имени...
информация	33 690	свойства_сущ(0000033690,"Неодушевленное",0,"Женский", "Единственное", "...
информации	33 691	свойства_сущ(0000033691,"Неодушевленное",0,"", "Множественное", "Имени...
информирование	33 692	свойства_сущ(0000033692,"Неодушевленное",0,"Средний", "Единственное", "...
информирования	33 693	свойства_сущ(0000033693,"Неодушевленное",0,"", "Множественное", "Имени...
информированность	33 694	свойства_сущ(0000033694,"Неодушевленное",0,"Женский", "Единственное", "...
информированности	33 695	свойства_сущ(0000033695,"Неодушевленное",0,"", "Множественное", "Имени...
информосома	33 696	свойства_сущ(0000033696,"Неодушевленное",0,"Женский", "Единственное", "...
информосомы	33 697	свойства_сущ(0000033697,"Неодушевленное",0,"", "Множественное", "Имени...
инфразвук	33 698	свойства_сущ(0000033698,"Неодушевленное",0,"Мужской", "Единственное", "...
инфразвуки	33 699	свойства_сущ(0000033699,"Неодушевленное",0,"", "Множественное", "Имени...

Рисунок 3.1 – Таблица «Словарь»

Таким образом, поддерживается иерархия, которая, при дальнейшей разработке подсистемы, выходящей за рамки данной ВКР, играет роль отношений для автоматического поиска и исправления ошибок в тексте, а так же для предложения или, в случае семантической интерпретации, предположения альтернатив анализируемому в тексте слову или фрагменту.

Соответствующим образом, данная иерархия позволяет наследовать семантическое значение слова на его словоформы, учитывая их контекст в зависимости от склонения, времени, спряжения и т.д. Данная концепция основана на том, что словами человек описывает окружающий мир, каждый объект и каждое явление которого могут быть каким-либо образом связаны между собой.

Синтаксическая конструкция может включать в себя как такую же син-

таксическую конструкцию уровнем ниже, так и слово, связанное с этой, второй, конструкцией. Пример представлен на рисунках 3.4 и 3.5.

Рисунок 3.2 – Форма записи слова в БД

Существительное	Код	Одушевленность	Имя собственное	Род	Число	Падеж
информация	0000033690	Неодушевленное		Женский	Единственное	Именительный
информаци	0000298013	Неодушевленное		Женский	Единственное	Родительный
информаци	0000298014	Неодушевленное		Женский	Единственное	Дательный
информацию	0000298015	Неодушевленное		Женский	Единственное	Винительный
информацией	0000298016	Неодушевленное		Женский	Единственное	Творительный
информаци	0000298017	Неодушевленное		Женский	Единственное	Предложный

Рисунок 3.3 – Склонения слова по падежам

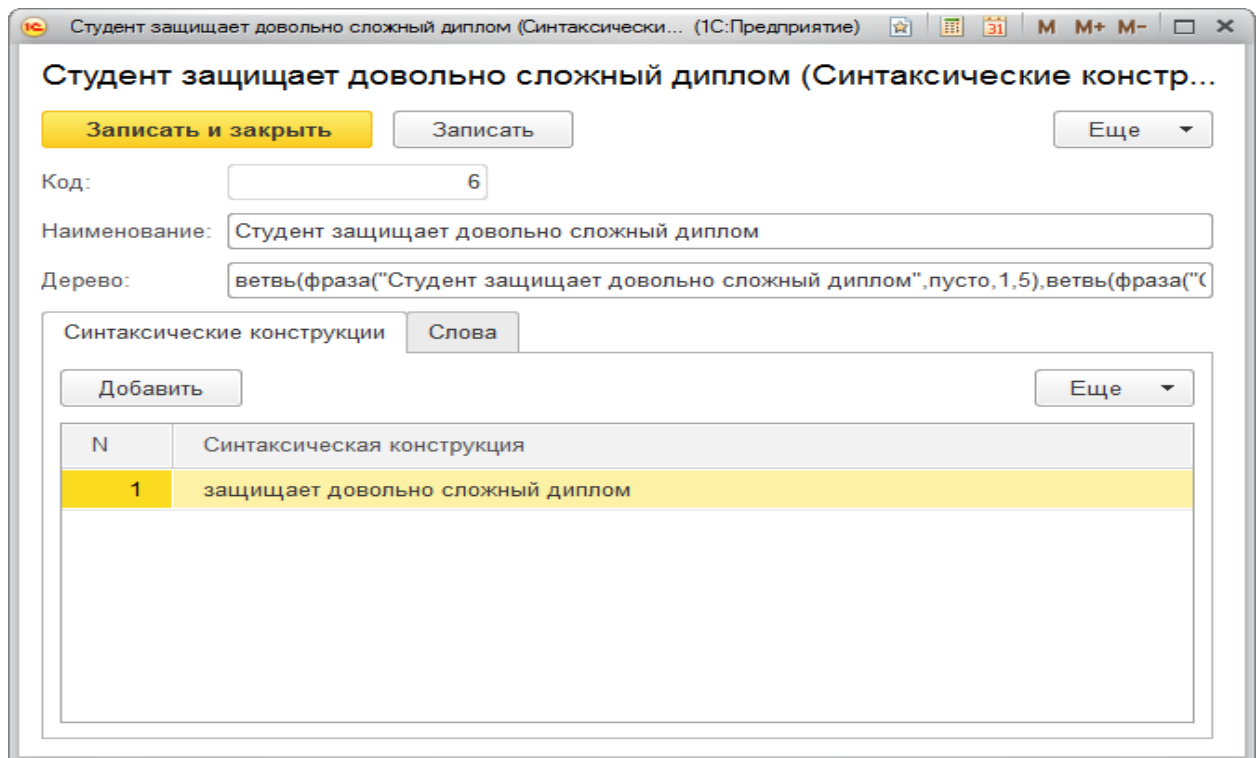


Рисунок 3.4 – Конструкция, включающая другую конструкцию

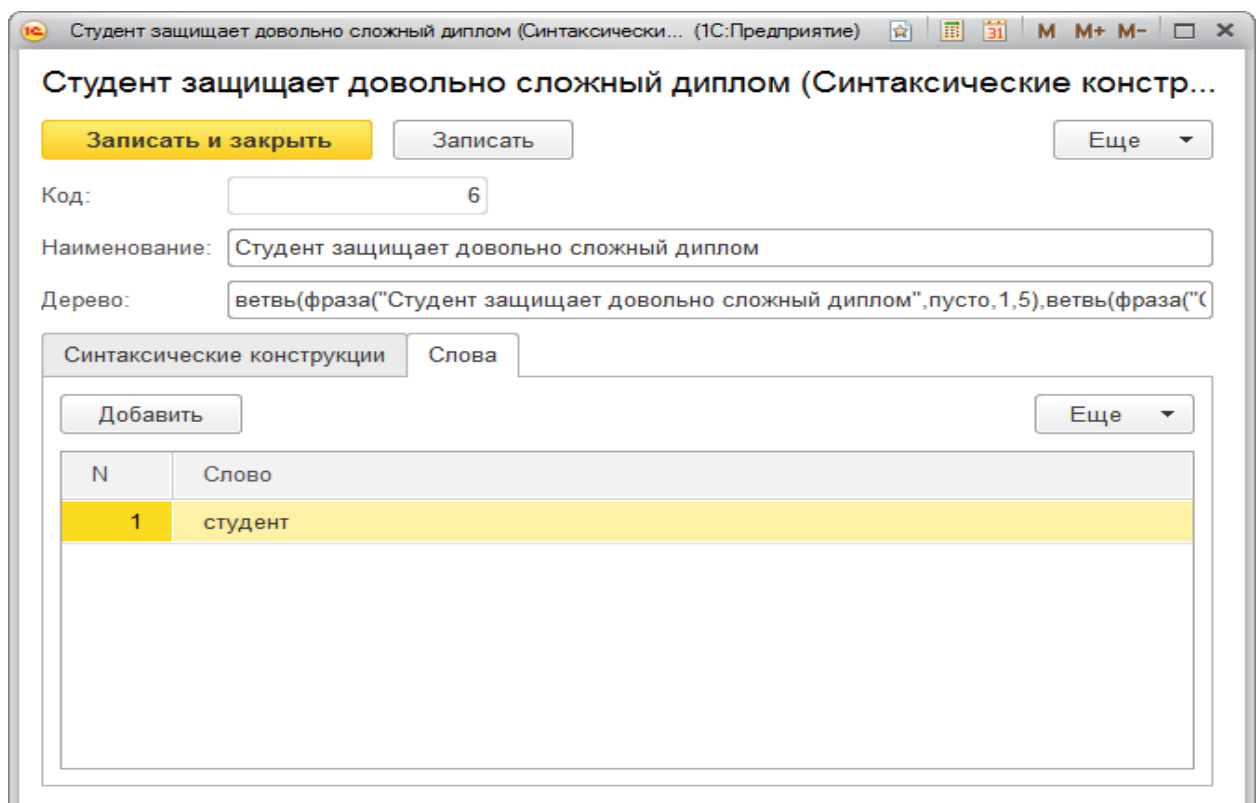


Рисунок 3.5 – Слово, включенное в ту же конструкцию

Конструкция, включенная в данную синтаксическую конструкцию, в

свою очередь, состоит из еще одних конструкций, более низкого уровня, и так до уровня слов – морфологических конструкций. Все данные, в вышеприведенных рисунках, являются ссылками на данные морфологической и синтаксической БД. Таким образом, все данные базы данных подсистемы взаимосвязаны между собой. На основе этих связей возможна дальнейшая разработка логических, семантических механизмов анализа и т.д.

3.2 Разработка синтаксического анализатора

Согласно тому же пункту 2.1, для разработки синтаксического анализатора используется среда Visual Prolog 5.2, на языке логического программирования Prolog. Как говорилось в пункте 1.1.1, данный язык оперирует фактами и правилами, т.е. предикатами – выражениями, имеющими значения лишь Ложь или Истина. Факты – это однозначно истинные предикаты, в то время, как правила – это предикаты, истинность которых напрямую зависит от наличия всех фактов и подправил, которые они в себя включают. Другими словами, правила требуют доказательств.

Синтаксис Prolog подразумевает разделение базы фактов и базы правил, причем, если правила необходимо создать и полностью описать на этапе разработки кода, то для фактов можно использовать сторонние базы, называемые базами фактов. Именно на этом принципе и построен, в данной подсистеме, механизм обмена данными в виде фактов между БД на 1С и БЗ на Prolog.

При создании синтаксических правил связи слов, словосочетаний и использования знаков препинания русского языка разработчик активно использует методическую и учебную литературу по синтаксису русского языка. Правила синтаксического разбора описываются в виде двух морфологических конструкций, каждая из которых имеет определенные признаки, уникальные для данного правила, либо, что чаще всего, данные признаки являются общими для этих двух конструкций, например: падеж, род, число. Сами признаки записаны в виде переменных, которые, при выполнении анализа, связываются с такими же признаками у фактов-конструкций из подключаемой базы фактов, если таковые связки возможны.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		42

Если все переменные первой и второй конструкций в правиле удалось связать и они соответствуют условиям правила (равенство признаков, определенный падеж одного из слов или любой другой, кроме него, время и т.д.), то данное правило принимает значение Истина.

Ниже представлен листинг одного из таких правил:

```
синтаксические_правила( Слово1, Слово2, Свойства1, Свойства2, X, Итоговое_слово ) :-
X=63, структура_из_слов( ветвь( фраза( _, Свойства1, _, _ ), _, _ ),
ветвь( фраза( _, свойства_мест( _, "Прилагательное", Род, "Единственное", Падеж, _, _
), _, _ ), _, _ )),
структура_из_слов( ветвь( фраза( _, Свойства2, _, _ ), _, _ ),
ветвь( фраза( _, свойства_сущ( _, _, _, Род, "Единственное", Падеж, _, _, _ ), _, _ ), _, _
)), concat( Слово1, " ", Промежуточное_слово ),
concat( Промежуточное_слово, Слово2, Итоговое_слово );
```

Предикат «синтаксические правила» производит поиск соответствия переменным, указанным в скобках, в предикатах, описанных после символа «:-», а так же проверяет эти предикаты на истинность. Морфологические конструкции представлены и воспринимаются программой как структуры-деревья, содержащие морфологические свойства слов. Символ «_» обозначает анонимную переменную, значение которой, в контексте данного правила, программу не интересует, а потому полностью игнорируется.

Листинг одной из таких конструкций представлен следующим образом:

```
структура_из_слов(
ветвь( фраза( "парень", свойства_сущ( 0000067833, "Одушевленное", 0, "Мужской",
"Единственное", "Именительный", "", 0, "" ), 1, 1 ),
пусто,
пусто )).
```

Схематично данная конструкция представлена на рисунке 3.6.

