

**Московский Государственный Технический Университет им. Н.Э. Баумана
Кафедра «Системы обработки информации и управления»**

**Расчетно-пояснительная записка
к дипломному проекту**

Лингвокультурный тезаурус русского языка

Студент-дипломник: Сиренко А.В.

Москва, 2007 г.

Реферат

Расчетно-пояснительная записка выполнена:

Страниц 120, рисунков 44, таблиц 44.

Дипломный проект посвящен разработке автоматизированной системы «Лингвокультурный тезаурус русского языка», которая предназначена работы с базой данных фигур знания и ассоциативной сети, а также моделирования пассивного режима когнайзера.

Расчетно-пояснительная записка содержит четыре части:

Анализ предметной области и постановка задачи:

- Лингвокультурный тезаурус;
- Активный и пассивный режимы когнайзера;
- Постановка задачи информационной технологии;

Методы и алгоритмы лингвокультурного тезауруса. В данной части расчетно-пояснительной записки были решены следующие задачи:

- Алгоритм поиска путей ассоциативной сети;
- Алгоритмы обращения к базе данных;
- Построение запроса выборки данных.

Реализация информационной технологии. В данной части расчетно-пояснительной записки были решены следующие задачи:

- Архитектура программного изделия;
- Выбор среды разработки программного изделия;
- Проектирование базы данных.
- Оптимизация базы данных
- Выбор способа доступа к базе данных
- Предварительная обработка материалов
- Реализация работы когнайзера

Моделирование и экспериментальные исследования. В данной части расчетно-пояснительной записки рассмотрены:

- Редактирование материалов ЛКТ;
- Выборка материалов по критерию;

- Моделирование работы когнитивера;

Эргономика и охрана труда. Данная часть посвящена:

- Характеристике влияния условий и режима труда на утомляемость и снижение производительности труда;
- Характеристика условий труда;
- Эргономические требования к рабочему месту;
- Психологические требования к рабочему месту;
- Расчет освещенности рабочего места;
- Схема рабочего места;

Организационно-экономическая часть. В данной части расчетно-пояснительной записки рассчитаны:

- Стадии и этапы разработки;
- Затраты на разработку системы;
- Обоснование разработки программного изделия;

Содержание

Нормативные ссылки	6
Определения, обозначения, сокращения	7
Введение	8
1. Анализ предметной области и постановка задачи.	9
1.1 Лингвокультурный тезаурус	9
1.2 Обзор аналогов и прототипов	10
RCO Semantic Network SDK	10
www.lexfn.com	11
Microsoft Excel	13
Заключение	14
1.3 Активный и пассивный режимы работы когнайзера	15
1.4 Постановка задачи информационной технологии	18
1.4.1 Назначение программного изделия	18
1.4.2 Функции программного изделия	18
1.4.3 Решаемые задачи	19
2. Методы и алгоритмы лингвокультурного тезауруса.	20
2.1. Алгоритм поиска путей ассоциативной сети.	20
2.1.1 Исходные данные	20
2.1.2 Задача поиска	23
2.1.3 Алгоритм решения	23
2.2 Алгоритмы обращения к базе данных.	32
2.3 Построение запроса выборки данных.	34
3. Реализация информационной технологии	39
3.1. Архитектура программного изделия	39
3.2. Выбор среды разработки программного изделия	40
3.3. Проектирование базы данных.	41
3.4. Оптимизация базы данных	46
3.4.1. Алгоритм построения оптимальной схемы БД	47
3.4.2. Алгоритм построения минимального покрытия множества функциональных зависимостей	47
3.4.3 Алгоритм построения замыкания множества атрибутов	48
3.4.5. Доказательство оптимальности схемы БД	48
3.5. Выбор способа доступа к базе данных.	54
3.5.1. Доступ через ODBC	54
3.5.2. Компоненты DAO	54
3.5.3. Технология ADO	55
3.5.4. Компоненты KADao	55
3.6. Предварительная обработка материалов	57
3.6. Реализация работы когнайзера.	62
3.7. Справочная система	63
4. Моделирование работы тезауруса	64
4.1. Редактирование материалов ЛКТ	66
4.1.1. Редактирование с помощью главной формы	66
4.1.2. Редактирование в табличном режиме	70
4.2. Выборка материалов по критерию	76
4.3. Моделирование работы когнайзера	80
5. Эргономика и охрана труда	86
5.1. Характеристика влияния условий и режима труда на утомляемость и снижение производительности труда	86
5.2. Характеристика условий труда	92

5.3. Эргономические требования к рабочему месту.	93
5.4. Психологические требования к рабочему месту.	97
5.5. Расчет освещенности рабочего места.	100
5.6. Схема рабочего места.	103
6. Организационно-экономическая часть.....	108
6.1. Цель организационно-экономической части.....	108
6.2. Стадии и этапы разработки.....	108
6.3. Определение затрат на разработку программного изделия.	109
6.3.1. Определение времени реализации проекта.	109
6.3.2. Определение финансовых затрат на разработку	110
6.4. Обоснование разработки программного изделия.	115
Заключение.....	116
Список литературы.....	117
Приложения.....	119
Приложение 1 Графическая часть дипломного проекта	119

Нормативные ссылки

При разработке дипломного проекта использовались следующие стандарты:

1. СН 181-70. Указания по проектированию цветовой отделки интерьеров производственных зданий.
2. ГОСТ 19.001-77. Единая система программной документации. Общие положения.
3. ГОСТ 19.502-78. Единая система программной документации. Описание применения. Требования к содержанию и оформлению.
4. ГОСТ 19.404-79. Единая система программной документации. Пояснительная записка. Требования к содержанию и оформлению.
5. ГОСТ 12.1.003-83 (СТ СЭВ 1930-79). Система стандартов безопасности труда. Шум. Общие требования безопасности.
6. ГОСТ 20886-85. Организация данных в системах обработки данных. Термины и определения.
7. ГОСТ 12.1.005-88. Общие санитарно-гигиенические требования к воздуху рабочей зоны.
8. ГОСТ 19781-90. Обеспечение систем обработки информации программное. Термины и определения.
9. ГОСТ Р ИСО/МЭК 9075-93. Информационная технология. Язык баз данных SQL с расширением целостности.
10. СНиП 23-05-95. Освещение промышленных и общественных зданий. Нормы проектирования. Свет. Нормы освещенности. Источники света.
11. ГОСТ 34.320-96. Информационные технологии. Система стандартов по базам данных. Концепции и терминология для концептуальной схемы и информационной базы.
12. ГОСТ 34.321-96 Информационные технологии. Система стандартов по базам данных. Эталонная модель управления данными.
13. ГОСТ Р ИСО/МЭК 12119-2000. Информационная технология. Пакеты программ. Требования к качеству и тестирование.
14. ГОСТ Р ИСО/МЭК ТО 12182-2002. Информационная технология. Классификация программных средств.
15. СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к ЭВМ и организации работы.

Определения, обозначения, сокращения

В расчетно-пояснительной записке применяют следующие термины:

Индекс	Объект базы данных, создаваемый с целью повышения производительности выполнения запросов.
База данных	Совокупность связанных данных, организованных по определенным правилам, предусматривающим общие принципы описания, хранения и манипулирования, независимая от прикладных программ. База данных является информационной моделью предметной области
Система управления базами данных	Комплекс программных и лингвистических средств общего или специального назначения, реализующий поддержку создания баз данных, централизованное управление и организацию доступа к ним различных пользователей в условиях принятой технологии обработки данных
Психолингвистика	Наука, исследующая обусловленность процессов речи и ее восприятия структурой соответствующего языка
Словоформа	Термин, обозначающий конкретное слово в конкретной грамматической форме.
Процессор базы данных	Набор динамических библиотек, функции которых позволяют не только обращаться к данным, но и управлять ими на стороне приложения. Процессор БД также предоставляет механизм блокировок при совместном доступе к базе данных.

В расчетно-пояснительной записке применяют следующие обозначения и сокращения:

БД	База данных
ЛКТ	Лингвокультурный тезаурус
ПО	Программное обеспечение
СУБД	Система управления базами данных

Введение

Естественный язык долгое время был для лингвистов неприкосновенным объектом изучения, и работы в области языковедения носили описательный характер. Язык рассматривался как некий самостоятельно развивающийся объект изучения. Рассматривались различные факты языка, проводилась систематизация, выделение его структур, схем функционирования. С развитием вычислительной техники, постановкой новых задач и накоплением знаний о языках появилась необходимость в лингвистических объектах, имеющих новые свойства. Позволяющих с другой стороны взглянуть на естественный язык. Не только на сложившиеся внешние признаки его функционирования, но и на принципы развития языка. Такими объектами могут быть порождающие синтезаторы, грамматики нового типа, новые типы словарей.

1. Анализ предметной области и постановка задачи.

1.1 Лингвокультурный тезаурус

Словари строятся на эмпирической основе. То есть для их составления необходимо обработать некий массив языковых данных, например, в текстовом виде. Если же исходных данных в необходимом количестве нет, то составление словаря по сложившейся методике невозможно. Создание же порождающего механизма, метода лингвистического конструирования могло бы помочь иначе взглянуть на эту задачу. Кроме того, язык функционирует и развивается в рамках окружающей его действительности и, с одной стороны, является ее отражением, а с другой, формирует наше представление о ней, так называемое «языковое сознание». Язык в данном случае – инструмент, посредством которого мы можем взглянуть на «сознание», на восприятие мира отдельным субъектом. Очевидно, что язык динамичен, но мы будем рассматривать относительно стабильную форму его существования в виде текстов.

Мы подошли к понятию Языковой Картины Мира. Она состоит из элементов, отражающих знания об окружающем мире (Единицы Знания о Мире - ЕЗМ). Эти элементы могут представлять собой различные языковые структуры: определения, пословицы, прецедентные тексты. Совокупность таких элементов создает разнородную картину мира. Языковое сознание имеет двойственную структуру, представляя собой некий механизм соединения знаний о языке со знанием о мире. Языковое Сознание (далее ЯС) можно представить в виде структуры:

<Языковое Сознание> = <Единица Знаний о Мире> + <Языковая Единица>

В этой зависимости языковое сознание представляет собой когнайзер, в котором происходит постоянное преобразование Единиц Знаний о Мире в Языковые Единицы и наоборот. Работу когнайзера по переходу от слова (знака) к знанию будем называть активным режимом работы, а от знания к знаку – пассивным.

1.2 Обзор аналогов и прототипов

Система ЛКТ является узкоспециализированным программным продуктом, поэтому не существует систем, повторяющих ее функциональность. В качестве аналогов и прототипов были выбраны:

- RSO Semantic Network SDK – инструментарий разработчика.
- www.lexfn.com – интернет-ресурс, работающий с англоязычной ассоциативной сетью.
- Microsoft Excel – самостоятельная работа исследователя с фигурами знания и ассоциативной сетью в качестве таблиц Excel.

RCO Semantic Network SDK

Средства библиотеки RCO Semantic Network позволяют автоматически анализировать содержание текстовых документов, представляя его в форме ассоциативной семантической сети. Ассоциативная семантическая сеть в программе представляет собой ориентированный граф, вершинами которого служат значимые темы, выделенные в анализируемом тексте, а дугами – связи между ними. С каждой вершиной связаны вес (значимость) и частота упоминания темы, а с каждой дугой – вес (сила) связи и частота подкрепления связи в тексте.

Помимо частоты упоминания в тексте, каждой теме присваивается вес от 1 до 100, отражающий ее значимость по отношению к другим темам. Пользователь может задать минимальный порог по весу, ниже которого темы не включаются в семантическую сеть.

Ассоциативные связи между темами выделяются на основе частоты их совместного появления в одном предложении. Пользователь может задать минимальный порог по частоте, ниже которого связи отбрасываются. В конечном представлении связь преобразуется в две противоположные по направленности дуги графа, которым присваиваются веса от 1 до 100, которые отражают условную вероятность упоминания первой темы совместно со второй – силу связи.

Дополнительно на каждую тему выдается тематический реферат, представляющий наиболее информативные фрагменты текста, в которых данная тема упоминалась. Общий реферат текста представляет компиляцию наиболее информативных фрагментов по ключевым темам. Подробность реферирования может настраиваться пользователем.

Ниже представлены выявленные свойства RCO Semantic Network:

Таблица 1 Свойства RCO Semantic Network

Свойство	Оценка	Примечание
Наличие ассоциативной сети	Да	
Наличие тезауруса	Нет	Тезаурус в понимании ЛКТ как совокупность фигур знания библиотекой не поддерживается.
Возможность пополнения базы данных	Да	
Реализация пассивного режима когнайзера	Отсутствует	
Переносимость системы	Определяется приложением	Данный продукт, являясь комплектом разработчика, функционирует в рамках более общего приложения, предоставляя свои функции.

www.lexfn.com

Интернет-ресурс предоставляет широкие возможности по поиску путей между знаками в ассоциативной сети. Система различает множество типов связей между словами, выполняет поиск не только непосредственно достижимых из начальной вершин сети, но и находит пути с ограничением по длине. На рисунке ниже изображена главная страница ресурса, а так же пример результата запроса.

Lexical FreeNet
Connected thesaurus

[Dictionary Search](#) [Cinema FreeNet](#) [Rhyme Zone](#) [Help](#) [Database info](#) [Technical notes](#) [Acknowledgments](#)

>> Type one or two words, concepts, or famous names into the boxes below:

Word 1: Word 2:

apple england

(required) (optional)

- Show related: Find words related to the first word
- Connection: Find connections between the words
- Show reachable: Find words reachable within links
- Intersection: Find words reachable from both words
- Rhyme coercion: Find related words that rhyme
- Spell check: Find words spelled similarly to the first
- Substring: Find words containing the first as a substring

Restrict results to: Common words Nouns Verbs Adjectives Adverbs

Рисунок 1 Главная страница

Shortest paths between *apple* and *england*

[Reverse the order of the query](#)

- Synonymous ↔
- Triggers →
- Generalizes →
- Specializes ←
- Comprises ←
- Part of →
- Antonym of ↔
- Rhymes ↔
- Sounds like ↔
- Anagram of ↔
- Occupation of ←
- Nationality of →
- Birth year of →
- Death year of →
- Bio triggers →
- Also known as ↔

apple → computer → information ↔ restoration

→ england

apple → computer → industry → samuel slater → england

apple → computer → screen ↔ queen → england

apple → computer → can ↔ throne → england

apple → computer → use → margaret louise sanger → england

apple → computer → node ↔ rhode → england

[More paths >>](#)

Word 1:

Word 2:

Рисунок 2 Результаты запроса

Среди сильных сторон системы можно выделить хорошую проработку ассоциативной сети, множество различных типов связей, объемную базу данных и высокую скорость выполнения запросов.

В системе отсутствует тезаурус и, как следствие, нет полноценной реализации пассивного режима когнайзера. Особенностью интернет-проекта в том, что он не доступен локально, то есть при отсутствии подключения к сети Интернет. Пополнением базы данных занимаются разработчики, вносить изменения самостоятельно не допускается. Система работает только в рамках английского языка, русскоязычного варианта не существует.

Таблица 2 Свойства lexfn.com

Свойство	Оценка	Примечание
Наличие ассоциативной сети	англоязычный вариант	
Наличие тезауруса	Нет	
Возможность пополнения базы данных	Нет	
Реализация пассивного режима когнайзера	Пути для пары знаков	Ограничено вследствие отсутствия тезауруса
Переносимость системы	Доступ через Интернет	

Microsoft Excel

В качестве третьей альтернативы системе ЛКТ был выбран метод ручной обработки данных ассоциативной сети и тезауруса, в случае работы с таблицами Microsoft Excel.

Возможностей поиска и редактирования данных достаточно для поиска фигур знания и осуществления редактирования. Вместе с тем самостоятельное выполнения этих действий может привести к ошибкам и порче данных.

Работа с ассоциативной сетью является простой только при нахождении непосредственно-достижимых вершин. Выполнение моделирования пассивного режима когнайзера операция выполняемая, но может быть крайне трудоемкой.

При таком методе работы с тезаурусом добавление фигур знания может повлечь ошибки внесения повторных записей, ошибок добавления лишних областей или способов. И тем не менее этот метод долгое время использовался исследователями ввиду отсутствия альтернатив. В таблице ниже представлены характеристики метода.

Таблица 3 Свойства обработки в Excel

Свойство	Оценка	Примечание
Наличие ассоциативной сети	Да	
Наличие тезауруса	Да	Тезаурус в понимании ЛКТ как совокупность фигур знания библиотекой не поддерживается.
Возможность пополнения базы данных	Да	
Реализация пассивного режима когнайзера	Трудоемко	
Переносимость системы	Да	Реализуется переносом документа Excel

Обработка данных в Excel важна не только как самостоятельное решение работы с базой данных тезауруса, но и как вспомогательное средство при работе с другими программами. Например при подготовке материалов для импорта в ЛКТ, или реализации пассивного режима когнайзера с использованием возможностей работы с ассоциативной сетью системы www.lexfn.com.

Заключение

В результате анализа аналогов и прототипов нельзя сделать вывод о неоспоримо более высоком качестве какого-либо из них, или превосходстве разработанной системы, так как они имеют разный круг поставленных задач, разную программную платформу и сферу применения. Целью данного сравнения являлся обзор имеющихся у исследователя средств в данной предметной области.

1.3 Активный и пассивный режимы работы когнитивера

Разработке лингвокультурного тезауруса (сокращенно ЛКТ) предшествовал ассоциативный эксперимент. В нем респонденту называлось некое слово (стимул) и его задачей было назвать приходящее при этом в голову слово. В результате формировалась пара слов «стимул-реакция», которые можно рассматривать как отражение Языкового Сознания, где стимул выступает в роли Языковой Единицы, а реакция в роли Единицы Знаний о Мире. Причем, как показала практика, связь между стимулом и реакцией двунаправлена. Если стимул «абзац» вызывает реакцию «текст», то, как правило, имеет место и обратная связь. Стимул «текст», скорее всего, имеет среди множества своих слов-реакций «абзац». Для этого необходимо проведение достаточно полного опроса. Таким образом, здесь не важно, что мы выберем за Языковую Единицу, а что за Единицу Знания о Мире. Ассоциативный эксперимент отражает активный режим работы когнитивера.

Для рассмотрения пассивного режима работы когнитивера используются материалы кроссвордов. В кроссворде человеку предлагается смысл разгадываемого слова в той или иной форме (дефиниции, синонима и т.д.). Этот смысл, чаще всего в соединении с самим разгадываемым словом, представляет собой не что иное, как элементарное знание о мире, а именно о знаке, который необходимо разгадать. Например:

- Оттенок на фоне какого-нибудь цвета – ОТЛИВ (Дескрипция).
- Применяют, когда штурм не удался – ОСАДА (Антоним).
- Очень мелкие цветные бусинки – БИСЕР (Дескрипция).

Таким образом, в материалы кроссвордов представляют в своеобразной форме Языковую Картину Мира. И, если предположения о пассивном и активном режимах работы когнитивера верны, то результаты разгадывания кроссворда должны найти свое подкрепление и в активном режиме работы когнитивера – ассоциативном эксперименте.

Знания об окружающем мире можно представить в виде так называемых Фигур Знания [Караулов, 2004]. Это элементарные когнитивные единицы, которыми мы оперируем в повседневной жизни. Они содержат как интенциональные

характеристики, а именно Знак, Способ, Смысл, описанные выше, так и экстенциональные, отражающие положение Фигуры знания в общем пространстве знаний. Такими параметрами являются когнитивная область и функция. Схематичное представление фигуры знания можно увидеть на рис.1

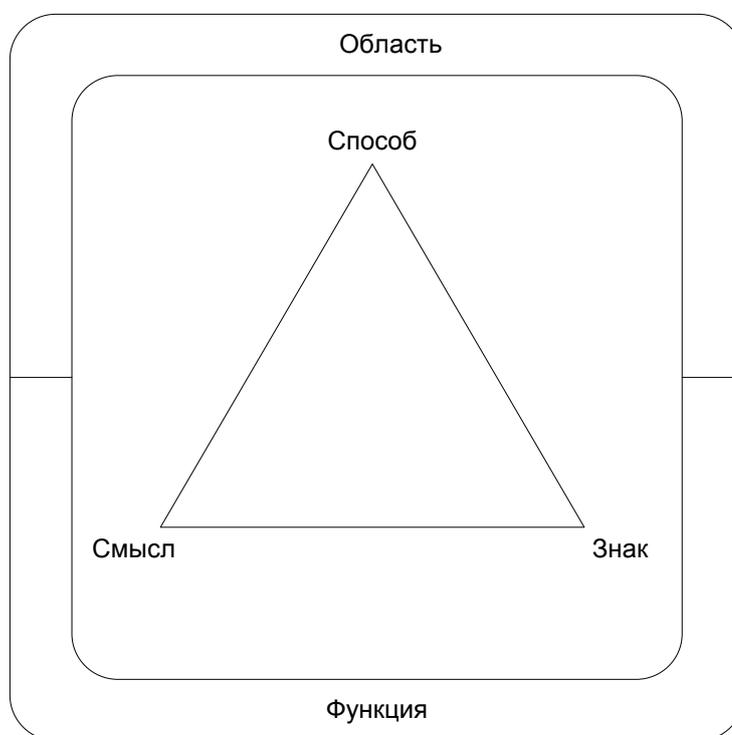


Рисунок 3 Фигура знания

Все знания мы можем разбить на категории, отражающие определенные стороны окружающего мира, на области. Знак может быть связан с Областью многообразными ассоциативными связями. К чему относится, с чем связано, частью чего является – вот далеко не полный список возможных вариантов. Строго говоря, составление списка возможных областей и разбиение фигур знания по ним - процесс субъективный, и каждый исследователь сделает это по-своему. В процессе наполнения базы данных эксперимента постепенно формируется множество возможных областей. При обработке достаточно большого количества понятий это множество перестает расти. Очевидно, одно и то же понятие может принадлежать нескольким областям, что свидетельствует о том, что области могут частично перекрываться, и разбиение на них весьма условно.

В активном режиме когнайзер переходит от Знака к Смыслу. При этом определяется Способ, которым Знак раскрывается через Смысл. В пассивном же режиме изначально имеется Смысл и, таким образом заданный, Способ. Имея эти

два параметра, когнайзер переходит к Знаку. В обоих режимах экстенционал фигуры знания определяется после интенционала. Казалось бы, связи между экстенциональными и интенциональными параметрами быть не должно, однако, некоторые зависимости можно проследить. Например, зависимости между Областью и Способом. В некоторых областях очевидно предпочтение определенным Способам выражения смысла. Например, Способ «Элемент множества» очень популярен в Географии. Также среди фигур знания, принадлежащих области «Авиация», многие имеют Функцию «Ретушь», так как знания в этой области не всегда важны в обыденной жизни. В области «Быт» наблюдаем обратную тенденцию.

Попробуем на примере проследить связь между активным и пассивным режимом в связи знака с областью. Важно заметить, что в ассоциативно-вербальной сети присутствует столь великое множество пар стимулов-реакций, что если мы будем искать связь между словами через сколь угодно множество промежуточных переходов, то на практике можем столкнуться с перемещением в огромном числе ветвлений и цепочек, где практически между любыми словами находится путь. Наша же задача стоит в нахождении минимальных путей. Рассмотрим следующую единицу знания пассивного режима когнайзера:

- Последняя буква кириллицы. 2. Фрейм. 3. ИЖИЦА. 4. Язык. 5. Рцт.

Рассмотрим обратный ассоциативный словарь, организованный от реакции к стимулу, включающий в себя все содержимое ассоциативного словаря. Ищем пары, в которых ИЖИЦА является реакцией. Таких пар не нашлось, значит, из смысловой формулы подбираем возможный аналог, это знак БУКВА. В обратном ассоциативном словаре БУКВА является реакцией на множество стимулов, среди которых есть стимул «язык», что подтверждает связь знака с областью в активном и пассивном режиме.

1.4 Постановка задачи информационной технологии

1.4.1 Назначение программного изделия

Программный продукт предназначен для демонстрации теоретических основ работы когнайзера, а также для проведения исследований тезауруса русского языка и ассоциативной сети. ЛКТ предназначен также для изменения и пополнения базы данных фигур знания, позволяя также выгружать данные в текстовый файл, для использования в других программах. Система разработана для распространения среди специалистов в области лингвистики и смежных областей.

1.4.2 Функции программного изделия

Информационная система должна выполнять следующие функции:

- Позволять пользователю изменять информацию о знаках, присутствующих в фигурах знания.
- Давать возможность пользователю изменять информацию о формулах смысла, связанных с определенным знаком.
- Предоставлять пользователю возможность изменять информацию об областях, к которым принадлежит фигура знания.
- Давать возможность пользователю изменять информацию о способе построения формулы смысла фигуры знания.
- Обеспечивать пользователю возможность изменять информацию о функции фигуры знания.
- Позволять выполнять поиск фигур знания по определенным критериям.
- Позволять осуществлять поиск путей между знаками в ассоциативной сети.
- Давать возможность искать пути от знаков, представляющих формулу смысла, к знаку, входящему в рассматриваемую фигуру знания.
- Осуществлять поиск путей минимальной длины в ассоциативной сети.
- Осуществлять поиск путей диапазона длин в ассоциативной сети.
- Осуществлять поиск путей определенной длины в ассоциативной сети.

1.4.3 Решаемые задачи

Система ЛКТ решает задачи:

- Хранения фигур знания и ассоциативной сети
- Изменения базы данных фигур знания
- Распространения базы данных тезауруса
- Импорта фигур знания в систему.
- Отбора фигур знания по определенному признаку
- Сортировки выборки фигур знания.
- Выгрузки фигур знания во внешний файл
- Моделирования пассивного режима работы когнайзера

2. Методы и алгоритмы лингвокультурного тезауруса.

2.1. Алгоритм поиска путей ассоциативной сети.

Для моделирования пассивного режима работы когнайзера необходимо осуществлять переход от формулы смысла к знаку, используя базу данных ассоциативного эксперимента. Для этого формула смысла разделяется на знаки в соответствии с алгоритмом поиска словоформ в ассоциативной сети. Целью данного алгоритма является представление формулы смысла в виде множества знаков, представленных в базе данных ассоциативного эксперимента.

Затем необходимо произвести поиск путей от каждого знака, входящего в формулу смысла к знаку, которого данная формула смысла описывает согласно базе данных кроссвордов. Возможно, часть формулы смысла не найдет своего отражения в базе данных ассоциативного эксперимента и дальнейшие вычисления к ней применяться не будут.

База данных ассоциативного эксперимента описывает направленный граф связей между знаками ассоциативной сети, поэтому поставленная задача является задачей поиска путей на графах, с отличиями, обусловленными особенностями результатов ассоциативного эксперимента.

2.1.1 Исходные данные

Исходными данными для алгоритма поиска являются:

- Множество исходных знаков – знаков, из которых будут искаться пути.
- Конечный знак – знак, к которому относится данная формула смысла согласно базе данных кроссвордов.
- База данных ассоциативного эксперимента

База данных ассоциативного эксперимента представляет таблицу следующей структуры:

Ассоциативная_сеть	
РК	Код
	Стимул Реакция Частота

Рисунок 4

Количество записей таблицы характеризуется десятками тысяч, что накладывает свои ограничения на алгоритм поиска путей.

Как показал анализ базы данных ассоциативного эксперимента, сеть представляет собой направленный граф. Общее количество связей 122.588. Количество знаков, являющихся стимулами, равно 6577. Количество реакций значительно превосходит эту величину и составляет 29623. Это объясняется тем, что респондент устанавливал реакцию не из множества стимулов, а произвольно. Поэтому возможных реакций значительно больше, за счет тех, которые никогда не являются стимулами.

Для ускорения доступа к конкретным сведениям в таблице базы данных применяются индексы. Индекс является структурой, которая упорядочивает значения в одном или нескольких столбцах таблицы базы данных, индекс помогает получить нужные сведения быстрее по сравнению с поиском по всем строкам таблицы.

Индекс обеспечивает указатели на значения данных, сохраненные в определенных столбцах таблицы, и упорядочивает эти указатели согласно заданному порядку сортировки. База данных использует индекс. Индекс сохраняется в базе данных вместе с таблицей, к которой он присоединен, или при сохранении схемы базы данных, в которую входит таблица.

Индекс для таблицы следует создавать только в том случае, если предполагается часто выполнять запросы по данным в индексированных столбцах. Индексы занимают место на диске и замедляют операции добавления, удаления и обновления строк. В большинстве ситуаций выигрыш в скорости за счет индексов перевешивает эти недостатки. Однако при частом обновлении данных в приложении или при ограничениях на дисковое пространство следует ограничивать количество индексов. Данная особенность указывает на необходимость применения индексирования к таблице ассоциативной сети, так как алгоритм когнайзера

предполагает частое обращение к ней для поиска данных, в то время как изменение содержимого ассоциативной сети не предполагается.

В базе данных Access существуют три типа индексов:

- Уникальный индекс
- Индекс первичного ключа
- Кластерный индекс

Уникальным индексом называют индекс, в котором две строки не могут иметь одинаковые значения индекса. В большинстве баз данных не допускается сохранение таблицы после создания уникального индекса, когда в существующих данных имеются повторяющиеся ключевые значения. База данных может также запрещать добавление данных, которые приведут к появлению в таблице повторяющихся ключевых значений.

Когда для таблицы в схеме базы данных создается первичный ключ, автоматически создается индекс первичного ключа, являющийся особым типом уникального индекса. Такой индекс требует уникальности каждого значения первичного ключа. Использование индекса первичного ключа в запросах также обеспечивает быстрый доступ к данным.

В кластерном индексе физический порядок строк в таблице совпадает с логическим (индексированным) порядком ключевых значений. Таблица может содержать только один кластерный индекс.

В работе когнайзера поиск ведется по полям «Стимул» и «Реакция» таблицы «Сеть». Уникальный индекс в данном случае неприменим, так как наименование стимулов и реакций могут совпадать. В целях ускорения выполняемых запросов к таблице ассоциативного эксперимента был создан кластерный индекс, включающий поля Стимул и Реакция.

