

Санкт-Петербургский государственный политехнический университет  
Физико-механический факультет  
Кафедра прикладной математики

## **КУРСОВАЯ РАБОТА**

**Тема:** *Разработка пользовательского интерфейса сайта для  
изучения неизвестных иностранных слов в субтитрах*

Выполнил студент гр. 5057/2

В.В. Руцкий

Преподаватель:

Н.Б. Суетова

Санкт-Петербург  
2010

## Содержание

Тема работы.....	3
Сравнение с аналогами.....	3
Целевая аудитория.....	4
Пользовательские истории.....	4
Анализ требований.....	4
Внешняя спецификация.....	5
Варианты использования.....	5
Диаграммы деятельности.....	5
Диаграммы состояний.....	9
Прототип.....	10
Руководство пользователя (TODO).....	14
План тестирования (TODO).....	14

## Тема работы

Данный сервис предназначен для людей со средним знанием английского языка, которые могут смотреть и понимать большую часть просматриваемой серии на иностранном языке, но ещё неизвестные для них 5-10% слов или выражений в серии не дают полностью понять ее содержимое.

Если зритель сможет увидеть сериал с субтитрами и подсматривание в словарь при обнаружении неизвестных слов, то это отвлекает его от ее содержания серии и не даёт погрузиться в просмотр. Чтение субтитров до просмотра серии приводит к раскрытию сюжета и просмотр эпизода сериала становится уже неинтересным.

Оптимальным для зрителя выходом было бы иметь словарь неизвестных для него слов или выражений для серии, которую он собирается посмотреть, перемешанных и приведённых к исходным формам так, чтобы никак не отражать содержимое серии. Тогда он мог бы изучить новые слова непосредственно перед просмотром серии и смотреть эпизод, зная практически весь используемый в серии набор иностранных слов.

Проектируемый сервис предназначен для составления такого словаря неизвестных слов.

Базовый вариант использования следующий:

- 1) пользователь загружает субтитры серии, которую хочет посмотреть, на сервис;
- 2) сервис строит словарь используемых слов и **выражений** в загруженном тексте субтитров;
- 3) сервис выделяет из словаря серии неизвестные для пользователя слова и предоставляет их список с переводом и произношением;
- 4) пользователь отмечает изученные или уже известные слова из полученного списка;
- 5) сервис запоминает слова, которые пользователь отметил как (уже) известные ему, и при следующем запросе не выдаёт их.

Возможны (и ожидаемы) вариации способа получения субтитров к серии (например хранить каталог серий и сериалов, для которых сервис уже подготовил словари), вариации способа определения неизвестных слов (например путём определения с помощью теста примерного уровня знаний иностранного языка или известных областей языка пользователя, и выдача слов, которые статистически малоизвестны для класса людей, которому принадлежит пользователь).

Ввиду нетривиальности выделения из произвольного текста выражений и слов в исходной форме, предполагается некоторый элемент совместной работы пользователей по составлению такого словаря или редактированию автоматически созданного словаря.

Также пользователи могут совместно определять малоизвестные слова и выражения — эта статистика может использоваться при определении новых слов для других пользователей.

## Сравнение с аналогами

Задачу, которую должен помочь решить проектируемый сервис, можно обобщить до следующей: конечному пользователю необходимо узнать незнакомые ему слова из иностранного текста и изучить их. Сервисов ни для решения поставленной задачи, ни для обобщённой задачи найдено не было, что, по видимому, связано со сложностью задачи и непрактичности тривиального решения (которое будет реализовано).

## Целевая аудитория

Данный сервис предназначен для следующей аудитории:

- русскоговорящие люди, изучающие английский язык,
- русскоговорящие люди, смотрящие фильмы и сериалы на иностранном языке.