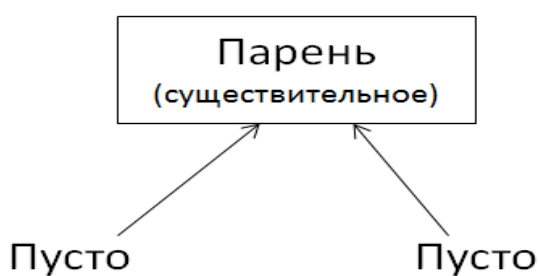


Рисунок 3.6 – Схема морфологической конструкции

Поскольку все конструкции в анализаторе представлены в виде деревьев, то и морфологические конструкции, несмотря на то, что являются ветвями самого низкого уровня, включают в себя две ветви, не содержащие никакого значения, что обозначено доменом «пусто».

Для построения синтаксических структур, на основе истинности синтаксических правил, написаны рекурсивные правила анализа, основанные на алгоритме восходящего парсинга. Суть алгоритма заключается в построении синтаксического дерева, чаще всего одного целого предложения, начиная от слов – нижнего уровня, и заканчивая данным предложением, являющимся самым верхним уровнем – корнем. Другими словами, анализ происходит снизу-вверх. Схематичное описание данного алгоритма представлено на рисунке 3.3.

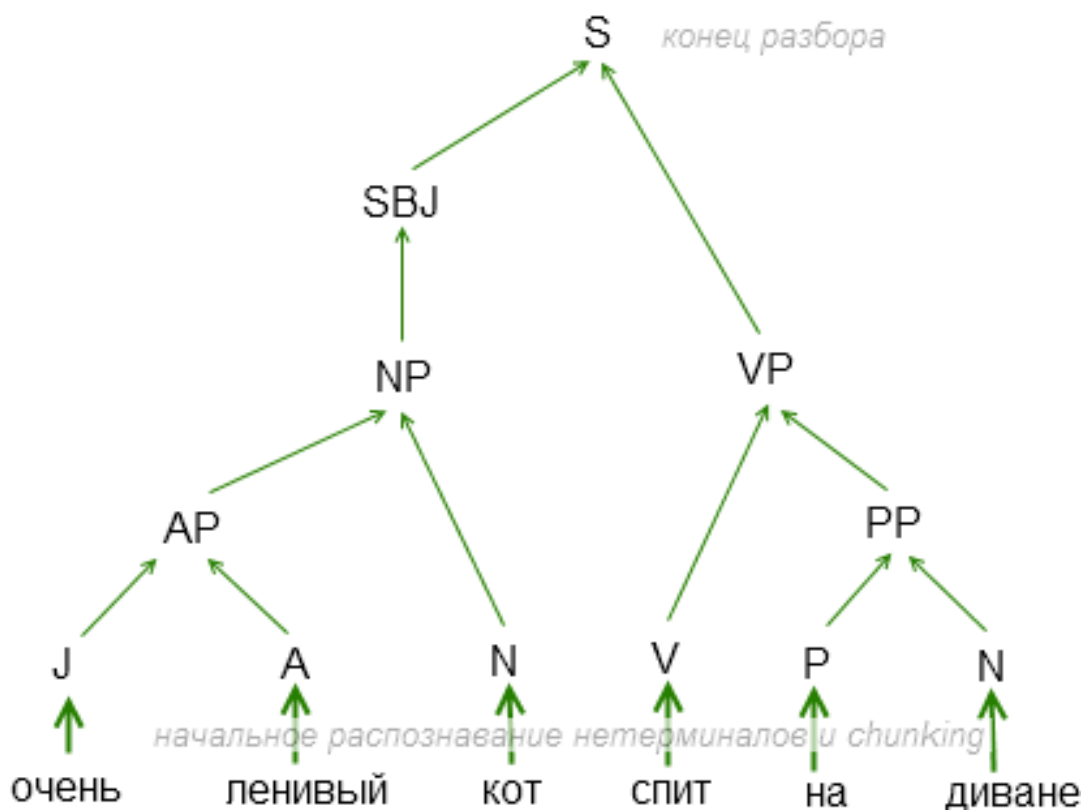


Рисунок 3.7 – Алгоритм восходящего парсинга

Вначале собираются конструкции-свойства, например группы прилагательного, группы деепричастия и т.д., а так же конструкции с предлогами и частицами. Затем данные конструкции связываются с двумя основными группами, составляющими главную часть предложения:

– группа существительного, местоимения или числительного в именительном падеже, которая может включать такие же части речи, но в склонениях, уточняющие конструкции-свойства;

– группа глагола.

В качестве примера приведем схему синтаксической конструкции проанализированного простого предложения «Парень читает интересную книгу» на рисунке 3.8.

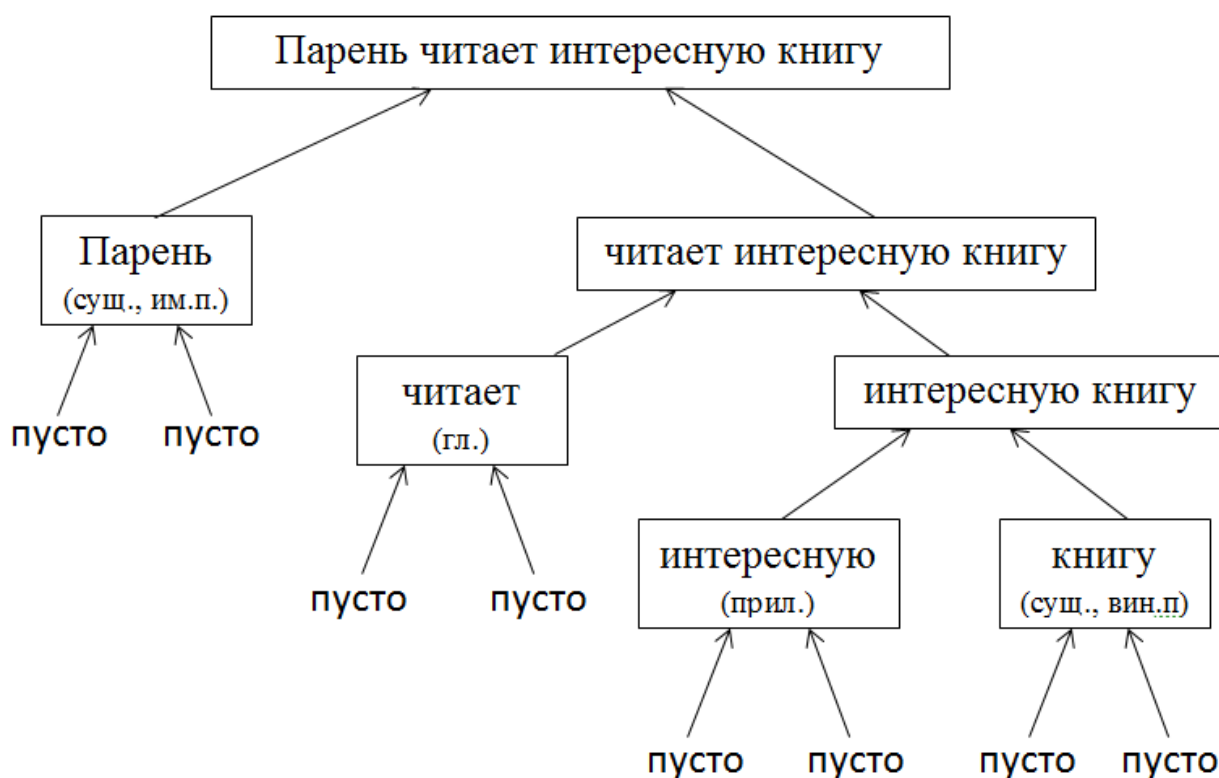


Рисунок 3.8 – Схема синтаксической конструкции

Ниже представлен листинг данной конструкции, в качестве входных или выходных данных синтаксического анализатора:

```

структура_из_слов(
  ветвь( фраза( "Парень читает интересную книгу", пусто, 1, 4 ),
    ветвь(фраза("Парень", свойства_сущ( 67833, "Одушевленное", 0, "Мужской",
      "Единственное", "Именительный", "", 0, "" ), 1, 1 ),
      пусто,
      пусто ),
    ветвь( фраза( "читает интересную книгу", пусто, 2, 4 ),
      ветвь( фраза( "читает", свойства_гл(461929, "Несовершенный", 1, 0, 0, 0, "", 0,
        "Настоящее", "", "Единственное", "", "", 0, "читать" ), 2, 2 ),
        пусто,
  
```

Изм.	Лист	№ докум.	Подп.	Дата

ВКР.135194.09.03.02.ПЗ

Лист

45

пусто),
 ветвь(фраза("интересную книгу", пусто, 3, 4),
 ветвь(фраза("интересную", свойства_прил(333886, "", 0, 0, "Женский",
 "Единственное", "Винительный", 0, 0, "интересная"), 3, 3),
 пусто,
 пусто),
 ветвь(фраза("книгу", свойства_сущ(321777, "Неодушевленное", 0,
 "Женский", "Единственное", "Винительный", "", 0, "книга"), 4, 4),
 пусто,
 пусто)))))).

Вышеприведенные структуры, затем, обратно поступают в модуль системы на 1С и производится их обратное преобразование и запись в синтаксическую БД.

3.3 Разработка подсистемы обмена данными БД с синтаксическим анализатором

Данная подсистема разработана на языке 1С и выполняет следующий алгоритм действий.

На вход системы подается текст, написанный пользователем вручную в строке ввода, либо построчно из текстового документа. Из текста циклично выделяются слова и знаки препинания, которые, затем, записываются в массив слов. Далее, в процедуре обработки этого массива происходит цикличный перебор комбинации слов, т.е. предложений, словосочетаний, вводных слов и т.д., с одновременным запросом в БД на наличие данной комбинации, начиная от самой длинной комбинации и заканчивая, в случае отсутствия таковых, самыми короткими - словами. Таким образом, если система знает целый фрагмент текста, например предложение, то логично, что нет необходимости проводить его повторный анализ, а вместо этого достаточно взять уже проанализированный фрагмент. Данный алгоритм является своеобразным механизмом самообучения подсистемы, который можно описать так: чем больше текста анализирует подсистема, тем меньше времени затрачивается на последующий анализ. И так до, практически, мгновенного анализа, в случае содержания в БД всех возможных комбинаций естественно-языковых структур.

В конечном итоге все найденные комбинации преобразуются в базу фак-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		46

тов и передаются в качестве входных данных синтаксическому анализатору.

После анализа, алгоритм которого описан в пункте 3.2, новые факты – синтаксические конструкции, передаются обратно в подсистему в качестве входных данных, и преобразуются с последующей их записью в БД.

В приложении Е представлена подробная диаграмма процессов преобразования входных данных в выходные в разработанной подсистеме синтаксического анализа текстовой информации. На рисунке Е.1 представлена подсистема в целом, а на рисунке Е.2 представлена декомпозиция системы с подробным описанием процессов преобразования входных данных в выходные функциональными компонентами подсистемы.

3.4 Руководство пользователя

При запуске системы, в первую очередь активируется форма авторизации пользователя, представленная на рисунке 3.8. Если пользователь уже занесен в базу администратором подсистемы, то ему необходимо выбрать из списка или ввести вручную свой логин, а так же соответствующий пароль. Запрос на регистрацию в системе пользователь должен отправлять непосредственно администратору.

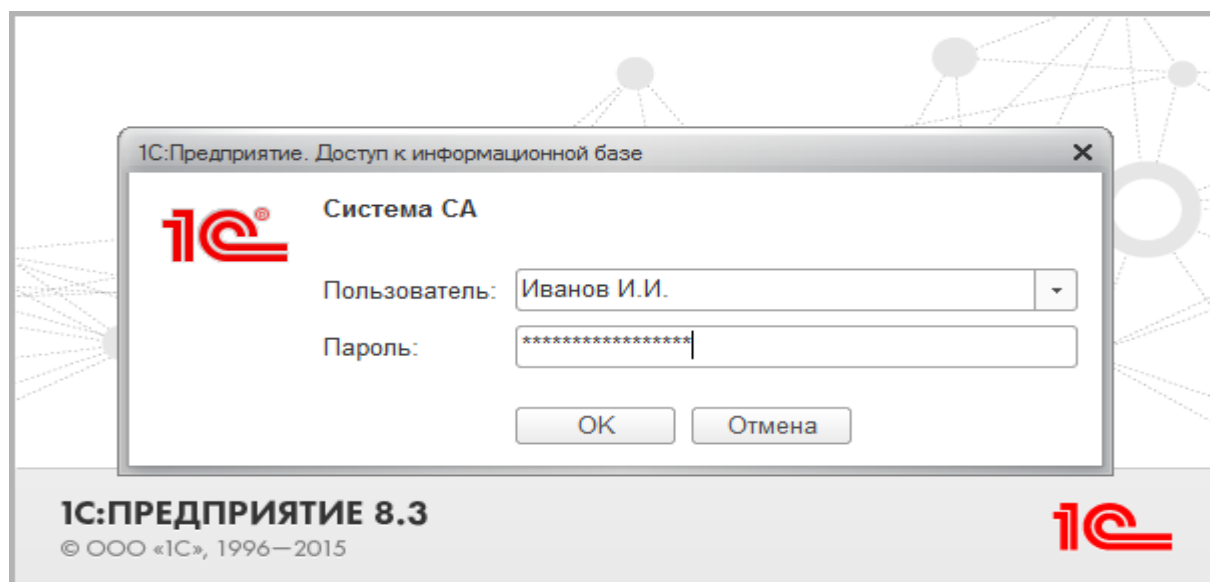


Рисунок 3.8 – Форма авторизации пользователя

После успешной авторизации запускается рабочий экран подсистемы, представленный на рисунке 3.9, который содержит сверху ссылки на морфоло-

гическую и синтаксическую БД, доступ к действиям в которой регулируется ролями, доступными пользователю. Механизм ролевого разграничения доступа подробно описан в пункте 4.2.