Создание индекса возможно в режиме конструирования таблицы «Сеть», открытием вкладки «Индексы».

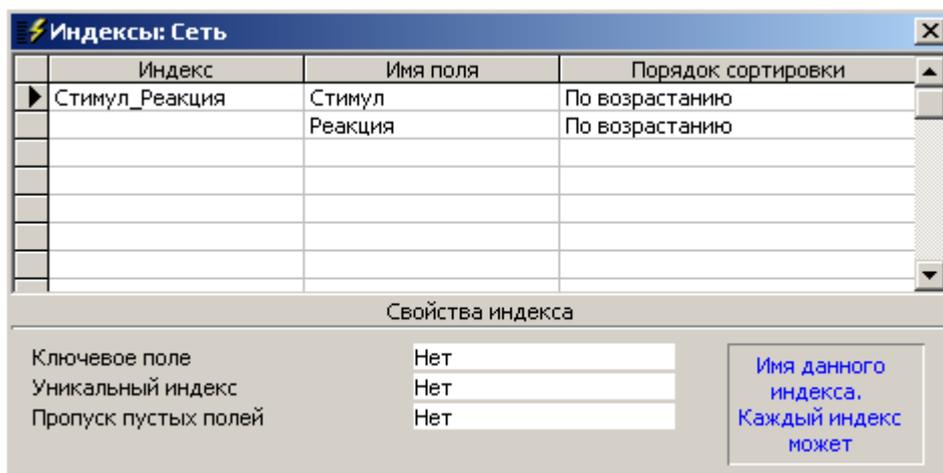


Рисунок 5 Создание индекса

Создание индекса возможно с помощью SQL-запроса вида:

```
CREATE INDEX Индекс
ON Сеть (Стимул, Реакция)
```

2.1.2 Задача поиска

Поиск путей от формулы смысла к знаку разделяется на поиск путей от каждого знака, входящего в формулу смысла. Таким образом, в основе задачи имеем задачу поиска путей между двумя знаками в ассоциативной сети. На результаты поиска пользователем могут накладываться ограничения, например, поиск только кратчайших путей, или поиск путей определенного диапазона длин. Эти условия влияют на критерий завершения работы алгоритма, а также на представляемые пользователю результаты.

Применяемый алгоритм поиска должен обладать следующими характеристиками:

- Определять длину пути.
- Восстанавливать сам путь при его нахождении.
- Должен быть не ресурсоемким.

Последний пункт означает минимальные затраты памяти и вычислительной мощности.

2.1.3 Алгоритм решения

Примененный в ЛКТ алгоритм поиска основан на алгоритме Дейкстра с тем отличием, что алгоритм Дейкстра оперирует путями между вершинами различных длин, а в ассоциативном эксперименте длина пути между соседними вершинами всегда равна единице.

Рассмотрим участок ассоциативной сети, схематично представленный на рисунке ниже.

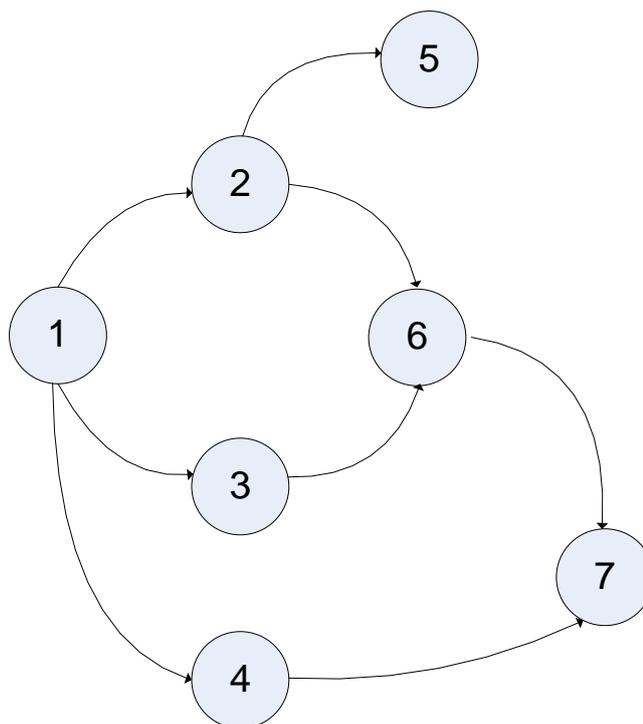


Рисунок 6

Необходимо найти пути от вершины 1 к вершине 7. В качестве дополнительного условия к алгоритму устанавливаем минимальную и максимальную длину путей, которые нам необходимо найти.

Maxlevel – верхняя граница искомых путей.

Minlevel – нижняя граница искомых путей.

Для каждой вершины в процессе поиска заполняется структура следующего вида:

Таблица 4

Наименование поля	Тип	Описание

Name	строка	наименование вершины
Level	целое число	Удаленность вершины от начальной
Up_element	указатель	Ссылка на вершину, из которой алгоритм перешел на данную
Old_elements	массив строк	Вершины, присутствующие в пути к данной вершине

Алгоритм поиска выглядит следующим образом:

1. Для начальной вершины заполняется структура.
Уровень = 0.
Указатель Up_element нулевой.
Список Old_elements пуст.
Текущей вершиной является начальная.
2. Если Level текущей вершины = Maxlevel тогда переход к пункту 8.
Иначе переход к пункту 3.
3. Рассматриваем список вершин, в которые можно попасть из текущей и которые не присутствуют в списке Old_elements текущей вершины.
4. Если в списке присутствует вершина, путь к которой мы ищем, выводим путь до текущей вершины с конечной вершиной как искомый путь. Обработка текущей вершины завершена. Переход к пункту 7.
5. Если список пуст, переход к пункту 8.
Иначе переход к пункту 6.
6. Для каждой вершины в списке заполняем структуру
Level = Уровень текущей вершины+1;
Up_element = Адрес текущей вершины;
Old_element = Old_element текущей вершины + Name текущей вершины.
Для каждой вершины начинаем работу алгоритма в пункте 2.
7. Если Level < Minlevel переход к пункту 8.
Иначе вывод списка Old_elements в качестве пути до текущей вершины.
Переход к пункту 8.

8. Выход.

В результате работы алгоритма, вершинам графа соответствуют множества `Old_element` и поля `Level`, см. Рисунок 7.

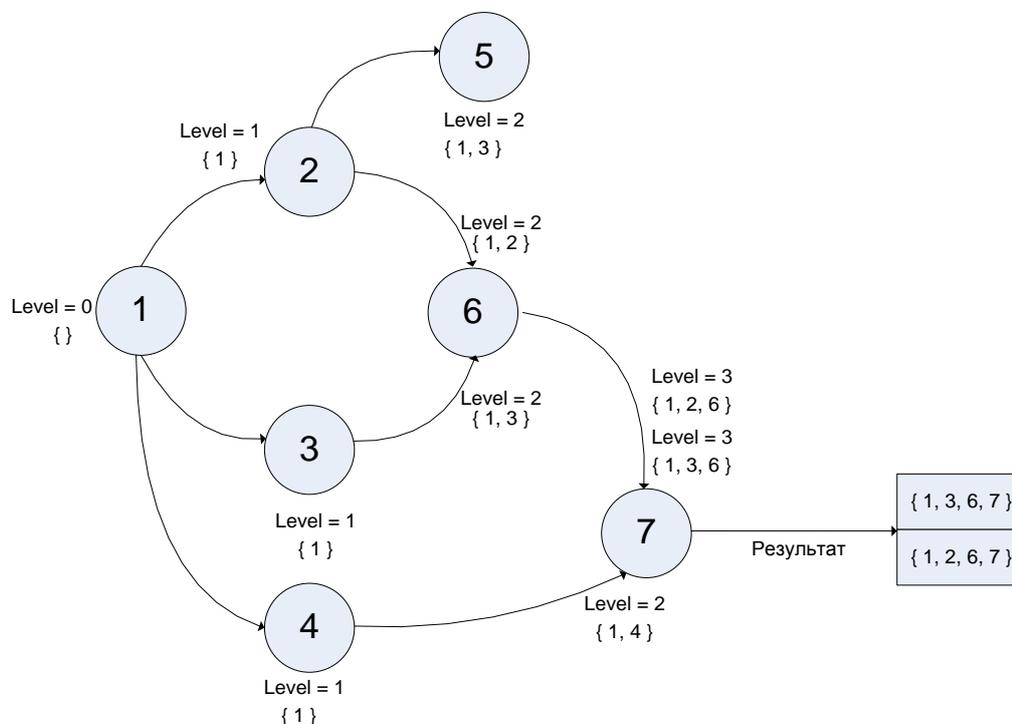


Рисунок 7

Так как необходимо искать не только кратчайшие, но и прочие пути к конечной вершине, применяем рекурсивный алгоритм. Переход от пункта 6 к пункту 2 может образовывать ветвление вычислений в случае, если образованный в пункте 3 список состоит более чем из одной вершины. На рисунке 3 это выражается в соответствии вершине 6 двух структур:

1. Level = 2
Old_elements = { 1, 2 };
2. Level = 2
Old_elements = { 1, 3};

Это соответствует множеству путей до вершины 6 от начальной вершины.

Исходный алгоритм Дейкстры имеет сложность $\Theta(n^2)$, но предназначен для поиска кратчайшего пути, в то время, как применяемый алгоритм обеспечивает

поиск всех путей заданного диапазона длин. Это требует ветвления вычислений, как это показано на рисунке 3. Каждая вершина графа может быть точкой ветвления вычислений на $n-1$. Таких вершин может быть $n-1$. В результате сложность алгоритма вырастает до $O(n^4)$. Произведем численную оценку роста сложности алгоритма относительно алгоритма Дейкстра, учитывая особенности ассоциативной сети.

Анализ ассоциативной сети показал, что в рассмотрении путей длиной более 5 нет необходимости в силу высокой степени связности. Если вершина не достигается за 5 шагов, она недостижима. Среднее количество реакций на стимул равно 18.78, при этом множество возможных реакции больше, чем множество стимулов. Это дает возможность предположить, что значительная часть связей стимул-реакция не имеют продолжения и реакция является листом ассоциативного дерева.

Для изучения свойств ассоциативной сети было проведено изучение свойств сети в двух случаях:

- При изменении количества стимулов при неизменном числе реакций каждого стимула.
- При изменении количества связей сети, в том числе при добавлении новых стимулов.

Для оценки характеристик сети при нарастании числа стимулов были произведены сокращения числа стимулов и расчет параметров:

- Количество связей
- Количество нелистьевых связей
- Процентное соотношение нелистьевых связей

Таблица 5 Изменение числа стимулов

Процент стимулов	Количество стимулов	Количество связей	Количество нелистьевых	Процент нелистьевых
------------------	---------------------	-------------------	------------------------	---------------------

			связей	связей
10	658	11005	701	6,36
20	1316	23788	3172	13,33
30	1974	36490	7328	20,08
40	2632	48560	12533	25,80
50	3290	60918	19649	32,25
60	3948	74305	28746	38,68
70	4606	86720	38630	44,54
80	5263	97857	49792	50,88
90	5920	109044	62322	57,15
100	6577	122059	78086	63,97

Ниже показана графическая интерпретация данной таблицы.

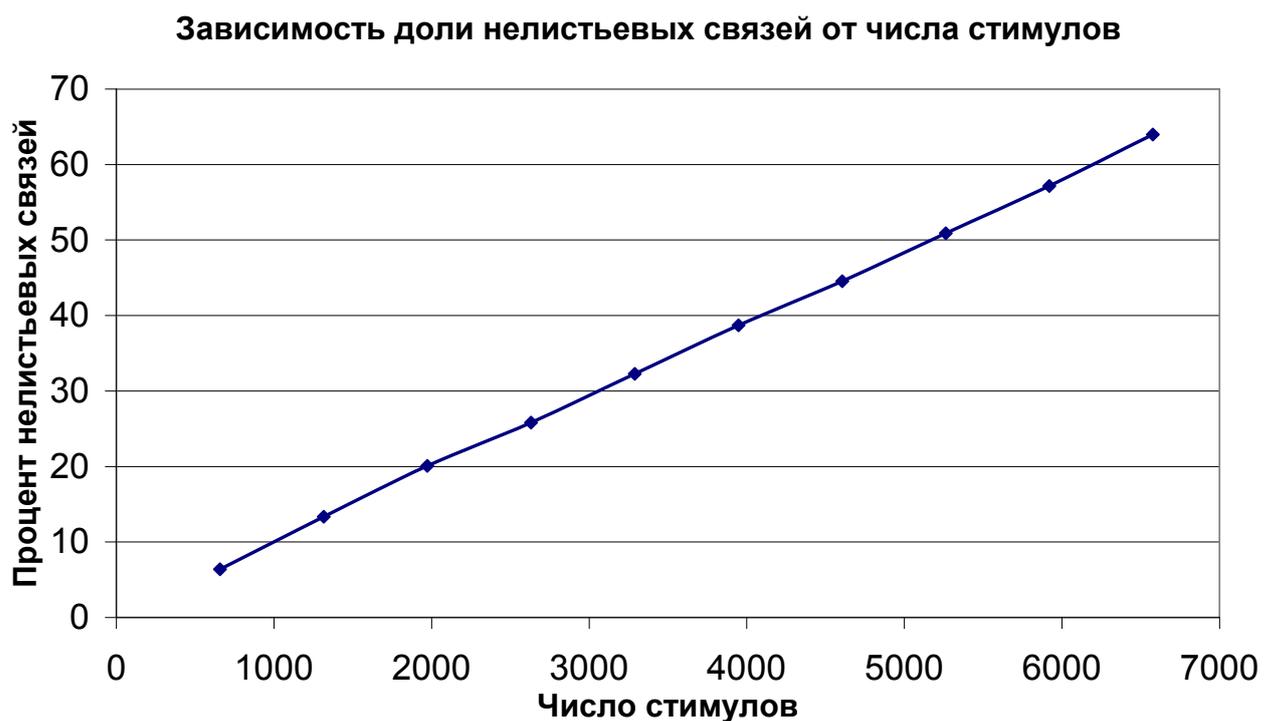


Рисунок 8 Изменение числа стимулов

На графике мы видим почти линейную зависимость процента нелистьевых связей от числа стимулов в диапазоне от 20 до 100 процентов стимулов исходной базы данных. Очевидно, что продолжив опрос респондентов и предложив им новые стимулы, мы можем уменьшить число листьев элементов сети.

Для оценки характеристик сети при нарастании числа связей были произведены сокращения числа связей и расчет параметров:

- Количество связей
- Количество стимулов
- Количество нелистьевых связей
- Процентное соотношение нелистьевых связей

На таблице ниже представлены результаты расчетов.

Таблица 6 Изменение числа связей

Процент связей	Количество стимулов	Количество связей	Нелистьевые элементы	Процент нелистьевых связей
1	932	1221	366	0,2997543
2	1521	2442	1013	0,414823915
3	1973	3663	1717	0,468741469
4	2343	4884	2491	0,51003276
5	2658	6105	3259	0,533824734
6	2929	7326	4009	0,547229047
7	3176	8547	4772	0,558324558
10	3838	12206	6934	0,568081272
20	5117	24412	14777	0,605317057
30	5740	36618	22706	0,620077557
40	6073	48824	30547	0,625655415
100	6577	122059	78086	0,639739798

На рисунках ниже приведены графические интерпретации содержимого таблицы.

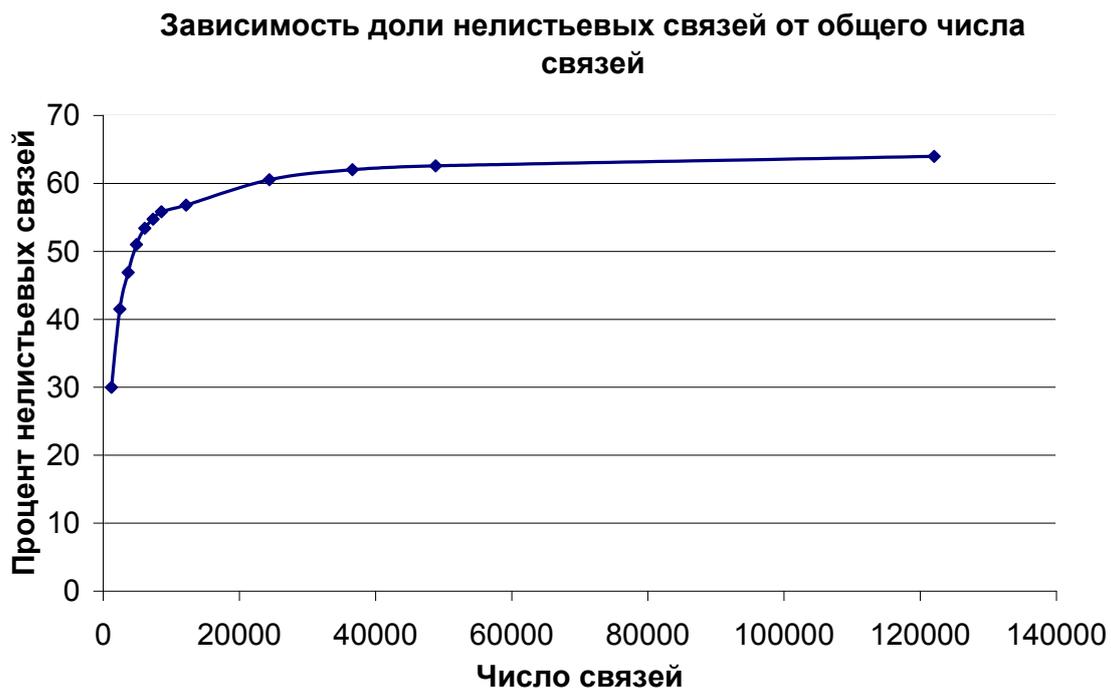


Рисунок 9 Зависимость доли нелистьевых связей

Судя по графику, при данном наборе стимулов, дальнейший опрос респондентов не приведет к значительному снижению доли листовых элементов сети.

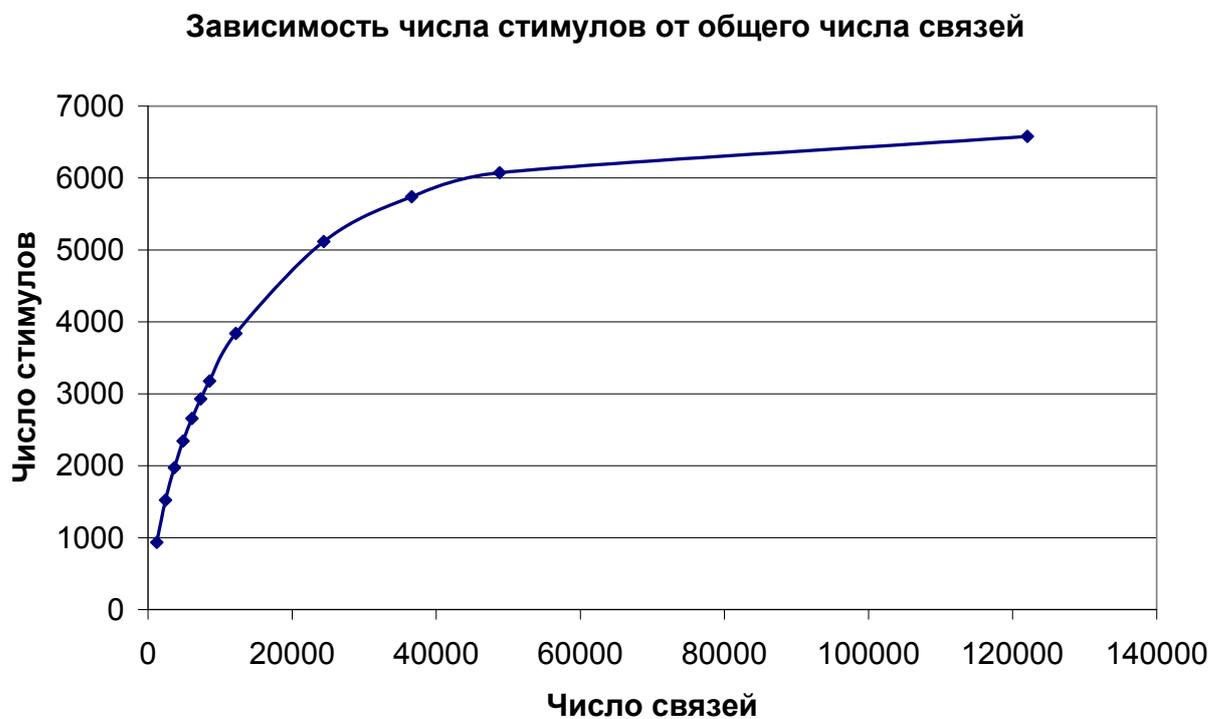


Рисунок 10 Зависимость числа стимулов от числа связей

На данном графике мы видим заполнение ассоциативной сети во время опроса респондентов с точки зрения появления в ней новых стимулов.

В данном случае была произведена более широкая выборка, для получения граничных характеристик при малом числе связей. Фактически эти цифры иллюстрируют состояние сети в процессе заполнения и могут служить основой выводов по структуре сети, а также по достаточности произведенного опроса.

Учитывая, что пути длиной более 5 в системе не рассматриваются, а также ограниченность количества реакций на стимул (что позволяет нам считать линейной зависимость между количеством вершин сети и ее связей), результирующая сложность алгоритма составляет $\Theta(k \cdot n^2)$, где k - константа, $k \ll n$.

2.2 Алгоритмы обращения к базе данных.

Для доступа к базе данных используются компоненты KADao.

Операции выполнения SQL-запросов и обработки результатов их выполнения осуществляются методами компонента TKADaoTable.

Операции с БД можно разделить на 2 категории:

1) Операции изменения БД.

Используются операторы INSERT, UPDATE, CREATE, DELETE;

После выполнения операции возможна проверка успешности произведенной операции.

2) Поисквые запросы

Используется оператор SELECT. Кроме возможной проверки успешности произведенного поискового запроса его результаты, как правило, подлежат дальнейшей обработке в программе.

Пример операции изменения БД:

Table - компонент для доступа к БД.

Деактивация компонента для изменения его параметров

```
Table.Active:=false;
```

Очистка строки SQL-запроса в компоненте

```
Table.SQL.Clear;
```

Заполнение строки SQL-запроса

```
Table.SQL.Add('DELETE FROM фигура_Область');
```

```
Table.SQL.Add(' WHERE Код_фигуры = 14 ');
```

```
Table.SQL.Add(' AND Код_области = 455');
```

Выполнение запроса

```
Table.ExecutesQL;
```

Пример поискового запроса:

AsWebTable – компонент для доступа к БД.

Деактивация компонента для изменения его параметров

```
AsWebTable.Active:=false;
```

Очистка строки SQL-запроса в компоненте

```
AsWebTable.SQL.Clear;
```

Заполнение строки SQL-запроса

```
AsWebTable.SQL.Add('SELECT DISTINCT Реакция, Стимул FROM  
Сеть_новой_структуры ');
```

```
Form_Calc.AsWebTable.SQL.Add('WHERE Стимул = 'Вода' AND Реакция <>  
'');
```

Активация компонента – ведет к выполнению запроса

```
AsWebTable.Active:=true;
```

Проверка наличия записей в результатах запроса

```
if (AsWebTable.RecordCount<1) then Exit;
```

Обработка результатов

```
for k:=1 to AsWebTable.RecordCount do
```

```
begin
```

```
AsWebTable.RecNo:=k; // перемещение по записям
```

```
S:= AsWebTable.FieldName('Реакция').AsWideString;
```

```
end;
```

2.3 Построение запроса выборки данных.

Система ЛКТ позволяет осуществлять выборку данных и поиск в базе данных тезауруса. Для этого используется форма запроса выборки данных, которая запускается выбором пункта «По параметрам» подменю «Запрос» главного меню. Форма содержит элементы управления в верхней части и таблицу результатов в нижней.

Далее на рисунке элементы управления.

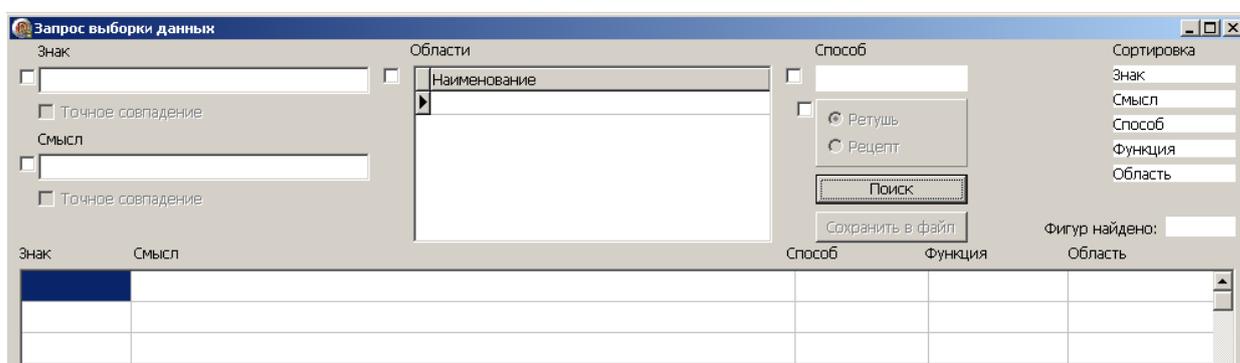


Рисунок 11 Параметры запроса выборки

Каждому параметру фигуры знания соответствует элемент управления, позволяющий конкретизировать его значение в результатах выборки. Рядом с элементами управления расположены переключатели, позволяющие включить/выключить его обработку.

Справа расположены индикаторы порядка сортировки результатов.

После нажатия кнопки «Поиск» формируется SQL-запрос. Рассмотрим его формирование.

Ob – строка, содержащая список областей, в которые должна входить фигура знания, через запятую.

Znak_use.Checked – используем ли отбор по знаку.

Znak_accur.Checked – точность отбора по знаку.

Smisl_use.Checked – используем ли отбор по формуле смысла.

Smisl_accur.Checked – точность отбора по формуле смысла.

Function_use.Checked – используем ли отбор по функции.

Sposob_use.Checked – используем ли отбор по способу.

Sposob_Kod.Caption – код способа для отбора.

retush.Checked – отбор ретушей.

recept.Checked – отбор рецептов.

Znak_value.Text – параметр отбора по знаку.

Smisl_value.Text – параметр отбора по формуле смысла.

s – строка, формируемый запрос.

1) Внесение постоянной части запроса

```
s:='SELECT Знаки.Наименование AS Знак, фигуры.Смысл AS Смысл,  
Способы.Наименование AS Способ, функции.Наименование AS функция,  
Области.Наименование AS Область FROM фигуры, Знаки, функции,  
фигура_функция, Способы, фигура_Способ, Области, фигура_Область  
WHERE ';
```

2) Формирование отбора по знаку

```
if(Znak_use.Checked=true) then  
begin  
if(Znak_value.Text='') then  
begin  
s:=s+'фигуры.Код_знака=NULL';  
end  
else  
begin  
if(Znak_accur.Checked=true) then  
s:=s+'Данные.Код_знака=Знаки.Код AND  
Знаки.Наименование = '''+Znak_value.Text+''''  
else  
s:=s+'фигуры.Код_знака=Знаки.Код AND  
Знаки.Наименование LIKE  
''*'+Znak_value.Text+'*''';  
end;  
end  
else s:=s+'фигуры.Код_знака=Знаки.Код';
```

3) Формирование отбора по формуле смысла

```
if(Smisl_use.Checked=true) then
  begin
    s:=s+' AND фигуры.Смысл ' ;
    if(Smisl_value.Text='') then s:=s+'= NULL '
    else
      if(Smisl_accur.Checked=true) then
        s:=s+'= '''+Smisl_value.Text+''''
      else s:=s+' LIKE '*'+Smisl_value.Text+'*''';
    end;
  end;
```

4) Формирование отбора по функции

```
if(Function_use.Checked=true) then
  begin
    if(retush.Checked=true) then
      begin
        s:=s+' AND фигуры.Код = фигура_функция.Код_фигуры;
        s:=s+' AND фигура_функция.Код_функции=функции.Код AND
        фигура_функция.Код_функции = 1 ' ;
      end;
    if(recept.Checked=true) then
      begin
        s:=s+' AND фигуры.Код = фигура_функция.Код_фигуры' ;
        s:=s+' AND фигура_функция.Код_функции=функции.Код AND
        фигура_функция.Код_функции = 2 ' ;
      end;
    end
  else
    s:=s+' AND фигура_функция.Код_функции=функции.Код And
    фигура_функция.Код_фигуры=фигуры.Код' ;
  end;
```

5) Формирование отбора по способу

```

if(Sposob_use.Checked=true) then
  begin
    if(Sposob_Kod.Caption='') then
      begin
        s:=s+' AND фигура_Способ.Код_способа=Способы.Код And
        фигура_Способ.Код_фигуры=Данные.Код';
      end
    else
      begin
        s:=s+' AND фигуры.Код = фигура_Способ.Код_фигуры';
        s:=s+' AND фигура_Способ.Код_способа='+
        Sposob_Kod.Caption +' AND Способы.Код =
        '+Sposob_Kod.Caption;
      end;
    end;
end;

```

```

if(Sposob_use.Checked=false) then
  s:=s+' AND фигура_Способ.Код_способа=Способы.Код And
  фигура_Способ.Код_фигуры=фигуры.Код';

```

6) Формирование отбора по областям

```

if(Oblasti_use.Checked=false) then
  s:=s+' AND фигура_Область.Код_фигуры=фигуры.Код And
  фигура_Область.Код_области=Области.Код';

```

```

if(ob<>'') then
  begin
    s:=s+' AND фигура_Область.Код_фигуры=фигуры.Код And
    фигура_Область.Код_области IN ('+ob+')';
  end

```

```

else

```

```
s:=s+' AND фигура_Область.Код_фигуры=фигуры.Код And
фигура_Область.Код_области=Области.Код' ;
s:=s+' AND фигура_Область.Код_фигуры=Данные.Код And
фигура_Область.Код_области=Области.Код' ;
s:=s+' ORDER BY ' ;
s:=s + Условие сортировки;
```

7) Запуск поискового запроса

```
Table.Active:=false;
Table.Close;
Table.SQL.Clear;
Table.SQL.Add(s);
Table.Active:=true;
```

3. Реализация информационной технологии

3.1. Архитектура программного изделия

Лингвокультурный тезаурус русского языка должен включать в себя данные ассоциативной сети и данные тезауруса – представленные в базе данных пятикомпонентные фигуры знания. Важным требованием к построению базы данных является переносимость. Система должна легко распространяться, без сложных процедур установки и настройки, а также работать на доступной и распространенной программно-аппаратной платформе.

