**МИНИСТЕРСТВО ВЫСШЕГО ОБРАЗОВАНИЯ РФ**

ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

#### Кафедра «Математическое обеспечение и применение ЭВМ»

**Экспериментальные методы оценки ПИ**

Реферат по дисциплине

«Проектирование человеко-машинного интерфейса»

##### Авторы работы Угроватов Д.В.

Лялин Н.С.

## Группа 16ВП1

##### Специальность 09.03.04

##### Руководитель работы к.т.н, доцент Шибанов С.В.

С.п. Дзюба Е.А.

Работа сдана «\_\_\_»\_\_\_\_\_2019г.

##### Работа защищена «\_\_\_»\_\_\_\_\_2019г.

##### Оценка \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Пенза, 2019г

**Оглавление**

[Введение 3](#_Toc9546068)

[Метод фокус-групп 4](#_Toc9546069)

[Лабораторное тестирование 7](#_Toc9546070)

[Удаленное модерируемое тестирование 7](#_Toc9546071)

[Удаленное немодерируемое тестирование 8](#_Toc9546072)

[A/B тестирование 8](#_Toc9546073)

[Экспертная оценка 9](#_Toc9546074)

[Карточная сортировка 11](#_Toc9546075)

[Практическая часть 13](#_Toc9546076)

[Список литературы 17](#_Toc9546077)

## **Введение**

Процесс оценки качества пользовательского интерфейса достаточно­ субъективен и трудно формализуем [5], однако можно утверждать, что качественный интерфейс должен обеспечивать эффективную и продуктивную работу пользователя. Существует также ряд критериев качества любого интерфейса [1,2, 4], четырьмя основными из которых являются: скорость работы пользователей, количество человеческих ошибок, скорость обучения и субъективное удовлетворение пользователей (соответствие интерфейса задачам пользователя является неотъемлемым свойством интерфейса).

Для анализа последнего из них широко используются такие эвристические методы, как анкетирование пользователей, метод фокус-групп и опрос экспертов [5], но получить однозначную и объективную оценку интерфейса по данному критерию, к сожалению, невозможно.

## **Метод****фокус-групп**

В основе этого метода лежит специальная форма интервью, проводимого в группе. Фокус-группа представляет собой группу пользователей или специалистов (обычно 7-10 человек), не знакомых с предлагаемым им для оценки интерфейсом и, как правило, являющиеся потенциальными или заинтересованными пользователями.

Работа фокус-группы может как предварять количественные исследования, так и проводиться после них.

В первом случае, на тестирование и обсуждение предлагается прототип интерфейса, подлежащего оценке, при этом основная задача фокус-группы – собрать первоначальные мнения об интерфейсе, проверить, насколько он соответствует ожиданиям, выяснить, что вызывает вопросы. Такое исследование позволяет сузить круг проблем и выдвинуть гипотезы для их дальнейшего решения.

Во втором случае, фокус-группы, как правило, направлены на уточнение данных количественного исследования, его дополнения за счет более подробной проработки полученной ранее информации. На тестирование и обсуждение предлагается рабочий вариант информационной системы с интерфейсом, подлежащим оценке. Такое тестирование позволяет выявить то, что было упущено на ранних этапах проектирования пользовательского интерфейса и получить предложения по улучшению интерфейса.

Работа фокус-группы ведется по составленному заранее сценарию, при этом вся беседа должна быть записана на видео или аудио носители для дальней расшифровки и анализа.

С помощью метода фокус-групп можно получить достаточно глубинную информацию об особенностях поведения конечных пользователей, которую просто невозможно выяснить другими методами. Этот метод позволяет и лучше понять пользователей – выявить волнующие их проблемы и пожелания.

Обычно анализ проводят несколько небольших и независимых фокус-групп (важно чтобы группы различались по своему составу, например, это могут быть группы опытных пользователей (технических специалистов), новичков и средних пользователей) такой подход позволяет выявить наиболее проблемные участки в интерфейсе и вместе с тем он позволяет провести оценку в очень короткий срок, не прибегая к масштабному тестированию:

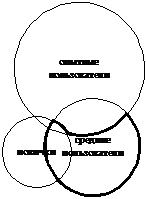


Рис. 2.  Типичное соотношение между числом проблемам (пожеланий) в различных группах пользователей

При этом вполне понятно, что в первую очередь нужно решать проблемы средних пользователей (так их абсолютное большинство) ­– рис. 2.

Полезно произвести несколько повторных оценок интерфейса (итераций) теми же фокус-группами уже после внесения в него изменений.

