ТРЕБОВАНИЯ К ВЫПОЛНЕНИЮ КУРСОВЫХ РАБОТ ПО НАПРАВЛЕНИЮ 45.03.03 "ФУНДАМЕНТАЛЬНАЯ И ПРИКЛАДНАЯ ЛИНГВИСТИКА"

1. Общие положения

1.1. **Курсовая работа (проект)** на первом году обучения и реализуется в рамках курса «Информатика и основы программирования». Порядок выполнения и требования для курсового проекта первого курса определяются сотрудником кафедры фундаментальной и прикладной лингвистики ГИ НГУ, преподающим данный предмет.

1.2. **Курсовая работа последующих лет обучения**, являясь одной из важнейших форм учебного процесса, представляет собой самостоятельную разработку студентом конкретной научной темы и может рассматриваться как этап подготовки выпускной бакалаврской работы. На втором курсе допускается написание курсовой работы реферативного характера, служащей для углублённого изучения определённой теоретической или прикладной области

1.2.1. Курсовые работы пишутся на втором и третьем курсах. Отметка за курсовую работу включена в план летней сессии. Студент допускается к экзаменационной сессии при условии получения зачета по курсовой работе. Студенты, не получившие в срок зачеты за курсовую работу, и не ликвидировавшие указанную академическую задолженность в сроки, установленные дирекцией ГИ, отчисляются из университета. Пересдача зачета по курсовой работе с оценки «неудовлетворительно» разрешается в установленном порядке.

1.2.2. Курсовая работа студентов 1 и 2 курсов оценивается недифференцированной оценкой («зачёт»), курсовая работа студента 3 курса выполнятся по направлению научной специализации и оценивается дифференцированным зачетом с отметками: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

1.2.3. Студент 2 и 3 курсов имеет право самостоятельно выбрать себе научного руководителя при условии согласия кафедры. Вместе с тем, потенциальный научный руководитель имеет право согласиться или отказаться руководить работой данного студента. Согласовать возможность выполнения курсовой работы под руководством выбранного руководителя студенту необходимо до 30 октября текущего года.

Тема курсовой работы определяется совместно научным руководителем и студентом. Тематика работы должна быть актуальной, соответствовать направлению образования, современному состоянию и перспективам развития науки о языке. При выборе тематики рекомендуется учитывать реальные теоретические и прикладные задачи данной области.

1.2.4. Руководителями курсовых работ являются наиболее квалифицированные преподаватели кафедры ФиПЛ. Руководителями могут быть также научные сотрудники институтов СО РАН, СО РАМН, квалифицированные сотрудники других организаций, заинтересованных в развитии и разработке лингвистических программ и ресурсов. Привлечение к руководству курсовыми работами может осуществляться на условиях совместительства или почасовой оплаты.

В случае выполнения курсовой работы в междисциплинарной области возможно назначение двух руководителей для одной курсовой работы, или привлечение консультантов.

1.2.5. Зачеты по курсовым работам по специализации на 3-м курсе выставляются на основе результатов защиты курсовых работ студентами на заседании кафедры, либо спецсеминара, либо комиссии в составе 2-х – 3-х преподавателей, но обязательно с участием научного руководителя.

2. Направления специализации

Тематика курсовых работ определяется направлениями научной работы сотрудников кафедры и может меняться в соответствии с изменением преподавательского состава кафедры и состава привлекаемых к руководству курсовыми и выпускными работами сторонних специалистов. Базовыми направлениями специализации для кафедры фундаментальной и прикладной лингвистики являются следующие типы междисциплинарных исследований естественного языка.

2.1. Теоретические или прикладные работы по математическим метаязыкам, предназначенным для моделирования различных аспектов естественного языка (формальные грамматики, логические языки, статистико-вероятностные средства и др.), в частности:

- исследование значимых для моделирования естественного языка свойств метаязыков,
- модификация имеющихся и разработка новых метаязыков для моделирования естественного языка,
- изучение адекватности моделей, основывающихся на различных метаязыках,
- исследование, разработка средств оценивания / сравнения метаязыков по сложности, трудоёмкости реализации, эффективности использования, адекватности,
- построение математической модели некоторого аспекта естественного языка,
- исследование алгоритмических возможностей реализации моделей.

Научные руководители направления специализации:

Одинцов Сергей Павлович, д.ф.-м.н., ведущий научный сотрудник Института математики СО РАН,

Савельев Лев Яковлевич, к.ф.-м.н., профессор ММФ НГУ, старший научный сотрудник Института математики СО РАН.

Саломатина Наталья Васильевна, к.ф.-м.н., научный сотрудник Института математики СО РАН,

Стукачев Алексей Ильич, доцент ММФ и ГИ НГУ, старший научный сотрудник Института математики СО РАН

Стукачёва Марина Викторовна, к.ф.-м.н., доцент ММФ НГУ,

Тимофеева Мария Кирилловна, д.филол.н., профессор ГФ НГУ, старший научный сотрудник Института математики СО РАН.