Базу данных предполагалось строить на основе MS Access. Причин для этого несколько:

- 1) MS Access установлен на подавляющем большинстве персональных компьютеров, что позволит работать с базой данных практически везде, то есть любому исследователю.
- 2) База данных Access представляет собой отдельный файл формата mdb и при переносе программного обеспечения эксперимента этот файл также может быть перемещен, без необходимости проведения каких-либо наладочных работ. Это дает мобильность.
- 3) Эта СУБД входит в пакет MS Office, поэтому имеет возможность импорта и экспорта данных в различном виде. Например, импорта данных в виде файла MS Excel, или экспорта базы данных в СУБД MS SQL Server, что важно как предположительный вариант модернизации базы данных в случае необходимости улучшения ее характеристик.
- 4) Access имеет собственные средства для обработки данных, выполнения запросов, генерации отчетов и выборки.

3.2. Выбор среды разработки программного изделия

В качестве среды разработки программы был применен Borland Delphi 7.0, вследствие ряда причин:

- 1) Легкость проектирования интерфейсов и экранных форм.
- 2) Простота используемого в Delphi 7.0 языка программирования Object Pascal.
- 3) Наличие большого числа доступных компонент.
- 4) Невысокая стоимость среды разработки.
- 5) Большое количество разработанных алгоритмов.

Кроме того Borland распространяет программу создания инсталляционных пакетов Install Shield Express.

Для подготовки справочной документации в виде chm-файла используется редактор HelpCruiser

3.3. Проектирование базы данных.

На этапе разработки инфологической структуры БД были установлены присутствующие в ней сущности, их свойства, а также связи между ними.

Таблица 7 Сущность Фигура

Сущность «Фигура»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Код знака(Внешний ключ)	INTEGER
Смысл	VARCHAR(255)

Таблица 8 Сущность Знак

Сущность «Знак»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Наименование	VARCHAR(50)

Таблица 9 Сущность Фигура_Область

Сущность «Фигура_Область»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Код фигуры (Внешний ключ)	INTEGER
Код области (Внешний ключ)	INTEGER

Таблица 10 Сущность Фигура_Способ

Сущность «Фигура_Способ»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Код фигуры (Внешний ключ)	INTEGER
Код способа (Внешний ключ)	INTEGER

Таблица 11 Сущность Фигура_Функция

Сущность «Фигура_Функция»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Код фигуры (Внешний ключ)	INTEGER
Код функции (Внешний ключ)	INTEGER

Таблица 12 Сущность Область

Сущность «Область»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Наименование	VARCHAR(50)

Таблица 13 Сущность Способ

Сущность «Способ»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Наименование	VARCHAR(50)

Таблица 14 Сущность Функция

Сущность «Функция»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Наименование	VARCHAR(50)

Таблица 15 Сущность Ассоциация

Сущность «Ассоциация»	
Название атрибута	Тип поля
Код (Первичный ключ)	INTEGER
Стимул	VARCHAR(50)
Реакция	VARCHAR(50)
Частота	INTEGER

Таблица 16 Связи между сущностями

Связанные сущности	Название связи	Тип связи
Знак-Фигура	Имеет	1:M
Фигура-Фигура_Область	Содержит связь	1:M
Фигура-Фигура_Способ	Содержит связь	1:M
Фигура-Фигура_Функция	Содержит связь	1:M
Область-Фигура_Область	Относится к связи	1:M
Область-Фигура_Способ	Относится к связи	1:M
Область-Фигура_Функция	Относится к связи	1:M

Даталогическая схема БД разработана в пакете Erwin 4.1 и представлена в соответствующем листе графической части.

Таблица 17

«Фигуры»	
Название атрибута	Тип поля
Код	INTEGER
Код_знака	INTEGER
Смысл	VARCHAR(255)

Таблица 18

«Знаки»	
Название атрибута	Тип поля
Код	INTEGER

Наименование	VARCHAR(50)
--------------	-------------

Таблица 19

«Фигура_Область»

Название атрибута	Тип поля
Код	INTEGER
Код_фигуры	INTEGER
Код_области	INTEGER

Таблица 20

«Фигура_Способ»

Название атрибута	Тип поля
Код	INTEGER
Код_фигуры	INTEGER
Код_способа	INTEGER

Таблица 21

«Фигура_Функция»

Название атрибута	Тип поля
Код	INTEGER
Код_фигуры	INTEGER
Код_функции	INTEGER

Таблица 22

«Области»

Название атрибута	Тип поля
Код	INTEGER
Наименование	VARCHAR(50)

Таблица 23

«Способы»

Название атрибута	Тип поля
Код	INTEGER
Наименование	VARCHAR(50)

Таблица 24

«Функции»

«Функции»	
Название атрибута	Тип поля
Код	INTEGER
Наименование	VARCHAR(50)

Таблица 25

«Сеть»

«Сеть»	
Название атрибута	Тип поля
Код	INTEGER
Стимул	VARCHAR(50)
Реакция	VARCHAR(50)
Частота	INTEGER

3.4. Оптимизация базы данных.

Оптимальной схемой БД будем считать схему, удовлетворяющую следующим требованиям:

- свойство соединения без потерь
- свойство сохранения зависимостей
- нахождение схемы отношений БД в третьей нормальной форме

Нахождение схемы отношений БД в третьей нормальной форме обеспечивает отсутствие в каждой схеме отношений аномалий:

- избыточности
- потенциальной противоречивости
- аномалии включенных записей

Свойство соединения без потерь:

Пусть $\rho = (R_1 \dots R_n)$ – схема базы данных,

$\bigcup_{i=1}^n R_i$ – универсальная схема отношений,

ρ – обладает свойством соединения без потерь, если для любого экземпляра r универсальной схемы отношений R имеет место:

$r = \text{Pr}_{R_1}(r) \blacktriangleright \blacktriangleleft \dots \blacktriangleright \blacktriangleleft \text{Pr}_{R_n}(r)$, где

$\text{Pr}_{R_n}(r)$ – проекция отношения r на множество атрибутов R_i ,

$\blacktriangleright \blacktriangleleft$ - естественное соединение.

Свойство сохранения зависимостей:

Определение:

Проекция множества F на R_i – это множество $X \rightarrow Y \subseteq F^+$, такое что $X, Y \subseteq R_i$.

Пусть $\rho = (R_1 \dots R_n)$ – схема базы данных,

F – множество функциональных зависимостей на R , где R – универсальная схема отношений.

Схема БД обладает свойством сохранения зависимостей, если справедливо равенство:

$$\left(\bigcup_{i=1}^n R_i(F) \right)^+ = F^+$$

Нахождение схемы отношений БД в 3NF.

Схема отношений R находится в третьей нормальной форме, если не существует ключа X , множества атрибутов $Y \subseteq R$ и непервичного атрибута $H \notin Y$, таких что:

$X \rightarrow Y$ – имеет место,

$Y \rightarrow H$ – имеет место,

$Y \rightarrow X$ – не имеет место.

3.4.1. Алгоритм построения оптимальной схемы БД

Исходные данные:

Пусть $R = (A_1 \dots A_n)$ – универсальная схема отношений,

F – множество функциональных зависимостей на R .

Алгоритм.

1. Положить $\rho = 0$ – множество схем отношений, которые образуют схему БД.

2. Определить G – минимальное покрытие для F .

3. Каждую зависимость $V \rightarrow W$ из G заменить на множество атрибутов VW .

Получившееся множество схем отношений обозначить через Q .

4. Если множество атрибутов $(A_1 A_2 \dots A_n) \subseteq Q$, то добавить в ρ схему отношений $R = (A_1 \dots A_n)$. Выйти из алгоритма. В этом случае оптимальная схема БД будет состоять только из одной схемы отношений R . Иначе перейти к пункту 5.

5. Добавить в ρ в качестве схем отношений те одиночные атрибуты, которые не вошли ни в одну из схем из Q .

6. Добавить в ρ все схемы отношений из Q .

3.4.2. Алгоритм построения минимального покрытия множества функциональных зависимостей

1. Каждую функциональную зависимость из F заменить на совокупность зависимостей, каждая из которых имеет один атрибут из правой части (это возможно

в силу правила декомпозиции). Полученное множество функциональных зависимостей обозначить G .

2. Для каждой функциональной зависимости $X \rightarrow A$ из множества G выполнить следующие действия:

А) Проверить, принадлежит ли данная зависимость замыканию $(G - (X \rightarrow A))^+$. Если да, то исключить зависимость из G . Для того чтобы убедиться, входит ли зависимость $X \rightarrow A$ в $(G - (X \rightarrow A))^+$ нужно построить замыкание атрибута X . Если $A \subset X^+$, то $X \rightarrow A \subset (G - (X \rightarrow A))^+$, т.е. зависимость может быть выведена из меньшего числа функциональных зависимостей.

Б) Для каждого собственного подмножества $Z \subset X$ проверить, принадлежит ли зависимость $Z \rightarrow A$ замыканию Z^+ . Если да, то заменить $X \rightarrow A$ на $Z \rightarrow A$. Повторить данный пункт для зависимости $Z \rightarrow A$.

3. Зависимости с одинаковой левой частью объединяются в одну функциональную зависимость.

3.4.3 Алгоритм построения замыкания множества атрибутов

1. $i := 0$, $X_0^+ = X$ (X_i^+ – замыкание множества атрибутов на i -ом шаге).
2. $X_{i+1}^+ = X_i^+ \cup V$, где V – множество атрибутов таких, что в F существуют функциональные зависимости $Y \rightarrow Z$, для которых $Y \subseteq X_i^+$, $V \subseteq Z$.
3. Если $X_{i+1}^+ = X_i^+$, то закончить выполнение алгоритма, иначе $i := i + 1$, перейти к пункту 2.

3.4.5. Доказательство оптимальности схемы БД

В таблицах ниже приведены условные обозначения полей сущностей.

Таблица 26 Сущность Фигура

Сущность «Фигура»	
Название атрибута	Обозначение
Код (Первичный ключ)	A
Код знака(Внешний ключ)	B
Смысл	A1

Таблица 27 Сущность Знак

Сущность «Знак»	
Название атрибута	Обозначение
Код (Первичный ключ)	В
Наименование	В1

Таблица 28 Сущность Фигура_Область

Сущность «Фигура_Область»	
Название атрибута	Обозначение
Код (Первичный ключ)	F
Код фигуры (Внешний ключ)	A
Код области (Внешний ключ)	K

Таблица 29 Сущность Фигура_Способ

Сущность «Фигура_Способ»	
Название атрибута	Обозначение
Код (Первичный ключ)	G
Код фигуры (Внешний ключ)	A
Код способа (Внешний ключ)	C

Таблица 30 Сущность Фигура_Функция

Сущность «Фигура_Функция»	
Название атрибута	Обозначение
Код (Первичный ключ)	H
Код фигуры (Внешний ключ)	A
Код функции (Внешний ключ)	D

Таблица 31 Сущность Область

Сущность «Область»	
Название атрибута	Обозначение
Код (Первичный ключ)	K
Наименование	K1

Таблица 32 Способ

Сущность «Способ»	
Название атрибута	Обозначение
Код (Первичный ключ)	C
Наименование	C1

Таблица 33 Сущность Функция

Сущность «Функция»	
Название атрибута	Обозначение
Код (Первичный ключ)	D
Наименование	D1

Таблица 34 Сущность Ассоциация

Сущность «Ассоциация»	
Название атрибута	Обозначение
Код (Первичный ключ)	E
Стимул	E1
Реакция	E2
Частота	E3

Сформируем множество функциональных зависимостей F:

$A \rightarrow B, A \rightarrow A_1,$

$B \rightarrow B_1,$

$F \rightarrow A, F \rightarrow K,$

$G \rightarrow A, G \rightarrow C,$

$H \rightarrow A, H \rightarrow D,$

$K \rightarrow K_1,$

$C \rightarrow C_1,$

$D \rightarrow D_1,$

$E \rightarrow E_1, E \rightarrow E_2, E \rightarrow E_3$

Положим $\rho = 0$.

Построим минимальное покрытие для F.

В множестве F нет функциональных зависимостей, у которых в правой части есть несколько атрибутов.

Определим, какие зависимости можно удалить из G.

Таблица 35 Минимальное покрытие G

Зависимость / вывод	Замыкание атрибута в левой части зависимости
$A \rightarrow B$ $B \notin X^+ \Rightarrow A \rightarrow B \in G$	0) $X_0^+ = A$ 1) $X_1^+ = X_0^+ \cup A_1 = AA_1$ $X^+ = AA_1$
$A \rightarrow A_1$ $A_1 \notin X^+ \Rightarrow A \rightarrow A_1 \in G$	0) $X_0^+ = A$ 1) $X_1^+ = X_0^+ \cup B = AB$ 2) $X_2^+ = X_1^+ \cup B_1 = ABB_1$ $X^+ = ABB_1$
$F \rightarrow A$ $A \notin X^+ \Rightarrow F \rightarrow A \in G$	0) $X_0^+ = F$ 1) $X_1^+ = X_0^+ \cup K = FK$ 2) $X_2^+ = X_1^+ \cup K_1 = FKK_1$ $X^+ = FKK_1$
$F \rightarrow K$ $K \notin X^+ \Rightarrow F \rightarrow K \in G$	0) $X_0^+ = F$ 1) $X_1^+ = X_0^+ \cup A = FA$ 2) $X_2^+ = X_1^+ \cup A_1 \cup B = FAA_1B$ 3) $X_3^+ = X_2^+ \cup B_1 = FAA_1BB_1$ $X^+ = FAA_1BB_1$
$G \rightarrow A$ $A \notin X^+ \Rightarrow G \rightarrow A \in G$	0) $X_0^+ = G$ 1) $X_1^+ = X_0^+ \cup C = GC$ 2) $X_2^+ = X_1^+ \cup C_1 = FCC_1$ $X^+ = FCC_1$
$G \rightarrow C$ $C \notin X^+ \Rightarrow G \rightarrow C \in G$	0) $X_0^+ = G$ 1) $X_1^+ = X_0^+ \cup A = GA$ 2) $X_2^+ = X_1^+ \cup A_1 \cup B = GAA_1B$

	$3) X_3^+ = X_2^+ \cup B_1 = GAA_1BB_1$ $X^+ = GAA_1BB_1$
$H \rightarrow A$ $A \notin X^+ \Rightarrow H \rightarrow A \in G$	$0) X_0^+ = H$ $1) X_1^+ = X_0^+ \cup D = HD$ $2) X_2^+ = X_1^+ \cup D_1 = HDD_1$ $X^+ = HDD_1$
$H \rightarrow D$ $D \notin X^+ \Rightarrow H \rightarrow D \in G$	$0) X_0^+ = H$ $1) X_1^+ = X_0^+ \cup A = HA$ $2) X_2^+ = X_1^+ \cup A_1 \cup B = HAA_1B$ $3) X_3^+ = X_2^+ \cup B_1 = HAA_1BB_1$ $X^+ = HAA_1BB_1$
$K \rightarrow K_1$ $K_1 \notin X^+ \Rightarrow K \rightarrow K_1 \in G$	$0) X_0^+ = K$ $1) X_1^+ = X_0^+$ $X^+ = K$
$C \rightarrow C_1$ $C_1 \notin X^+ \Rightarrow C \rightarrow C_1 \in G$	$0) X_0^+ = C$ $1) X_1^+ = X_0^+$ $X^+ = C$
$D \rightarrow D_1$ $D_1 \notin X^+ \Rightarrow D \rightarrow D_1 \in G$	$0) X_0^+ = D$ $1) X_1^+ = X_0^+$ $X^+ = D$
$E \rightarrow E_1$ $E_1 \notin X^+ \Rightarrow E \rightarrow E_1 \in G$	$0) X_0^+ = E$ $1) X_1^+ = X_0^+ \cup E_2 \cup E_3 = EE_2E_3$ $X^+ = EE_2E_3$
$E \rightarrow E_2$ $E_2 \notin X^+ \Rightarrow E \rightarrow E_2 \in G$	$0) X_0^+ = E$ $1) X_1^+ = X_0^+ \cup E_1 \cup E_3 = EE_1E_3$ $X^+ = EE_1E_3$

$E \rightarrow E_3$	0) $X_0^+ = E$
$E_3 \notin X^+ \Rightarrow E \rightarrow E_3 \in G$	1) $X_1^+ = X_0^+ \cup E_1 \cup E_2 = EE_1E_2$
	$X^+ = EE_1E_2$

Объединяем зависимости с одинаковой левой частью

$$G = \{A \rightarrow BA_1, B \rightarrow B_1, F \rightarrow AK, G \rightarrow AC, H \rightarrow AD, C \rightarrow C_1, D \rightarrow D_1, K \rightarrow K_1, E \rightarrow E_1E_2E_2\}.$$

4. Добавляем в схему отношений все элементы из множества Q.

Полученное в процессе алгоритма множество таблиц полностью совпадает с исходным набором таблиц базы данных, следовательно, спроектированная схема базы данных является оптимальной.

Инфологическая и даталогическая схемы базы данных представлены в приложении.

3.5. Выбор способа доступа к базе данных.

Для доступа к базе данных Access из программы, разрабатываемой в Delphi, существует несколько технологий, наиболее предпочтительными из которых являются:

- 1) Доступ через ODBC
- 2) Компоненты DAO
- 3) Технология ADO
- 4) Компоненты KADao

3.5.1. Доступ через ODBC

ODBC (Open Database connectivity) – открытое соединение с базами данных – комплекс элементов, обеспечивающих стандартных средств взаимодействия с источниками данных при помощи синтаксиса SQL.

Целью создания технологии было предоставление разработчикам программного обеспечения единообразного доступа к различным источникам данных. Это реализуется с помощью Диспетчера драйверов ODBC и драйверов ODBC.

Диспетчер драйверов ODBC (ODBC Driver Manager) представляет собой библиотеку dll, выступающую в роли шлюз, при помощи которого приложения взаимодействуют с СУБД.

Драйвер ODBC (ODBC Driver) – компонент, взаимодействующий с СУБД, при обращении к нему диспетчера драйверов.

Для настройки ODBC используется утилита ODBC Data Source Administrator. Она отображает установленные в системе драйвера и позволяет ассоциировать необходимый драйвер с источником данных. Это усложняет переносимость системы.

3.5.2. Компоненты DAO

DAO (Data Access Objects) — объекты для доступа к данным — технология доступа к данным компании Microsoft, разработанная для работы с СУБД Microsoft Jet и позволяющая обращаться к базе данных Access.

Преимуществом в сравнении с ODBC-технологией является доступ ко всем функциональным возможностям процессора Jet, являющегося ядром СУБД Access. Физически DAO реализован в виде иерархии объектов и коллекций объектов, упакованных в библиотеку dll.

3.5.3. Технология ADO

ActiveX Data Objects — интерфейс программирования приложений для доступа к данным, разработанный компанией Microsoft и основанный на технологии компонентов ActiveX. ADO позволяет представлять данные из разнообразных источников (реляционных баз данных, текстовых файлов и т. д.) в объектно-ориентированном виде.

Среди сильных сторон ADO можно указать отсутствие сложной структуры и единообразную иерархию объектов.

3.5.4. Компоненты KADao

KADao – компоненты доступа к базе данных Microsoft Access, использующие технологию DAO. Имеет простую систему классов, в значительной степени соответствующих таковой в ADO.

В результате разработки была обнаружена проблема совместимости компонент ADO с компонентом доступа Microsoft Data Access Components, распространяемым со многими продуктами Microsoft.

В результате для получения доступа к базе данных в ЛКТ используется набор компонент KaDAO.

TKADaoDataBase – компонент подключения к базе данных.

Установленные свойства компонента:

Таблица 36 Свойства TKADaoDataBase

Свойство	Значение	Описание
DatabaseName	Tes.mdb	Путь к файлу базы данных
EngineType	dbUseJet	Тип поставщика соединения
Version	3.6	Используемая версия компонента

TKADaoTable – компонент выполнения запросов к базе данных и обработки результатов

3.6. Предварительная обработка материалов.

В системе ЛКТ реализована возможность пополнения БД тезауруса новыми фигурами знания из таблицы Excel определенного формата. Первоначально заполнение БД осуществлялось именно таким образом.

Исходная таблица базы данных тезауруса, представленная в текстовом документе MS Word, была переведена в MS Excel, а затем перенесена в базу данных MS Access в таблицу «Данные_исходные» с помощью функции Импорт в контекстном меню СУБД Access.

Таблица имеет следующий вид:

Таблица 37 Исходная таблица

Код	Знак	Смысл	Область	Способ	Функция

Будем именовать эту таблицу далее таблицей фигур знания. Необходимо внести пояснения по ее составу.

- КОД имеет числовой формат и служит для идентификации записей.
- ЗНАК содержит в текстовом виде слово, которое в кроссворде зашифровано в формуле смысла.
- СМЫСЛ представляет собой текст, на который ориентируется отгадывающий человек, то есть разъяснение понятия, представленной в той или иной форме.
- ОБЛАСТЬ является перечислением тех областей знания, к которым относится данное понятие. Каждое понятие принадлежит как минимум одной области. Области в поле перечислены через запятую, пробелы либо другие небуквенные символы.

- СПОСОБ представляет собой структуру формулы смысла, через который загадано слово. Это может быть дефиниция, прецедентный текст, метафора или множество других способов.
- ФУНКЦИЯ определяет так называемую «ценность» понятия. С точки зрения ценности различают знание-рецепт и знание-ретушь. Знание-рецепт несет в себе важную, существенную для исследователя информацию, необходимую ему в повседневной жизни. Знание-ретушь содержит дополнительные, менее существенные данные, которые в некоторой степени можно не рассматривать при анализе текста. Исследователь, определяя, к какому типу знание относится, руководствуется собственными представлениями о необходимости этого знания для него самого, для его понимания окружающего мира. Рассмотренная выше таблица имеет ряд недостатков.

Например, поле СПОСОБ может принимать одно из некоторого набора возможных значений. На этапе пополнения БД количество вариантов способов не известны. Запись в поле СПОСОБ наименования способа ведет не только к избыточности объемов хранимых в памяти данных, но и является потенциальным источником ошибок. Поэтому возможные области выделяются в отдельную таблицу, с которой устанавливается связь.

Аналогично можно поступить с полем ФУНКЦИЯ.

Особенностью областей является то, что, во-первых, области образуют некоторое конечное множество, во-вторых, любое понятие может принадлежать одной либо нескольким областям.

Схема БД тезауруса представлена на рисунке 1.

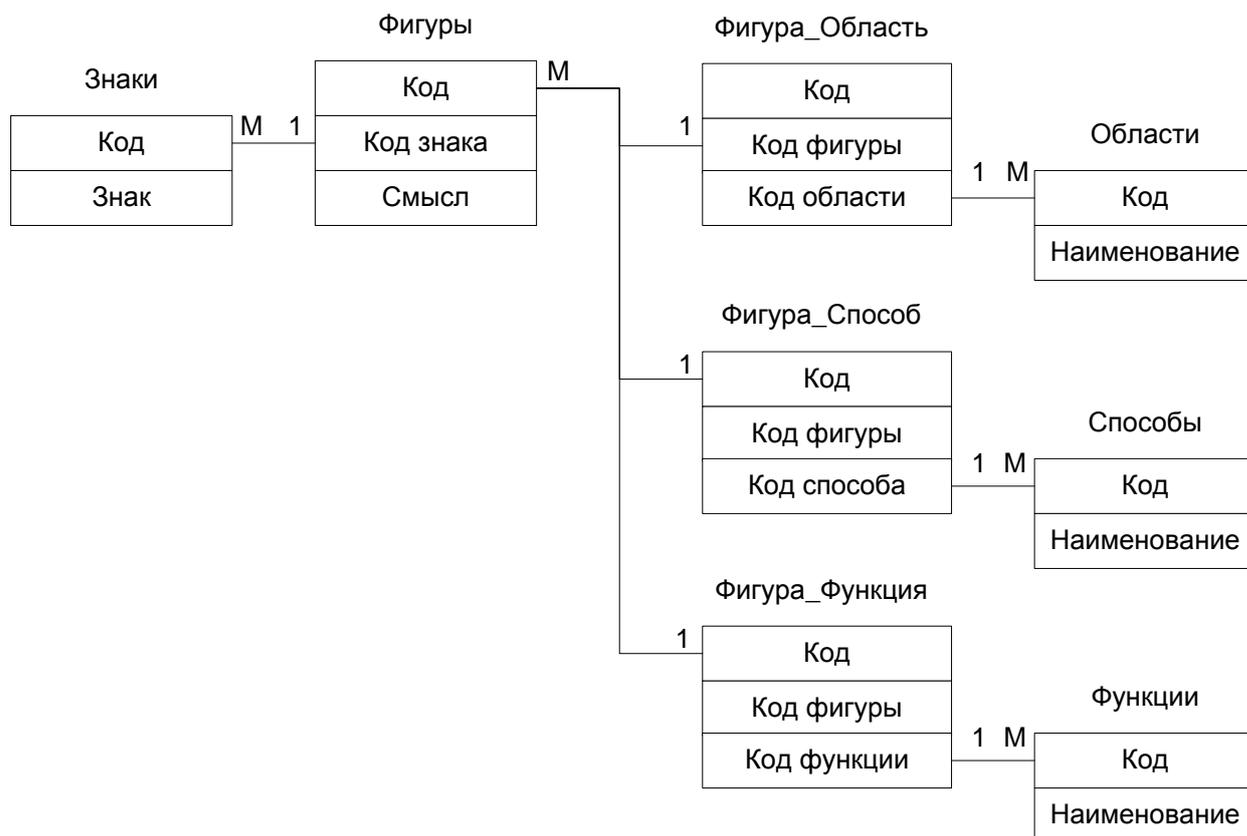


Рисунок 12 Представление фигур знания в тезаурусе

Заполнение таблиц Знаки, Фигуры, Фигура_Область, Фигура_Способ, Фигура_Функция, Области, Способы по исходной таблице не может выполняться полностью в автоматическом режиме. Предположим, в поле Область исходной таблицы мы встречаем обозначение, которое не присутствует среди имеющихся в таблице областей наименований. Либо нам необходимо добавить новую область и создать связь, либо это одна из имеющихся областей, просто иначе обозначенная в исходных текстах, либо в наименовании которой допущена ошибка. Поиск же в БД имеющихся областей и предложение их пользователю автоматизировано.

Для автоматизации разбора исходной таблицы была разработана форма, изображенная на рисунке ниже.

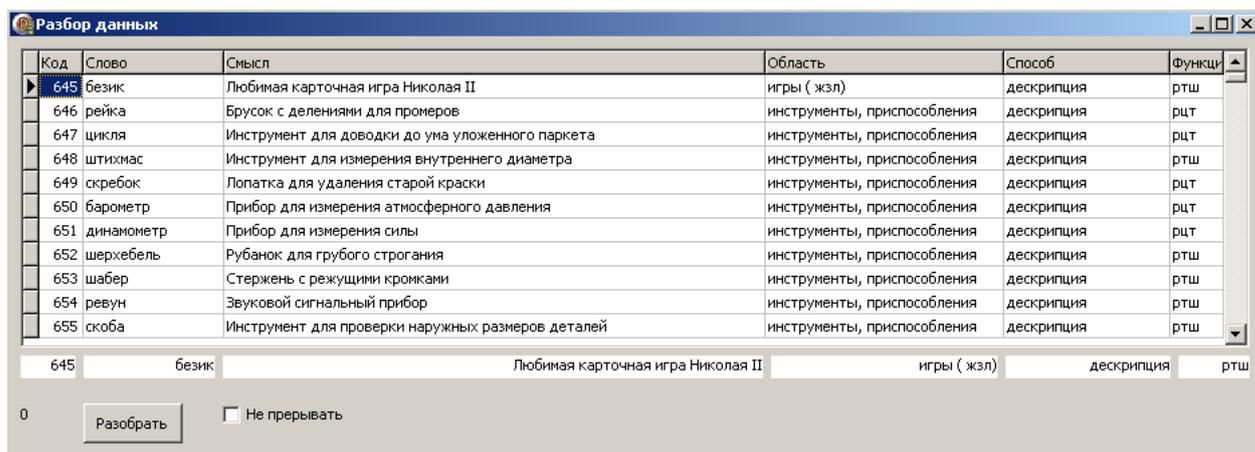


Рисунок 13 Форма разбора исходных данных

В таблице представлены еще не обработанные записи. Ниже в текстовых полях представлен состав активной записи, которая будет обработана при нажатии кнопки «Разобрать».

Разбор очередной записи исходной таблицы и заполнение результирующих таблиц представляет собой последовательность из следующих этапов:

- 1) Если среди Знаков уже есть содержимое поля «Знак» исходной таблицы, Фигура знания связывается с имеющимся знаком. Иначе создается новый Знак.
- 2) Среди Областей осуществляется поиск элементов, наименования которых встречаются в поле «Область» исходной таблицы. В случае если в записи исходной таблицы всего одна область, и наблюдается полное совпадение с какой-либо областью таблицы областей, связь устанавливается автоматически, без участия пользователя. Иначе предлагается вручную выбрать необходимые области, в том числе с возможностью создать новые. Пример такого диалога можно увидеть на рисунке ниже.