В рабочей области экрана слева размещается строка для ввода пользователем фраз и предложений. При нажатии на кнопку «Анализ» выполняется соответствующая обработка события, поэтапно проводящая, в автоматическом режиме, морфологический анализ. Затем выполняется запуск синтаксического анализатора, командная строка которого показана на рисунке 3.10. По окончании работы анализатора, в итоговых окнах справа, выводится результат анализа как в виде дерева значений в верхнем правом окошке, так и в виде факта на языке Prolog (для возможности отладки) в нижнем правом окошке. Результат выполнения анализа предложения, введенного пользователем, представлен на рисунке 3.11.

В то время, как в рабочем окне пользователю представлен визуальный результат работы подсистемы, в синтаксической БД имеются иерархически зависимые записи результатов, представленные в приложении Д. На рисунках приложения можно увидеть как саму синтаксическую ветвь проанализированного предложения, так и ее левую и правую ветви нижнего подуровня в табличных частях «Синтаксические конструкции» и/или «Слова».

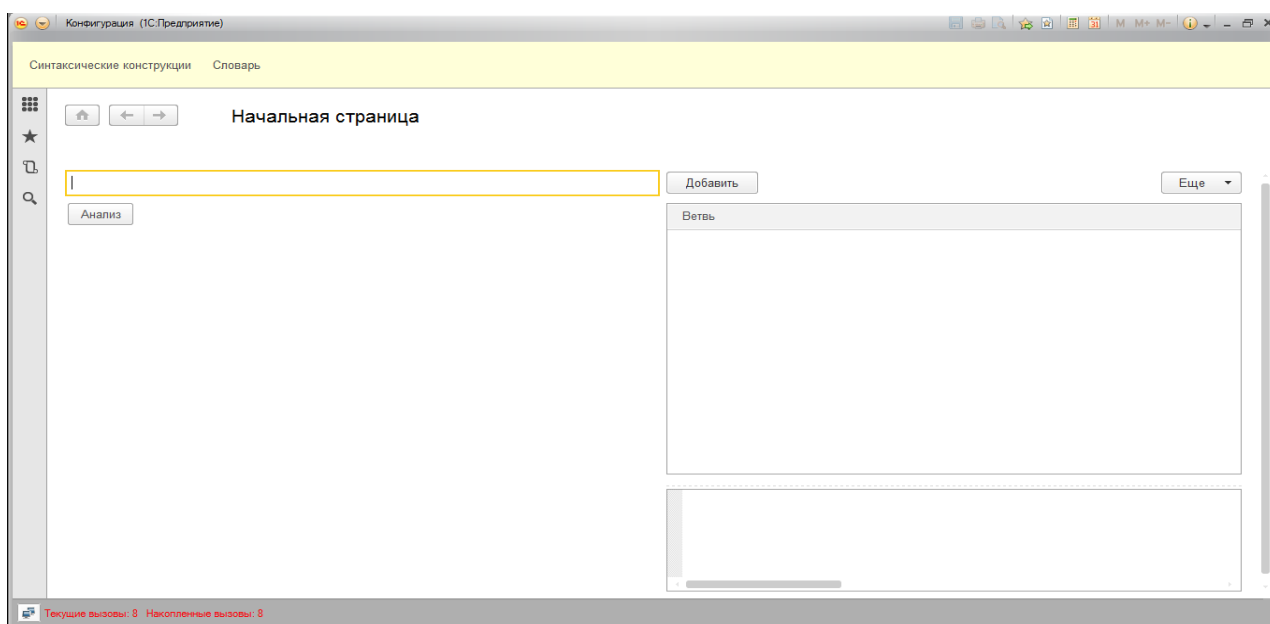


Рисунок 3.9 – Рабочий экран подсистемы

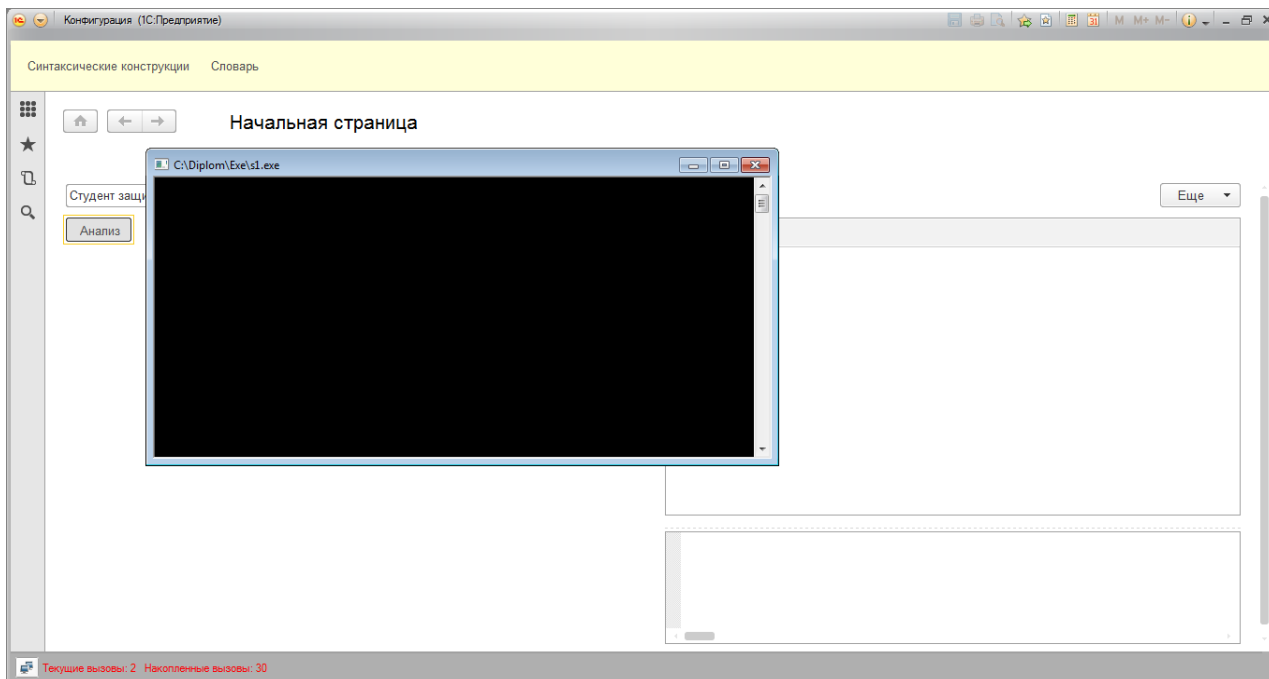


Рисунок 3.10 – Командная строка работающего exe-приложения синтаксического анализатора

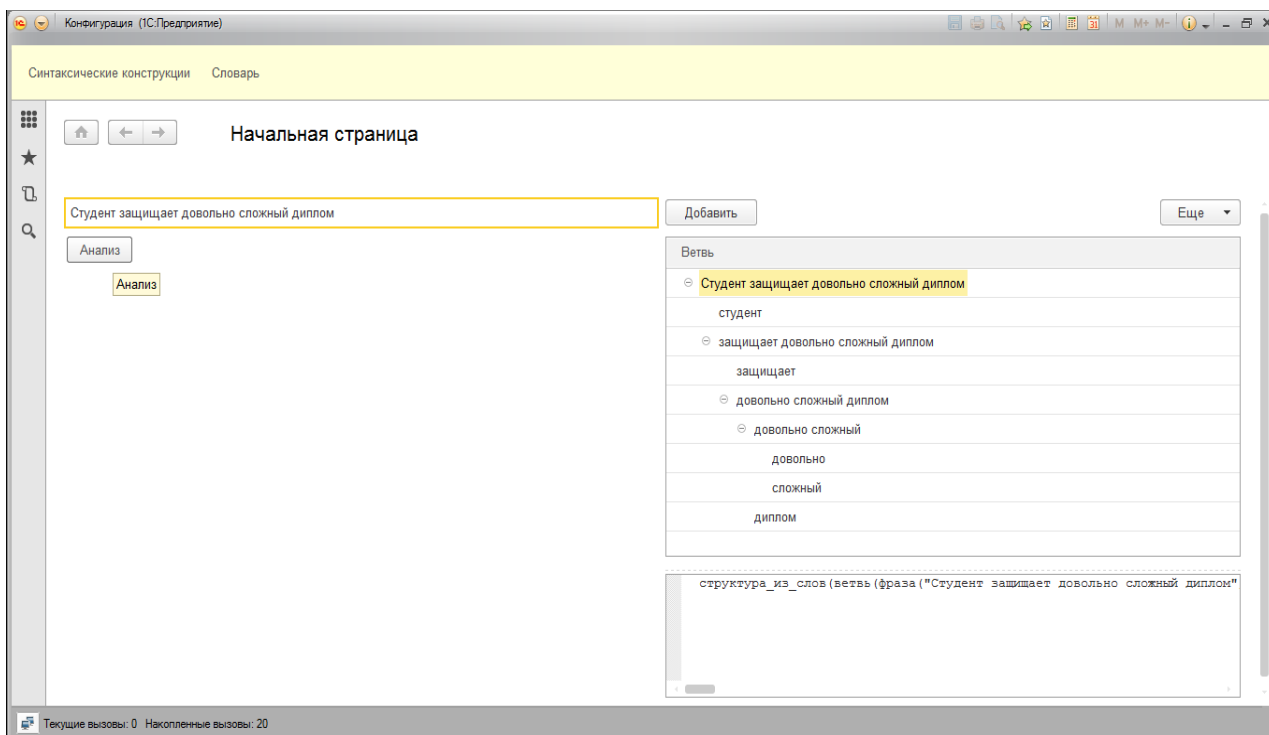


Рисунок 3.11 – Результат выполнения синтаксического анализа введенного пользователем предложения

Таким образом, на каждом из рисунков представлены следующие структуры:

– на рисунках Д.1 и Д.2 – левая и правая ветви конструкции 1-го уровня (самого верхнего), соответственно;

– на рисунках Д.3 и Д.4 – левая и правая ветви конструкции 2-го уровня, соответственно;

– на рисунках Д.5 и Д.6 – левая и правая ветви конструкции 3-го уровня, соответственно;

– и на рисунке Д.7 – левая и правая ветви конструкции 4-го уровня (самого нижнего).

Данные синтаксические конструкции, записанные в БД, теперь могут быть использованы в качестве входных данных синтаксического анализатора, без необходимости их повторного построения.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		50

4 ОПИСАНИЕ СЕРВИСОВ ОБЕСПЕЧЕНИЯ ИБ

Проблема обеспечения информационной безопасности является одной из наиболее актуальных для всех информационных систем. Под угрозой могут находиться:

- авторские права;
- данные или сведения, содержащиеся в системе и являющиеся как корпоративными, так и личными;
- правильное и полное функционирование системы.

В данной главе описываются механизмы обеспечения информационной безопасности разрабатываемой подсистемы.

4.1 Угрозы информационной безопасности

Рассмотрим подробно возможные угрозы ИБ, в общем, и для подсистемы синтаксического анализа, в частности.

Анализ негативных последствий возникновения и осуществления угроз предполагает обязательную идентификацию возможных источников угроз, уязвимостей, способствующих их проявлению и методов реализации. В связи с этим угрозы экономической и информационной безопасности можно классифицировать по нескольким критериям:

- по важнейшим составляющим информационной безопасности (доступность, целостность, конфиденциальность), против которых направлены угрозы в первую очередь;
- по компонентам информационных систем и технологий (данные, программно-аппаратные комплексы, сети, поддерживающая инфраструктура), на которые угрозы непосредственно нацелены;
- по способу осуществления (случайные или преднамеренные действия, события техногенного или природного масштаба);
- по локализации источника угроз (вне или внутри информационной технологии или системы).

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		51

На рисунке 4.1 представлена одна из возможных моделей классификаций угроз ИБ.