2.2. Компьютерные прикладные системы обработки языковых данных.

Создание и реализация алгоритмов обработки языковых данных для решения различных прикладных задач:

- сравнительное исследование эмпирических возможностей различных компьютерных средств моделирования естественного языка, анализ условий, режимов использования применительно к разным типам языковых данных и разным задачам,
- разработка компьютерных систем, имитирующих разные виды языковой деятельности человека (автоматический перевод, поиск информации, морфологический, синтаксический, семантический анализ и синтез текста,

распознавание и синтез устной речи, автоматическое реферирование, автоматический анализ и распознавание стилей, программы-собеседники, коммуникация с роботами, визуализация текста и т.д.),
- создание, развитие, поддержка компьютерных программ и электронных коллекций, облегчающих работу лингвиста (корпусов текстов, электронных словарей, различных языковых баз данных).

Научные руководители направления специализации:

Гриф Михаил Геннадьевич, д.т.н., профессор, заведующий кафедрой автоматизированных систем управления Новосибирского государственного технического университета,

Емельянов Павел Геннадьевич, к.ф.-м-н., с.н.с. лаборатории смешанных вычислений Института систем информатики им. А.П.Ершова СО РАН, доцент кафедры программирования ММФ,

Леман Анна Яковлевна, магистр техники и технологии по направлению «информатика и вычислительная техника», аналитик,

Мирошниченко Любовь Александровна, к.т.н., старший научный сотрудник Института математики СО РАН,

Саломатина Наталья Васильевна, к.ф.-м.н., научный сотрудник Института математики СО РАН,

2.3. Теоретические или прикладные работы, выполняемые на стыке лингвистики и физиологии / медицины / когнитивной науки, в частности:
- экспериментальные исследования в области нейрофизиологии языковой деятельности,
- изучение когнитивных основ языковой деятельности.

Научные руководители направления специализации:

Савостьянов Александр Николаевич, к.биол.н., д.филос.н., профессор кафедры ФиПЛ, факультета информационных технологий НГУ, зав. группой когнитивной нейролингвистики Института физиологии СО РАМН, зав. лабораторией биологических маркеров социального поведения человека ГИ НГУ

Бочаров Андрей Викторович, к.б.н., лаборатория механизмов регуляции памяти НИИ физиологии СО РАМН.

2.4. Теоретические или прикладные работы, выполняемые на стыке лингвистики и психологии.

Научный руководитель направления специализации: Можейкина Людмила Борисовна, к.психол.н., доцент факультета психологии НГУ.

2.5. Исследования в области прикладной лингвистики, проводимые на материале русского, английского или других естественных языков.

Научный руководитель направления: Снытникова Наталья Ивановна, к.пед.н., доцент ФИЯ НГУ.

3. Примеры тематик курсовых проектов и курсовых работ

3.1. Примеры тематик курсовых проектов

3. Автоматическая категоризация рекламных объявлений

В рамках проекта должен быть разработан модуль, выделяющий признаки, характерные для различных групп объявлений (на основании выборки реальных объявлений), и модуль, в автоматизированном режиме относящий введенное объявление к определенной группе (группам).

2. Чат-бот, эмулирующий живого собеседника (2 темы)
В рамках проекта должен быть разработан чат-бот, эмулирующий диалог на определенную тему с живым собеседником в режиме вопрос-ответ.

3. Автоматизированное извлечение из текстов фактографической информации, связанной с датами событий.
В рамках проекта должно быть разработано ПО, позволяющее выделить из введенного текста информацию о событиях, связанных с какой-либо датой.

4. Автоматизированное извлечение из текстов фактографической персональной информации. Извлечение биографической информации.
В рамках проекта должно быть разработано ПО, позволяющее выделить из введенного текста биографическую информацию.

5. Преобразование файлов LaTeX в файлы HTML. Преобразование файлов HTML в файлы LaTeX.
В рамках проекта должен быть разработан модуль, преобразовывающий тексты LaTeX в тексты с HTML-разметкой, и модуль, преобразовывающий тексты с HTML-разметкой в LaTeX. Полноценного перевода из одной разметки в другую не требуется.

6. Автоматизированная проверка орфографии с использованием aspell
В рамках проекта требуется разработать ПО, которое осуществляет проверку орфографии введенного текста в автоматизированном режиме с использованием aspell.

7. Автоматизированная проверка орфографии с использованием Яндекс.Спеллер
В рамках проекта требуется разработать ПО, которое осуществляет проверку орфографии введенного текста в автоматизированном режиме с использованием Яндекс.Спеллер.

8. Проверка текстов на заимствования (определение степени схожести двух текстов). Система «Антиплагиат»
В рамках проекта требуется разработать ПО, позволяющее в автоматизированном режиме определить степень схожести двух текстов (не менее двух метрик схожести).