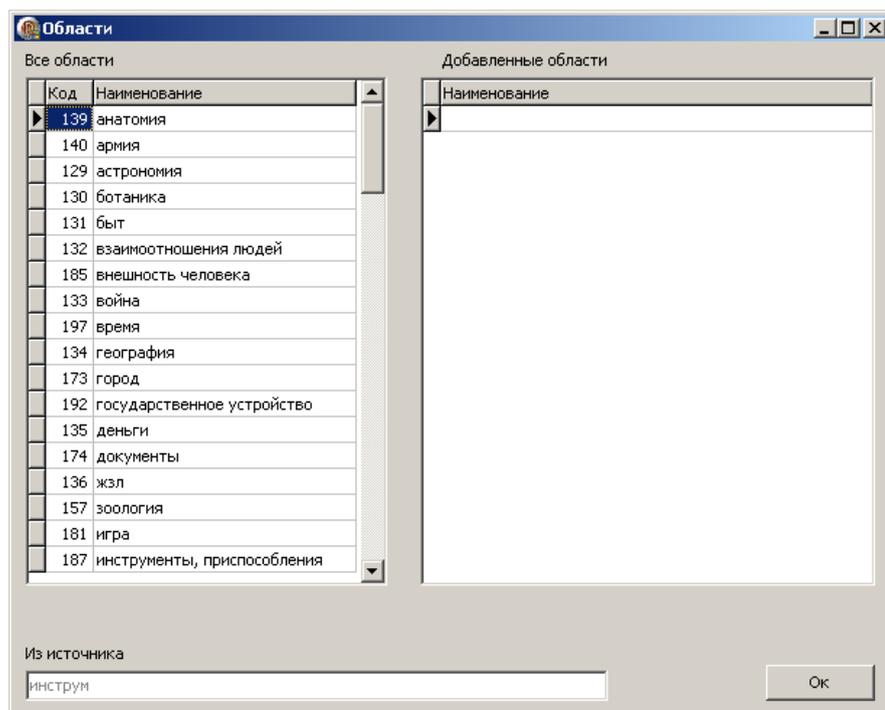


Рисунок 14 Установка областей фигуры

3) Производится разбор поля Способ, аналогично полю Область.

Стоит заметить, что если осуществить разбор поля автоматически не удастся, пользователю предлагается выбрать Способ из имеющихся в базе данных, либо ввести новый.

4) Заполнение поля Функция.

Обработав таким образом все записи исходной таблицы, переводим их к схеме базы данных тезауруса.

3.6. Реализация работы когнайзера.

Лингвокультурный тезаурус должен осуществлять пассивный режим когнайзера, то есть переход от формулы смысла к знаку фигуры знания, используя ассоциативную сеть.

Знаки, составляющие формулу смысла, могут находиться в произвольных словоформах, поэтому интерфейс когнайзера позволяет не только разбить формулу смысла на знаки, но и осуществить их правку, с целью максимального использования возможностей ассоциативной сети.

Процедура обработки фигуры знания в когнайзере состоит из следующих этапов:

- 1) Запуск когнайзера для фигуры знания.
- 2) Правка словоформ формулы смысла для поиска соответствующих знаков ассоциативной сети.
- 3) Отключения не значимых для поиска знаков, или поиск по которым вести не надо.
- 4) Установка ограничений на длину путей.
- 5) Запуск работы когнайзера.
- 6) Получение результатов расчета.

Пример всех этапов работы когнайзера отражен в главе «Моделирование и экспериментальные исследования»

3.7. Справочная система.

Справочную информацию о системе можно получить, выбрав пункт Справка -> Помощь главного меню.

Подсистема справки реализована в виде документа chm-формата, внешний вид которого представлен на рисунке ниже.

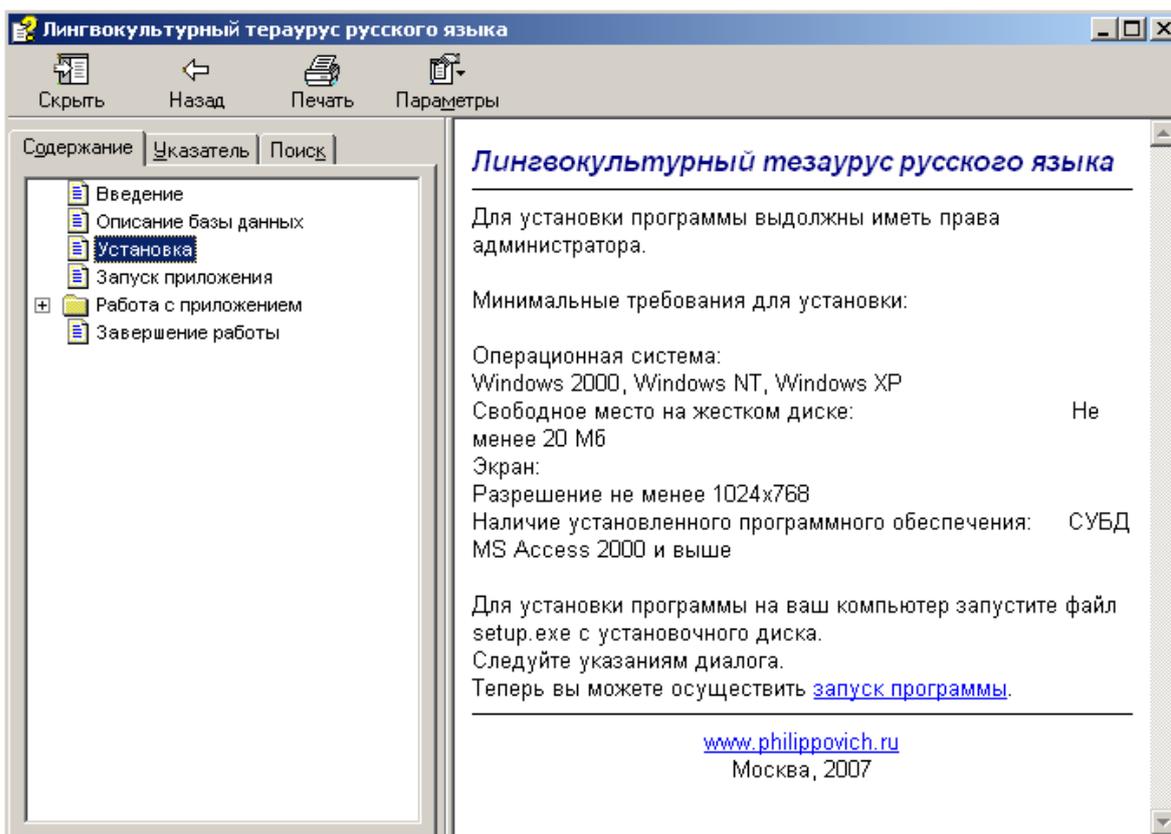


Рисунок 15 Справочная система

4. Моделирование работы тезауруса

Система ЛКТ реализована в виде оконного приложения. Стандартные элементы управления и отображения информации дают возможность пользователю быстро научиться использовать ее функциональные возможности. При запуске программы пользователю открывается главная форма. Ниже приведен граф диалога пользователя, в котором отражены дальнейшие возможные действия.

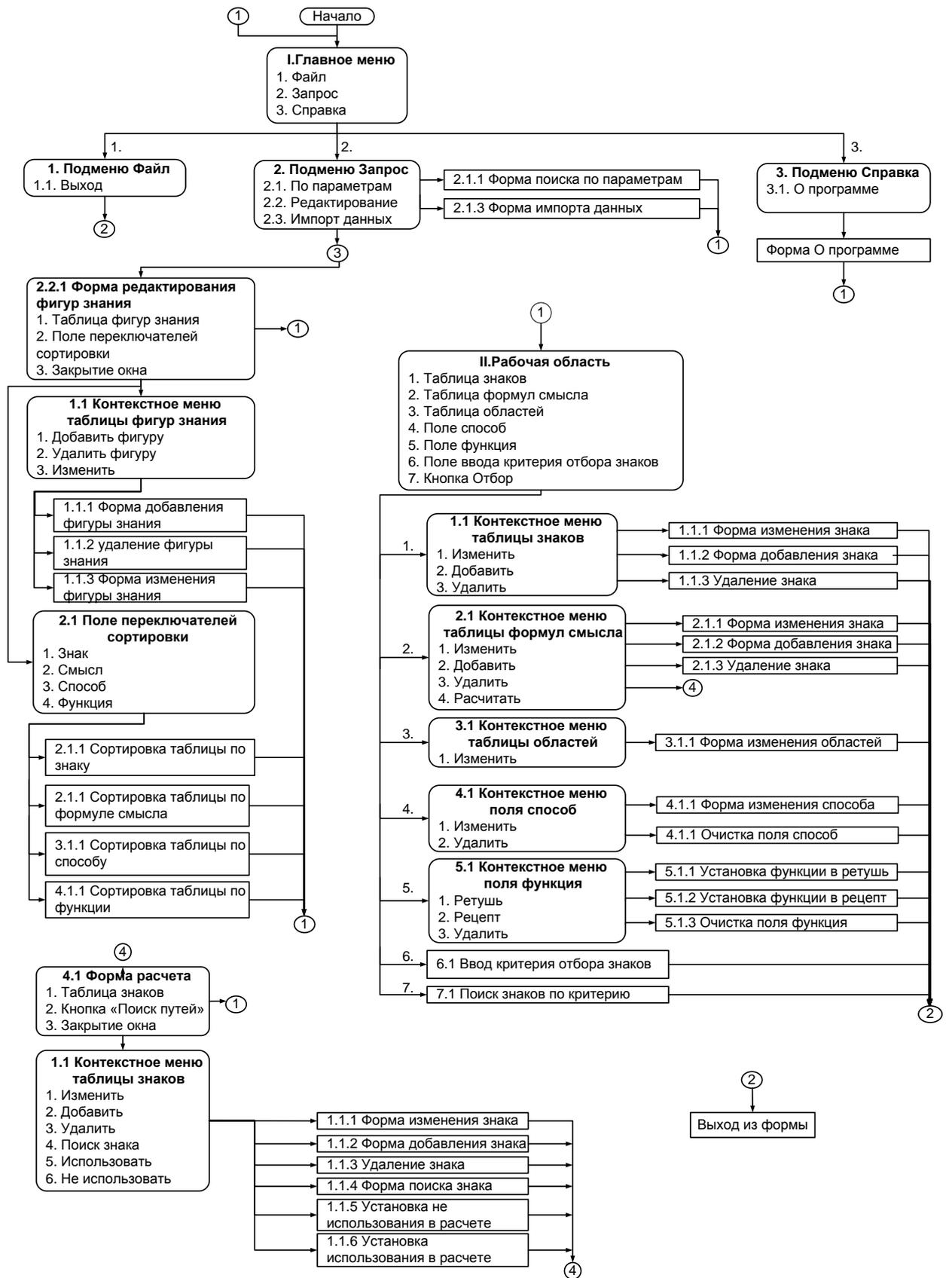


Рисунок 16 Граф диалога пользователя

4.1. Редактирование материалов ЛКТ.

Система позволяет редактировать фигуры знания, представленные в тезаурусе. Редактирование содержимого ассоциативной сети не предполагается. Для изменения содержимого тезауруса можно воспользоваться главной формой ЛКТ, либо войти в табличный режим представления с помощью команды главного меню Запрос-> Редактировать.

4.1.1. Редактирование с помощью главной формы

Интерфейс главной формы отражает пятикомпонентную структуру фигур знания.

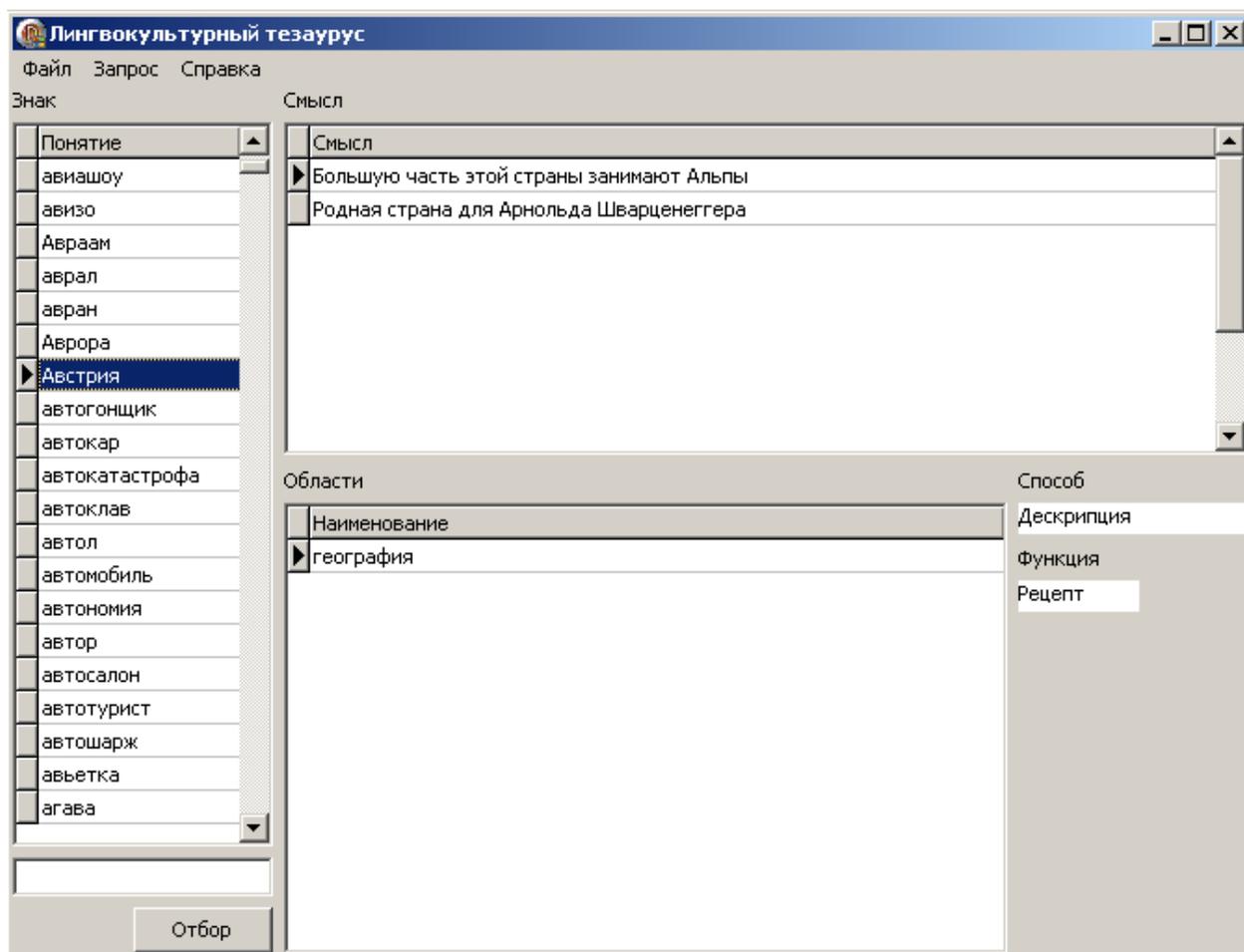


Рисунок 17 Главная форма

Список знаков в правой части экрана содержит знаки всех фигур знания. Если в базе имеется несколько фигур знания, отражающих один знак, то в списке ему соответствует одна запись. Слева от списка знаков находится список формул смысла, соответствующих выбранному в списке знаков элементу. Количество записей списка формул смысла показывает количество фигур знания, относящихся к текущему знаку. Например, на Рисунок 17 знаку «Австрия» соответствуют две формулы смысла:

- Большую часть этой страны занимают Альпы
- Родная страна для Арнольда Шварцнегера

Список Области в нижней части формы отображает список областей, к которым относится фигура знания, соответствующая активной формуле смысла. В данном случае это единственная область «География».

Поле Способ показывает способ представления знака через формулу смысла. Поле Функция содержит характеристику значимости фигуры знания и, в случае, если не пуста, имеет значение Ретушь или Рецепт.

Каждому из описанных выше элементов управления соответствует контекстное меню.

Осуществим добавление в базу данных тезауруса фигуры знания

Андорра - Государство в Восточных Пиренеях – География, Дескрипция, Ретушь.

В контекстном меню списка знаков выбираем пункт «Добавить». В открывшемся окне вводим наименование знака «Андорра» и нажимаем кнопку Ок (см. Рисунок 18 Добавление знака).



Рисунок 18 Добавление знака

Появляется сообщение о том, что данный знак уже присутствует в базе. Новый знак добавляться не будет. Текущий элемент списка знаков – «Андорра». Этому знаку соответствует формула смысла «Европейское княжество, все население которого легко разместится на трибунах футбольного стадиона». В контекстном меню списка формул смысла выбираем пункт «Добавить».

Открывшееся диалоговое окно позволяет ввести формулу смысла добавляемой фигуры знания (см. Рисунок 19).



Рисунок 19 Добавление формулы смысла

Список областей созданной фигуры знания пуст. Для списка областей выбираем пункт «Изменить» контекстного меню. Откроется окно редактирования списка областей, изображенное на рисунке ниже.

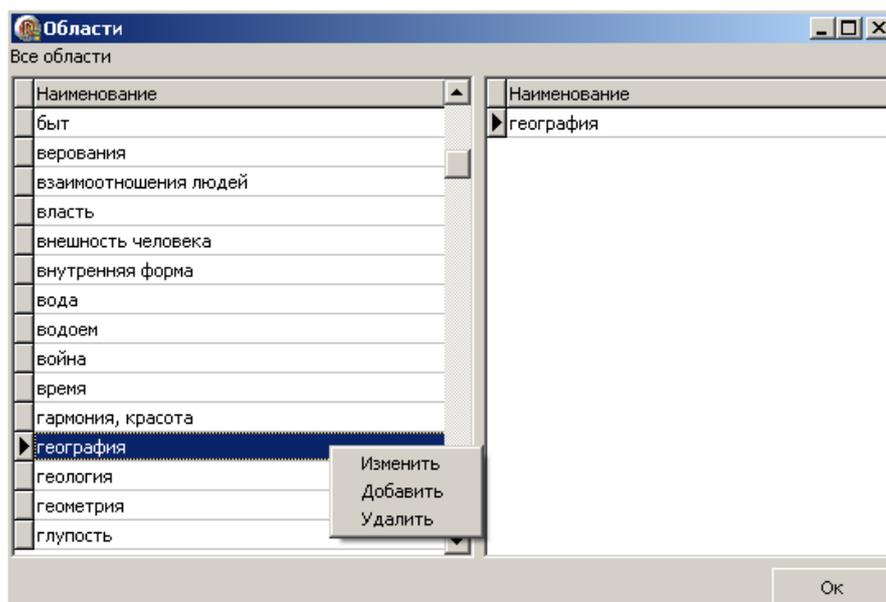


Рисунок 20 Редактирование списка областей

Рабочая область формы содержит две таблицы: слева расположена таблица областей, присутствующих в системе, а справа список областей, к которым относится текущая фигура знания. Список областей в системе можно редактировать, вызвав контекстное меню соответствующей таблицы. В данном случае область «География» уже присутствует в системе. Дважды щелкнув по данному элементу списка добавляем его в список областей редактируемой фигуры знания. После нажатия кнопки Ок изменения вступят в силу.

На вышестоящем рисунке показано контекстное меню таблицы областей, содержащее три позиции:

- Изменить
- Добавить
- Удалить

Пункт «Изменить» служит для изменение написания текущей области. При его выборе пользователю предлагается диалоговое окно.

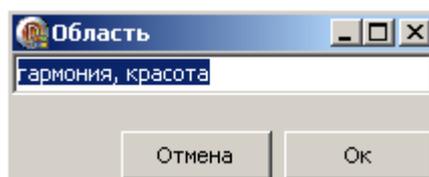


Рисунок 21 Изменение областей

При внесении изменений в написание области все имеющиеся с ней связи в таблице «Фигура_Область» сохраняются.

Пункт «Добавить» служит для добавления новой области. Форма добавления аналогична описанной выше форме добавления.

Пункт «Удалить» удаляет область из списка областей и все связи таблицы «Фигура_Область»

Следующий этап редактирования – установка значения Способ фигуры знания значение Описание. Для этого вызываем пункт «Изменить» контекстного меню поля Способ. Открывается форма выбора способа, содержащая список способов, присутствующих в системе. Этот список можно редактировать аналогично списку областей, с помощью контекстного меню. Для выбора способа необходимо щелкнуть мышью по нужному элементу списка (см. Рисунок 22).

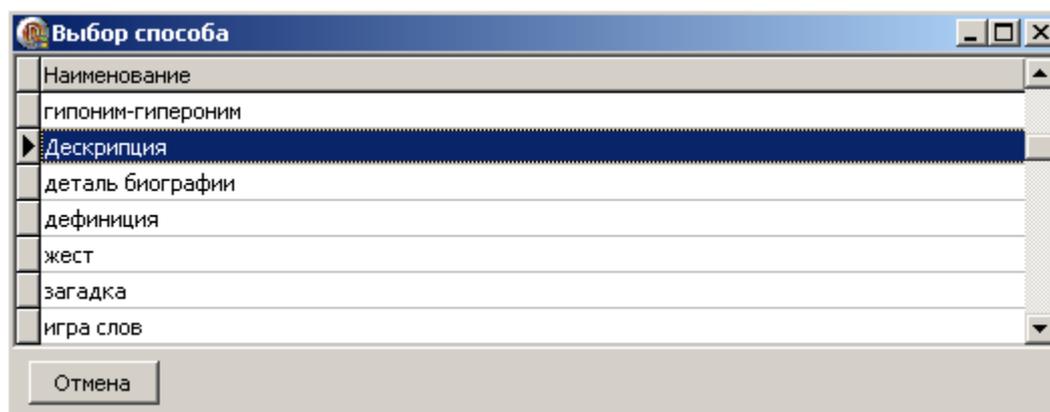


Рисунок 22 Редактирование способа

Последний элемент фигуры знания, который не был установлен – Функция.
Для этого используем пункт «Ретушь» контекстного меню поля «Функция».

В результате произведенных действий в тезаурусе присутствует новая фигура знания, а в главной форме можно наблюдать ее содержимое.

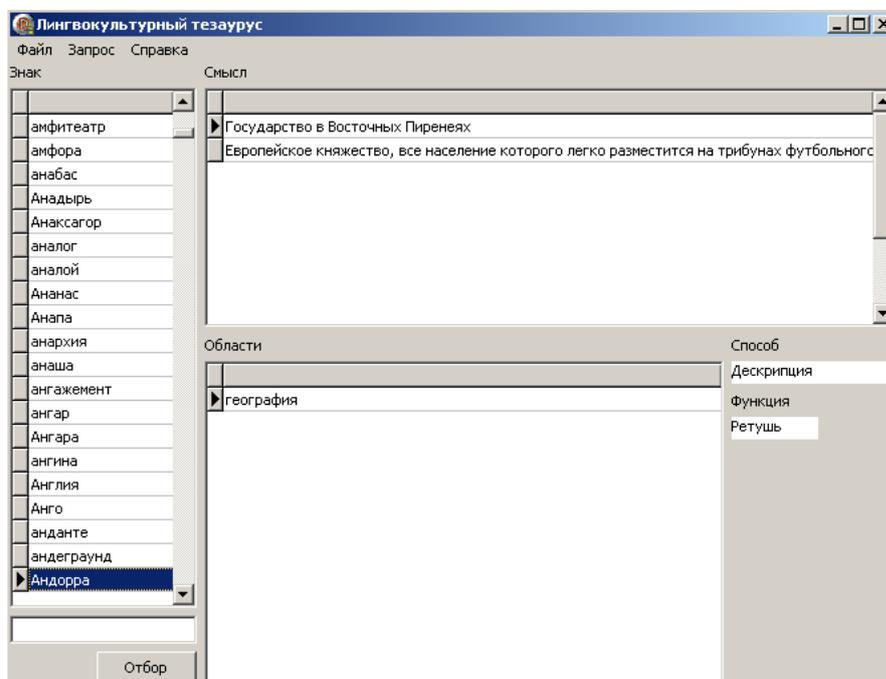


Рисунок 23 Новая фигура знания

4.1.2. Редактирование в табличном режиме

Главная форма обеспечивает полную функциональность в качестве инструмента редактирования базы данных фигур знаний.

Редактирование в табличном режиме предназначено для тех пользователей, которым привычнее и интуитивно более понятен такой режим представления фигур знания, вместо структурированного в главной форме.

Для перехода в табличный режим редактирования необходимо в главном меню выбрать пункт Запрос -> Редактирование.

Откроется форма, изображенная на рисунке ниже.

Знак	Смысл	Способ	Функция	Область
ассортимент	Товарное разнообразие	перифраза	Рецепт	торговля
Астана	Столица Казахстана	перифраза	Рецепт	география
астра	Любимый дачный цветок	Множество	Рецепт	ботаника
астра	По старинной легенде этот цветок вырос из пылинки, упавшей со звезды	Дескрипция	Рецепт	мифология
астра	Садовый махровый цветок	Множество	Рецепт	ботаника
Астрахань	Арбузный город в дельте Волги	Дескрипция	Рецепт	география
астролог	Мастер составлять гороскопы	суждение	Рецепт	астрономия, верс
астроном	Ученый, страдающий звездной болезнью	Метафора	Рецепт	астрономия
асфальт	Серое покрытие дорог	Дескрипция	Рецепт	дорога
атака	Наиболее решительный момент в поступательных действиях войск	дефиниция	Рецепт	война, армия
атака	Нападение	синоним	Рецепт	война, армия
атака	Стремительный натиск, сопровождаемый традиционным "ура!"	фрейм	Рецепт	война, армия
Атакама	Самая безводная пустыня на земле	Дескрипция	Ретушь	география
атаман	Казачий капитан	перифраза	Рецепт	война, армия
Атас	Хит группы "Любе"	Множество	Ретушь	поп-арт
ателье	Где работает киногероиня Ирины Купченк из фильма "Одинокая женщина желает познакомиться"	прец. текст	Ретушь	кино, tv
Атлант	Когда-то он держал небо на своих плечах	Дескрипция	Рецепт	мифология
Атлант	Первый шейный позвонок	Дескрипция	Ретушь	физиология, анат
атлас	Много карт под одной обложкой	Дескрипция	Рецепт	география
атлас	Много карт, но не колода	омоним	Рецепт	язык

Рисунок 24 Табличный режим редактирования

В нижней части формы расположены настройки сортировки. Данные, отображаемые в таблице, сортируются согласно выбранному параметру.

Центральную часть формы занимает таблица данных тезауруса. Каждая строка таблицы представляет собой одну фигуру знания. Для произведения изменений с отдельной фигурой знания, или базой тезауруса используется контекстное меню таблицы, имеющее следующий вид:

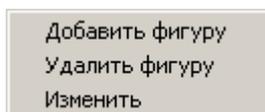


Рисунок 25 Контекстное меню

Пункт «Удалить фигуру» служит для удаления из таблиц базы данных тезауруса всех записей, относящихся к текущей фигуре знания.

Пункт «Добавить фигуру» служит для добавления фигур знания в базу данных тезауруса. При выборе данного пункта меню пользователю открывается форма создания фигуры знания, изображенная на рисунке ниже.

Фигура знания

Знак

Способ

Функция

Формула смысла

Области

Наименование

Рисунок 26 Добавление фигуры

Форма включает в себя пять элементов ввода, соответствующих составу фигуры знания:

- Поле ввода знака
- Список выбора способа
- Список выбора функции
- Поле ввода формулы смысла
- Список областей

Поле ввода знака обязательно для заполнения. В случае, если знак создаваемой фигуры знания уже присутствует в системе, будет использоваться имеющаяся запись таблицы «Знаки». Рядом с полем расположена кнопка «Выбрать знак». В случае, если пользователь создает фигуру для знака, уже присутствующего в системе, есть возможность не вводить его наименование, а выбрать из уже имеющихся. Форма выбора знака, открывающаяся при нажатии данной кнопки, представлена на рисунке ниже.

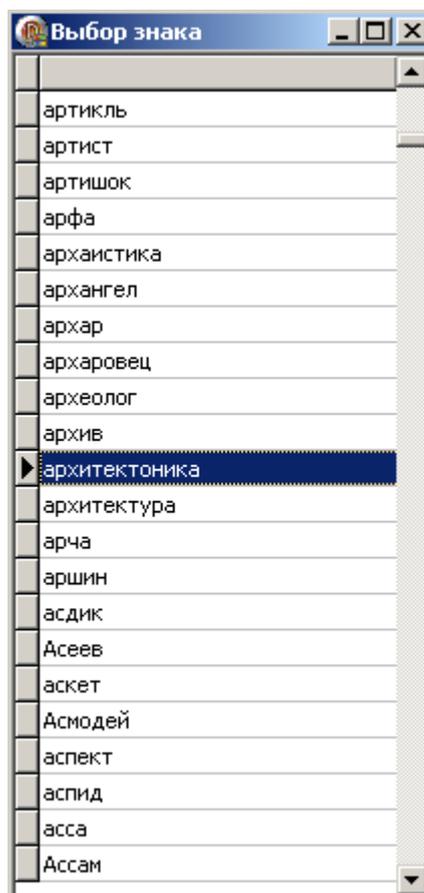


Рисунок 27 Выбор знака

Для выбора знака достаточно щелкнуть дважды по соответствующей строке таблицы. Либо закрыть окно, в случае отказа от выбора.

Список выбора способа содержит все способы, присутствующие в системе. Добавление новых способов через эту форму невозможен. Для этих целей необходимо воспользоваться главной формой системы.

Список выбора функции содержит возможные варианты функций фигуры знания: ретушь и рецепт.

Поле ввода формулы смысла представляет собой поле ввода текста. Для создания фигуры знания необходим ввод какого-либо текста в это поле.

Список областей представляет собой таблицу, в которой представлены области, к которым относится создаваемая фигура знания. Первоначально этот список пуст, затем он заполняется. Для добавления в него областей, необходимо вызвать контекстное меню списка и выбрать единственно-присутствующий в нем пункт «Изменить».