## Пользовательские истории

- Русский студент Боб учится на программиста. Он любит смотреть американские юмористические сериалы на оригинальном языке, т. к. в переводах часто теряются шутки. Боб изучал английский в школе и читал большое количество технической документации на английском языке — он плохо знает грамматику английского языка, но имеет неплохой словарный запас. Смотря серию на английском языке он понимает большую часть речи, и ему лень прерываться на то, чтобы найти в английском скрипте непонятое слово или выражение, а затем посмотреть его перевод. Он бы очень хотел иметь список неизвестных ему слов перед просмотром серии, тогда он был бы мотивирован к изучению новых слов и мог бы смотреть сериалы с большей пользой.
- Русский дядюшка Сэм любит демократию, а также изучать Английский язык. Недавно он узнал про хороший метод изучения иностранного языка и иностранной культуры: через просмотр иностранных телевизионных передач на родном для них языке. Но специально подготовленных передач, которые были бы отсортированы в порядке увеличения сложности, со словарём новых и сложных слов, очень мало. Да и не всегда они ему интересны, а вот некоторые сериалы он видел уже на русском языке, они ему очень понравились, и он был бы рад пересмотреть их в оригинале. Ему бы пригодился сервис, который бы подготавливал словарь для произвольного сериала, а также оценивал его сложность.

## Анализ требований

Сайт должен представлять собой каталог англоязычных сериалов, где для каждого эпизода хранятся субтитры на английском языке, а также подготовленный словарь слов и выражений, используемых в эпизоде.

Для каждого зарегистрированного пользователя, должна храниться статистика известных ему слов и выражений.

Пользователь, авторизовавшись на сайте, может выбрать интересующий его эпизод сериала и получить список новых для него слов и выражений из данного эпизода. Для каждого слова из списка должны приводиться переводы и транскрипции. В процессе ознакомления со новыми словами серии пользователь отмечает слова как известные, далее сервис не должен показывать их в списках новых слов для данного пользователя.

Для привилегированных пользователей должна быть предоставлена возможность добавления нового эпизода. Для нового эпизода пользователь предоставляет файл субтитров, по которому сервис автоматическим образом строит словарь. При необходимости, пользователь должен иметь возможность редактировать словарь эпизода.

# Внешняя спецификация

## Варианты использования

На Рис. 1 приведена диаграмма вариантов использования сервиса.