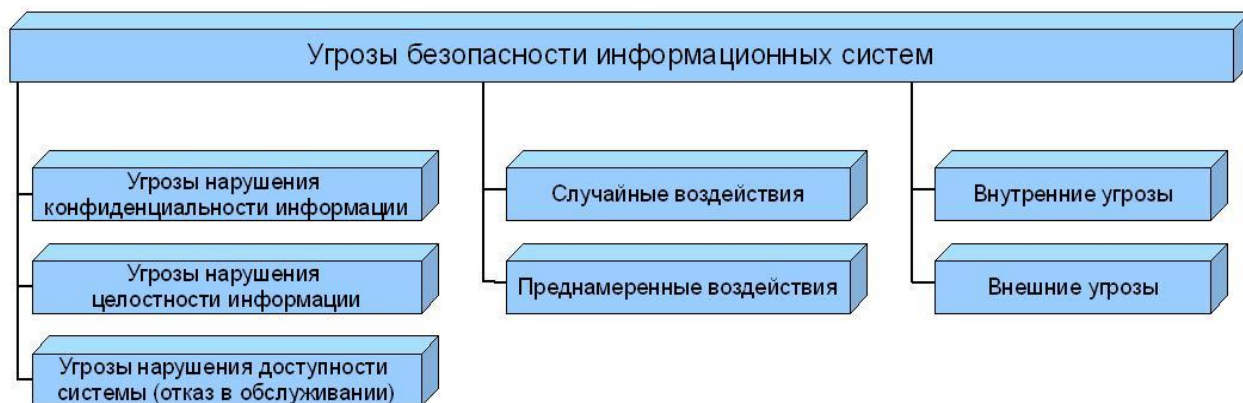


Рисунок 4.1 – Классификация угроз ИБ

У каждой информационной системы есть хотя бы один пользователь, который, в свою очередь, может быть источником следующих угроз:

- намеренная (встраивание логической бомбы, которая со временем разрушит программное ядро или приложения) или непреднамеренная потеря или искажение данных и информации, «взлом» системы администрирования, кража данных и паролей, передача их посторонним лицам и т.д.;

- нежелание пользователя работать с информационной системой (чаще всего проявляется при необходимости осваивать новые возможности или при расхождении между запросами пользователей и фактическими возможностями и техническими характеристиками) и намеренный вывод из строя её программно-аппаратных устройств;

- невозможность работать с системой в силу отсутствия соответствующей подготовки (недостаток общей компьютерной грамотности, неумение интерпретировать диагностические сообщения, неумение работать с документацией и т. п.).

Очевидно, что эффективный способ борьбы с непреднамеренными ошибками – максимальная автоматизация и стандартизация, информационных процессов, использование устройств «защита от дурака» (Fool Proof Device), регламентация и строгий контроль действий пользователей. Необходимо также

следить за тем, чтобы при увольнении сотрудника его права доступа (логического и физического) к информационным ресурсам аннулировались.

Основными источниками внутренних системных отказов являются:

– невозможность работать с системой в силу отсутствия технической поддержки (неполнота документации, недостаток справочной информации и т. п.);

– отступление (случайное или умышленное) от установленных правил эксплуатации;

– выход системы из штатного режима эксплуатации в силу случайных или преднамеренных действий пользователей или обслуживающего персонала (превышение расчетного числа запросов, чрезмерный объем обрабатываемой информации и т. п.);

– ошибки конфигурирования системы;

– отказы программного и аппаратного обеспечения;

– разрушение данных;

– разрушение или повреждение аппаратуры.

По отношению к поддерживающей инфраструктуре рекомендуется рассматривать следующие угрозы:

– нарушение работы (случайное или умышленное) систем связи, электропитания, водо- и теплоснабжения, кондиционирования;

– разрушение или повреждение помещений;

– невозможность или нежелание обслуживающего персонала и/или пользователей выполнять свои обязанности (гражданские беспорядки, аварии на транспорте, террористический акт или его угроза, забастовка и т. п.).

Опасны, разумеется, стихийные бедствия, такие как: наводнения, землетрясения, ураганы, а так же события, являющиеся результатом техногенных катастроф, например, пожары, взрывы, обрушения зданий и т.д. По статистике, на долю огня, воды и тому подобных источников угроз, среди которых самый

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		53

опасный – сбой электропитания, приходится 13-15% потерь, нанесенных производственным информационным системам и ресурсам.

Так как информационная система данной ВКР разработана на платформе 1С:Предприятие 8.3, то опишем используемую в ней модель разграничения доступа – ролевую.

4.2 Ролевая модель разграничения доступа в системе 1С:Предприятие 8

Разграничение доступа в 1С:Предприятии 8 построено на трех понятиях – права, роли и пользователи.

Разграничение доступа в 1С:Предприятии 8 задается в конфигурации и не изменяется в режиме Предприятия – таким образом, все объекты 1С:Предприятия, соответствующие некоторому объекту метаданных, например, документы одного типа, защищены одинаково. Для разграничения доступа защищаемые объекты метаданных обладают списком прав, иначе говоря, – списком контролируемых способов доступа. Включаемые в список права определяются типом объекта метаданных. Существуют как общеупотребительные права, например, чтение или просмотр, так и права, уникальные для объектов метаданных, например, для документов это проведение.

Права в списке могут неявно делиться на обычные и интерактивные. Интерактивные права проверяются только в случае, если доступ к объекту осуществляется в интерактивном режиме Предприятия, и отображение и редактирование информации выполняется системными средствами, относящимися к конкретным объектам метаданных. К этим системным средствам относятся формы, табличные поля и поля ввода, обладающие расширениями. Именно расширения перечисленных интерактивных элементов обеспечивают контроль прав соответствующих этим расширениям объектов. В случае если расширение отсутствует, отсутствует и контроль прав – так, например, если во встроенном языке считать данные с помощью запроса, выгрузить их в таблицу значений и отобразить эту таблицу в табличном поле, то проверка интерактивных прав выполняться не будет.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		54

В отличие от интерактивных, обычные права проверяются во всех случаях (обращения из встроенного языка, Automation и т.д.) и включают в себя аналогичные интерактивные права - как, например, «Удаление» и «Интерактивное удаление». Устанавливая такие права в «запрещено», разработчик полностью отключает определяемый этими правами доступ к объектам. В случае если доступ к объектам все же требуется, как в вышеописанном примере для запроса данных, можно запретить только интерактивный доступ.

Проиллюстрируем вышеприведенные правила на примере разрабатываемой подсистемы. Администратор подсистемы должен иметь доступ ко всему имеющемуся функционалу для администрирования подсистемы, выявлению и исправлению ошибок в подсистеме, а так же, в качестве разработчика, возможность дополнять и изменять функционал системы. Поэтому его роли принадлежат права на совершение любых действий над объектами подсистемы. Администратору баз данных, который, в случае данной подсистемы, занимается администрированием морфологической и синтаксической БД, т.е. исправлением, добавлением и пометкой на удаление новых данных. Следовательно, его роли принадлежат права на чтение, запись и установка пометки на удаление. Непосредственным удалением помеченных данных, в целях безопасности, занимается администратор, имеющий на это полные права. И, наконец, пользователь подсистемы, роли которого принадлежат права только на ввод входных данных и интерактивный запуск анализа текста. Доступ к БД ограничивается лишь правом на чтение определенных данных из нее – слов и результатов анализа.

Отметим, что список прав не несет никакой информации по разрешению, либо запрещению того или иного способа доступа – эта информация хранится в ролях.

Роли являются объектами метаданных и содержат установки прав всех объектов метаданных в конфигурации – каждое право (как контролируемый способ доступа) может быть разрешено или запрещено. Как объекты метаданных роли редактируются в Конфигураторе и позволяют установить нужные значения для любых прав, но с учетом зависимостей между правами – напри-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		55

мер, запрещение права на чтение справочника установит в «запрещено» и все остальные права. В предметной области роли обычно (но не обязательно) соответствуют должностям, поэтому установки прав в них соответствуют служебным обязанностям сотрудников, занимающих эти должности. По умолчанию, все права объекта метаданных (как при создании нового объекта, так и при создании роли) устанавливаются в «разрешено». Как права, так и роли относятся к предметной области решений на основе 1С:Предприятия 8 – такой подход позволяет разрабатывать «коробочные» решения, не ориентируясь на конкретные внедрения. В отличие от прав, определяющих доступ к отдельным объектам конфигурации, роли определяют доступ к решению в целом, что позволяет управлять разграничением доступа на уровне администрирования системы, не вдаваясь в детали устройства решения.

Как установка прав, так и роли могут проверяться во встроенном языке 1С:Предприятия 8 – для этого служат соответственно функции ПравоДоступа() и РольДоступна(). Списки прав для объектов фиксированы и не изменяются, поэтому для проверки возможности некоторого специфического способа доступа, например, возможности интерактивного изменения результатов анализа до занесения их в БД можно использовать два различных подхода:

– проверять доступность уже существующей роли например, «Пользователь»;

– создать отдельную роль и затем проверять во встроенном языке доступность этой роли при доступе к объекту, например, «Синтаксические Конструкции».

При внедрении этих решений в систему добавляются пользователи, доступ которых к объектам системы и требуется разграничить. Для этого каждому пользователю присваивается одна или несколько ролей в соответствии с их служебными обязанностями, к примеру, Администратор БД исполняет также обязанности Пользователя. При запросе пользователем некоторого доступа к некоторому объекту системы, для каждой роли, присвоенной пользователю, по

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		56

объекту определяется установка соответствующего права. Если хотя бы для одной роли это право установлено, то пользователю дается разрешение на указанный доступ к указанному объекту. В противном случае доступ запрещается. Такой способ «сложения» установок прав в разных ролях позволяет упростить добавление прав при адаптации готового решения к специфике внедрения – для добавления пользователю права на доступ достаточно создать новую роль и в ней установить необходимое право в «разрешено». При этом уже существующие объекты конфигурации не изменяются, что упрощает ее дальнейшую поддержку разработчиком.

В приложении Г описывается метод ролевого разграничения доступа. На рисунках Г.1 и Г.2 приведен список прав для ролей «Администратор» и «Пользователь», соответственно.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		57

5 БЕЗОПАСНОСТЬ И ЭКОЛОГИЧНОСТЬ

В данном разделе рассматриваются правила и методы обеспечения безопасности жизнедеятельности пользователей, работающих с разработанной подсистемой. Опираясь на нормативный сборник основных аспектов БЖД – СанПиН 2.2.2/2.4.1340-03, будет проведен анализ следующих 3 положений: безопасность, экологичность и защита от ЧС.

5.1 Требования безопасности, предъявляемые к работе с ЭВМ

5.1.1 Эргономические требования

Организация рабочего места должна учитывать антропометрические, сенсорные, биомеханические и психофизиологические характеристики человека.

Для снижения утомления и напряжения рук оператора ПК в процессе работы выполняются следующие требования по планировке рабочего места:

– клавиатура располагается так, чтобы руки оператора ПК были направлены вниз, т. е. уровень клавиатуры должен быть ниже уровня локтевого сустава;

– на кресле предусматриваются подлокотники, чтобы создать опору для локтевых суставов;

– компьютерная мышь располагается на одном уровне с клавиатурой, чтобы руки в процессе работы находились на одном уровне, что избавит от усталости в плечах;

– у клавиатуры предусматриваются специальные упоры для поддержания кистей рук.

В соответствии с санитарными нормами на одно рабочее место, оснащенное компьютером, предусматривается не менее 4,5 кв. м.

Оптимальное размещение компьютерной аппаратуры и документации в зонах досягаемости:

– дисплей размещается в зоне а (в центре);

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		58

- системный блок размещается в предусмотренной нише стола;
- клавиатура размещается в зоне г/д;
- «мышь» размещается в зоне в справа;
- документация, необходимая при работе должна располагаться в зоне легкой досягаемости ладони, а литература, неиспользуемая постоянно – в выдвижных ящиках стола.

Помещения должны быть оборудованы защитным заземлением в соответствии с техническими требованиями по эксплуатации.

Не следует размещать рабочее место с ЭВМ вблизи силовых кабелей и вводов, высоковольтных трансформаторов, технологического оборудования, создающего помехи в работе с ЭВМ.

Рассмотрим эргономические требования к организации рабочего места.

Первый и самый главный фактор - это нагрузка на зрение. Именно из-за нагрузки на зрение через непродолжительное время у ребенка (или другого пользователя) возникает головная боль и головокружение. Если работать на компьютере достаточно долго, то зрительное переутомление может привести к устойчивому снижению остроты зрения. Однако заметим сразу, не компьютер является основной причиной развития близорукости. Огромную роль в этом играет наследственность, телевизор, чтение в темноте. При грамотной постановке дела нагрузка на зрение от компьютера может быть существенно снижена.

Во-первых, вредны подергивания изображения из-за низкой частоты вертикального обновления или из-за низкого качества развертки монитора.