Выбор этого пункта открывает пользователю форму выбора областей, состоящую из двух таблиц. В таблице слева представлены все имеющиеся в

системе области. Для добавления к создаваемой фигуре областей, необходимо осуществить двойное нажатие на соответствующей записи таблицы. При этом область появится в таблице справа. В случае, если область была добавлена по ошибке, ее можно удалить двойным нажатием мышью на соответствующей записи правой таблицы.

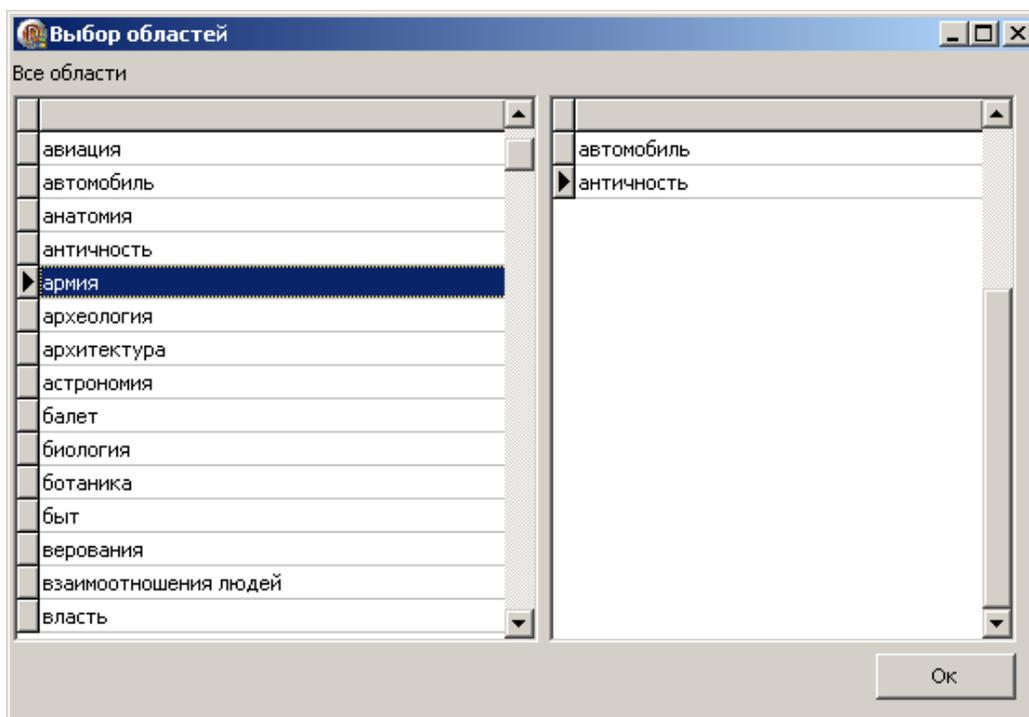


Рисунок 28 Выбор областей

Нажав кнопку «Ок», принимаем произведенные изменения. Закрытие формы отменяет изменения.

Для добавления фигуры знания необходимо нажать кнопку «Ок» формы добавления. Сама форма закроется, а в базе данных тезауруса будут сделаны необходимые изменения.

Для изменения имеющейся фигуры знания необходимо выбрать пункт «Изменить» контекстного меню (см. Рисунок 25) После этого откроется форма изменения фигуры знания, совпадающая по структуре с ранее описанной формой добавления фигур. Отличие формы изменения в том, что первоначально она не пуста, а заполнена согласно изменяемой фигуре знания. Внешний вид формы на показан ниже на примере изменения одной из фигур знания.

Знак	абзац	Выбрать знак
Способ	перифраза	
Функция	Рецепт	
Формула смысла	Красная строка	
Области	Наименование	
	письменный текст	

Рисунок 29 Изменение фигуры

Для фиксации изменений необходимо нажать кнопку «Ок». В случае отказа от изменений необходимо закрыть форму изменения фигуры знания.

- Поля порядка сортировки
- Поле количества найденных фигур
- Кнопка поиска
- Кнопка сохранения результатов

Рядом с полями ввода первых пяти параметров расположены элементы ввода «Выбор» (checkbox) , имеющие два режима: включен и выключен. Если это поле не активно, то параметр не будет учитываться в запросе вне зависимости от того, какое значение имеет рядом расположенное поле.

Поле свойств знака задает требования к знакам обираемых фигур. По умолчанию будут отобраны знаки, в состав которых входит подстрока, задаваемая данным полем. Для того, чтобы выбрать те фигуры, знаки которых полностью совпадают с содержимым поля, необходимо установить галочку в поле «Точное совпадение».

Поле свойств смысла задает требования к формуле смысла отбираемых фигур знания. Так же, как и в случае с полем свойств знаков, имеется возможность включить или выключить учет этого параметра в запросе, а также искать точные совпадения (по умолчанию ведется поиск подстроки).

Таблица свойств областей задает список областей, к какой-либо из которых должна принадлежать фигура знания, чтобы попасть в результаты запроса. Таблица может не учитываться в запросе, для чего есть элемент управления «Выбор» справа от таблицы.

Поле свойств способа задает требования к способу представления формулы смысла отбираемых фигур знания. Так же, как и в случае прочими свойствами фигуры знания, имеется возможность включить или выключить учет этого параметра в запросе.

Поле выбора функции принимает два возможных значения: ретушь и рецепт, согласно описанию предметной области. Пользователь может включить или выключить учет этого параметра в запросе.

Поля порядка сортировки соответствуют порядку сортировки результатов поиска, при этом чем выше параметр, тем раньше он расположен в порядке сортировки. Для изменения порядка необходимо вызвать контекстное меню полей сортировки, представленное на рисунке ниже.

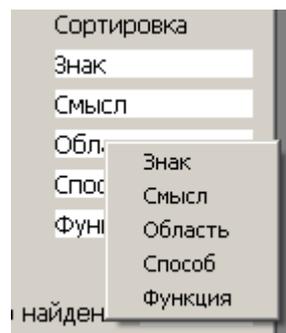


Рисунок 31 Контекстное меню сортировки

Меню предлагает выбрать содержимое текущей позиции сортировки. При этом текущее содержимое перемещается на место, которое занимало новое значение текущей позиции. То есть новое и прежнее значения текущей позиции меняются местами в списке. Данный порядок сортировки будет учитываться как при выводе результатов, так и при сохранении результатов в табулированный файл.

Поле количества найденных фигур является индикатором, отражающим количество фигур знания, удовлетворивших условиям поиска.

Кнопка поиска служит для формирования SQL-запроса к базе данных, его последующего выполнения и вывода результатов пользователю. Кнопка сохранения результатов до выполнения запроса не активна. После его выполнения нажатие кнопки приведет к сохранению результатов поиска в табулированный файл «выгрузка.txt»

На рисунке ниже приведен пример выполнения отбора фигур знания с областями «Античность», «Археология» или «Верования», имеющих функцию Ретушь. Результаты отсортированы в порядке Область, Знак, Смысл, Способ, Функция.

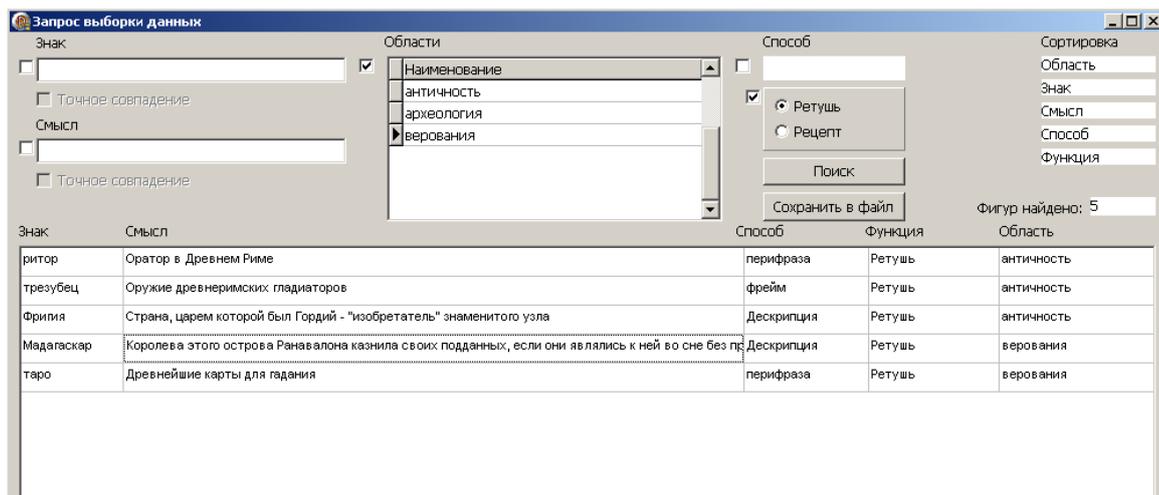


Рисунок 32 Результаты поиска

После выполнения поиска кнопка «Сохранить в файл» стала активной. Воспользовавшись ей, получим результаты поиска в виде табулированного файла, который можно открыть текстовым редактором, или Microsoft Excel.

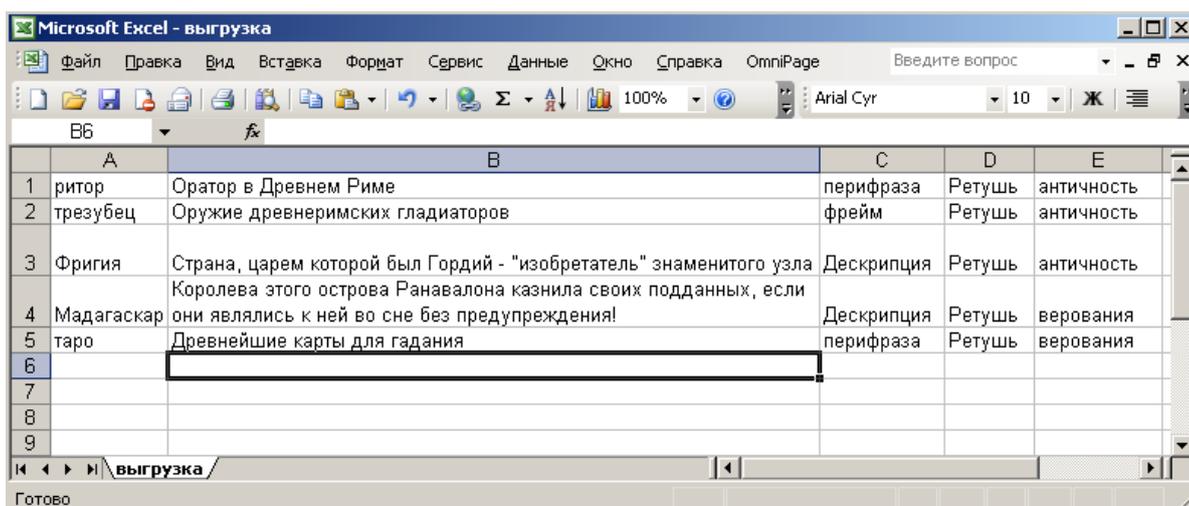


Рисунок 33 Результаты выгрузки

Для выхода из формы поиска закройте форму.

4.3. Моделирование работы когнайзера

Моделирование пассивного режима работы когнайзера осуществляется из главной формы. Для этого необходимо выбрать фигуру знания, для которой будет производиться расчет. Фигура знания в главной форме однозначно идентифицируется через формулу смысла, поэтому для начала работы когнайзера необходимо вызвать контекстное меню списка формул смысла, сделав в этом списке активной нужную позицию. Действие проиллюстрировано на рисунке ниже.

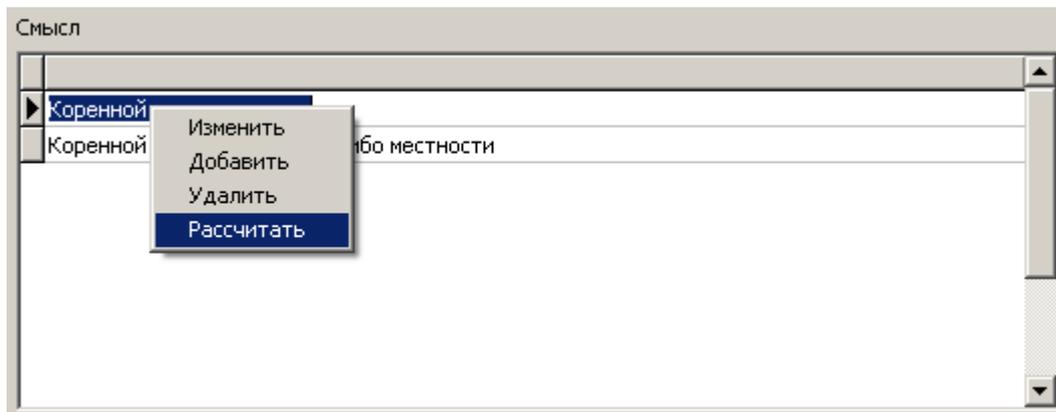


Рисунок 34 Выбор расчета в меню формулы смысла

В результате выбора пункта «Рассчитать» пользователю открывается форма когнайзера.

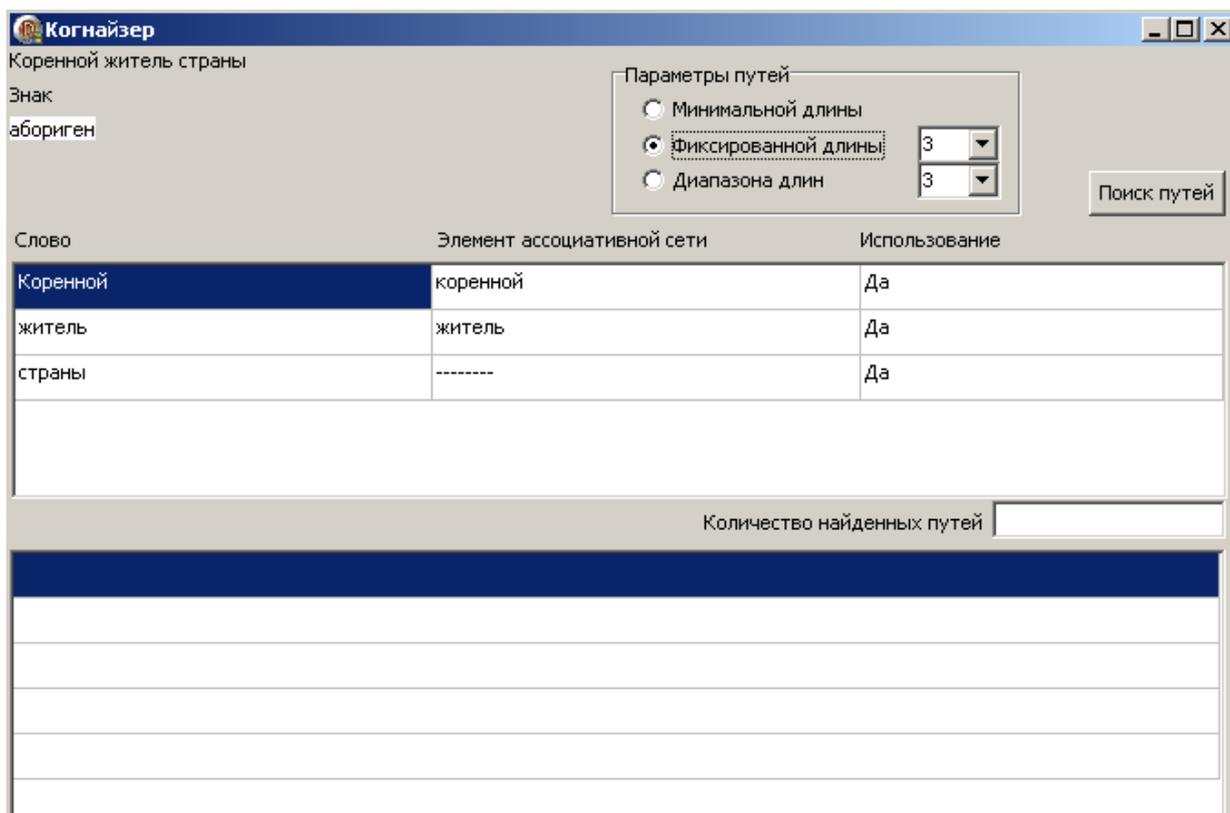


Рисунок 35 Начало работы когнайзера

В верхней части окна расположено написание исходной формулы смысла. Ниже расположен знак фигуры знания. Он используется в качестве конечной вершины искомых путей. Очевидно, что в случае, если знак фигуры знания отсутствует в системе, пути искать бессмысленно. Поэтому в случае, если знак фигуры отсутствует среди реакций ассоциативной сети, он обозначен красным цветом. Если это произошло, либо пользователь решил принять другой знак за конечный, необходимо вызвать контекстное меню поля Знак и выбрать в нем единственный пункт «Изменить». После этого открывается окно выбора знака, изображенное на рисунке ниже.

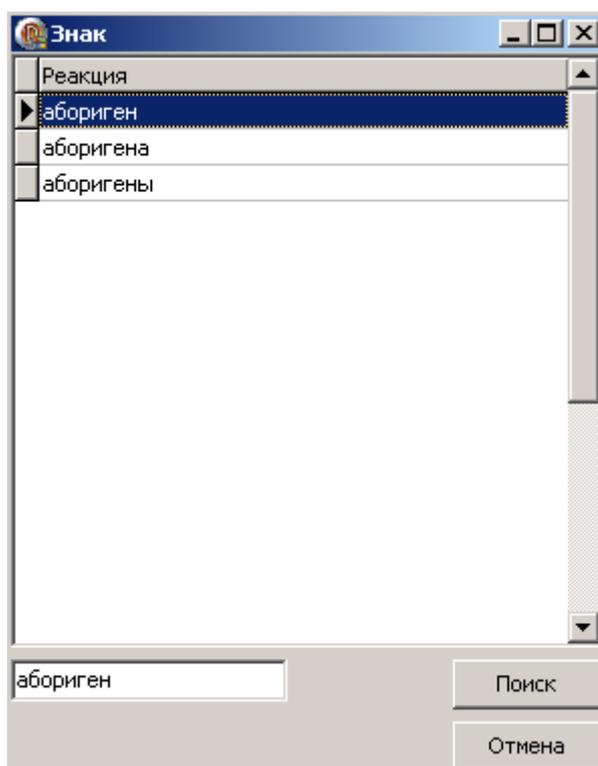


Рисунок 36 Выбор знака

В верхней части окна расположен список вершин, подходящих пользователю. Для выбора других вершин необходимо ввести в поле слева начало написания необходимой вершины и нажать кнопку «Поиск» при этом в списке допустимых вершин отобразятся те реакции ассоциативной сети, которые имеют совпадающее с запросом начало. Двойное нажатие соответствующей строки таблицы производит выбор знака в качестве конечной вершины пути и закрытие формы выбора. Для закрытия формы без изменений служит нажатие кнопки «Отмена».

В правой части формы когнайзера расположены элементы установки ограничений на длину путей.

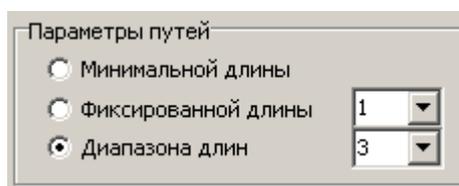


Рисунок 37 Установка длин путей

Группа переключателей состоит из трех позиций:

- Пути минимальной длины
- Пути фиксированной длины
- Пути диапазона длин

В случае поиска путей фиксированной длины, список длин путей, расположенный с правой стороны, позволяет установить длину путей в диапазоне от 1 до 5.

При поиске путей диапазона длин, верхний выпадающий список указывает нижнюю границу длины путей, а нижний – верхнюю.

В средней части окна расположена таблица найденных в формуле смысла знаков. Таблица включает в себя поля:

- Слово

Выделенная в отдельную словоформу часть формулы смысла

- Элемент ассоциативной сети

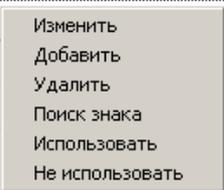
Соответствующая слову вершина ассоциативной сети. Пока элемент ассоциативной сети не сопоставлен слову, в работе когнайзера он учитываться не будет.

- Использование

Позволяет принудительно исключить из рассмотрения какое-либо слово.

Для настройки исходных вершин поиска таблица имеет контекстное меню.

Слово	Элемент ассоциативной сети	Использование
Коренной	коренной	Да
житель	житель	Да
страны	-----	Да



The context menu is open over the 'коренной' cell in the 'Элемент ассоциативной сети' column. It contains the following options: Изменить, Добавить, Удалить, Поиск знака, Использовать, and Не использовать.

Рисунок 38 Меню редактирования исходных вершин

Пункт меню «Изменить» предлагает пользователю ввести новое обозначение слова с помощью формы, приведенной ниже.



The dialog box is titled 'Изменить знак' and contains a text input field with the word 'Коренной'. At the bottom, there are two buttons: 'Отмена' and 'Ок'.

Рисунок 39 Изменение слова

После изменения словоформы поиск элемента ассоциативной сети выполняется снова.

Пункт «Добавить» выполняет включение в поиск заданной пользователем словоформы путем аналогичной изменению формы.

«Удалить» позволяет удалить словоформу из таблицы и таким образом не включать ее в рассмотрение когнайзера.

«Поиск знака» необходим в случае, если изменение словоформы не приводит к поиску подходящего знака. В этом случае возможен подбор подходящего элемента сети с помощью формы, аналогичной показанной на рисунке «Выбор знака». Отличие в списке доступных, вершин, которые выбираются из числа стимулов, а не реакций.

В третьей колонке таблицы показано использование словоформ для поиска путей. На рисунке выше все словоформы используются для поиска. Для каждой записи таблицы можно отключить ее использование в работе когнайзера с помощью пункта меню «Не использовать». Включение осуществляется пунктом «Использовать».

Нажатие кнопки «Поиск путей» ведет к запуску алгоритма когнайзера и поиску путей, при завершении которого пользователю представляются результаты в виде таблицы в нижней части окна.

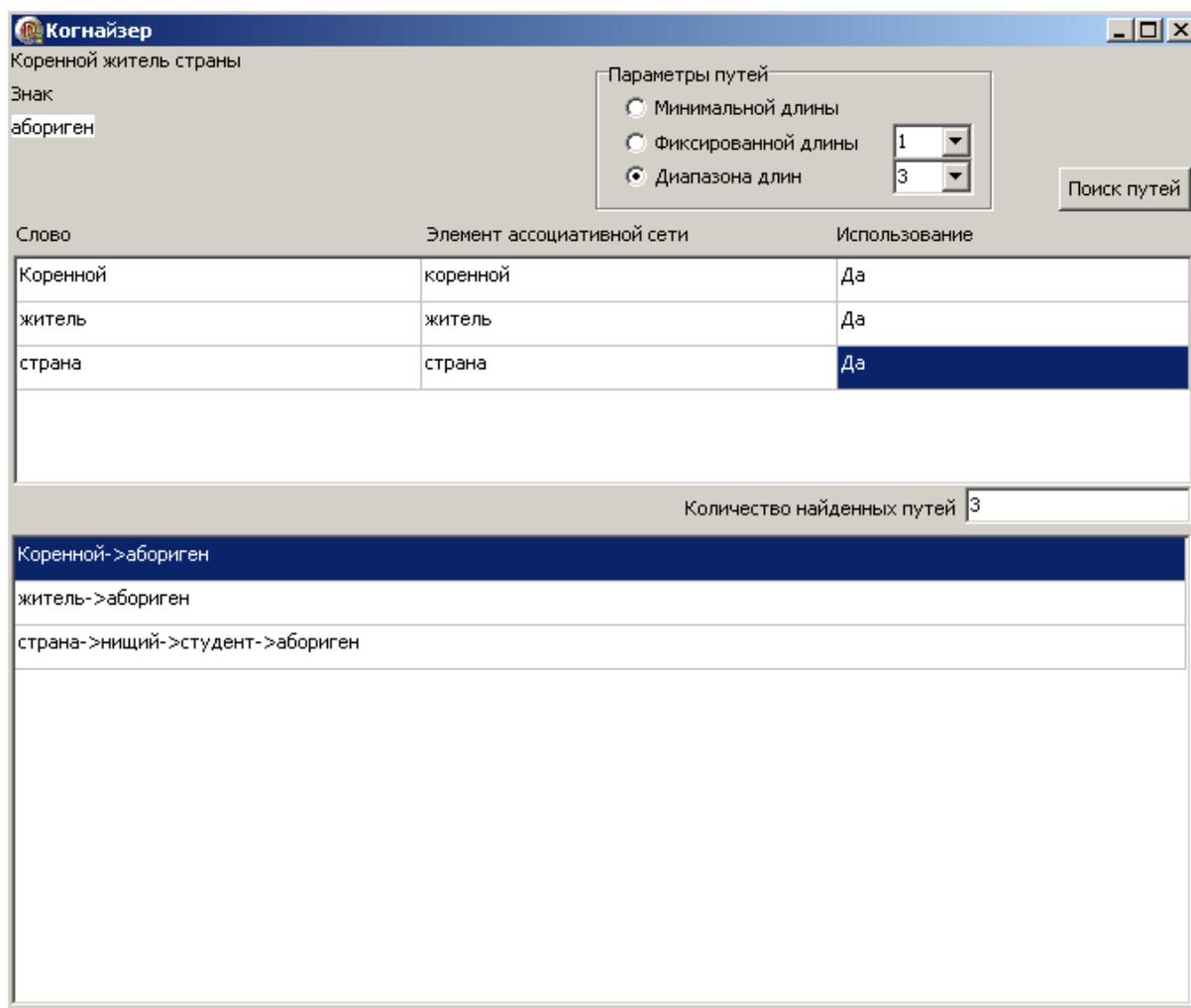


Рисунок 40 Результат расчета

На рисунке представлен результат расчета фигуры знания

Абориген – Коренной житель страны.

5. Эргономика и охрана труда

5.1. Характеристика влияния условий и режима труда на утомляемость и снижение производительности труда.

На сегодняшний день работа конструкторов, программистов, расчетчиков в основном связана с использованием персональных компьютеров (ПЭВМ), поэтому в дальнейшем будем рассматривать автоматизированное рабочее место.

При работе с персональным компьютером важную роль играет соблюдение правильного режима труда и отдыха. В противном случае появляется значительное напряжение зрительного аппарата, головная боль, раздражительность, нарушение сна, усталость и болезненные ощущения в глазах, пояснице, в области шеи и руках.

В таблице 1 представлены сведения о перерывах, которые необходимо делать при работе на компьютере, в зависимости от продолжительности рабочей смены, видов и категорий трудовой деятельности с ПЭВМ (установлено в СанПиН 2.2.2/2.4-1340-03 «Гигиенические требования к персональным электронно-вычислительным машинам и организации работы»).

Таблица 38 Время регламентированных перерывов при работе на компьютере

Категория работы с ПЭВМ	Уровень нагрузки за рабочую смену при видах работы с ПЭВМ			Суммарное время регламентированных	
	Группа А, кол. знаков	Группа Б, кол. знаков	Группа В, часов	При 8-час. смене	При 12-час. смене
I	до 20 000	до 15 000	до 2,0	30	70
II	до 40 000	до 30 000	до 4,0	50	90
III	до 60 000	до 40 000	до 6,0	70	120

Данные правила обязательны к исполнению при использовании вышеуказанных Санитарных правил и норм. Если же фактические условия труда не соответствуют требованиям Санитарных правил и норм, время регламентированных перерывов следует увеличить на 30%.

Основным фактором, вызывающим утомление, является интегральная экстенсивностная напряженность деятельности - нагрузка. Нагрузка характеризуется не только своей абсолютной величиной, но и рядом других показателей:

- интенсивность нагрузки, то есть ее распределение во времени;
- статический или динамический характер нагрузки;
- постоянный и ритмический характер нагрузки;
- функциональные возможности организма человека.

Функциональные возможности организма зависят от антропометрических параметров человека, а также от степени тренированности.

Условия труда оказывают значительное влияние на утомляемость и могут способствовать ее развитию. Сами по себе в допустимых пределах они не вызывают утомления, но являются фоном для воздействия нагрузки. Условия труда включают в себя множество проявлений окружения работника:

- 1) микроклимат
- 2) использование техники
- 3) нарушение режима труда и отдыха.

К первой группе относятся параметры окружающего воздуха как по его составу, так и по температурным показателям и показателям давления: пониженное содержание кислорода во вдыхаемом воздухе или изменение его парциального давления, повышенное содержание углекислого газа, некомфортная температура воздуха, влажность, изменение барометрического давления и запыленность.

Вторая группа параметров связана с использованием технических средств, что может привести к разносторонним изменениям условий работы. Среди причин, входящих в эту группу, следует назвать действие механических сил, ведущих к вибрации, тряске, воздействию электромагнитных колебаний, шумов и ультразвука, изменение освещенности, неудобство рабочей позы и многое другое, с чем приходится сталкиваться людям, работающим с ПЭВМ.

К третьей группе относятся факторы, связанные в основном с нарушением режима труда и отдыха: недостаточность времени на восстановление сил после утомления, неправильное использование перерывов между работой, непродуманное планирование работы и отдыха.

В соответствии с СанПиН 2.2.2/2.4-1340-03 в производственных помещениях, в которых работа с использованием ПЭВМ является вспомогательной, температура,

относительная влажность и скорость движения воздуха на рабочих местах должны соответствовать действующим санитарным нормам микроклимата производственных помещений.

В производственных помещениях, в которых работа с использованием ПЭВМ является основной (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.) и связана с нервно-эмоциональным напряжением, должны обеспечиваться оптимальные параметры микроклимата для категории работ 1а и 1б в соответствии с действующими санитарно-эпидемиологическими нормативами микроклимата производственных помещений. На других рабочих местах следует поддерживать параметры микроклимата на допустимом уровне, соответствующем требованиям указанных выше нормативов.

В помещениях, оборудованных ПЭВМ, проводится ежедневная влажная уборка и систематическое проветривание после каждого часа работы на ПЭВМ.