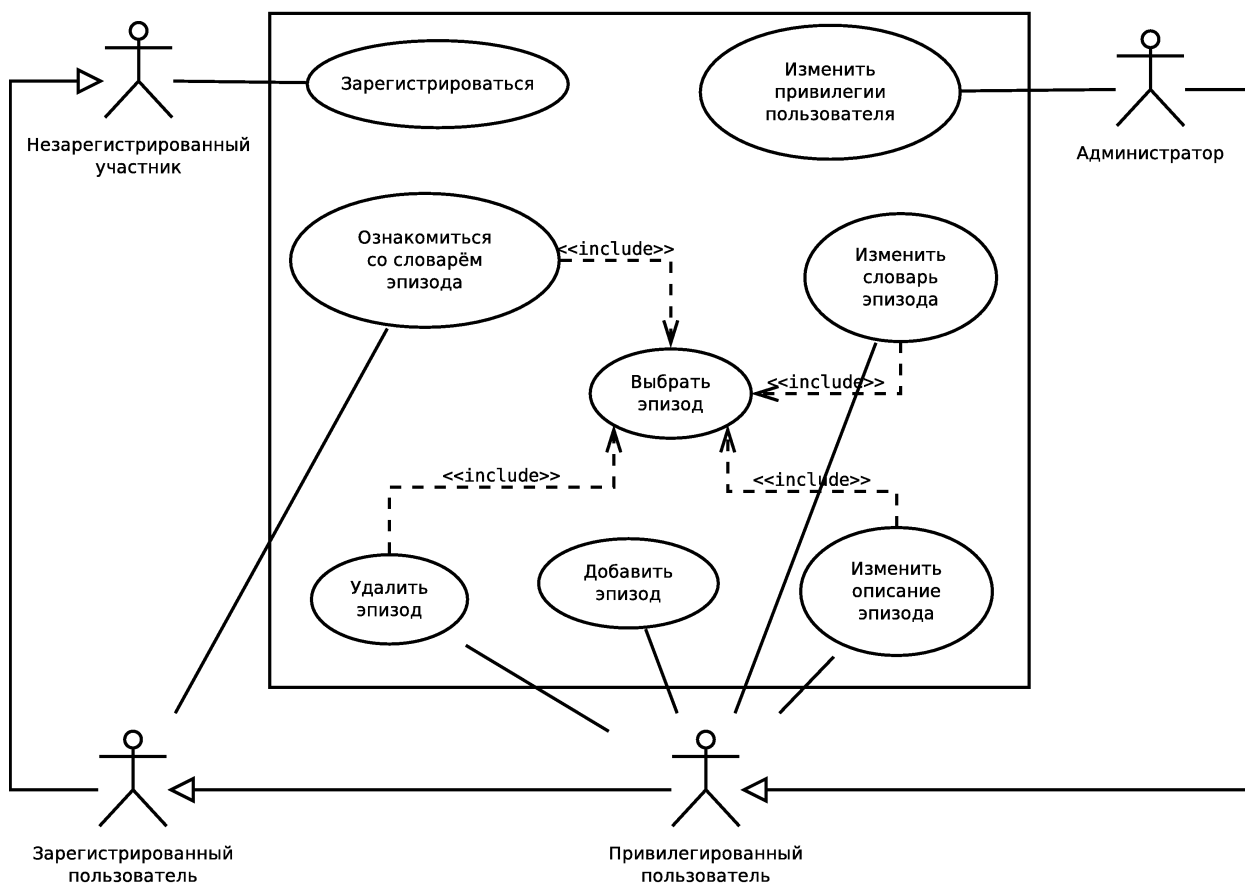


Рис. 1: Диаграмма вариантов использования

## Диаграммы деятельности

Регистрация, изменение привилегий пользователя, изменение описания эпизода и удаление эпизода являются тривиальными вариантами использования.

Все эпизоды находятся в узлах дерева иерархии сериалов (дерево разделяет эпизоды по категориям, например жанр → название сериала → номер сезона), а выбор эпизода представляет собой нахождение узла с интересующим эпизодом в этом дереве.

На Рис. 2 приведена диаграмма деятельности при добавлении эпизода.

На Рис. 3 приведена диаграмма деятельности при изменении словаря эпизода.

На Рис. 4 приведена диаграмма деятельности при ознакомлении со словарём эпизода.