Во-вторых, один из важных факторов утомляемости глаз – это содержание пользовательского интерфейса.

Интерфейс подсистемы должен быть понятным и удобным, не должен быть перегружен графическими элементами и должен обеспечивать быстрое отображение экранных форм.

Расположение управляющих элементов, таких как кнопки, ссылки и интерактивные формы ввода должно быть согласовано и последовательно. Так

логично располагать кнопку отправки формы внизу под формами ввода, поскольку, заполнив их, пользователь окажется именно там. Кнопки управления модульным содержимым, таким как сворачивающиеся блоки, не должны пропадать при их динамическом изменении, должны оставаться на одном месте, чтобы пользователь интуитивно знал, куда обратить свое внимание.

Информационные сообщения и сигналы должны акцентировать внимание пользователя, чтобы избежать вероятности ошибки при работе с подсистемой. Все сообщения располагаются в одной области и выделяются цветом, соответствующему смысловому содержанию.

Интерфейс, изображенный на рисунке 3.9, соответствует вышеприведенным нормам, а так же является довольно простым для восприятия, так как функционал разработанной подсистемы не предусматривает серьезного интерактивного взаимодействия с пользователем.

Второй по важности фактор – это психическая нагрузка. Компьютер требует не меньшей сосредоточенности, чем вождение автомобиля. В случае разработанной подсистемы синтаксического анализа текстов не требуется уделять много времени работе с ней, т.к. большинство функций выполняется в автоматическом режиме и с высокой скоростью, что позволяет получить результат за минимальное количество времени. По причине направленности работы подсистемы лишь на ее самообучение посредством синтаксического анализа текстовой информации так же не требуется освоения пользователем каких-либо профессиональных навыков для работы с подсистемой. Весь интерфейс подсистемы прост и интуитивно понятен.

5.1.2 Требования электробезопасности

Вся вычислительная техника и многие другие виды оборудование представляют для человека с точки зрения поражения электротоком потенциальной опасности, так как в процессе эксплуатации или при проведении профилактических работ человек может коснуться частей оборудования, находящихся под напряжением.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		60

Перед началом работы следует убедиться в отсутствии свешивающихся со стола или висящих под столом проводов электропитания, в целостности вилки и провода электропитания, в отсутствии видимых повреждений аппаратуры и рабочей мебели, в отсутствии повреждений и наличии заземления приэкраного фильтра.

Помещение, в котором находится рабочее место, относится к помещениям без повышенной опасности. В нормальных условиях электрические сети достаточно безопасны, так как они надежно изолированы, и окружающая среда практически не разрушает изоляцию. В помещении должны быть выполнены следующие мероприятия по защите людей от поражения электрическим током:

- защитное заземление – заземляются системный блок и защитный экран дисплея;

- части оборудования и проводка, находящиеся под напряжением, надежно защищены и расположены в недоступных местах;

- защитное зануление.

Благодаря заземлению и почти полному отсутствию открытых металлических частей на современной компьютерной технике, опасность возникновения зарядов статического электричества очень мала.

5.2 Экологичность

ФЗ № 89 от 24.06.1998 г. является основным в вопросах регулирования обращения с отходами производства и потребления с целью предотвращения вредного воздействия отходов на здоровье человека и окружающую среду.

При утилизации ЭВМ владелец обязан обратиться в организацию, имеющую лицензию на непосредственную утилизацию такого рода отходов.

Утилизация информации на носителях и компонентах ЭВМ может производиться, как с помощью блокирования доступа к информации на носителях, ее затиранию, а также дополнительным включением маскирующей информации.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		61

Наличие в компонентах ЭВМ драгоценных металлов накладывает на организацию определенные требования и обязанности по утилизации данных компонентов.

Эти требования регулируются законодательством в соответствии с ФЗ № 41. Несоблюдение вышеописанных требований может повлечь за собой административную ответственность для владельца ЭВМ.

5.3 Требования к обеспечению пожарной безопасности при работе с ЭВМ

Пожарная безопасность – состояние объекта, при котором исключается возможность пожара, а в случае его возникновения предотвращается воздействие на людей опасных его факторов и обеспечивается защита материальных ценностей.

Для ликвидации пожара в помещении, где расположено рабочее место, должны применяться первичные средства пожаротушения.

Первичные средства пожаротушения – это устройства, инструменты и материалы, предназначенные для локализации и ликвидации загорания на начальной стадии. К ним относятся огнетушители, внутренний пожарный кран, вода, песок, кошма, асбестовое полотно, ведро, лопата и др.

Огнетушители подразделяются на следующие виды конструкций:

- порошковые (ПСБ, ПФ, ОП);
- пенные (ОХП- 10);
- углекислотные (ОУ-2, ОУ-5).

Все огнетушители обязательно должны периодически проходить проверку на исправность и отсутствие дефектов, препятствующих их корректному и полному применению. Так же средства пожаротушения обязательно должны располагаться в максимально доступном месте, для своевременного их применения в случае возникновения пожароопасной ситуации

Хранение технических средств должно осуществляться в закрытых контейнерах для предотвращения накопления пыли в их составных частях.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		62

При эксплуатации ЭВМ и оргтехники необходимо проверять целостность токопроводящих кабелей, вилки и розетки, отсутствие повреждений аппаратуры.

Компоненты ЭВМ должны иметь функцию самоотключения при повышении температуры входе неисправности систем охлаждения и кондиционирования. Для предотвращения перегрева.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		63

ЗАКЛЮЧЕНИЕ

В ходе выполнения выпускной квалификационной работы была проанализирована предметная область экспертных систем и выявлена проблематика пополнения входящих в их структуру баз знаний. В результате было выбрано одно из возможных решений данной проблемы в виде создания механизмов автоматизированной обработки естественного языка, на основе алгоритмов которых была разработана подсистема синтаксического анализа текстовой информации, выполняющая два начальных этапа этой обработки – морфологический и синтаксический анализ текстов на русском языке и записывающая результаты, последнего, в синтаксическую БД.

Разработка подсистемы включала в себя этапы проектирования и, затем, физической реализации всех ее функциональных компонентов, начиная от баз данных и заканчивая синтаксическим анализатором.

Согласно классификации и определению экспертных систем, представленных в пункте 1.1.1, а именно наличию базы знаний, включающей в себя до 1000 простых синтаксических правил и правил вывода результатов анализа, разработанная подсистема может использоваться как простая экспертная система, имитирующей действия человека-лингвиста при синтаксическом разборе текста.

В своем начальном этапе подсистема содержит морфологическую БД, включающую более 4 млн. слов русского языка, на основе которой уже способна выполнять синтаксический анализ огромного количества текстов, представленных в виде простых предложений русского языка, и запоминать результаты этого анализа посредством их записи в синтаксическую БД.

На данном этапе подсистема находится на завершающей стадии доработки и отладки, с включением анализа сложных предложений и абзацев. Так же, для полноценной автоматизации работы подсистемы, в будущем, планируется разработать механизм автоматического пополнения морфологической БД, используя, например, интернет-словари.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		64

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1 Адаменко, А. Н. Логическое программирование и Visual Prolog / Адаменко, А. Н., Кучуков, А. М. – СПб.: БХВ-Петербург, 2003. – 992 с.

2 Афонин, В.Л. Интеллектуальные робототехнические системы / Афонин В.Л., Макушкин В.А. – М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – 222 с.

3 Беляков, Г.И. Безопасность жизнедеятельности. Охрана труда: Учебник для бакалавров / Г.И. Беляков. – М.: Юрайт, 2013. – 572 с.

4 Википедия [Электронный ресурс] : офиц. сайт. – 15.01.2001 – Режим доступа : https://ru.wikipedia.org/wiki/IBM_Watson. – 05.06.2017.

5 Википедия [Электронный ресурс] : офиц. сайт. – 15.01.2001 – Режим доступа : https://ru.wikipedia.org/wiki/Синтаксический_анализ. – 25.05.2017.

6 Википедия [Электронный ресурс] : офиц. сайт. – 15.01.2001 – Режим доступа : https://ru.wikipedia.org/wiki/Управление_доступом_на_основе_ролей. – 02.06.2017.

7 Википедия [Электронный ресурс] : офиц. сайт. – 15.01.2001 – Режим доступа : https://ru.wikipedia.org/wiki/Поиск_в_глубину. – 15.05.2017.

8 Жукова, Т.Е. Русский язык. Морфология и синтаксис: практикум для студентов технических университетов и иностранных студентов, изучающих русский язык на продвинутом этапе / Т.Е. Жукова, Н.Г. Блохина, И.С. Иванова – Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2014. – 164 с.

9 Интеллектуальные информационные системы и технологии: учебное пособие / Ю.Ю. Громов [и др.]. – Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2013. – 244 с.

10 Качановский, Ю.П. Основные технические, программные и организационные меры защиты информации при работе с компьютерными системами: методические указания к проведению лабораторной работы по курсу «Инфор-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		65

матика» / Ю.П. Качановский, А.С. Широков – Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2014. – 24 с.

11 Громов, Ю.Ю. Представление знаний в информационных системах: учебное пособие / Ю.Ю. Громов. – Тамбов: Тамбовский государственный технический университет, ЭБС АСВ, 2012. – 169 с.

12 Профессиональная разработка в системе «1С:Предприятие 8»: в 2 т. / 2-е изд. – М.: 1С-Паблишинг, 2012 – 5000 экз. – Т. 1. – 690 с.

13 Профессиональная разработка в системе «1С:Предприятие 8»: в 2 т. / 2-е изд. – М.: 1С-Паблишинг, 2012 – 5000 экз. – Т. 2. – 683 с.

14 Радченко, М.Г. 1С:Предприятие 8.3 Практическое пособие разработчика. Примеры и типовые приемы / М.Г. Радченко, Е.Ю. Хрусталева. – М.: ООО «1С-Паблишинг», 2013. – 874 с.

15 Рыженкова, Т.В. Синтаксис словосочетания и простого предложения: учебное пособие для студентов высших учебных заведений, обучающихся по направлению подготовки бакалавра «Педагогическое образование» / Т.В. Рыженкова – М.: Прометей, 2014. – 80 с.

16 СанПиН 2.2.2/2.4.1340-03 [Электронный ресурс]. – Режим доступа: <http://docs.cntd.ru/document/901865498>. – 25.05.2017

17 Симакова, Н.Н. Организация рабочих мест с персональными электронно-вычислительными машинами (ПЭВМ) [Электронный ресурс]: учебное пособие / Н.Н. Симакова – Электрон. текстовые данные. – Новосибирск: Сибирский государственный университет телекоммуникаций и информатики, 2013. – 78 с. – Режим доступа: <http://www.iprbookshop.ru/55486.html>. – ЭБС «IPRbooks»

18 Советов, Б.Я. Представление знаний в информационных системах: учебник для студ. высш. учеб. заведений / Б.Я. Советов, В.В. Цехановский, В.Д. Чертовской. – М. : Издательский центр «Академия», 2011. – 144 с.