Уровни положительных и отрицательных аэроионов в воздухе помещений, где расположены ПЭВМ, должны соответствовать действующим санитарно-эпидемиологическим нормативам.

Содержание вредных химических веществ в воздухе производственных помещений, в которых работа с использованием ПЭВМ является вспомогательной, не должно превышать предельно допустимых концентраций вредных веществ в воздухе рабочей зоны в соответствии с действующими гигиеническими нормативами. Содержание вредных химических веществ в производственных помещениях, в которых работа с использованием ПЭВМ является основной (диспетчерские, операторские, расчетные, кабины и посты управления, залы вычислительной техники и др.), не должно превышать предельно допустимых концентраций загрязняющих веществ в атмосферном воздухе населенных мест в соответствии с действующими гигиеническими нормативами.

Содержание вредных химических веществ в воздухе помещений, предназначенных для использования ПЭВМ во всех типах образовательных учреждений, не должно превышать предельно допустимых среднесуточных концентраций для атмосферного воздуха в соответствии с действующими санитарно-эпидемиологическими нормативами.

Объем помещений, в которых размещены работники вычислительных центров, не должен быть меньше $19,5 \text{ м}^3/\text{человека}$ с учетом максимального числа

одновременно работающих в смену. Нормы подачи свежего воздуха в помещения, где расположены компьютеры, приведены в таблице ниже.

Таблица 39 Параметры микроклимата для помещений, где установлены компьютеры

Температура, С°	Относительная влажность, %	Абсолютная влажность, г/м3	Скорость движения
19	62	10	<0.1
20	58	10	<0.1
21	55	10	<0.1

Для обеспечения комфортных условий используются как организационные методы (рациональная организация проведения работ в зависимости от времени года и суток, чередование труда и отдыха), так и технические средства (вентиляция, кондиционирование воздуха, отопительная система).

Акустический шум оказывает вредное воздействие на организм человека. Работа в условиях длительного шумового воздействия ведет к утомлению, раздражительности, головным болям, понижению аппетита, головокружению, боли в ушах. Под воздействием шума снижается концентрация внимания, нарушаются физиологические функции, появляется усталость в связи с повышенными энергетическими затратами и нервно-психическим напряжением, ухудшается речевая коммутация. Все это снижает работоспособность человека и его производительность, качество и безопасность труда. Длительное воздействие интенсивного шума (более 80 дБ) на слух человека приводит к его частичной или полной потере.

В соответствии с СанПиН 2.2.2/2.4-1340-03. В производственных помещениях при выполнении основных или вспомогательных работ с использованием ПЭВМ уровни шума на рабочих местах не должны превышать предельно допустимых значений, установленных для данных видов работ в соответствии с действующими санитарно-эпидемиологическими нормативами.

При выполнении работ с использованием ПЭВМ в производственных помещениях уровень вибрации не должен превышать допустимых значений

вибрации для рабочих мест (категория 3, тип "в") в соответствии с действующими санитарно-эпидемиологическими нормативами.

Шумящее оборудование (печатающие устройства, серверы и т.п.), уровни шума которого превышают нормативные, должно размещаться вне помещений с ПЭВМ.

Уровень шума на рабочем месте математиков-программистов и операторов видеоматериалов не должен превышать 50 дБА, а в залах обработки информации на вычислительных машинах - 65 дБА. Для снижения уровня шума стены и потолок помещений, где установлены компьютеры, могут быть облицованы звукопоглощающими материалами. Уровень вибрации в помещениях вычислительных центров может быть снижен путем установки оборудования на специальные виброизоляторы.

Электромагнитное и ионизирующее излучения.

Исчерпывающих данных относительно опасности воздействия излучения от мониторов на работающих с компьютерами не существует и исследования в этом направлении продолжаются.

Допустимые значения параметров неионизирующих электромагнитных излучений от монитора компьютера представлены в табл. 3.

Максимальный уровень рентгеновского излучения на рабочем месте оператора компьютера обычно не превышает 10 мкбэр/ч, а интенсивность ультрафиолетового и инфракрасного излучений от экрана монитора лежит в пределах 10...100 мВт/м².

Таблица 40 Допустимые значения параметров неионизирующих электромагнитных излучений

Наименование параметра	Допустимые
Напряженность электрической составляющей электромагнитного поля на расстоянии 50 см от поверхности видеомонитора	10 В/м
Напряженность магнитной составляющей электромагнитного поля на расстоянии 50 см от поверхности видеомонитора	0,3 А/м
Напряженность электростатического поля не должна превышать: для взрослых пользователей	20 кВ/м

для детей дошкольных учреждений и учащихся средних специальных и высших учебных заведений	15 кВ/м
---	---------

Для снижения воздействия этих видов излучения рекомендуется применять мониторы с пониженным уровнем излучения, устанавливать защитные экраны, а также соблюдать регламентированные режимы труда и отдыха.

5.2. Характеристика условий труда.

Работа с ПЭВМ характеризуется значительным умственным напряжением и нервно-эмоциональной нагрузкой операторов, высокой напряженностью зрительной работы и достаточно большой нагрузкой на мышцы рук при работе с клавиатурой ЭВМ. Большое значение имеет рациональная конструкция и расположение элементов рабочего места, что важно для поддержания оптимальной рабочей позы человека-оператора.

В процессе работы с компьютером необходимо соблюдать правильный режим труда и отдыха. В противном случае у персонала отмечаются значительное напряжение зрительного аппарата с появлением жалоб на неудовлетворенность работой, головные боли, раздражительность, нарушение сна, усталость и болезненные ощущения в глазах, в пояснице, в области шеи и руках.

В соответствии со СанПиН 2.2.2/2.4-1340-03 все виды трудовой деятельности, связанные с использованием компьютера, разделяются на три группы:

группа А: работа по считыванию информации с экрана ВДТ или ПЭВМ с предварительным запросом;

группа Б: работа по вводу информации;

группа В: творческая работа в режиме диалога с ЭВМ.

5.3. Эргономические требования к рабочему месту.

В соответствии с требованиями ГОСТ 12.2.032-78 в УЭС используются рабочие столы с высотой рабочей поверхности 725 мм, а также рабочие кресла с подъемно-поворотным устройством.

Пространственные параметры рабочего места

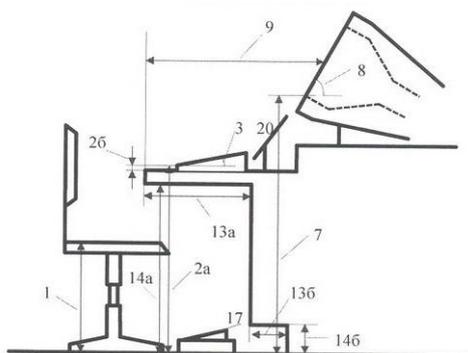


Рисунок 41 Рабочее место оператора

Усл. обозн.	Параметр	Способ измерения параметра	Степень необход. регул.	Знач. парам., мм	Примечание
1	Высота сиденья	От пола до верхней плоскости сиденья	Необходима	400-500	Регулируемый параметр
2а	Высота сиденья	От пола до нижнего ряда клавиатуры	Желательна	600-750 620-700	Диапазон регулировки Для нерегулируемой высоты
2б	То же (от уровня стола)	От базовой поверхности и до нижнего	Желательна	Около 20	При необходимости и клавиатуру можно

		ряда клавиатуры			встроить в поверхность стола
3	Угол наклона клавиатуры	От горизонтали	Возможна	(7-15)	Зависит от высоты клавиатуры
4	Ширина основной клавиатуры	Определяется оптимальной зоной моторного поля согласно ГОСТ 12.2.032-78	Возможно	Не более 400	Дополнительные блоки и клавиши, нечасто используемые, могут быть размещены в зоне легкой досягаемости за пределами основной клавиатуры
5	Глубина основной клавиатуры	То же	Возможно	Не более 200	То же
6	Удаленность	От переднего края стола до нижнего	Возможна	Не менее 80-100	При неподвижном креплении клавиатуры
7	Высота экрана	От пола до нижнего края окна	Желательна	950-1000 970-1050	Диапазон регулировки При отсутствии регулировки
8	Угол наклона экрана	От вертикали	Желательна	(0-30 =15) опт	Зависит от высоты экрана относительно глаз

9	Удаленность экрана от края стола	От переднего края стола до экрана дисплея	Желательна	500-700 500	Диапазон регулировки При отсутствии регулировки
10	Высота поверхности для записи	От пола	Желательна	670-860 725	Диапазон регулировки При отсутствии регулировки
11	Площадь поверхности для записи	Определяется оптимальной зоной моторного поля	Желательна	600x400 900x600	Прямоугольной формы Треугольной формы
12	Угол наклона поверхности для записи	От горизонтали	Возможна	(0-10)	Для постоянной письменной работы Для периодических записей
13а	Глубина пространства для ног на уровне: коленей	От переднего края стола	Возможна	Не менее 400	
13б	ступней			Не менее 600	
	Высота пространства для ног на уровне:	От пола	Возможна	Не менее	Для нерегулируемой высоты стола

14а	коленей			500	
14б	ступней			Не менее 250	
15а	Ширина пространства для ног на уровне: коленей		Возможна	Не менее 500	Параметр необходим при наличии боковых тумб или аналогичного оборудования
15б	ступней			Не менее 250	
16	Высота подставки для ног	От пола до передней части подставки	Желательна	50-130 80	Регулируемая величина Нерегулируемая
17	Угол наклона подставки для ног	От горизонтали	Желательна	(0-25) (15)	Диапазон регулировок Нерегулируемая величина
18	Глубина подставки для ног	От переднего до заднего края	Желательна	400	
19	Ширина подставки для ног		Желательна	300	
20	Пюпитр – подставка для документов	От горизонтали	Желательна	15 - 20	

5.4. Психологические требования к рабочему месту.

Для реализации потенциала оператора необходимо, чтобы его труд был хорошо организован, чтобы у него не было перерывов в работе по организационно-техническим причинам, чтобы поручаемая ему работа соответствовала профессиональной подготовке и уровню квалификации, чтобы работник не отвлекался на выполнение несвойственных ему функций (работ), тем более не соответствующих уровню квалификации, чтобы ему были созданы нормальные санитарно-гигиенические условия труда, обеспечивающие и нормальный уровень интенсивности труда и т.д.

Большую роль играет и социально-психологическая обстановка, способствующая взаимодействию исполнителей в процессе работы, появлению стимулов к высокопроизводительной и эффективной работе. Важным представляется также надлежащее выполнение работником своих производственных и служебных обязанностей, строгое соблюдение им дисциплины (трудовой, производственной, технологической), трудовая активность и творческая инициатива.

Под воздействием шума нарушения в работе ряда органов и систем организма человека могут вызвать негативные изменения в эмоциональном состоянии человека вплоть до стрессовых.

Под эмоциональными понимаются состояния, вызванные переживанием человеком его отношения к внешнему миру и к самому себе и характеризующиеся изменениями количественных и качественных параметров ответов на сигналы внешней среды. Таким образом, эмоциональное состояние тесно связано с индивидуальной семантической значимостью поступающей к человеку информации и являются как бы коррекцией, вносимой человеком в ответ, определяемый только информационной структурой раздражителя.

При изучении эмоциональных реакций различают две его формы - эмоциональное напряжение и эмоциональную напряженность. Эмоциональное напряжение характеризует степень мобилизации функций организма для наиболее успешного выполнения той или иной деятельности и связано с волевым актом, направленным на эту деятельность. В тех случаях, когда наступает динамическое рассогласование между объективной значимостью ситуации и ее субъективной

оценкой и появляются связанные с этим отрицательные изменения в двигательных и психических функциях, наступает состояние эмоциональной напряженности. При этом наблюдается и снижение устойчивости ряда психических функций. Момент перехода эмоционального напряжения в эмоциональную напряженность определяет так называемую эмоциональную устойчивость. Чем меньше эмоциональная устойчивость, тем скорее при меньших значениях эмоционального фактора развивается состояние эмоциональной напряженности. Эмоциональная устойчивость является показателем, очень тесно связанным с таким свойством личности, как уровень тревожности, она очень низка у лиц с высоким уровнем тревожности.

Следующее качество - эмоциональная возбудимость - определяет быстроту развития того или иного эмоционального состояния, т.е. это качество очень близко к тому, которое характеризует эмоциональную устойчивость.

Факторы, определяющие развитие эмоциональных состояний:

На основании данного представления о сущности эмоциональных реакций становится ясным, что их развитие определяют две группы факторов - внешние и внутренние.

Внешние эмоциогенные факторы:

К ним относятся прежде всего так называемые экстремальные факторы, т.е. такие, физические или информационные характеристики которых ведут к развитию крайней степени напряжения физиологических и психологических функций с полным истощением всех физиологических резервов.

К этой же группе факторов относятся и те, которые характеризуются очень высокой значимостью, хотя сами по себе факторы не являются экстремальными. Знак возникающей эмоциональной реакции и сила ее развития в этом случае, как правило, определяется сочетанием ряда внутренних по отношению к человеку факторов.

Внутренние эмоциогенные факторы:

Сами по себе эти факторы не являются эмоциогенными, они лишь тому или иному внешнему фактору необходимую степень эмоциональности. К этим факторам относятся такие, как характеристики нервной деятельности, темперамент, уровень тревожности, ригидность личности, и т.п. - они, как правило, определяют уровень реакции.

Можно активно управлять поведением человека - оператора, вводя те или иные эмоциогенные факторы. Например, для поддержания высокого уровня бдительности при монотонной работе наблюдения и контроля воздействие на эмоциогенную сферу является наиболее эффективным.

Если в работе возникает эмоциональная напряженность, то здесь усилия должны быть направлены на перевод этого состояния в состояние эмоционального напряжения.

В производственной среде цвет используется как средство информации и ориентации, как фактор психологического комфорта и как композиционное средство. Цвет оказывает влияние на работоспособность человека, на утомление, ориентировку, реакцию. Холодные цвета (голубой, зеленый, желтый) действуют успокаивающе на человека, теплые цвета (красный, оранжевый) действуют возбуждающе. Темные цвета оказывают угнетающее действие на психику.

При выборе цвета, цветовом оформлении интерьера нужно руководствоваться указаниями по рациональной цветовой отделке поверхностей производственных помещений и технологического оборудования ГОСТ 26568-85 и ГОСТ 12.4.026-76 ССБТ.

Цветовое решение интерьера характеризуется цветовой гаммой, цветовым контрастом, количеством цвета и коэффициентами отражения. Цветовая гамма - это совокупность цветов, принятая для цветового решения интерьера. Она может быть теплой, холодной и нейтральной. Цветовой контраст - это мера различия цветов по их яркости и цветовому тону. Цветовой контраст может быть большим, средним и малым.

Количество цвета - это степень цветового ощущения, зависящая от цветового тона, насыщенности цвета объекта и фона, от соотношения их яркостей и угловых размеров.

При выборе цветового решения интерьеров нужно учитывать категорию работы, ее точность, санитарно-гигиенические условия. Значительная роль в интерьере принадлежит выбору коэффициентов отражения (Р) поверхностей.

Потолки помещений окрашиваются в белый цвет или близкие к белому цвету. В светлые тона окрашиваются перекрытия. Нижняя часть стен окрашивается в спокойные тона (светло-зеленый, светло-синий).

5.5. Расчет освещенности рабочего места.

Выбираем светильники типа ОД с газоразрядными лампами;

Тип проводки - закрытая в строительных конструкциях под штукатуркой, провода - АППВ, выключатель нормального исполнения;

Светильники расположены параллельными рядами;

Характеристика выполняемой работы - разряд IV, полразряд – «в» (контраст - большой, фон - светлый). Минимальная освещённость от комбинированного освещения $E_{min.K} = 400$ лк, общее освещение $E_{min.O} = 200$ лк.

Система освещения - комбинированная: общее равномерное плюс местное;

Потребная освещённость при комбинированном освещении газоразрядными лампами от светильников общего освещения $E_{O.K} = 200$ лк, от местного – $E_{M.K} = 150$ лк; Необходимый коэффициент запаса $K_3 = 1.6$;

Наиболее выгодное отношение расстояния между светильниками L_{CB} к высоте

подвески светильников h_{CB} : $\gamma = L_{CB} / h_{CB} = 1.6$;

$L_{CB.д} = \gamma * h_{CB} = 1.6 * 2.20 = 3,52$ м; $L_{CB.ш} = 1.2$ м;

Расстояние между светильниками по ширине примем равным длине светильника плюс 0.05 м;

Расстояние от стены до первого ряда светильников:

$L_1 = 0.3 * L_{CB.д} = 0.3 * 3.52 = 1.056$ м;

Расстояние между крайними рядами по ширине помещения:

$L_2 = b - 2 * L_1 = 8 - 2 * 1.056 = 5.89$ м;

Число рядов, которое можно расположить между крайними рядами по ширине помещения:

$N_{CB.ш} = L_2 / L_{CB.ш} - 1 = 5.89 / 1.2 - 1 = 3.9$;

Общее число рядов светильников по ширине:

$N_{CB.ш.O} = N_{CB.ш} + 2 = 4 + 2 = 6$;

Расстояние между крайними рядами светильников по длине помещения:

$L_3 = a - 2 * L_1 = 9 - 2 * 1.056 = 6,888$ м;

Число светильников, которое можно расположить между крайними рядами по

длине: $N_{\text{СВ.Д}} = L_3 / L_{\text{СВ.Д}} - 1 = 6,888 / 3,52 - 1 = 1;$

Общее число рядов светильников по длине:

$N_{\text{СВ.Д.О}} = N_{\text{СВ.Д}} + 2 = 1 + 2 = 3;$

Общее число рядов светильников, которые необходимо установить по длине и ширине: $N_{\text{СВ.ОБЩ}} = N_{\text{СВ.Д.О}} + N_{\text{СВ.Ш.О}} = 6 * 3 = 18;$

Коэффициенты отражения от стен ($\rho_{\text{СТ}}$) и потолков ($\rho_{\text{ПОТ}}$) - по окраске стен и потолков:

$\rho_{\text{СТ}} = 56\%, \rho_{\text{ПОТ}} = 73\%;$

Коэффициент z учитывающий равномерность освещения в зависимости от типа светильников и отношения γ : $z = 1.15;$

Площадь пола освещаемого помещения:

$S_{\text{П}} = a * b = 9 * 8 = 72 \text{ кв.м};$

По длине a и ширине b помещения, и высоте подвески светильников $h_{\text{СВ}}$ находим

показатель помещения

$\phi = (a * b) / (a + b) * h_{\text{СВ}} = 72 / (2.2 * 17) = 1.93;$

Коэффициент использования светового потока: $\eta_{\text{И}} = 0.5;$

Расчётный (потребный) световой поток одной лампы:

$F_{\text{Л.РАСЧ}} = (E_{\text{О.К}} * K_3 * z * S_{\text{П}}) / (N_{\text{СВ.ОБЩ}} * \eta_{\text{И}}) = 150 * 1.6 * 1.15 * 72 / (18 * 0.5) = 2008 \text{ лм};$

По напряжению в сети $U_{\text{С}}$ и световому потоку одной лампы $F_{\text{Л.РАСЧ}} = 2008 \text{ лм}$ по справочным таблицам (ГОСТ 2239-70) определяем необходимую мощность электролампы ЛД40-4 $W_{\text{Л}} = 40 \text{ Вт}$. В каждом светильнике имеется лампа ЛД40-4 со световым потоком $F_{\text{Л.ТАБЛ}} = 2340 \text{ лм};$

Действительная освещённость:

$E_{\text{ДЕЙСТВ}} = F_{\text{Л.ТАБЛ}} * N_{\text{СВ.ОБЩ}} * \eta_{\text{И}} / (K_3 * z * S_{\text{П}}) = 2340 * 18 * 0.5 / (1.6 * 1.15 * 72) = 158,97 \text{ лк.}$

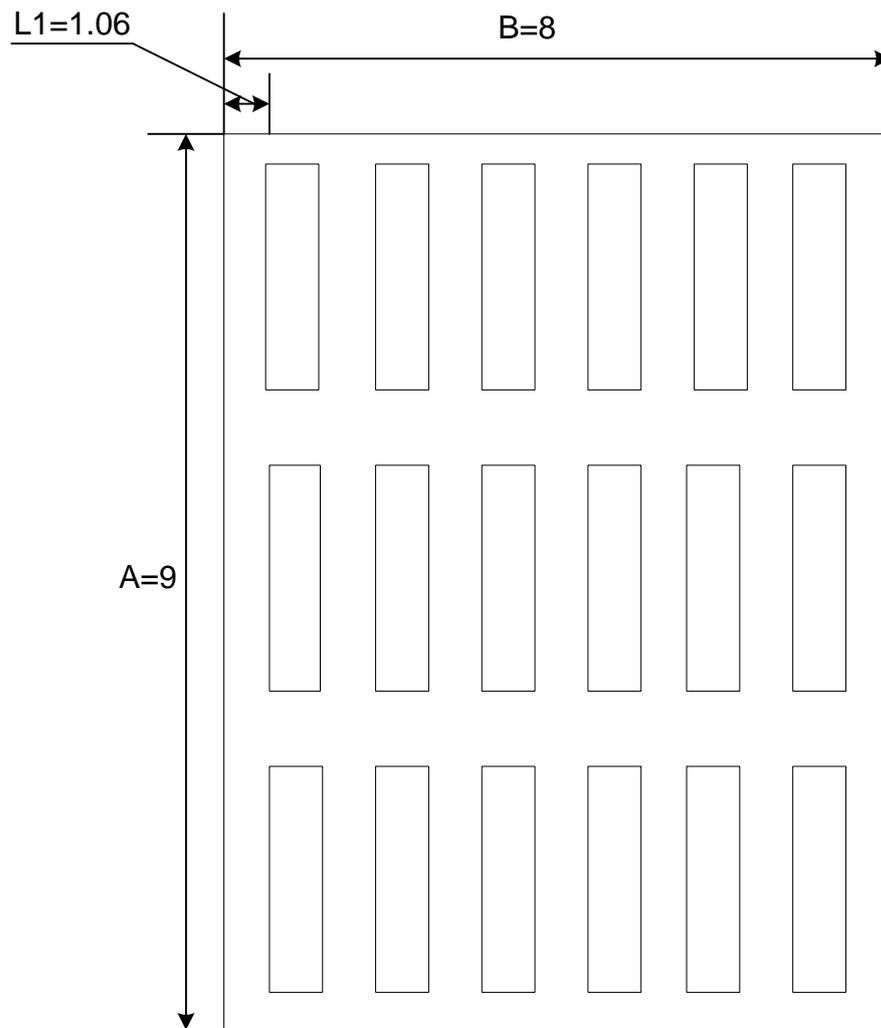


Рисунок 42 Схема расположения светильников

5.6. Схема рабочего места.

Проектирование рабочих мест, снабженных ПЭВМ, относится к числу важных проблем эргономического проектирования в области вычислительной техники.

Рабочее место и взаимное расположение всех его элементов должно соответствовать антропометрическим, физическим и психологическим требованиям. Большое значение имеет также характер работы. В частности, при организации рабочего места оператора должны быть соблюдены следующие основные условия: оптимальное размещение оборудования, входящего в состав рабочего места и достаточное рабочее пространство, позволяющее осуществлять все необходимые движения и перемещения.

Эргономическими аспектами проектирования видеотерминальных рабочих мест, в частности, являются: высота рабочей поверхности, размеры пространства для ног, требования к расположению документов на рабочем месте (наличие и размеры подставки для документов, возможность различного размещения документов, расстояние от глаз пользователя до экрана, документа, клавиатуры и т.д.), характеристики рабочего кресла, требования к поверхности рабочего стола, регулируемость элементов рабочего места.

Главными элементами рабочего места программиста являются стол и кресло. Основным рабочим положением является положение сидя. Рабочая поза сидя вызывает минимальное утомление. Рациональная планировка рабочего места предусматривает четкий порядок и постоянство размещения предметов, средств труда и документации. То, что требуется для выполнения работ чаще, расположено в зоне легкой досягаемости рабочего пространства.

Моторное поле - пространство рабочего места, в котором могут осуществляться двигательные действия человека.

Максимальная зона досягаемости рук - это часть моторного поля рабочего места, ограниченного дугами, описываемыми максимально вытянутыми руками при движении их в плечевом суставе.

Оптимальная зона - часть моторного поля рабочего места, ограниченного дугами, описываемыми предплечьями при движении в локтевых суставах с опорой в точке локтя и с относительно неподвижным плечом.

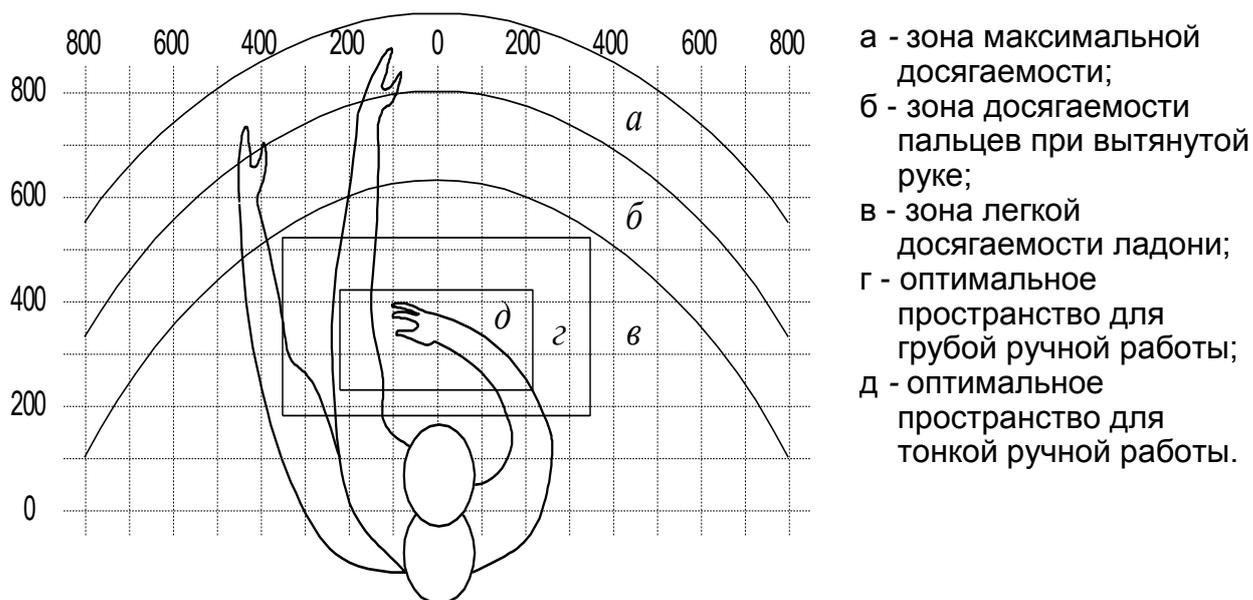


Рисунок 43 Зоны досягаемости рук в горизонтальной плоскости.

Оптимальное размещение предметов труда и документации в зонах досягаемости:

ДИСПЛЕЙ размещается в зоне **а** (в центре);

СИСТЕМНЫЙ БЛОК размещается в предусмотренной нише стола;

КЛАВИАТУРА - в зоне **г/д**;

«МЫШЬ» - в зоне **в** справа;

СКАНЕР в зоне **а/б** (слева);

ПРИНТЕР находится в зоне **а** (справа);

ДОКУМЕНТАЦИЯ: необходимая при работе - в зоне легкой досягаемости ладони – **в**,
а в выдвижных ящиках стола - литература,

неиспользуемая постоянно.

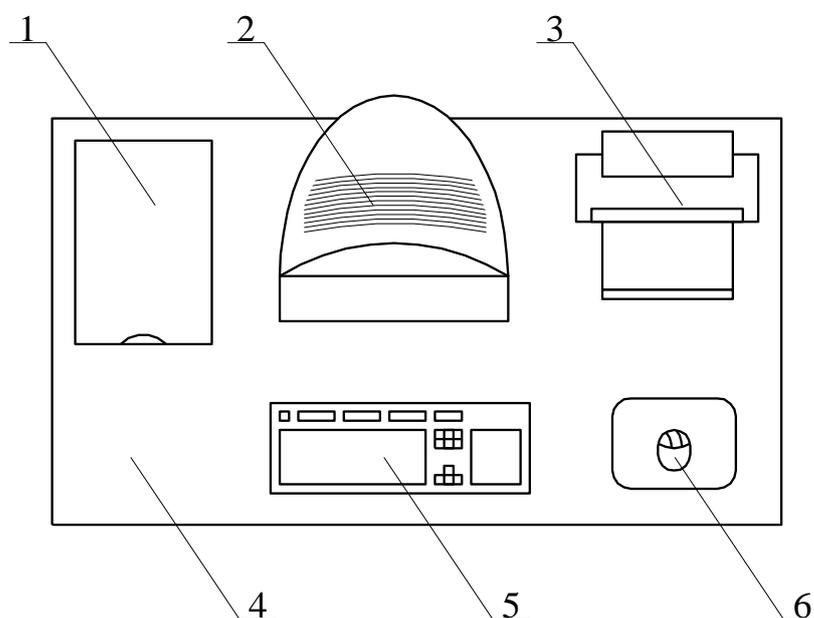


Рисунок 44Размещение основных и периферийных

На рисунке выше показан пример размещения основных и периферийных составляющих ПК на рабочем столе программиста.

1 – сканер, 2 – монитор, 3 – принтер, 4 – поверхность рабочего стола, 5 – клавиатура, 6 – манипулятор типа «мышь».

Для комфортной работы стол должен удовлетворять следующим условиям:
высота стола должна быть выбрана с учетом возможности сидеть свободно, в удобной позе, при необходимости опираясь на подлокотники;
нижняя часть стола должна быть сконструирована так, чтобы программист мог удобно сидеть, не был вынужден поджимать ноги;
поверхность стола должна обладать свойствами, исключающими появление бликов в поле зрения программиста;
конструкция стола должна предусматривать наличие выдвижных ящиков (не менее 3 для хранения документации, листингов, канцелярских принадлежностей).
Большое значение придается характеристикам рабочего кресла. Так, рекомендуемая высота сиденья над уровнем пола находится в пределах 420-550 мм. Поверхность сиденья мягкая, передний край закругленный, а угол наклона спинки - регулируемый.