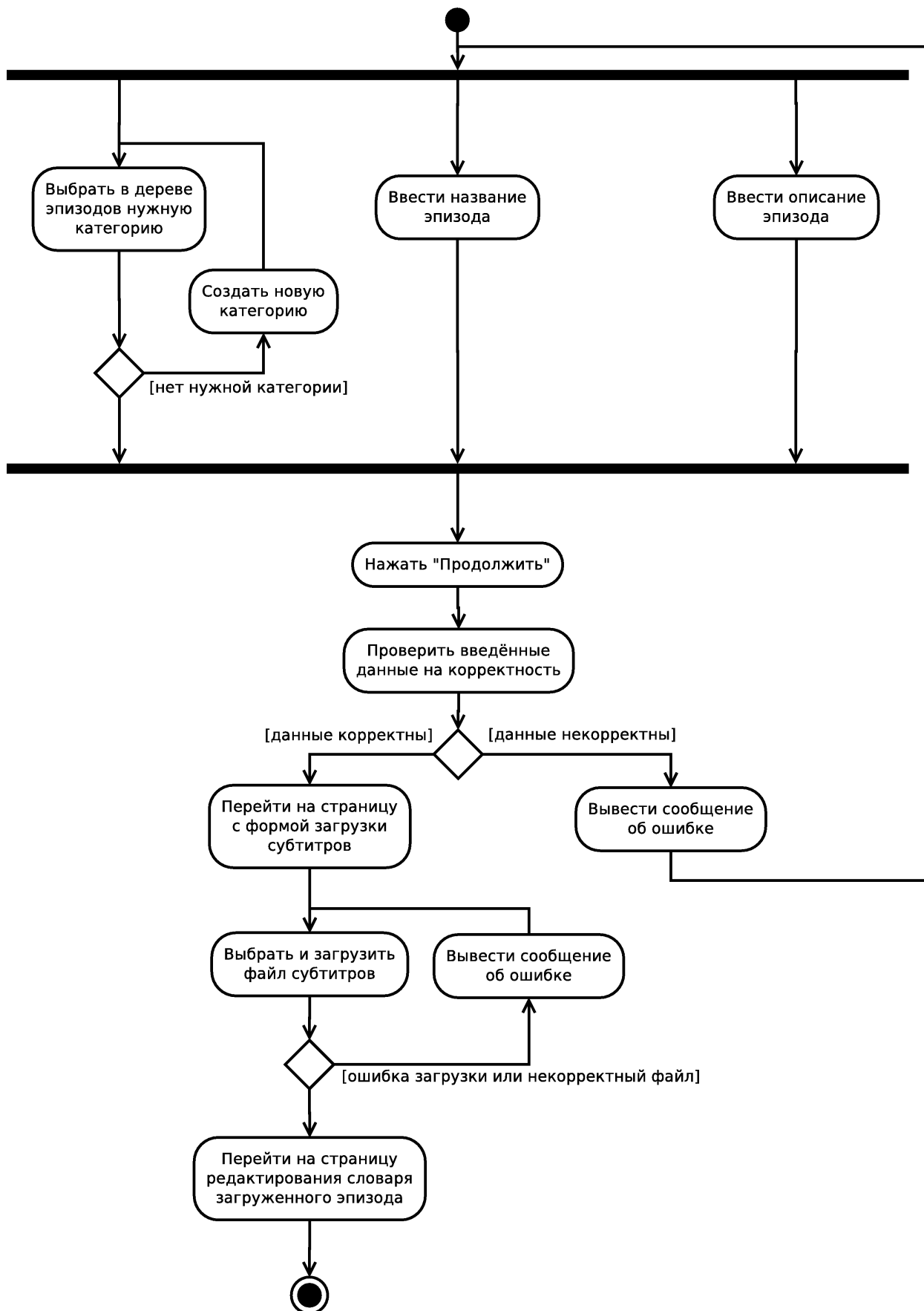


Рис. 2: Диаграмма деятельности при добавлении нового эпизода

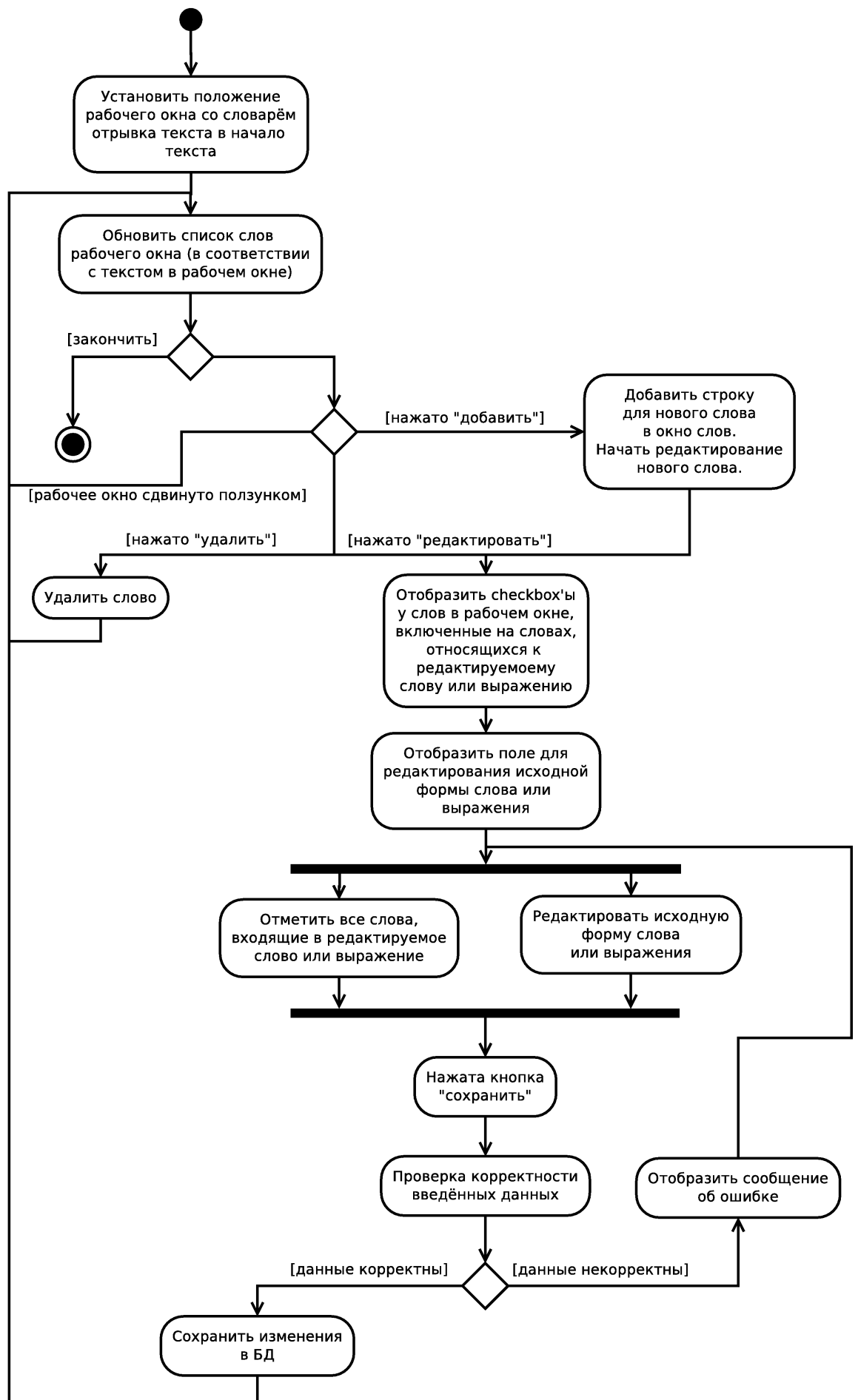


Рис. 3: Диаграмма деятельности при изменении словаря эпизода

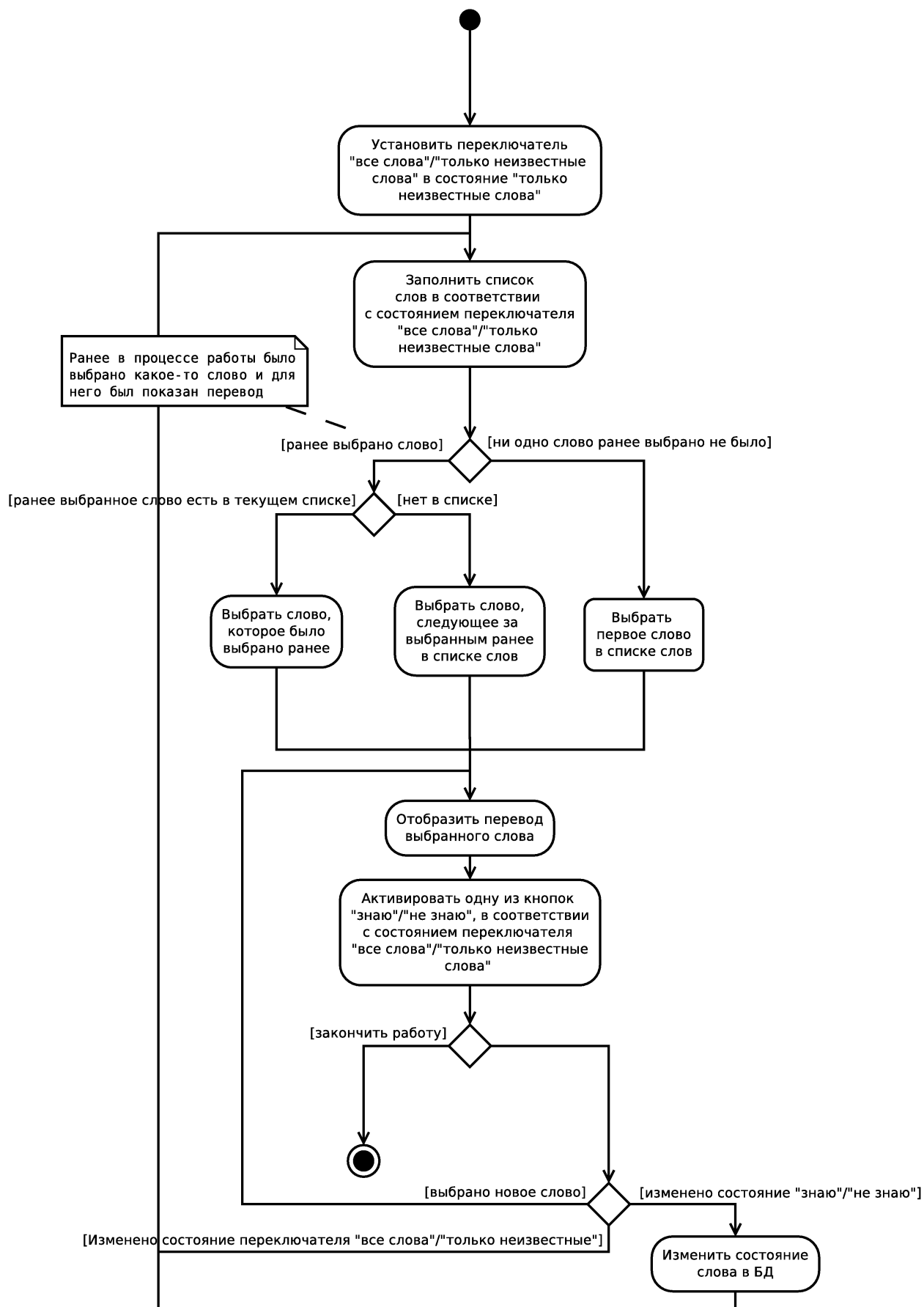