9. Простая текстовая игра в жанре Interactive Fiction (http://en.wikipedia.org/wiki/Interactive_fiction)
В рамках проекта требуется разработать текстовую игру в жанре Interactive Fiction

10. Программа, упрощающая создание книг-игр (http://ru.wikipedia.org/wiki/Книга-игра)
В рамках проекта требуется разработать ПО, позволяющее в автоматизированном режиме создавать книги-игры.

11. Программа, упрощающая создание корпуса текстов. Разметка текстов. Поиск по корпусу.
В рамках проекта должен быть разработан модуль, позволяющий в автоматизированном режиме размечать тексты, и модуль, позволяющий осуществлять поиск по размеченным текстам.

12. Оценка психологического состояния пользователя с помощью комплекса вопросов. Программа должна позволять нелинейное движение по списку вопросов т.е. выбирать вопросы в зависимости от ответов (реакции) пользователя.

13. Анализ и сравнение поисковой выдачи Google и Яндекс для уточняющих (семантически близких) запросов. Требуется сформулировать гипотезу о результатах выдачи для различных вариантов запросов и написать программу для проверки гипотезы..

14. Текстовая игра, в которой управление движением и другими действиями пользователя осуществляется с помощью текстовых команд. Требуется реализовать переход минимум между пятью локациями и использование минимум 10 предметов.

15. Оценка уровня популярности автора текста, размещенного на сайте, на основе анализа комментариев и ссылок на текст. Требуется выбрать сайт или коллективный блог, в котором доступно комментирование, и написать программу, выполняющую анализ положительных и отрицательных отзывов (комментариев).

16. Программа, осуществляющая перевод текста на язык жестов. Требуется разработать программу, преобразовывающую текст в последовательность изображений на языке жестов, а также возможность передачи слов дактильным алфавитом. Программа должна распознавать не менее 50 слов.

17. Генератор твитов на основе истории сообщений пользователя.

18. Генератор сказок. Требуется разработать генератор сказок, использующий минимум два различных сюжета и минимум три различных персонажа.

19. Оценка типа восприятия пользователя (аудиал, визуал, кинестетик) на основе анализа записей блогов и текстов, размещенных на сайтах.Требуется разработать программу, оценивающую степень соответствия пользователя каждому типу восприятия по минимум трем критериям.

20. Оценка схожести стихотворного произведения с произведениями известных авторов. Требуется сформировать базу стихотворных текстов не менее 10 авторов, не менее 10 произведений каждого автора. Результатом работы программы должна быть оценка степени схожести введенного текста с одним или несколькими авторами из базы.

21. Транскрибирование текста. Программа должна выводить фонетическую транскрипцию введенного текста.

22. Выявление плагиата на основе статистического анализа текста. Требуется сформировать базу текстов для сравнения (не менее 20 текстов по тематике) и сформулировать не менее 5 признаков схожести текстов. Результатом работы программы должна быть информация о степени схожести текстов по каждому признаку и по совокупности признаков.

23. Выявление терминологических изменений в нормативно-правовых документах разных периодов. Требуется сформировать базу документов из минимум 50 документов одной тематики (например, «Энергосбережение») за последние 40 лет и написать программу, выполняющую статистический и лексикологический анализ этих документов. Результатом работы программы должны быть данные о произошедших за указанный

период терминологических изменениях.

Для всех тем подразумеваются следующие условия:
1. Часть курсовой должна быть написана на языке perl, часть – на python.
2. В отчет о курсовой должно входить техническое задание в соответствии с ГОСТ 34.
3. В отчет о курсовой должно быть включено обоснование выбора используемых алгоритмов.

3.2.Примеры тематик курсовых работ

Формирование и апробация словаря индикаторов для автоматического поиска определений в научных текстах (предметная область - катализ)

Формирование словаря индикаторов для аспектов содержания научных текстов (предметная область - катализ).

ЭЭГ реакции при распознавании русских, английских и китайских предложений у носителей русского языка, изучающих китайский язык

Русские заимствования в английском языке

Наблюдения над русскими длительными конструкциями

Применение ЭЭГ-связанной речедвигательной методики для выявления речевых нарушений у школьников младшего возраста

Методы автоматического определения тональности

Заполнение валентностей существительных некоторыми классами прилагательных и существительными

Поведенческие реакции при распознавании синтаксической структуры и эмоциональной окраски предложений на родном языке у русских и тувинцев

Возрастная динамика мозговой активности у младших школьников в условиях выполнения речедвигательных заданий

Трудности, с которыми сталкиваются студенты Новосибирского государственного университета при составлении текстов различных жанров (научного и официально-делового стилей)

Приёмы речевой манипуляции при просьбах

Фоносемантика: противоречия и перспективы развития

Психологические особенности двусмысленных фраз

Различия разговорной речи юношей и девушек в ситуациях вежливого общения

Построение модели перевода даты с одного языка на другой

Русский язык жестов

Скрытые марковские модели

Анализ смысловых отношений между фрагментами гипертекста (на материале новостных сайтов)