Качество пользовательского интерфейса можно косвенно оценить через следующие формальные соотношения:

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image006.gif,

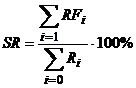
где:

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image008.gif– количество найденных ошибок, проблем и т.п. в интерфейсе

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image010.gif– количество проблем для которых предложено подходящее решение

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image012.gif– фактор продуктивности работы фокус группы

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image014.gif– номер итерации

,

где:

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image018.gif– общее количество проблем, для которых предложено подходящее решение (за все итерации)

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image020.gif– общее количество проблем, для которых предлагалось повторное решение (за все итерации), т.е. таких, первоначальное решение для которых оказалось ошибочным или недостаточным

http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image022.gif– общая оценка неудовлетворенности качеством интерфейса (ясно, что при большом числе повторных изменений она стремится к 100%, что говорит о плохой проработанности интерфейса, его противоречивости).

Опыт показывает [6], что в успешных проектах наиболее типична зависимость рис. 3.

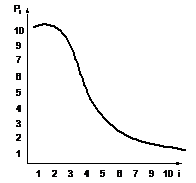


Рис. 3. Типичная зависимость числа ошибок по итерациям

Недостатком метода является то, что пользователи обычно не замечают удачных интерфейсных решений, так как таковые воспринимаются как естественные и не привлекают к себе внимания; поэтому важно с большой осторожностью относиться к изменениям в тех частях интерфейса относительно которых не было никаких комментариев пользователей.

## **Лабораторное тестирование**

Проводится при личном присутствии респондента и модератора. В исследуемом интерфейсе респондент выполняет задания, которые дает ему модератор. Техническое оснащение лаборатории позволяет производить запись сессий или регистрацию движений глаз (eye-traking).

Возможно проводить тестирование пользователей на рабочем месте. Это важно, когда для тестируемого продукта большое значение играет контекст использования — тестирование проводится в рамках заданного контекста и позволяет получить более достоверные данные.

Лабораторное тестирование позволяет работать с множеством качественных метрик: вопросами и затруднениями респондента, его эмоциональным состоянием, ожиданиями и проч. Мы можем тщательно проконтролировать качество выборки: соответствие целевой аудитории и другим требованиям. Все это делает лабораторное тестирование одним из самых эффективных и информативных методов. Но есть и минусы. Во-первых, это дорого. Во-вторых, это требует значительных усилий для респондента, а потому сильной мотивации к участию. Не стоит забывать и про хоторнский эффект: участники эксперимента действуют более усердно благодаря осознанию своей причастности к эксперименту.

## **Удаленное модерируемое тестирование**

Принцип тот же, что и у лабораторного, однако коммуникация между модератором и респондентом осуществляется удаленно, респондент выполняет задания через удаленный рабочий стол.

Данный метод накладывает ограничения на техническое обеспечение тестирования: мы работаем с тем оборудование, которое имеется у респондента. Главный плюс данного метода — вариативная география выборки. Мы можем исследовать аудиторию, которая находятся практически в любой точке мира. Среди других особенностей: возможность контролировать качество выборки и сложности с наблюдением за действиями респондентов и невербальными проявлениями респондентов.

## **Удаленное немодерируемое тестирование**

В данном методе задания формируются в одной из специализированных систем и рассылаются респондентам для самостоятельного выполнения.

Системы фиксируют данные по основным метрикам:

* выполнение заданий;
* пути пользователя по сайту;
* время на выполнение;
* ответы на вопросы;
* тепловые карты движения мыши по сайту.

Мы также можем задать пользователям интересующие нас вопросы.

Этот метод требует большого количества участников — около 100 на каждую целевую группу — и дает в первую очередь статистические данные согласно метрикам. Он дешевле и проще в реализации, чем описанные выше методы. Однако здесь из-за размера выборки осложнен контроль ее качества.

## **A/B тестирование**

Здесь несколько версий контента с незначительными различиями показываются большим группам пользователей, фиксируются и сравниваются данные по интересующим метрикам (целевые действия, время, глубина просмотра и проч.).

Качество тестирование определяют:

* простые различия между версиями — можно понять, что именно повлияло на поведение пользователей;
* внушительная выборка респондентов — часто ориентируются на 1000 респондентов для каждой версии, однако это число зависит от задач;
* однородность выборок для каждой версии.

В противном случае мы получим некорректные данные или не сможем их верно интерпретировать.