19 Шаньгин, В.Ф. Информационная безопасность и защита информации [Электронный ресурс] / В.Ф. Шаньгин – Электрон. текстовые данные. – Сара-

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		66

тов: Профобразование, 2017. – 702 с. – Режим доступа:
<http://www.iprbookshop.ru/63594.html>. – ЭБС «IPRbooks»

20 Шрайнер, П.А. Основы программирования на языке Пролог [Электронный ресурс] / Шрайнер П.А. – М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – 213 с. – Режим доступа:
<http://www.iprbookshop.ru/52194.html>.— ЭБС «IPRbooks»

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		67

ПРИЛОЖЕНИЕ А

Технологическая сеть проектирования

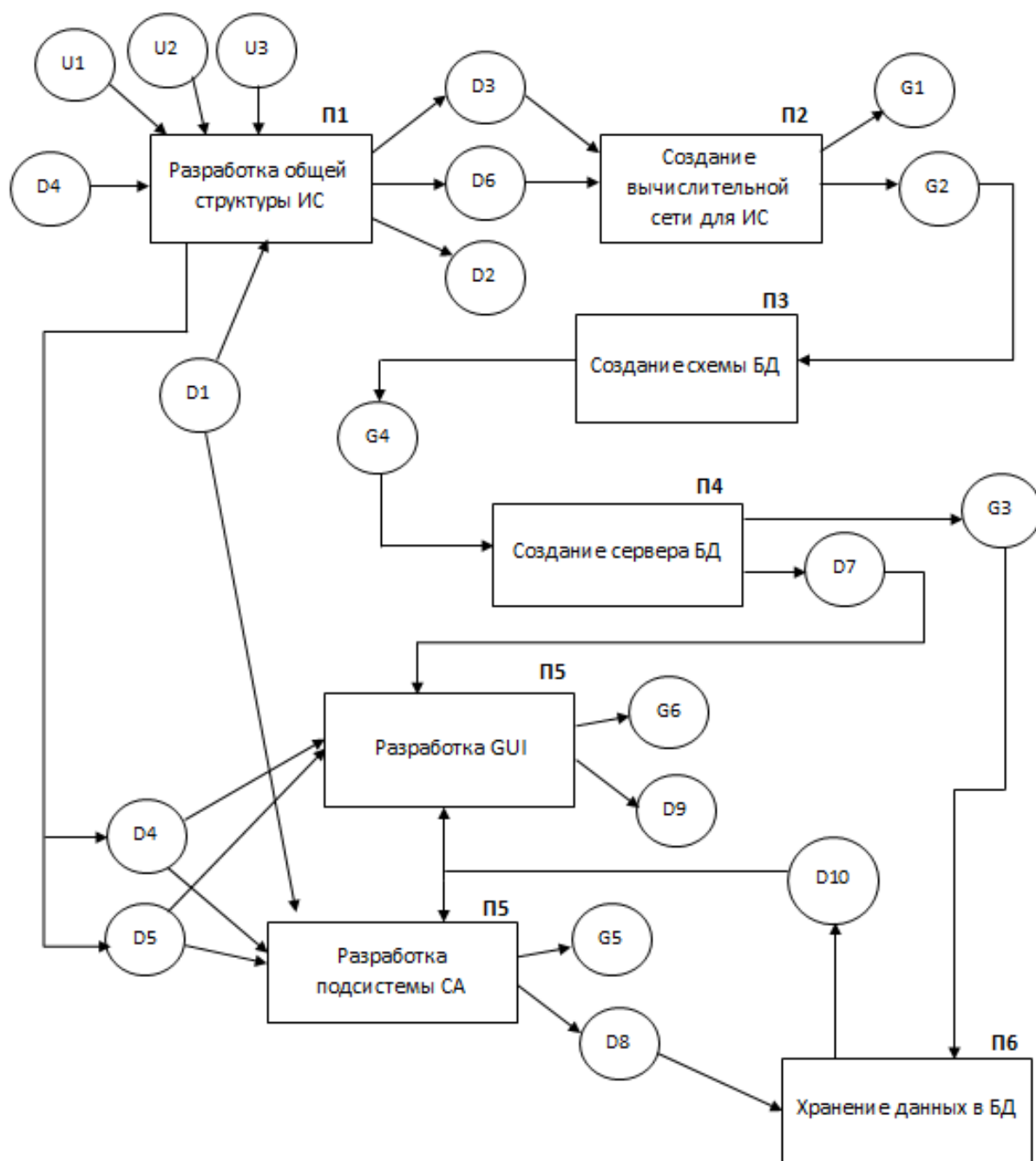


Рисунок А.1 – Технологическая сеть проектирования

УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

- D1- описание предметной области;
- D2- описание выбранного сервера БД;

Продолжение ПРИЛОЖЕНИЯ А

- D3- описание выбранной конфигурации технических средств и сетевой ОС;
- D4- техническое задание;
- D5- описание выбранных программных средств разработки ИС;
- D6- функциональные структуры ИС;
- D7- права доступа к различным категориям доступа;
- D8- проанализированные и преобразованные данные;
- D9- сопровождающая документация;
- D10- хранимые данные;
- U1- универсум сет. ОС и технических платформ;
- U2- универсум серверов;
- U3- программные средства разработки ИС;
- G1- вычислительная сеть;
- G2- СУБД;
- G3 - созданная БД;
- G4 - SQL описание БД с управляющими элементами;
- G5- подсистема синтаксического анализа;
- G6- графический интерфейс.

					<i>ВКР.135194.09.03.02.ПЗ</i>	<i>Лист</i>
<i>Изм.</i>	<i>Лист</i>	<i>№ докум.</i>	<i>Подп.</i>	<i>Дата</i>		69

ПРИЛОЖЕНИЕ Б

Логическая модель БД

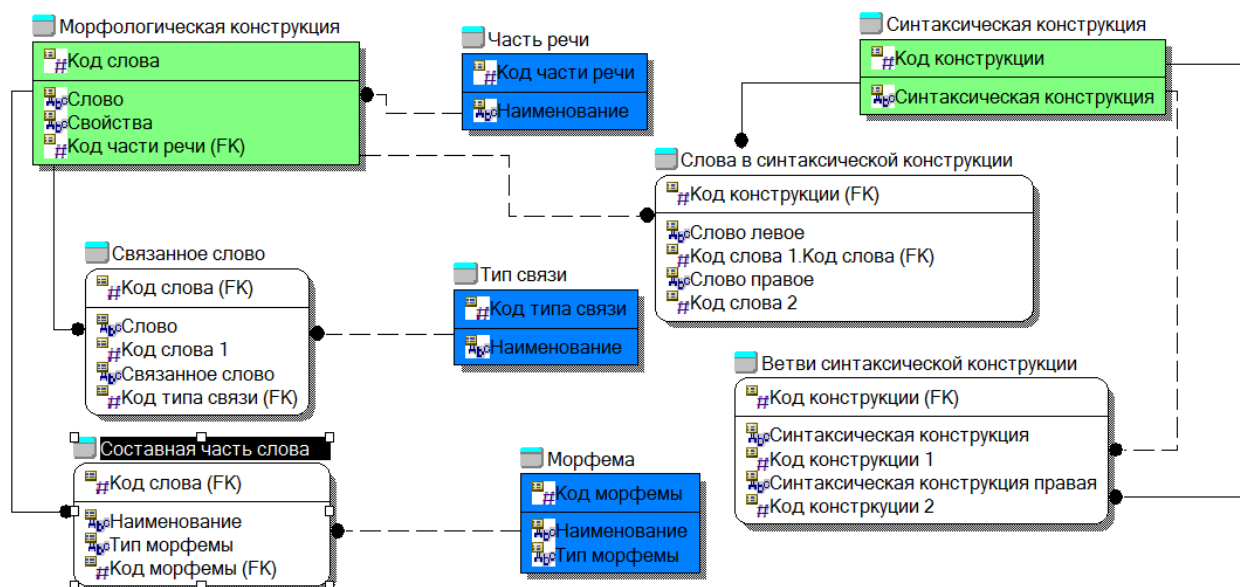


Рисунок Б.1 – Логическая модель БД

Изм.	Лист	№ докум.	Подп.	Дата

ВКР.135194.09.03.02.ПЗ

Лист

70

ПРИЛОЖЕНИЕ В

Физическая модель БД

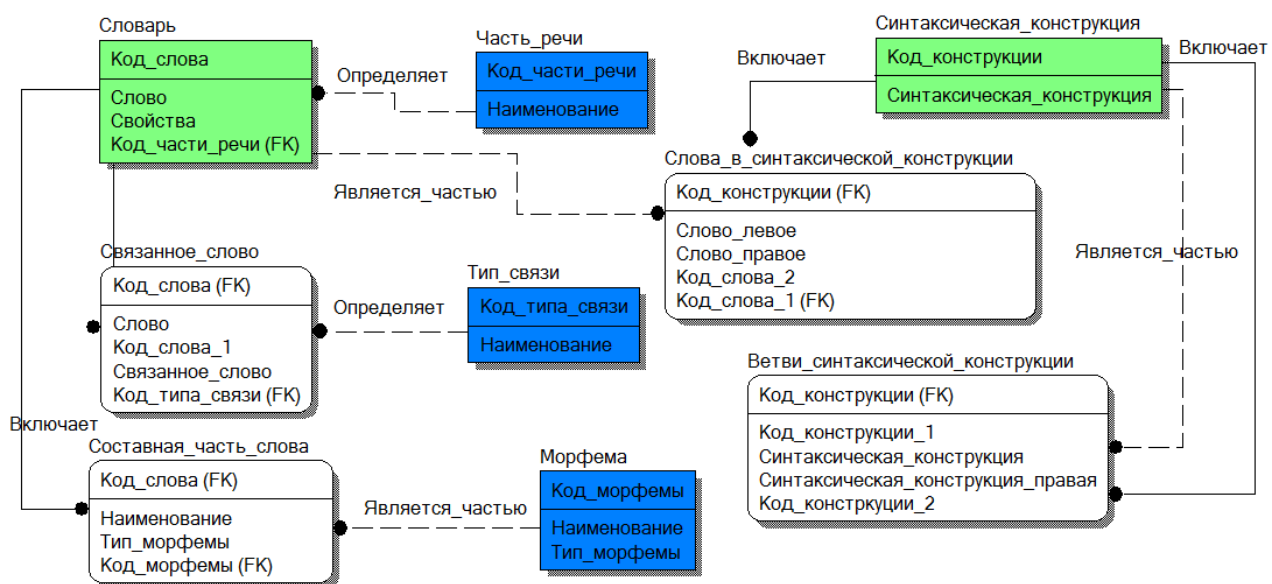


Рисунок В.1 – Физическая модель БД

ПРИЛОЖЕНИЕ Г

Ролевое разграничение доступа

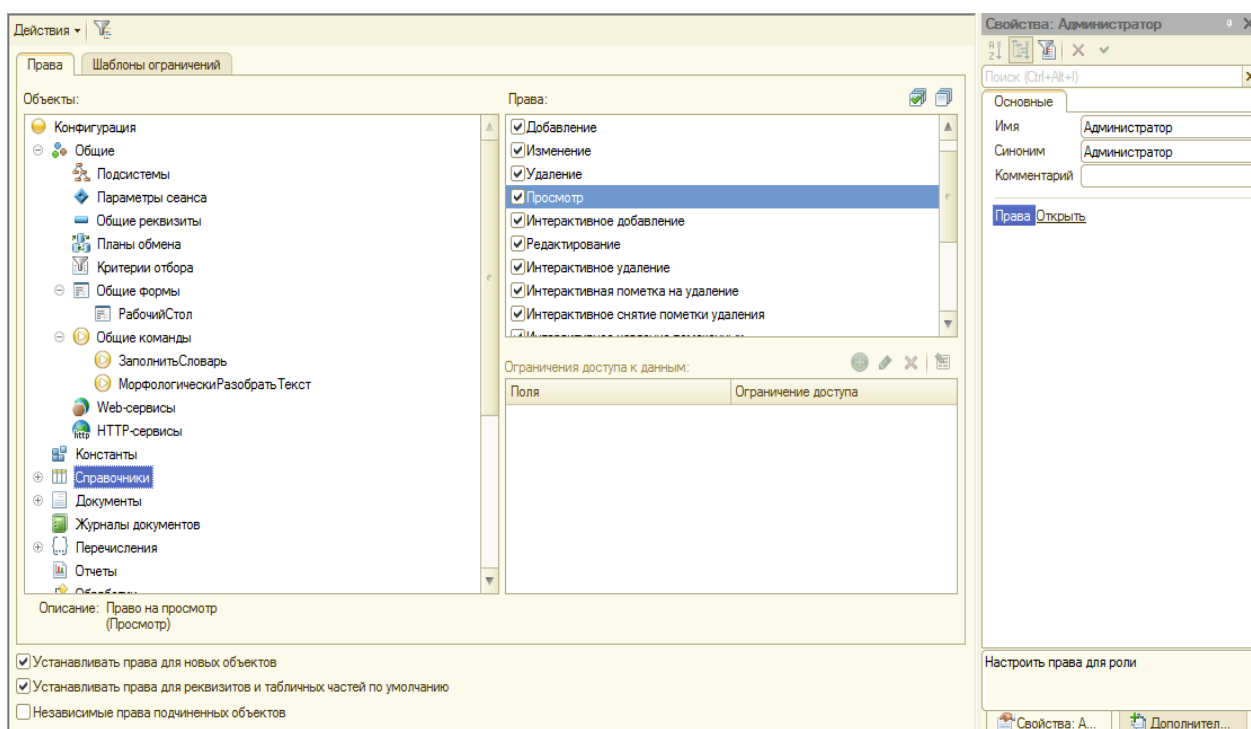


Рисунок Г.1 – Список прав для роли «Администратор»