Необходимо предусматривать при проектировании возможность различного размещения документов: сбоку от видеотерминала, между монитором и клавиатурой и т.п. Кроме того, в случаях, когда видеотерминал имеет низкое качество изображения, например заметны мелькания, расстояние от глаз до экрана делают больше (около 700 мм), чем расстояние от глаза до документа (300-450 мм). Вообще при высоком качестве изображения на видеотерминале расстояние от глаз пользователя до экрана, документа и клавиатуры может быть равным.

Положение экрана определяется:

- расстоянием считывания (0,6...0,7 м);
- углом считывания, направлением взгляда на 20° ниже горизонтали к центру экрана, причем экран перпендикулярен этому направлению.

Должна также предусматриваться возможность регулирования экрана:

- по высоте +3 см;
- по наклону от -10° до $+20^\circ$ относительно вертикали;
- в левом и правом направлениях.

Большое значение также придается правильной рабочей позе пользователя. При неудобной рабочей позе могут появиться боли в мышцах, суставах и сухожилиях.

Требования к рабочей позе пользователя видеотерминала следующие:

- голова не должна быть наклонена более чем на 20° ,
- плечи должны быть расслаблены,
- локти - под углом $80^\circ \dots 100^\circ$,
- предплечья и кисти рук - в горизонтальном положении.

Существенное значение для производительной и качественной работы на компьютере имеют размеры знаков, плотность их размещения, контраст и соотношение яркостей символов и фона экрана. Если расстояние от глаз оператора до экрана дисплея составляет 60...80 см, то высота знака должна быть не менее 3 мм, оптимальное соотношение ширины и высоты знака составляет 3:4, а расстояние между знаками – 15...20% их высоты. Соотношение яркости фона экрана и символов - от 1:2 до 1:15.

Во время пользования компьютером медики советуют устанавливать монитор на расстоянии 50-60 см от глаз. Специалисты также считают, что верхняя часть видеодисплея должна быть на уровне глаз или чуть ниже. Когда человек смотрит прямо перед собой, его глаза открываются шире, чем когда он смотрит вниз. За счет этого площадь обзора значительно увеличивается, вызывая обезвоживание глаз. К

тому же если экран установлен высоко, а глаза широко открыты, нарушается функция моргания.

Создание благоприятных условий труда и правильное эстетическое оформление рабочих мест на производстве имеет большое значение как для облегчения труда, так и для повышения его привлекательности, положительно влияющей на производительность труда.

6. Организационно-экономическая часть

6.1. Цель организационно-экономической части

Целью организационно-экономической части является оценка затрат на создание лингвокультурного тезауруса русского языка и обоснование необходимости создания программного продукта.

6.2. Стадии и этапы разработки

Разработка системы «Лингвокультурный тезаурус русского языка» состоит из семи стадий:

- 1) Формирование требований
- 2) Разработка концепции
- 3) Техническое задание
- 4) Эскизный проект
- 5) Технический проект
- 6) Рабочая документация
- 7) Ввод в эксплуатацию

Данный состав сформирован в соответствии с ГОСТ 34.601-90 «Автоматизированные системы. Стадии создания». ГОСТ допускает исключение стадии «Эскизный проект», отдельных этапов работ на всех стадиях в случае, если это предполагается особенностями проекта. [ГОСТ 34.601, 1990]

Таблица 41 Стадии и этапы разработки

Стадии	Этапы работ	Трудоемкость	
		Часы	Проценты
1. Формирование требований	1.1. Обследование объекта и обоснование необходимости создания	30	4.5
	1.2. Формирование требований	60	9.0

	пользователя		
	1.3. Оформление заявки на разработку	12	1.8
2. Разработка концепции	2.1. Изучение объекта	15	2.2
	2.2. Проведение необходимых научно-исследовательских работ	30	4.5
3. Техническое задание	3.1. Разработка и утверждение технического задания	20	3.0
4. Технический проект	4.1. Подготовка наполнения БД	60	9.0
	4.2. Разработка проектных решений по системе и ее частям	360	53.6
5. Рабочая документация	5.1. Разработка рабочей документации на систему и ее части	70	10.4
6. Ввод в эксплуатацию	7.1. Подготовка системы к вводу в действие	15	2.3
Итого:		672	

Продолжительность рабочего дня 8 часов при пятидневной рабочей неделе.

6.3. Определение затрат на разработку программного изделия.

6.3.1. Определение времени реализации проекта.

Для определения времени реализации проекта вычислим общее количество дней, требующихся на разработку проекта. Для расчета количества рабочих дней, требующихся на разработку проекта используется следующая формула:

$$T_{PD} = \frac{T_{час}}{t_{рд}},$$

где $T_{час}$ – общее время на разработку в часах, $t_{рд}$ – коэффициент, показывающий количество рабочих часов в одном дне. Для дальнейших расчетов примем $t_{рд} = 8$ час. Тогда $T_{рд} = T_{час} / 8 = 672 / 8 = 84$ дней.

Для определения времени реализации проекта требуется перевести рабочие дни в календарные дни (КД). Для перевода используется следующая формула:

$$T_{кд} = \frac{T_{рД} (1 + d)}{g},$$

где d - доля дополнительных работ, порученных другой группе работников попутно с основной работой (от 0,1 до 0.3). В нашем случае проект ведётся самостоятельно, $d = 0$, g - коэффициент перевода(в зависимости от выходных и праздничных дней) – 0,7.

$$T_{кд} = T_{рД} / g = 84 / 0,7 = 120 \text{ дней}$$

6.3.2. Определение финансовых затрат на разработку

6.3.2.1. Оплата машинного времени

В расчет оплаты машинного времени включаются оплата за пользование компьютером с установленным на нем программным обеспечением для нужд проектирования АСУ. Трудоемкость разработки АСУ с использованием персонального компьютера составляет в среднем 70% от общих временных затрат. Для разработки АСУ представляется компьютер следующей конфигурации:

- Процессор INTEL Celeron Mobile 1.5;
- Оперативная память 512 Мб;
- Объем жесткого диска – 40 Гб;
- Дисковод для чтения и записи CD дисков;
- Дисплей LCD 15" с разрешением матрицы 1024x768 точек;
- Лазерный принтер Samsung ML-1710P;

Стоимость технического оснащения рабочего места составила 21000 руб. На техническое оснащение рабочего места установлен срок полезного использования 3 года. Таким образом, можно вычислить сумму ежегодного износа по техническому оснащению рабочего места по следующей формуле:

$$C_m = 21000 / 3 = 7000 \text{ руб./год}$$

Программное обеспечение, устанавливаемое на одном рабочем месте, приведено в таблице ниже.

Таблица 42. Программное обеспечение рабочего места

Наименование ПО	Стоимость за лицензию (руб.)
Microsoft Windows XP Home Russian	2133
Microsoft Office Standart 2003	8325
Microsoft Visio 2007 Standart	6452
Borland Delphi 7 Professional	1233
ИТОГО:	18143

На программное обеспечение установлен срок, в течение которого списывается его стоимость – 5 лет. Таким образом, можно вычислить сумму ежегодного списания стоимости указанного программного обеспечения по следующей формуле:

$$\text{Спо} = 18143 / (5) = 3629 \text{ руб/год}$$

Тогда сумму затрат на оплату машинного времени за время разработки проекта можно рассчитать по следующей формуле (при расчетах использовался коэффициент, показывающий отношение количества работ, выполняемых с использованием компьютера к общему количеству работ, в данном случае, коэффициент равен 0,7):

$$\text{Сэвм} = (\text{Спо} + \text{См}) * 0,7 = (7000 + 3629) * 0,7 = 7440 \text{ руб/год.}$$

$$\text{За 4 месяца разработки затраты Сэвм} = 7440 / 3 = 2480.$$

6.3.2.2. Затраты на материалы

В статье учитываются суммарные затраты на материалы, приобретаемые для разработки данной АСУ (Таблица 3) и вспомогательные материалы.

Таблица 43. Затраты на материалы

Наименование	Количество	Сумма (руб)
Бумага формата А4 плот. 80 г/см ² , белизна 90% - для оформления дипломного проекта.	1 пачка (500 листов)	200
Картридж для струйного принтера ML-1710	1 штука	2000
Лазерные диски CD-R	5 шт.	100

Канцелярские товары (ручки, карандаши и т.д.)	-	250
Итого:		2550

6.3.2.3. Затраты на оплату труда

Затраты на оплату труда разработчика состоят из затрат на основную ($ЗП_{ОСН}$) и дополнительную ($ЗП_{ДОП}$) заработную плату, отчислений на единый социальный налог (ЕСН) и отчислений в фонд обязательного социального страхования от несчастных случаев на производстве (0,2%):

$$ЗП = ЗП_{ОСН} + ЗП_{ДОП} + ЕСН .$$

6.3.2.3.1. Основная заработная плата

Основная заработная плата рассчитывается по формуле:

$$ЗП_{ОСН} = T_D \times L_{СР.ДН} ,$$

где T_D - трудоемкость разработки, календарные дни;

$L_{СР.ДН}$ - среднедневной заработок разработчика.

Зарботная плата разработчика АСУ составляет 9000 руб.

Тогда среднедневной заработок разработчика находится по формуле:

$$L_{СР.ДН} = \frac{L_0}{F} ,$$

где L_0 – среднемесячная заработная плата разработчика;

F - среднее количество рабочих дней в месяце. F вычисляется по следующей формуле:

$$F = \sum N_{РАБ} / n = 21,$$

где $N_{РАБ}$ - количество рабочих дней в месяце,

n – число месяцев. В данном случае $n=4$

Тогда, $L_{СР.ДН} = 9000 / 21 = 428,57$ руб. и расходы на основную зарплату разработчика составят:

$$ЗП_{ОСН} = 428,57 * 84 = 36000 \text{ руб.}$$

6.3.2.3.2. Дополнительная заработная плата

Дополнительная заработная плата составляет примерно 20% от основной заработной платы, отсюда:

$$ЗП_{\text{доп}} = ЗП_{\text{осн}} * 20 / 100 = 36000 * 0,20 = 7200 \text{ руб.}$$

6.3.2.3.3. Расчет отчислений на единый социальный налог

От общей суммы основной и дополнительной заработной платы 26% отчисляется в следующие фонды:

- в пенсионный фонд (20%);
- в фонд социального страхования (2,9%);
- в фонд обязательного медицинского страхования (3,1%);

В итоге получаем: $C_H = (43200) * 26 \% = 11232 \text{ руб.}$

6.3.2.3.4. Расчет отчислений в фонд обязательного социального страхования от несчастных случаев на производстве

Отчисления в фонд обязательного социального страхования от несчастных случаев на производстве (0,2%) рассчитываются от общей суммы основной и дополнительной заработной платы:

$$C_{\text{отч}} = (43200) * 0,002 = 86 \text{ руб.}$$

6.3.2.4. Расчет накладных расходов

Накладные расходы непосредственно связаны с организационно-техническим обслуживанием помещений: электроснабжением, работой административного персонала и включают в себя командировочные расходы, плату за аренду помещений и т.д. Они составляют (30 – 50) % от суммы основной и дополнительной заработной платы. В итоге получаем:

$$H_P = 0,4 * (43200) = 17280 \text{ руб.}$$

6.3.2.5. Услуги сторонних организаций

В процессе разработки проекта сторонними организациями были оказаны услуги по распечатке 11 листов формата А1 на плоттере. Сумма затрат составила $11 \cdot 100 = 1100$ руб.

6.3.2.6. Общая сумма затрат на разработку проекта

Общая сумма затрат на разработку проекта определяется суммированием по статьям затрат и отражается в таблице 4.

Таблица 44. Затраты на разработку проекта

Статья затрат	Стоимость, руб.
1. Оплата машинного времени	2480
2. Затраты на материалы	2550
3. Затраты на оплату труда – всего	54518
3.1. Основная заработная плата разработчика	36000
3.2. Дополнительная заработная плата	7200
3.3. Единый социальный налог (26%)	11232
3.4. Отчисления в фонд обязательного социального страхования от несчастных случаев на производстве (0,2%)	86
4. Накладные расходы	17280
5. Услуги сторонних организаций	1100
Общая сумма затрат на разработку проекта	77928

6.4. Обоснование разработки программного изделия.

Лингвокультурный тезаурус русского языка представляет собой программный продукт, предполагаемый к использованию в фундаментальных исследованиях в таких областях как лингвистика, семиотика, а также получения доступа к результатам ассоциативного эксперимента в исследовательских целях. Работа выполнена при поддержке гранта РГНФ № *06-04-03803в* в рамках конкурса 2006-8 гг. «Образ России в современном мире».

Заключение

В процессе реализации дипломной работы были выполнены все требования технического задания. В процессе разработки решены задачи:

- изучения предметной области;
- анализа возможных вариантов решения поставленных задач;
- проектирования системы;
- реализации системы с учетом выбранных средств;
- доказательства оптимальности схемы базы данных;
- экономического обоснования проекта.

Пути возможных исследований:

- Определение вероятного знака по формуле смысла
- Методы представления базы фигур знания, меняющейся по структуре.
- Увеличение скорости работы когнитивера

Список литературы

- Арчер, Уайтчепел, 2005 Том Арчер, Эндрю Уайтчепел. Visual C++.NET. Библия пользователя. –М.: Издательский дом «Вильямс», 2005. – 1216с.
- Григорьев, Ревунков, 2002 Ю.А. Григорьев, Г.И. Ревунков. Банки данных. Учебник для вузов. – М., Изд. МГТУ им. Н.Э. Баумана, 2002. – 320 с.
- Дейт, 2001 Дейт, К., Дж. Введение в системы баз данных, 7-е издание.: Пер. с англ. –М.: Издательский дом «Вильямс», 2001.- 1072 с.
- Информатика, 2007 1. Информатика. Анализ сложности рекурсивных алгоритмов. [Электронный ресурс] / - Электрон. дан. – М., 2007 – режим доступа: http://it.kgsu.ru/ТИ_7/salg_007.html
- Караулов, 1981 Ю.Н.Караулов. Лингвистическое конструирование и тезаурус русского языка. –М.: Издательство «Наука», 1981.
- Караулов, Филиппович, 2005 Ю.Н.Караулов, Ю.Н.Филиппович. Лингвокультурологический тезаурус русского языка. –М.: 2005.
- Караулов, 2004 Ю.Н.Караулов. Концептография языковой картины мира. Статья 1. Первый этап «восхождения» к образу мира: от элементарных фигур знания к предметно-референтным областям культуры// Проблемы прикладной лингвистики. Выпуск 2. Сборник статей./ Отв. ред. Н.В.Васильева. –М.: «Азбуковник», 2004. – 400с.
- Кэнту, 2001 Кэнту М. Delphi 5 для профессионалов. – СПб.: Питер, 2001. – 944 с.
- Маклаков, 1999 Маклаков С.В. BPWin и ERWin: CASE - средства для разработки информационных систем. – М., Изд. Диалог – МИФИ, 1999. – 256 с.

Черкасова, 2004

Г.А.Черкасова. Формальная модель ассоциативного исследования.// Проблемы прикладной лингвистики. Выпуск 2. Сборник статей./ Отв. ред. Н.В.Васильева. –М.: «Азбуковник», 2004. – 400с.

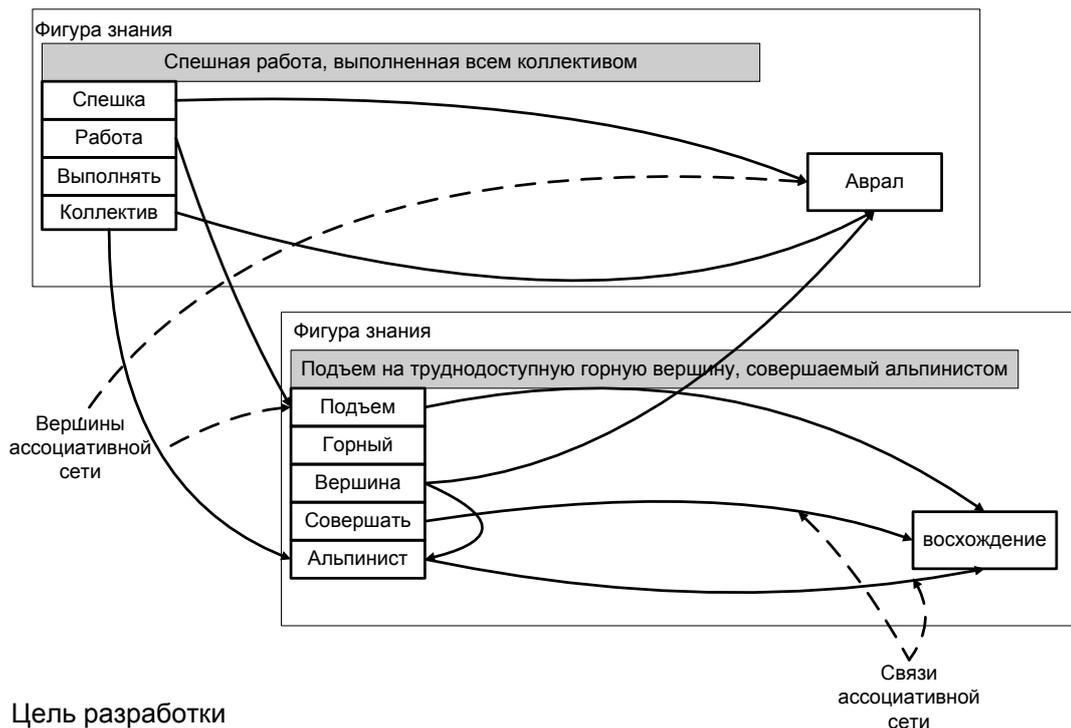
Эргономика, 2007

2. Эргономика [Электронный ресурс] / - Электрон. дан. – М., 2007 – режим доступа: <http://www.ixbt.com/peripheral/ergonomic.html>

Приложения

Приложение 1 Графическая часть дипломного проекта

Лингвокультурный тезаурус русского языка



Цель разработки

Построение лингвокультурного тезауруса русского языка на основе баз данных фигур знания и ассоциативного эксперимента, а также моделирование в нем пассивного режима работы когнитивера через нахождение путей в ассоциативной сети

Назначение системы

Система предназначена для использования специалистами в области психолингвистики, в качестве реализации пятикомпонентной структуры лингвокультурного тезауруса русского языка и моделирования пассивного режима когнитивера

Состав решаемых системой задач

1. Импорт данных тезауруса из таблицы MS Access
2. Редактирование фигур знания.
3. Поиск фигур знания по критерию.
4. Сортировка фигур знания.
5. Автоматизация разделения формулы смысла на знаки.
6. Имитация пассивного режима работы когнитивера поиском путей в ассоциативной сети.

Сравнительная характеристика аналогов и прототипов

	Наличие ассоциативной сети	Наличие тезауруса	Возможность пополнения базы данных	Реализация пассивного режима когнитивера	Переносимость системы
RCO Semantic Network SDK	да	нет	да	отсутствует	определяется приложением
www.lexfn.com	Англоязычный вариант	нет	нет	пути для пары знаков	доступ через Интернет
Microsoft Excel	да	да	да	трудоемко	да
Система ЛКТ	да	да	да	есть	да

Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Описание предметной области

Множество областей, соответствующих предметным областям в реальности, устанавливается экспертом и может динамически меняться
Сложность структуры представлена возможностью отнесения фигуры знания к нескольким областям.



Яблоня – Из плодов какого дерева делают шипучее вино кальвадос? (Дерево, Быт)

Способ отражает структуру формулы смысла и метод рассуждений, позволяющих перейти от нее к соответствующему знаку. Эта оценка производится на основе экспертного мнения и практически не формализуема. Установлено 49 способов задания смысла:

1. Аббревиатура
2. Анаграмма
3. Антоним
4. Ассоциация
4. Вид
5. Описание
6. Множество
7. Метафора
8. Прецедентный текст
9. Словосочетание
10. Синонимия
- ...



Функция отражает ценность знания для создания целостной картины мира.
Устанавливается на основе экспертных оценок.
Имеет два возможных значения:

Ретушь – знание второстепенно и не входит в картину мира.

Рецепт – входит в состав картины мира.

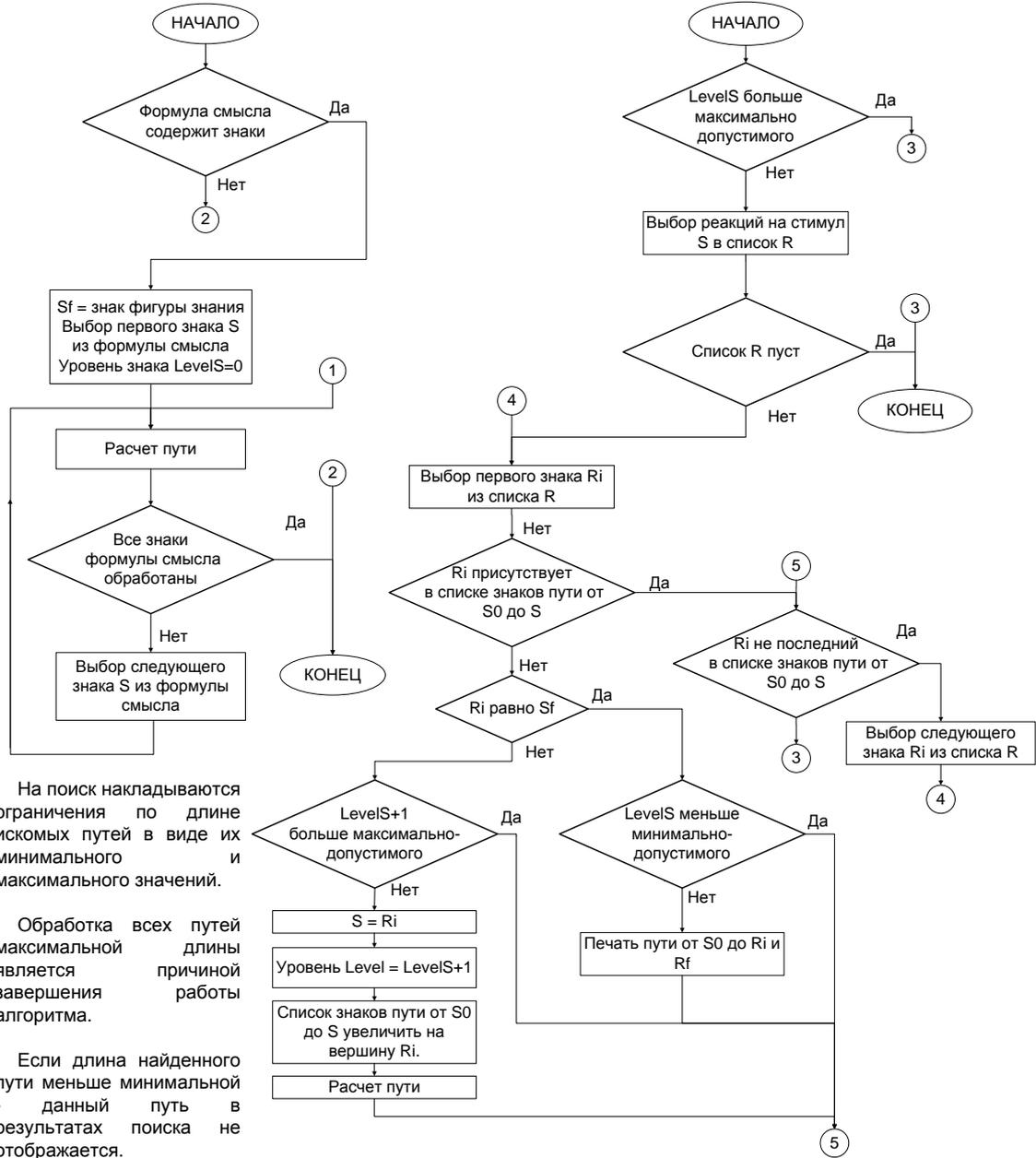
Алгоритм работы когнайзера

Алгоритм реализует поиск путей между вершинами ассоциативной сети.

Поиск осуществляется рекурсивно. Обрабатываемой вершине соответствует структура:

Наименование поля	Тип	Описание
Name	строка	наименование вершины
Level	целое число	Удаленность вершины от начальной
Old_elements	массив строк	Вершины, присутствующие в пути к данной вершине

Расчет пути



На поиск накладываются ограничения по длине искомых путей в виде их минимального и максимального значений.

Обработка всех путей максимальной длины является причиной завершения работы алгоритма.

Если длина найденного пути меньше минимальной - данный путь в результатах поиска не отображается.

Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Анализ алгоритма когнайзера

Без учета рекурсивности функции средняя сложность алгоритма равна $O(n^2)$, в худшем случае с учетом рекурсии $O(n^4)$

Оценка среднего значения сложности требует учета вероятностных законов, которые являются следствием характеристик ассоциативной сети.

Ассоциативная сеть содержит 6600 стимулов

Среднее количество реакций на стимул равно 18.78

Количество дуг сети, реакции в которых не являются стимулами других знаков равно 50 000. При обработке алгоритмом такой связи, дальнейшая ее рекурсивная обработка вестись не будет.

if(pEI^.level=maxlevel) then Exit; Проверка на превышение длины пути $O(1)$

Запрос реакций на стимул

<pre>AsWebTable.Active:=false; AsWebTable.SQL.Clear; AsWebTable.SQL.Add('SELECT DISTINCT Реакция, Стимул FROM Сеть_новой_структуры '); AsWebTable.SQL.Add('WHERE Стимул = '"+doubleC(pEI^.name)+"' AND Реакция <> ''"); AsWebTable.Active:=true;</pre>	$O(n)$
<pre>if (AsWebTable.RecordCount<1) then Exit; SetLength(kids, AsWebTable.RecordCount); AsWebTable.RecNo:=1;</pre>	$O(1)$

Заполнение массива реакций

<pre>for k:=1 to AsWebTable.RecordCount do begin AsWebTable.RecNo:=k; kids[k-1]:=AsWebTable.FieldByName('Реакция').AsWideString; end;</pre>	$O(1)$
---	--------

for k := 0 to Length(kids) - 1 do $O(n^2)$

<pre>begin temp:=kids[k]; if(check(temp, pEI^.old)<>1) then Обработка нахождения конечного знака if(temp=Znak.Caption) then $O(1)$ begin if ((pEI.level+1)< minlevel) then Exit; min_find:=1; print(pEI); Exit; end else Продолжение поиска для дочернего узла $O(n^2)$ begin if ((pEI.level+1)> maxlevel) then Exit; new(newpEI); newpEI^.name:=temp; newpEI^.level:=pEI^.level+1; newpEI^.up:=pEI; newpEI^.downNum:=0; newpEI^.down:=nil; SetLength(newpEI^.old,Length(pEI^.old)+1); for t:=1 to Length(pEI^.old) do begin temp2:=pEI^.old[t-1]; newpEI^.old[t-1]:=temp2; end; x:=Length(newpEI^.old); temp2:=pEI^.name; newpEI^.old[x-1]:=temp2; fillNext(newpEI, minlevel, maxlevel); end end</pre>	$O(n^2)$
---	----------

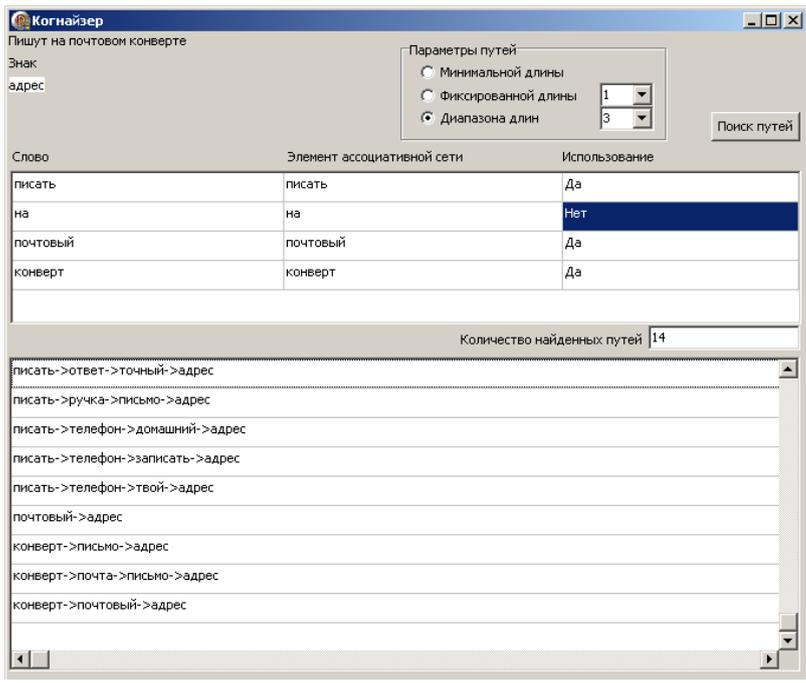
Ассоциативная сеть обладает высокой степенью связности. При поиске кратчайшего пути в случае его существования, достаточно анализировать пути длиной до 5.

При анализе реакций знака дальнейшей обработке подвергаются реакции, ранее в пути к стимулу не встречавшиеся. Это снижает сложность алгоритма. Без учета подобных вершин, среднее количество связей, не являющихся листовыми равно 13.