Рис. 4: Диаграмма деятельности при ознакомлении со словарём эпизода



## Диаграммы состояний

На Рис. 5 приведена диаграмма состояний привилегий пользователя.

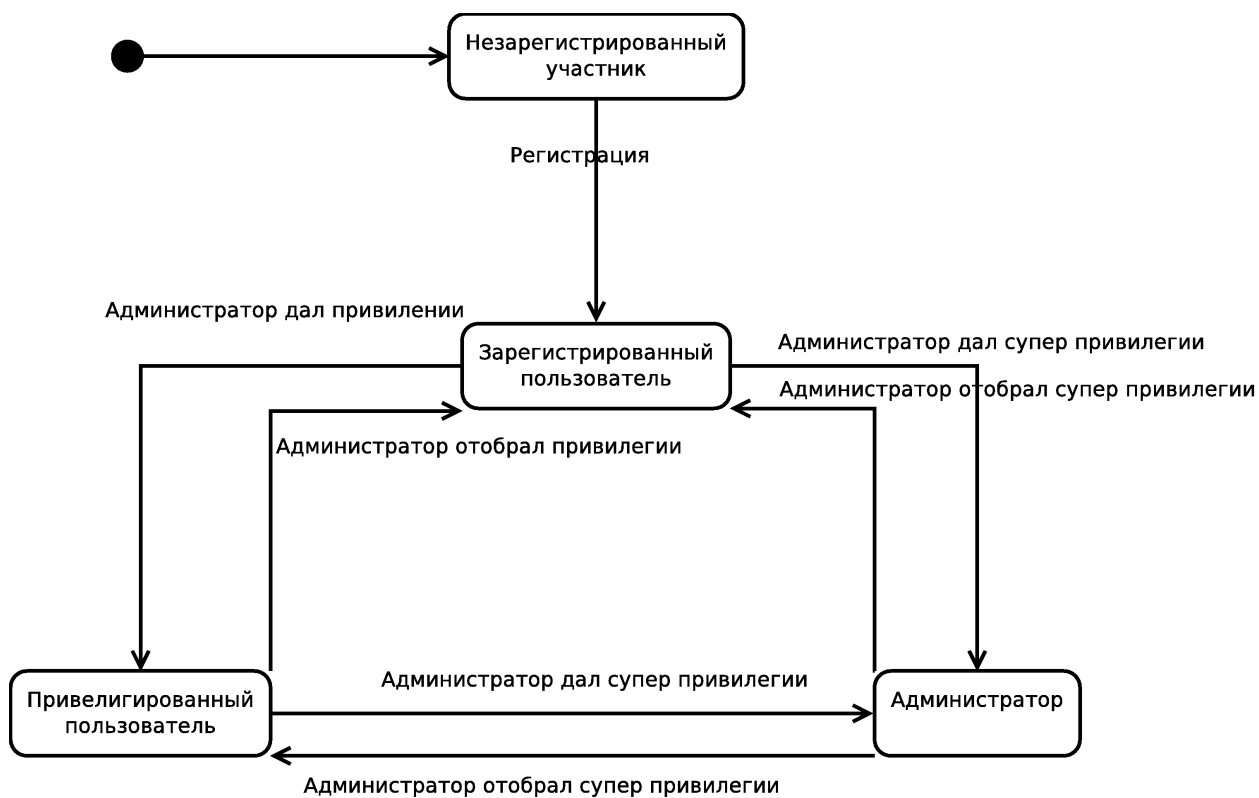


Рис. 5: Диаграмма состояний привилегий пользователя

## Прототип

На Рис. 6 приведён прототип страницы выбора эпизода.

На Рис. 7 приведён прототип страницы редактирования привилегий пользователя.

На Рис. 8 и Рис. 9 приведены прототипы страниц добавления эпизода.