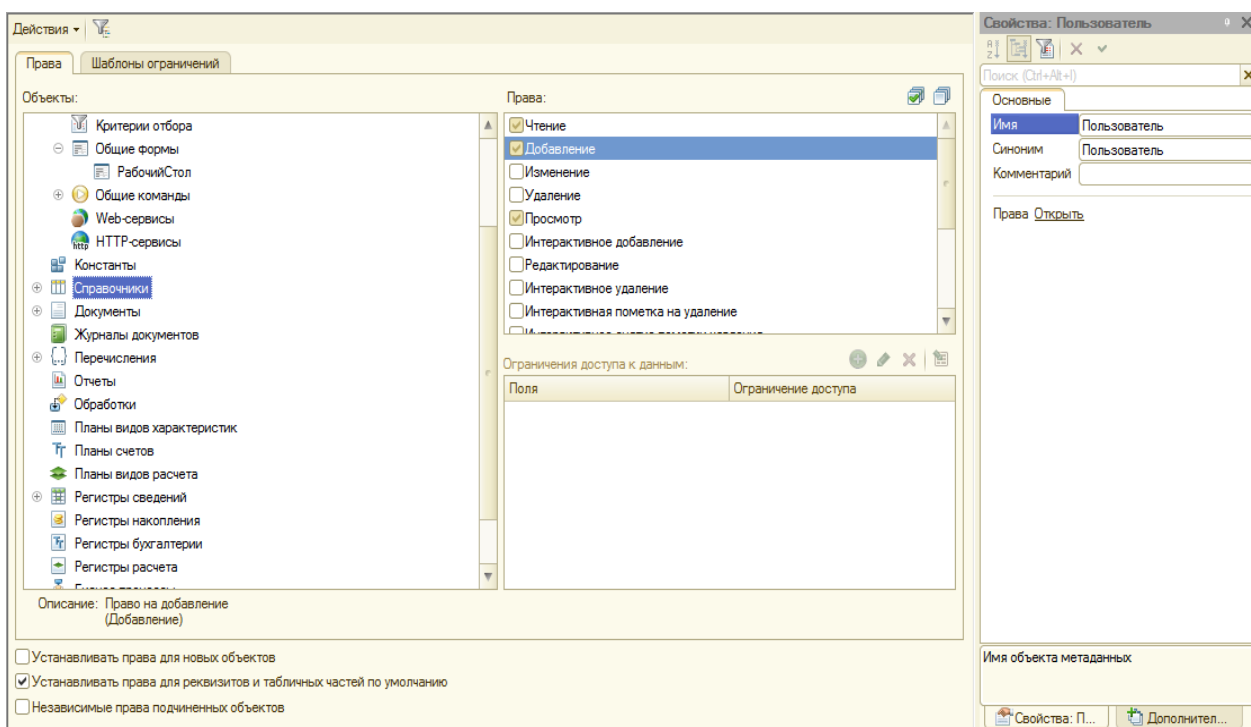


Рисунок Г.2 – Список прав для роли «Пользователь»

ПРИЛОЖЕНИЕ Д

Представление синтаксического дерева в БД

Студент защищает довольно сложный диплом (Синтаксические констр...

Записать и закрыть Записать Еще ▾

Код:

Наименование:

Дерево:

Синтаксические конструкции Слова

Добавить Еще ▾

N	Слово
1	студент

Рисунок Д.1 – Левая ветвь конструкции 1-го уровня

Студент защищает довольно сложный диплом (Синтаксические констр...

Записать и закрыть Записать Еще ▾

Код:

Наименование:

Дерево:

Синтаксические конструкции Слова

Добавить Еще ▾

N	Синтаксическая конструкция
1	защищает довольно сложный диплом

Рисунок Д.2 – Правая ветвь конструкции 1-го уровня

Продолжение ПРИЛОЖЕНИЯ Д

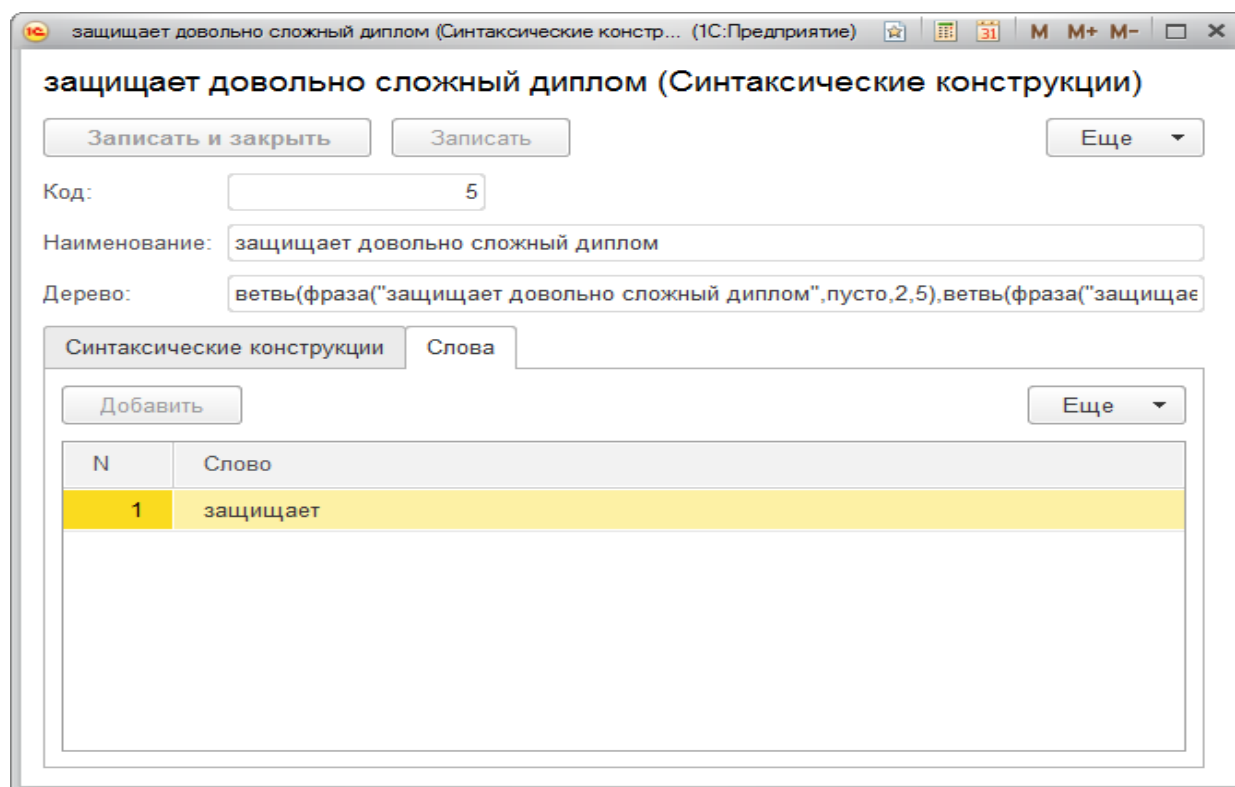


Рисунок Д.3 – Левая ветвь конструкции 2-го уровня

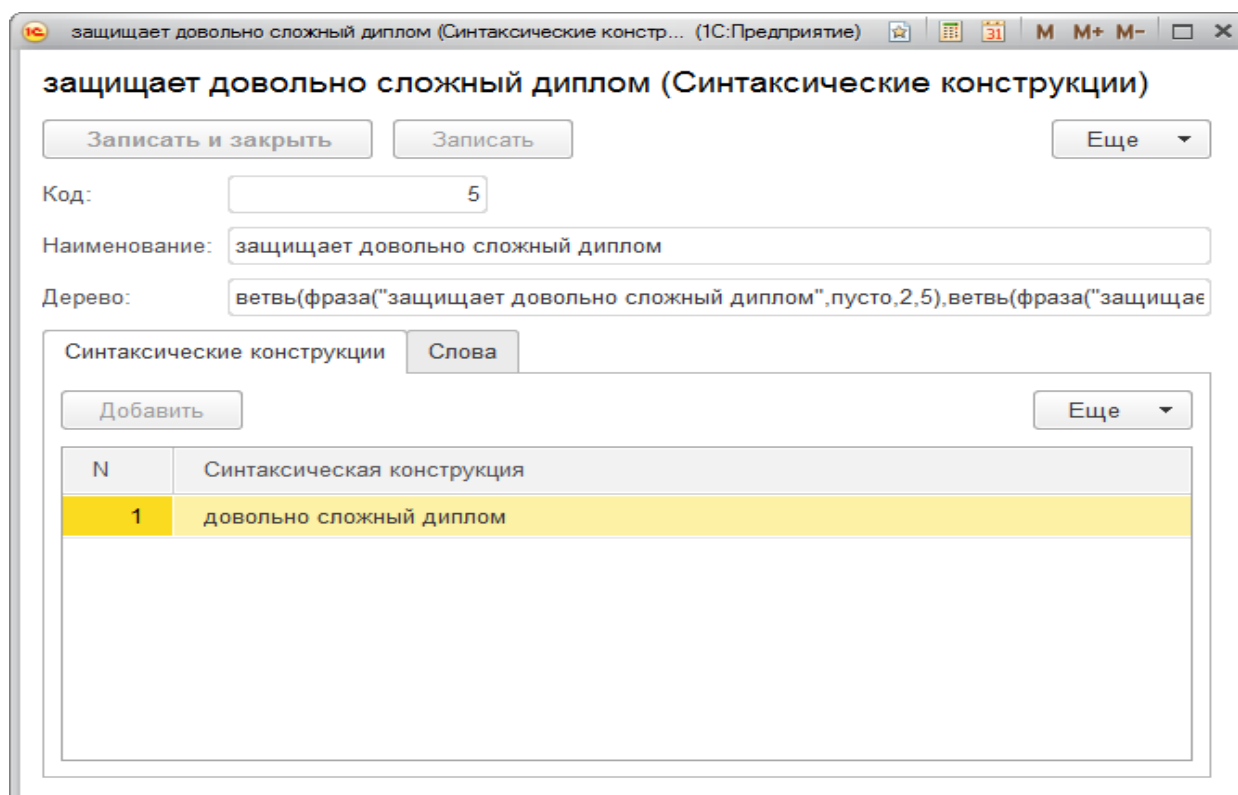


Рисунок Д.4 – Правая ветвь конструкции 2-го уровня

Продолжение ПРИЛОЖЕНИЯ Д

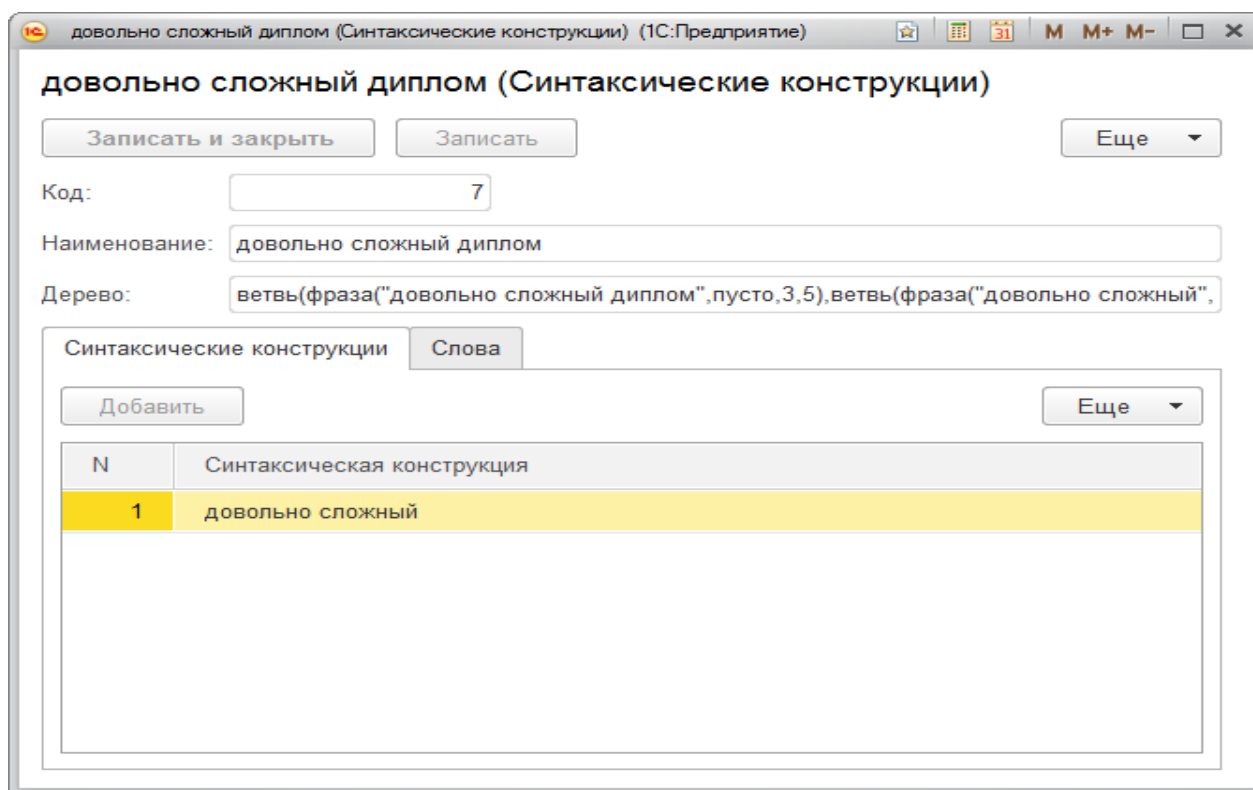


Рисунок Д.5 – Левая ветвь конструкции 3-го уровня

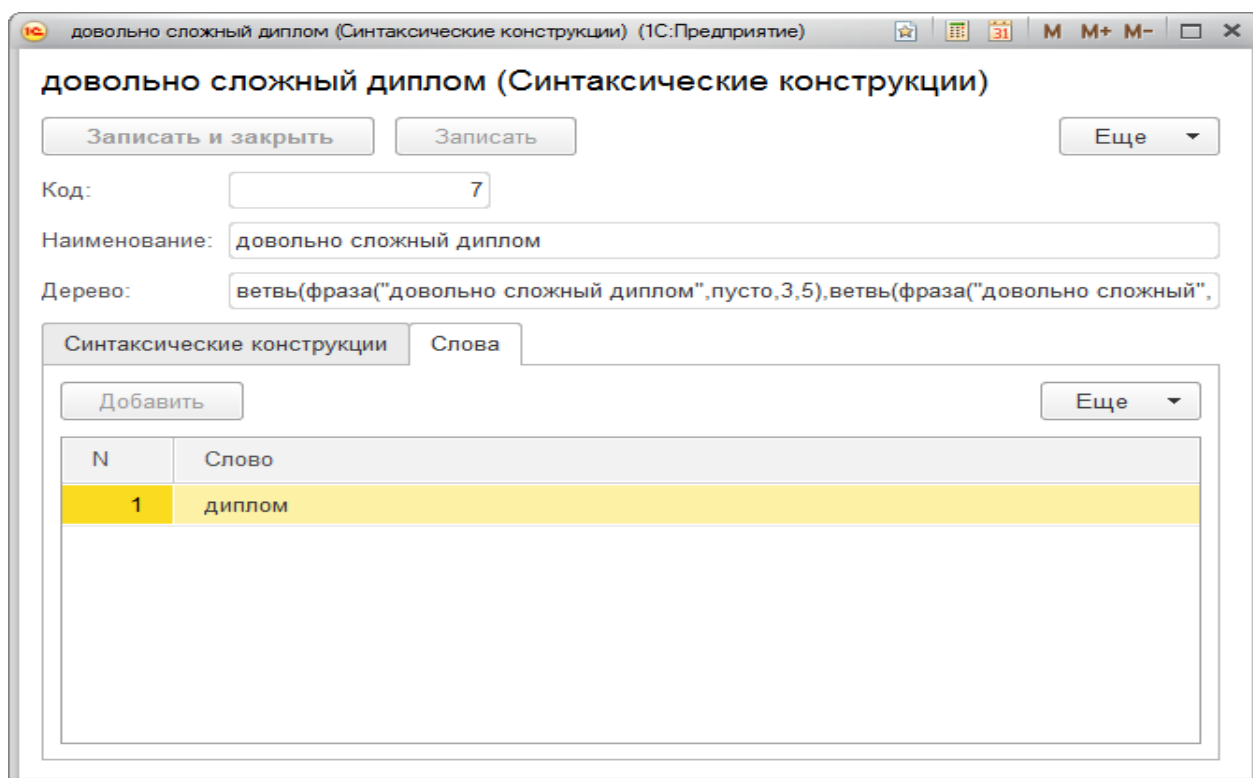


Рисунок Д.6 – Правая ветвь конструкции 3-го уровня

Продолжение ПРИЛОЖЕНИЯ Д

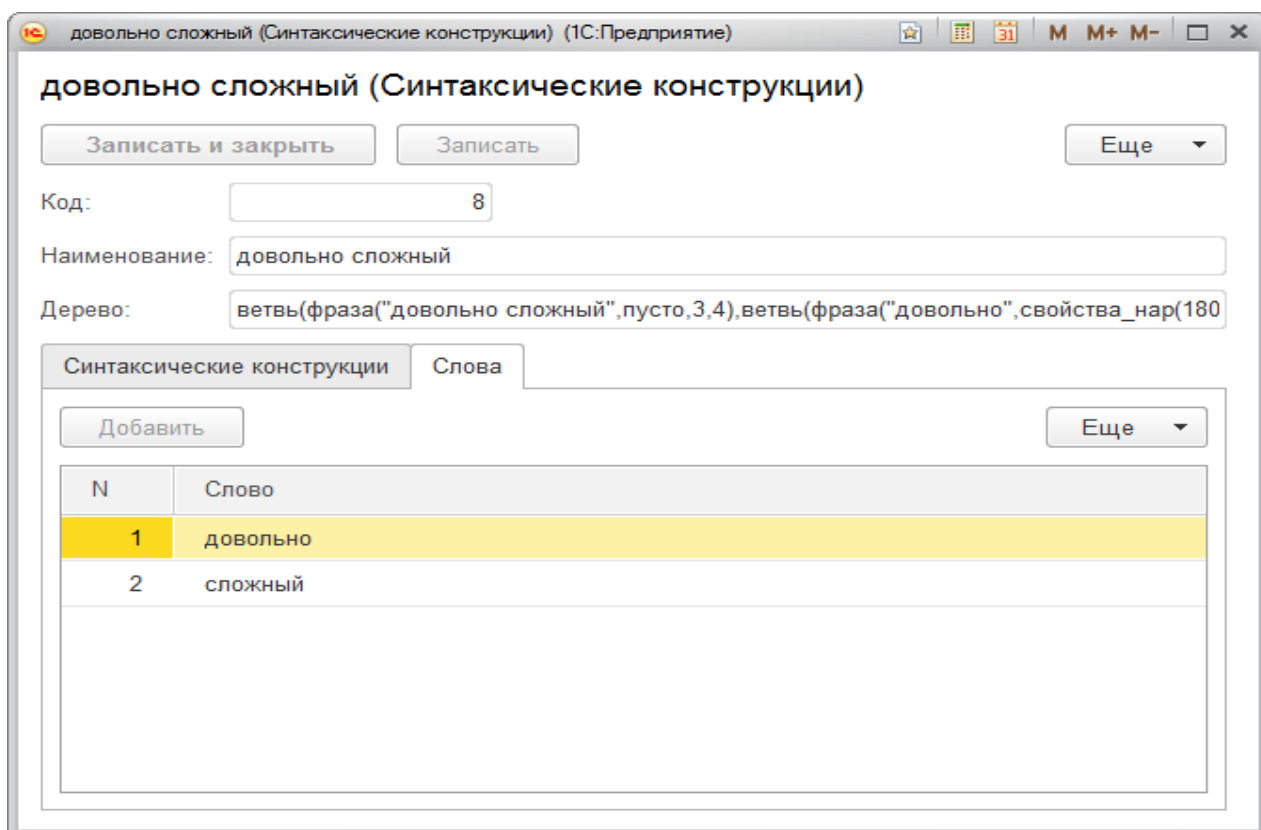


Рисунок Д.7 – Левая и правая ветви конструкции 4-го уровня

ПРИЛОЖЕНИЕ Е

Диаграмма процессов в подсистеме синтаксического анализа текстовой информации

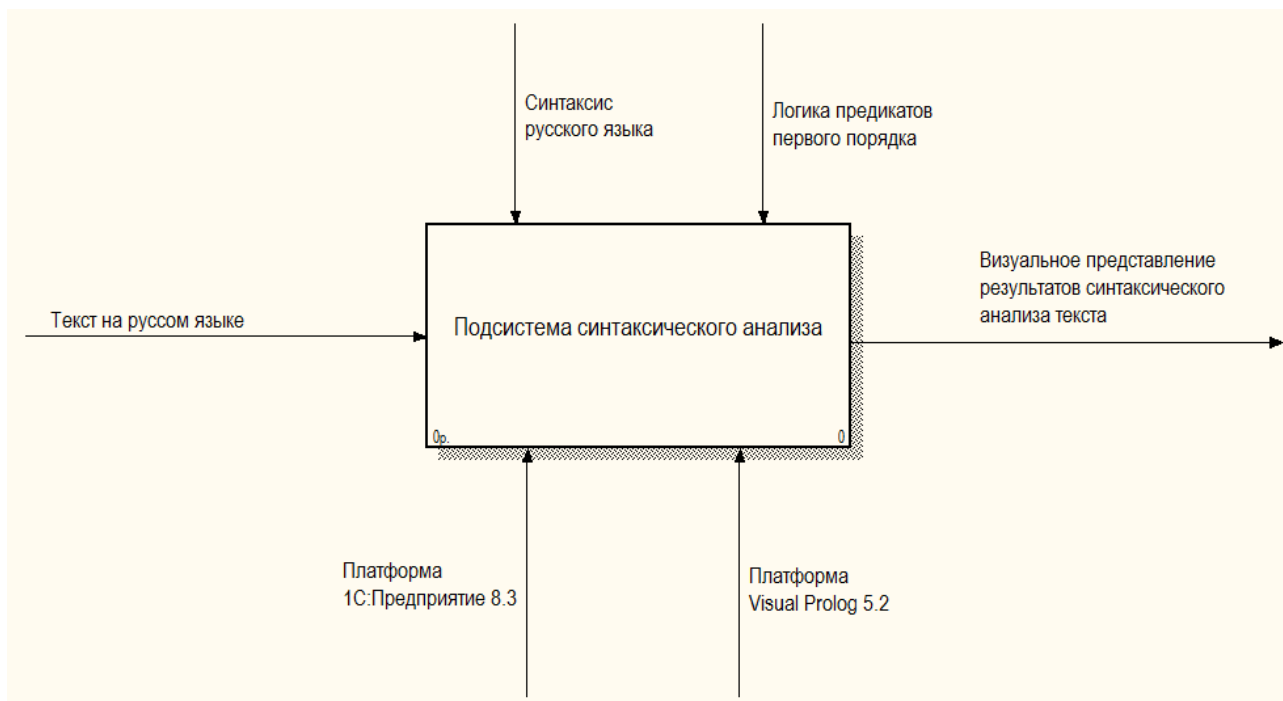


Рисунок Е.1 – Представление входных и выходных данных в подсистеме синтаксического анализа текстовой информации

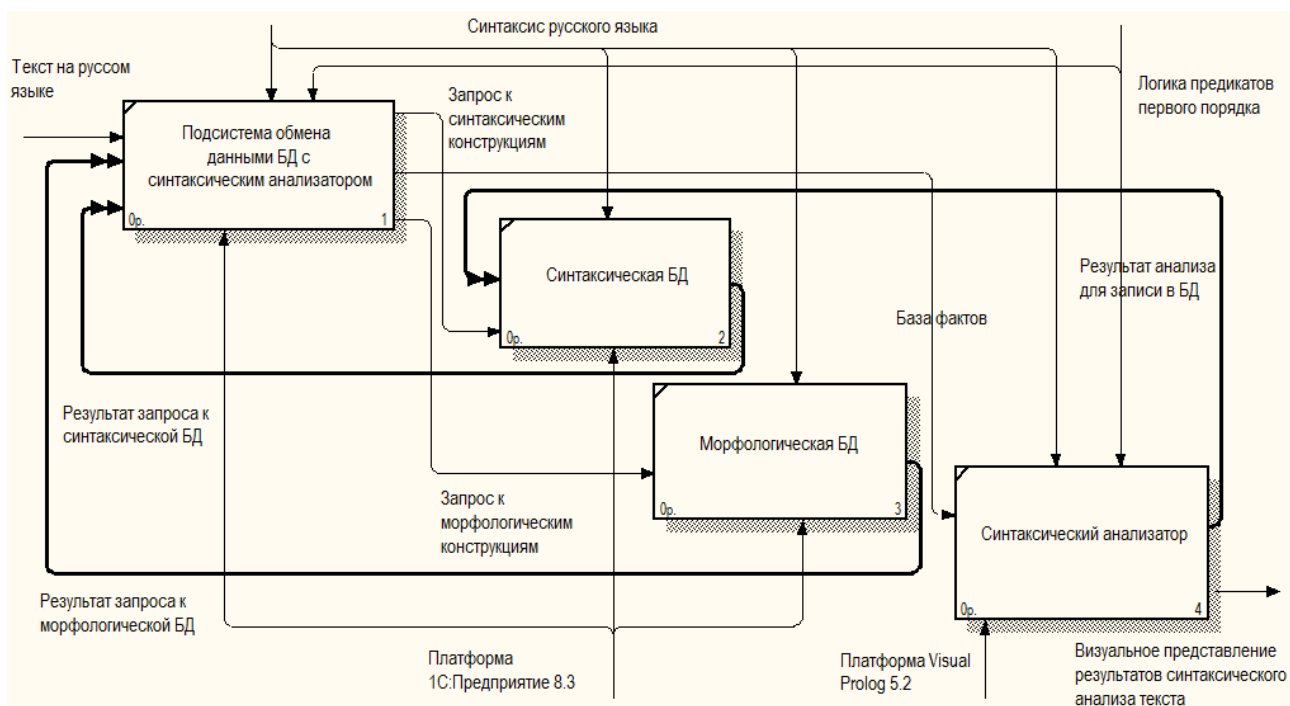


Рисунок Е.2 – Декомпозиция подсистемы на ее функциональные компоненты