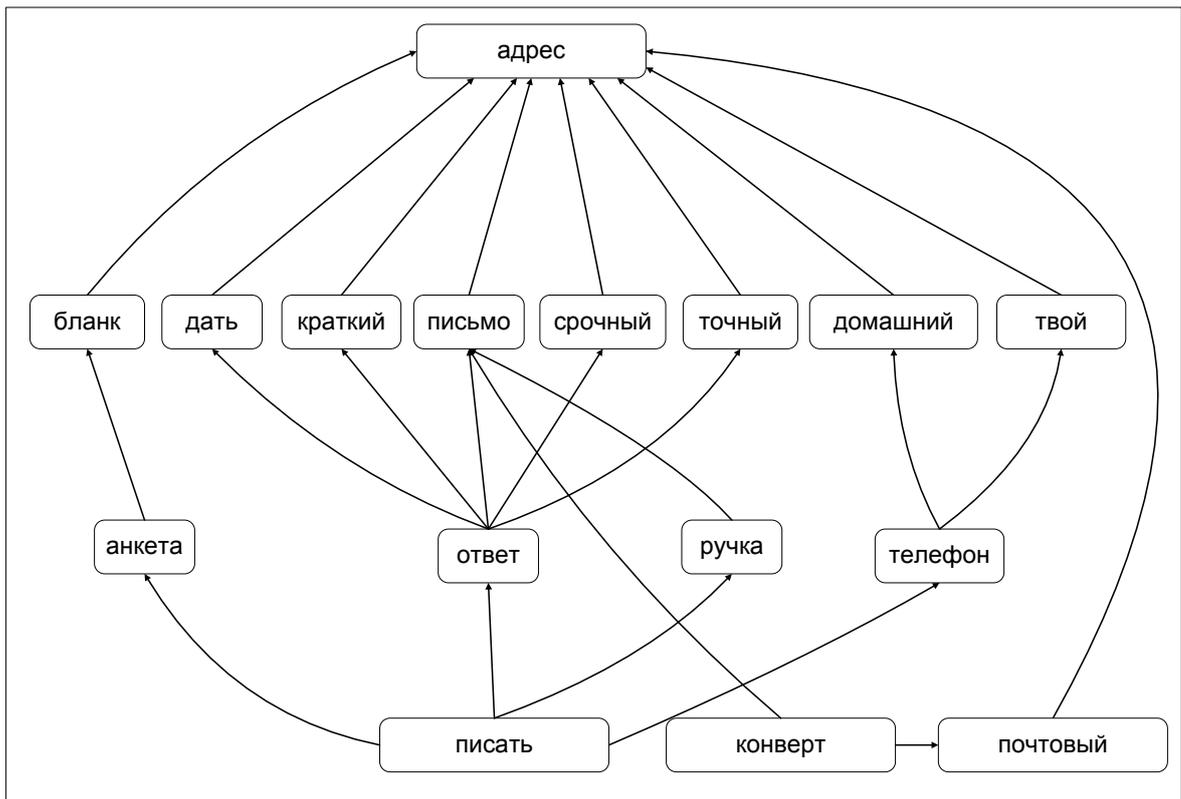
При нахождении путей длиной до 5 функция будет выполнена $13^5 = 371293$ раз. Так как эти соотношения получены на основе анализа свойств ассоциативной сети, незначительно зависящих от размера сети, результирующая сложность алгоритма будет равна $O(k*n^2) = O(n^2)$.

Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Пример работы когнайзера



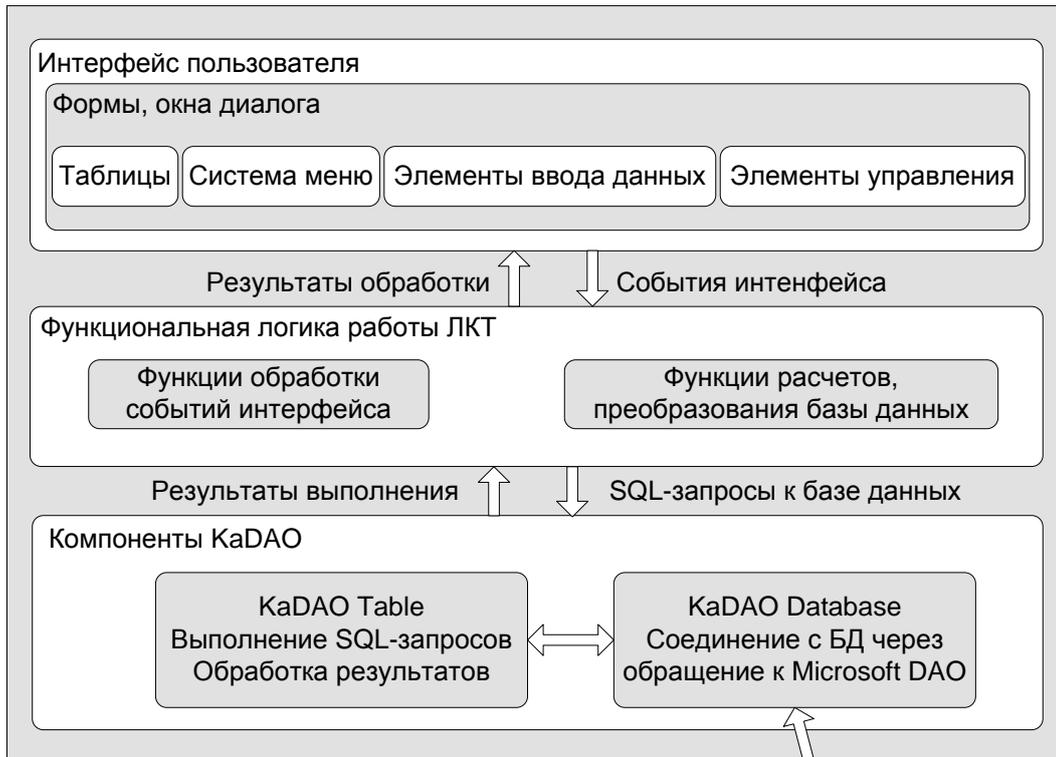
Экранная форма работы когнайзера



Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Структура программы

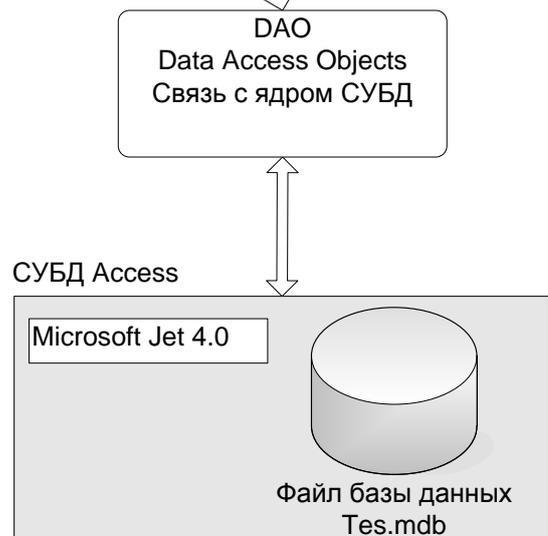
Приложение лингвокультурного тезауруса русского языка



KaDAO – компоненты для среды разработки Borland Delphi, предоставляющие высокоуровневые функции доступа к базе данных посредством обращения к объектам DAO

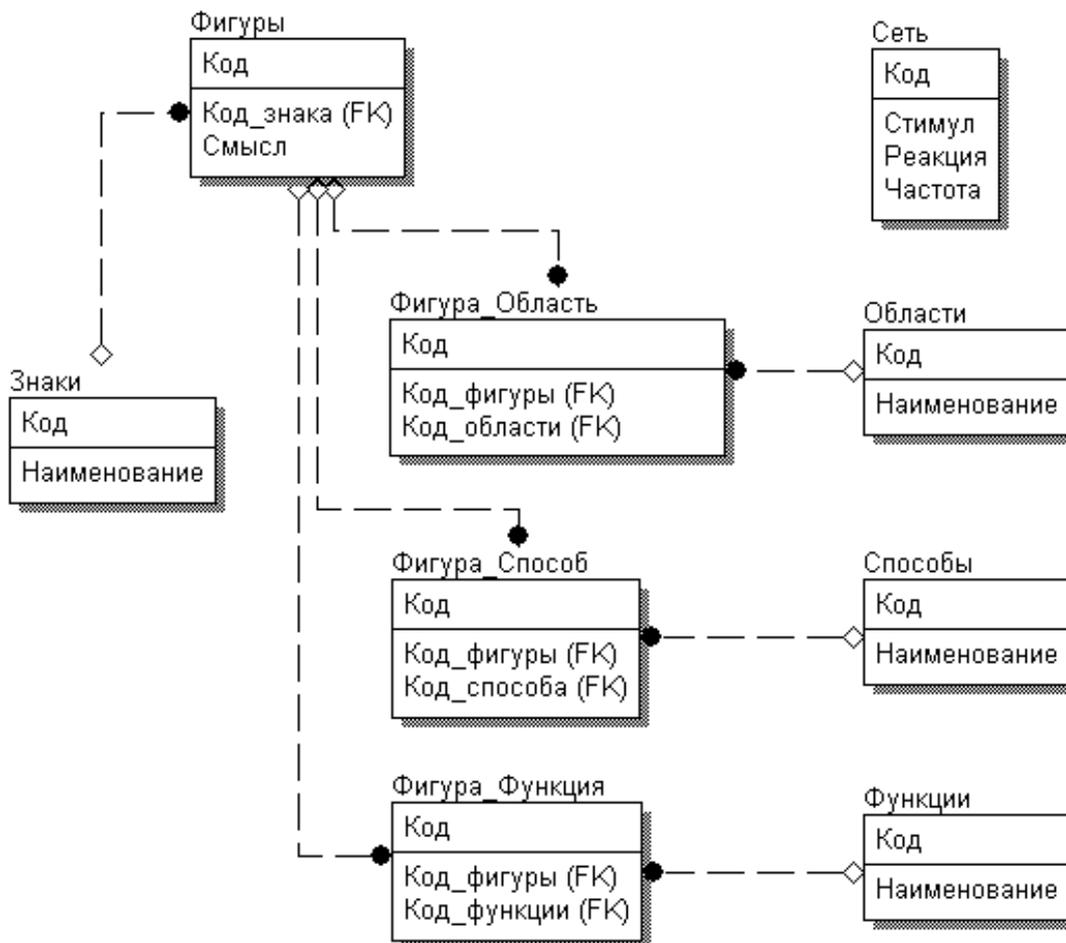
Microsoft DAO (Data Access Objects – объекты доступа к данным) – компоненты, позволяющие использовать функции процессора Jet

Microsoft Jet – ядро базы данных. Компонент Microsoft Access, выполняющий загрузку и сохранение данных в базах данных пользователей и в системных базах данных.

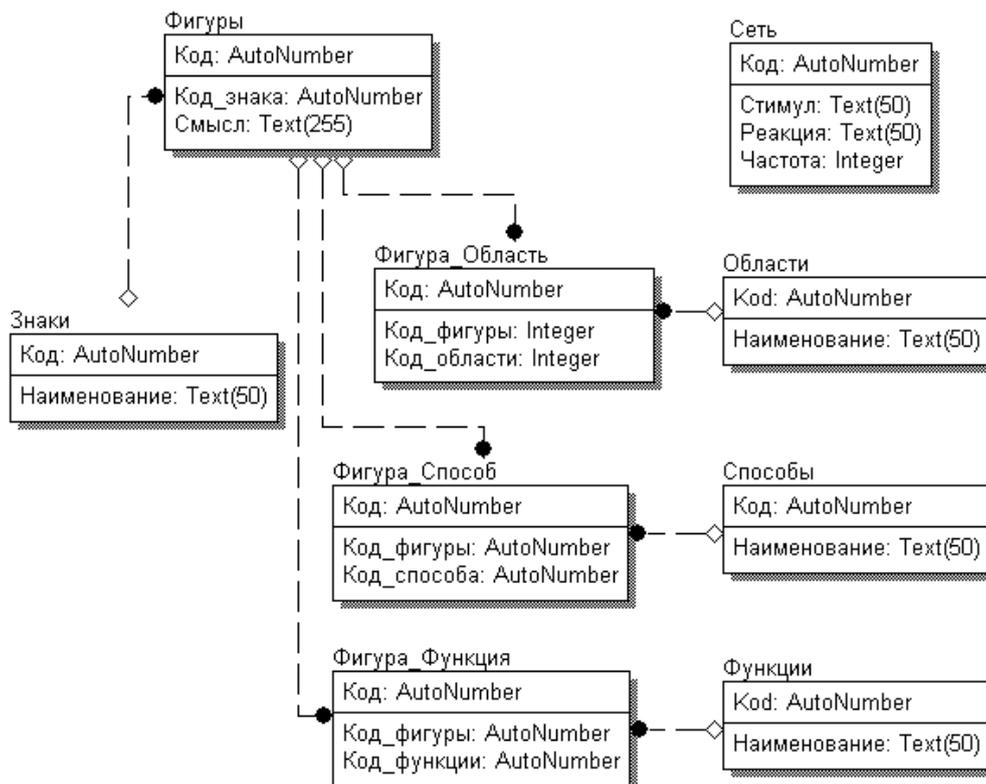


Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Даталогическая и инфологическая схемы базы данных



Инфологическая схема базы данных



Даталогическая схема базы данных

Добавление данных в лингвокультурный тезаурус русского языка

Подготовка материала

Код	Слово	Смысл	Область	Способ	Функция
1	Авеста	Древний иранский религиозный памятник	Религия	Множество	Ретушь
2	Адидас	Известная фирма спортивной одежды и обуви	Спорт	Множество	Рецепт
3	Ангара	Балет Андрея Эшпая	Искусство	Гипероним	Рецепт
4	Арго	Корабль, везший руно	Миф	Суждение	Ретушь
5	Асель	Балет Владимира Власова	Искусство	Множество	Ретушь
6	Асса	Фильм с участием В. Цоя	Кинематография	Дескрипция	Ретушь
7	Баретки	Рассказ Михаила Зощенко	Литература	Дескрипция	ретушь
8	Вольво	Марка шведских автомобилей	транспорт	Гипоним	ретушь
9	Выхино	Станция московского метро	Москва	Множество	Рецепт
10	Гроза	Пьеса Александра Островского	Литература	Множество	Ретушь
11	Дантон	Фильм с участием Жерара Депардье	кино, ТВ	Синонимия	ретушь
12	Идиот	Роман о князе Мышкине	Литература	Гипероним	Рецепт
13	Изабо	Опера Пьетро Масканы	Музыка	Дескрипция	рецепт
14	Канаста	Карточная игра	игра	Дефиниция	рецепт

Материалы для добавления размещаются в таблице «Импорт» базы данных тезауруса при помощи импорта данных из таблиц Microsoft Excel, либо других приложений, перенос из которых поддерживает Microsoft Access. Материалы представляют собой таблицу, включающую поля: Код, Слово, Смысл, Область, Способ, Функция.

Перенос в таблицы тезауруса

Для использования фигур импортируемой таблицы в тезаурусе приводим их к схеме таблиц ЛКТ с помощью формы «Разбор данных» меню Запрос -> Разбор данных

Код	Слово	Смысл	Область	Способ	Функция
1	Авеста	Древний иранский религиозный памятник	Религия	Множество	Ретушь
2	Адидас	Известная фирма спортивной одежды и обуви	Спорт	Множество	Рецепт
3	Ангара	Балет Андрея Эшпая	Искусство	Гипероним	Рецепт
4	Арго	Корабль, везший руно	Миф	Суждение	Ретушь
5	Асель	Балет Владимира Власова	Искусство	Множество	Ретушь
6	Асса	Фильм с участием В. Цоя	Кинематография	Дескрипция	Ретушь
7	Баретки	Рассказ Михаила Зощенко	Литература	Дескрипция	ретушь
8	Вольво	Марка шведских автомобилей	транспорт	Гипоним	ретушь
9	Выхино	Станция московского метро	Москва	Множество	Рецепт
10	Гроза	Пьеса Александра Островского	Литература	Множество	Ретушь
11	Дантон	Фильм с участием Жерара Депардье	кино, ТВ	Синонимия	ретушь

1 Авеста Древний иранский религиозный памятник Религия Множество Ретушь

0 Разобрать

Код	Наименование
225	курение
218	лес
144	литература
195	лошадь
154	марка
317	математика
164	материал
165	медицина
166	межгосударственные отношения
267	международные отношения
214	мера
249	металл
196	минерал
250	минералогия
145	мифология
251	мода
238	мораль
269	Москва

Из источника: литература

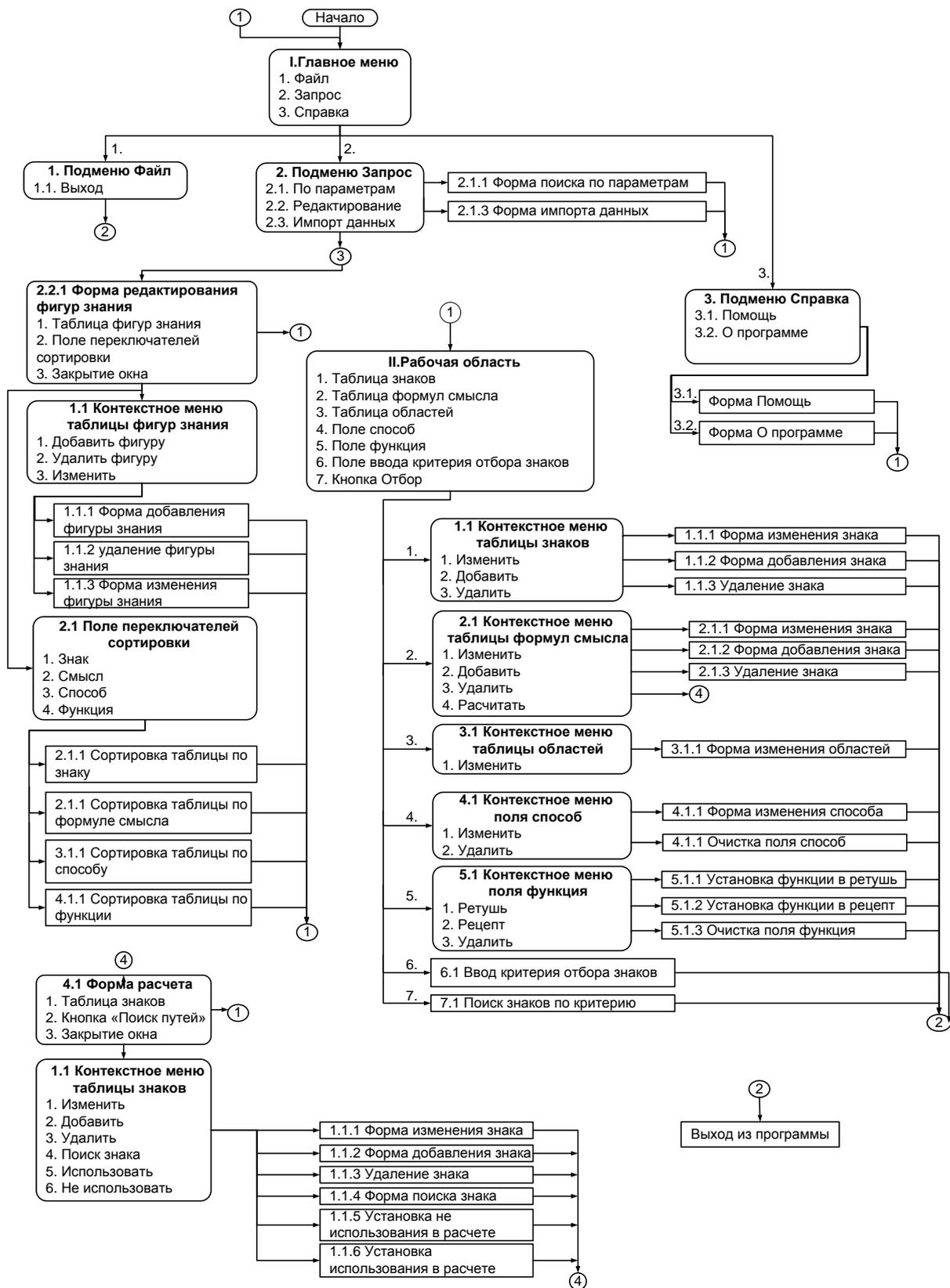
В процессе разбора записи исходной таблицы осуществляется:

1. Поиск знака среди имеющихся. В случае их наличия проверка присутствия аналогичной формулы смысла. Иначе создание нового знака.
2. Определение способа представления смысла на основе имеющегося в базе данных множества способов.
3. Определение функции фигуры знания
4. Разбор поля «Области» на множество областей, присутствующих в таблице «Области» базы данных тезауруса.

В случае, если в тезаурусе не найден подходящий способ или не вся строка «Области» разобрана, пользователь самостоятельно выбирает необходимые из базы данных тезауруса. Возможно ее пополнения новым способом или областью.

Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

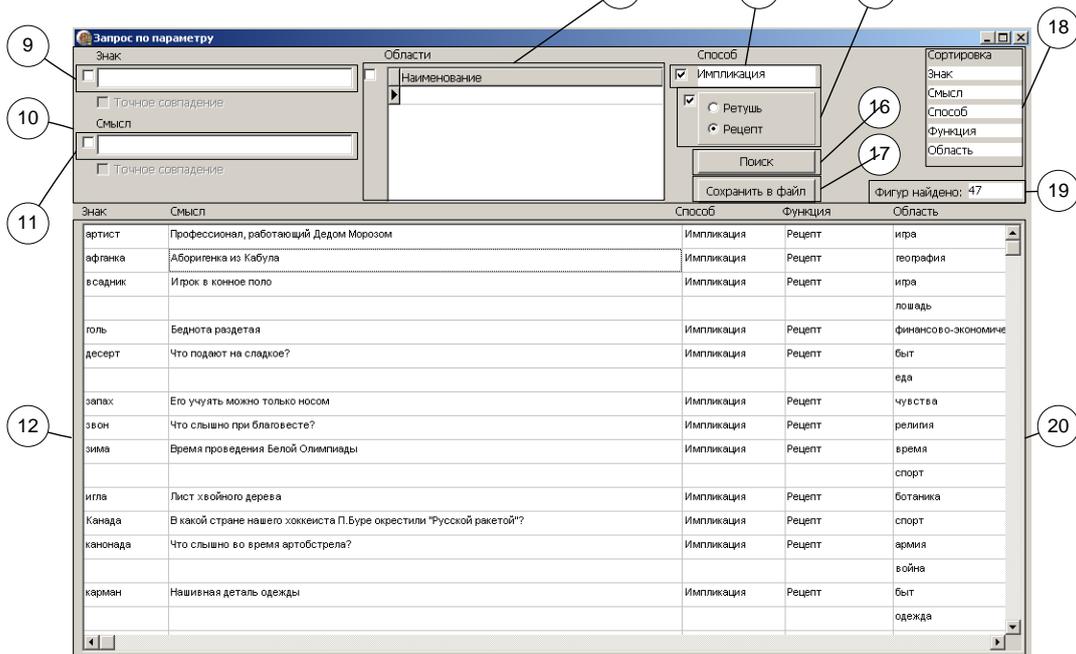
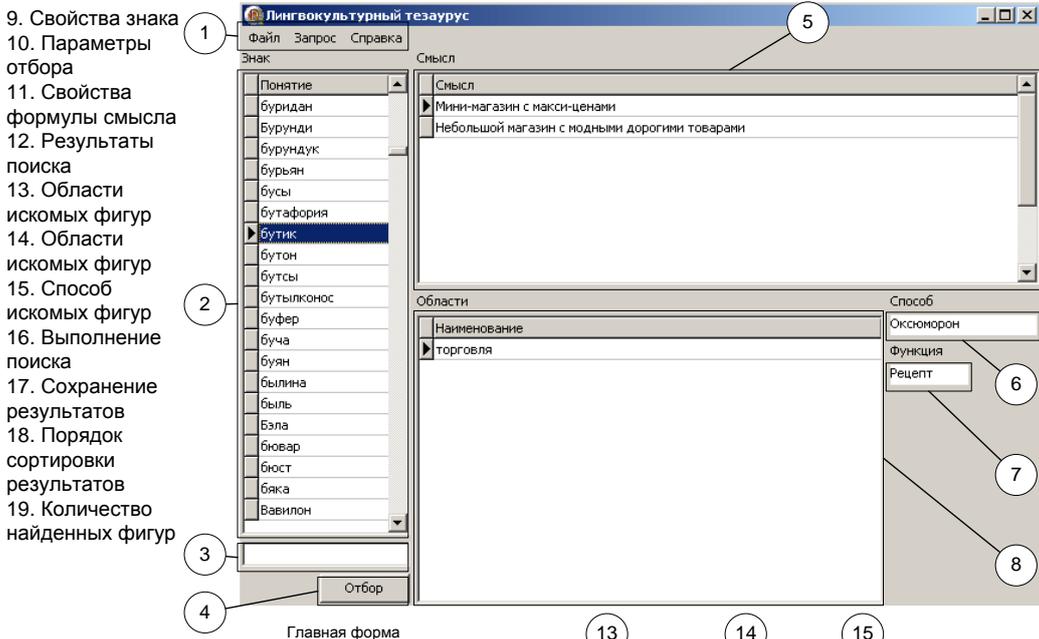
Граф диалога пользователя



Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Экранные формы

- | | | |
|--------------------------------|--|---|
| 1. Главное меню | 4. Запуск отбора знаков | 7. Функция текущей формулы смысла |
| 2. Знаки | 5. Формулы смысла фигуры знания | 8. Области принадлежности фигуры знания |
| 3. Поле критерия отбора знаков | 6. Способ представления текущей формулы смысла | |

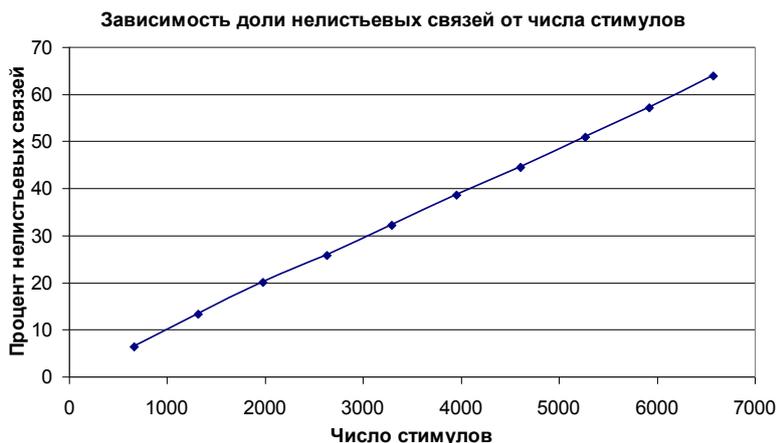


Поиск по критериям

Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

Анализ ассоциативной сети

Были произведены выборки определенного числа стимулов в ассоциативной сети. Затем производился подсчет нелистьевых связей выборок. Эксперимент показал, что зависимость между количеством стимулов и процентом нелистьевых связей на интервале от 10 до 100 процентов стимулов почти линейна.



В рамках второго эксперимента производились выборки различного числа связей ассоциативной сети. Зависимость числа стимулов от общего числа связей иллюстрирует процесс охвата стимулов при проведении опроса

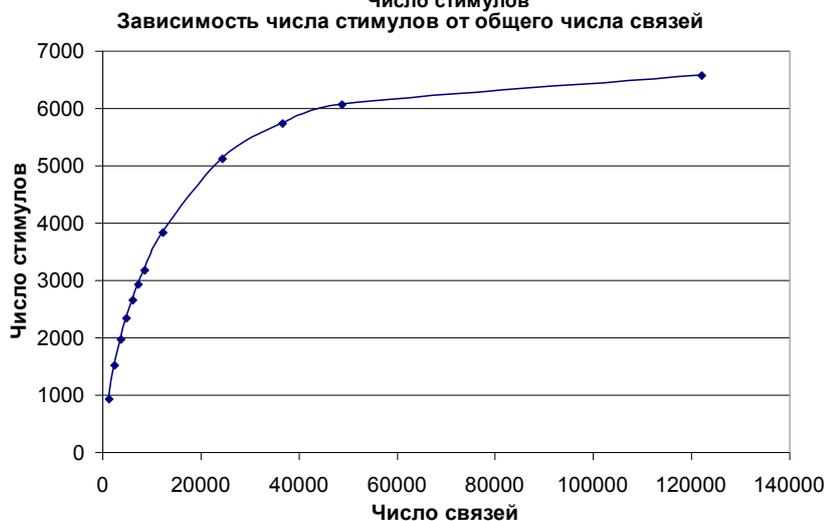
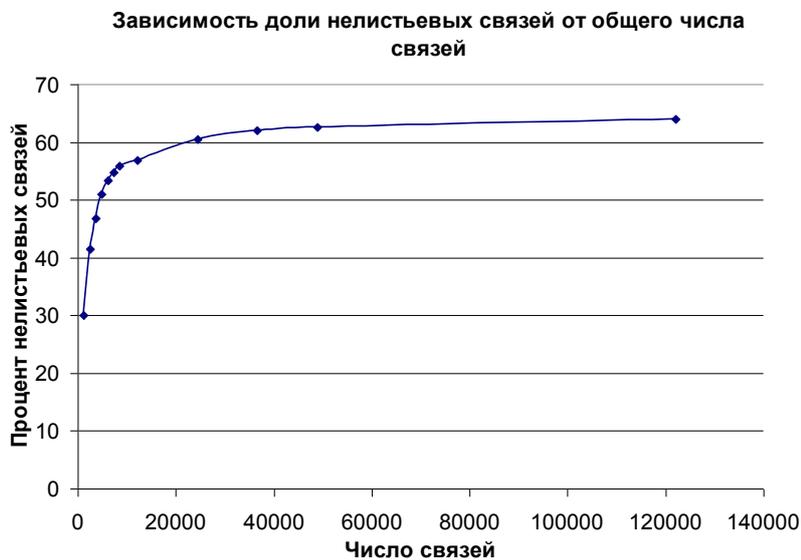


График зависимости доли нелистьевых элементов от общего числа связей показывает, что дальнейший опрос респондентов с данным набором стимулов в малой степени уменьшит процентное соотношение нелистьевых связей сети, так как график стремится к горизонтальной асимптоте при увеличении числа связей. Дальнейший опрос позволит уточнить частотные характеристики связей.



Выполнил: _____ (Сиренко А.В.) Утвердил: _____ (Филиппович Ю.Н.)