На Рис. 10 приведён прототип страницы изменения словаря эпизода.

На Рис. 11 приведён прототип страницы ознакомления пользователем со словарём эпизода.

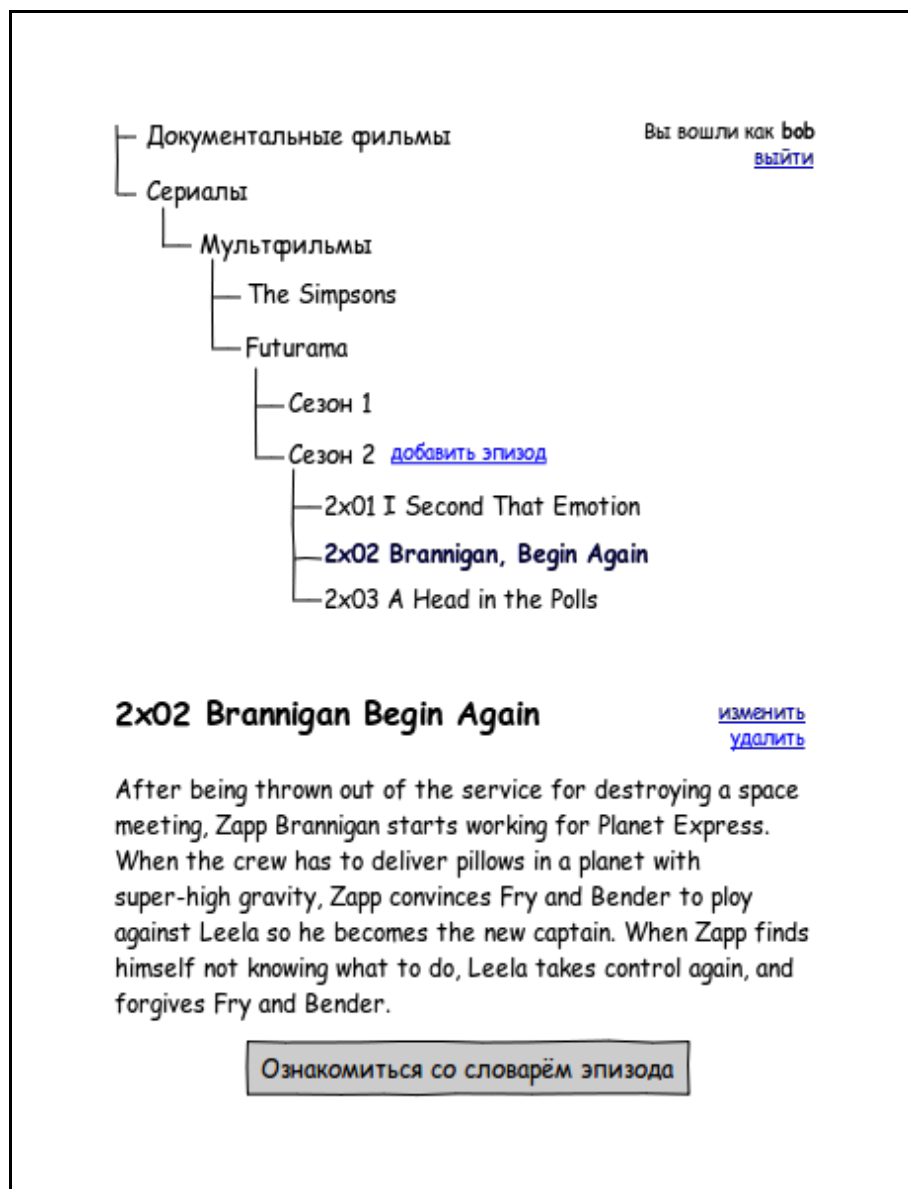


Рис. 6: Прототип страницы выбора эпизода

## Изменить привилегии

Вы вошли как **admin**

[Выйти](#)

Имя пользователя	Адм.	Доб. эп.	Ред. эп.	Уд. эп.	Ред. словарь
admin	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
moderator	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
bob	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
sam	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Сохранить

Рис. 7: Прототип страницы редактирования привилегий

Добавить эпизод

Вы вошли как **moderator**  
[Выйти](#)

Документальные фильмы

Сериалы

Мультфильмы

The Simpsons

Futurama

Сезон 1

**Сезон 2**

2x01 I Second That Emotion

2x02 Brannigan, Begin Again

2x03 A Head in the Polls

[добавить категорию](#)

Название:

Описание:

While on a ski trip, Fry begins to feel nostalgic for 20th century Christmases. To cheer him up, the rest of the Planet Express staff decide to decorate for what is now called Xmas...

Продолжить

Рис. 8: Прототип страницы добавления эпизода. Заполнение описания

Загрузить субтитры

Вы вошли как **moderator**  
[Выйти](#)

Выберите файл с субтитрами для «2x04 Xmas Story»:

Рис. 9: Прототип страницы добавления эпизода. Загрузка субтитров

Изменить словарь эпизода

Вы вошли как **moderator**  
[Выйти](#)

**2x04 Xmas Story**

Hello, there.  
Nothing like a warm fire  
and a Super Soaker of  
fine cognac really puts you  
in the Christmas spirit.

nothing

+

like

+

warm

+

fire

+

fine

+

✖

✖

✖

✖

✖

✖

✖

✖

✖

✖

Рис. 10: Прототип страницы изменения словаря эпизода

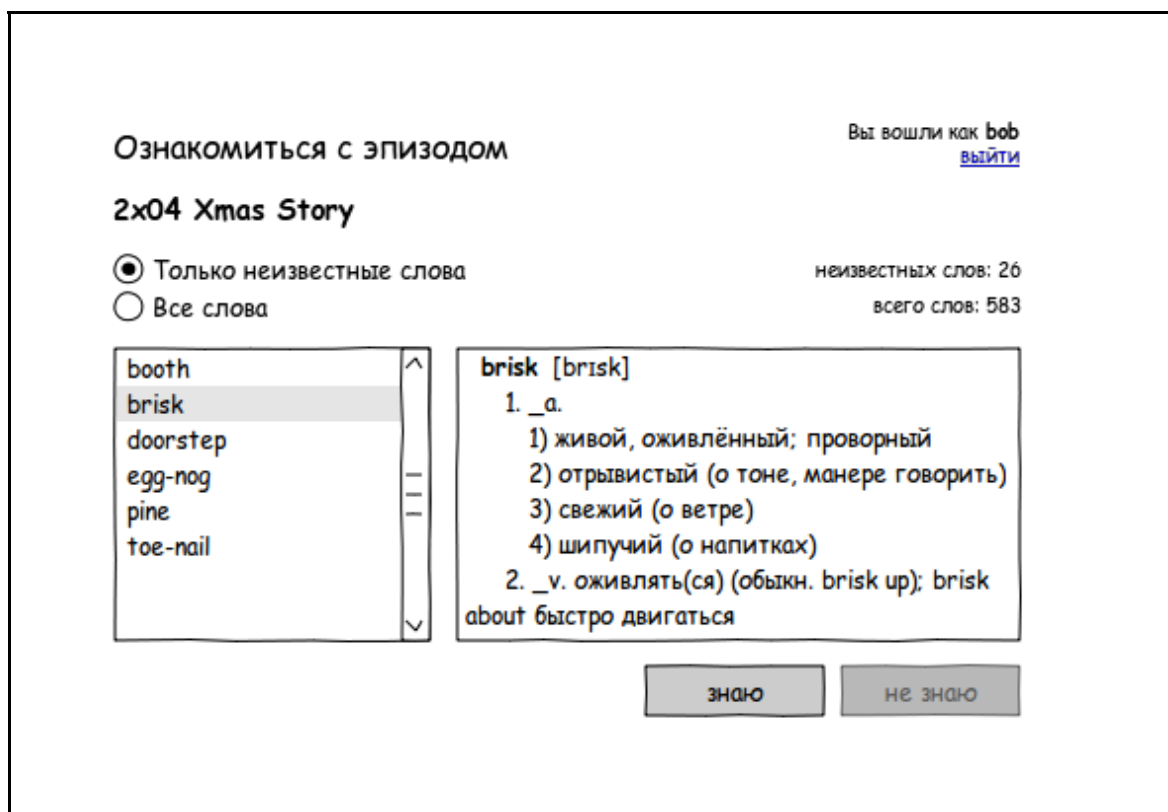


Рис. 11: Прототип страницы ознакомления со словарём эпизода

## Руководство пользователя (TODO)

## План тестирования (TODO)