Существуют также методы ЮТ, в которых в привлечении пользователей нет необходимости. Среди них экспертная оценка и численные методы.

## **Экспертная оценка**

Метод экспертной оценки качества интерфейса заключается в исследовании, насколько анализируемый интерфейс соответствует известным правилам, рекомендациям и методикам. В ходе такой оценки выявляются несоответствия и противоречия, которые и должны быть устранены.

Перед проведением оценки эксперт составляет список правил в порядке их важности, которые должны быть соблюдены. В этот список входят как рекомендации поставщика ОС и инструментальных средств, так и наработанные в данной предметной области типовые решения. При оценке проверяют насколько тот или иной интерфейс соответствует списку требований.

Данный метод во многом полагается на опыт, компетентность и профессионализм проводящих анализ специалистов.

Экспертная методика оценки определена и в [[ГОСТ 28195-89]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#84) и [[ГОСТ Р ИСО/МЭК 9126-93]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#85). Так, эти стандарты определяют показатели качества ПО и методики их оценки, данные показатели позволяют оценить качество ПО в целом, и в том числе удобство использования.

Так, [[ГОСТ Р ИСО/МЭК 9126-93]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#85) вводит понятие практичности ПО – Набор атрибутов, относящихся к объему работ, требуемых для использования и индивидуальной оценки такого использования определенным или предполагаемым кругом пользователей.

А также следующие характеристики практичности:

* понятность (усилия пользователя по пониманию общей логической концепции ПО и ее применимости);
* обучаемость (усилия пользователя по обучению применению ПО);
* простота использования (усилия пользователя по эксплуатации и оперативному управлению ПО).

В соответствии с [[ГОСТ Р ИСО/МЭК 9126-93]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#85) для ПО диалога конечного пользователя наиболее важны характеристики практичности, а пользователя могут интересовать следующие вопросы:

* Имеются ли требуемые функции в программном обеспечении?
* Является ли программное обеспечение удобным для использования?
* Насколько эффективно программное обеспечение?
* Насколько надежно программное обеспечение?

Основываясь на методике, изложенной в [[ГОСТ 28195-89]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#84) и используя современный набор характеристик качества по [[ГОСТ Р ИСО/МЭК 9126-93]](http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/sp_lit/sp_lit.htm#85) можно предложить следующую процедуру количественной оценки качества:

1. Выбирается категориальная шкала оценки характеристик качества (например целочисленные коэффициенты 0..7, где 0 – качество не удовлетворительно, 7 – предельно достижимый уровень на современном этапе развития отрасли): http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image056.gif.
2. Назначаются количественные значения весовых коэффициентов http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image058.gif характеристик качества (они зависят от потребностей покупателей и сегмента рынка, на который ориентировано ПО), причем http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image060.gif.
3. Определяются (по введенной шкале) количественные значения характеристик качества http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image062.gif продуктов–аналогов, того же функционального назначения, с такими же основными параметрами, подобной структуры и применяемые в тех же условиях эксплуатации.
4. Назначаются количественные значения базовых характеристик качества http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image064.gif– они должны соответствовать современному уровню качества и прогнозируемый мировой уровень (это может быть средний уровень характеристик качества по продуктам аналогам или более высокое значение, учитывающие тенденции развития рынка и технологий).
5. Определяются (по введенной шкале) количественные значения характеристик качества http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image066.gif анализируемого ПО.
6. Производится расчет взвешенной суммы (интегрального показателя качества): http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image068.gif, где http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image070.gif.
7. Качество ПО определяется путем сравнения полученного значения http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image072.gif с соответствующим базовым значением http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image074.gif и индивидуального сравнения с продуктами-аналогами http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/ponomarev2/ponomarev2_files/image076.gif

Уровень качества производимо ПО должен соответствовать базовому уровню и иметь потенциал для роста до предельно достижимого уровня.

## **Карточная сортировка**

Этот метод — палочка-выручалочка для информационной структуры любого сайта и помощник в обнаружении проблем с навигацией. Если ресурс плохо продуман, это ведет к тому, что посетителям трудно понять, в каком разделе каталога находится то, что они ищут, или что скрывается за названием сервиса. В результате запутавшийся пользователь покидает сайт.

Суть карточной сортировки: представителям целевой аудитории выдается набор бумажных карточек с названиями пунктов существующего каталога, а также несколько пустых карточек на случай, если тестируемые захотят вписать что-то свое. Затем участникам сортировки предлагают разложить карточки в группы и дать каждой группе общее название. Такой метод называется **открытой карточной сортировкой**.

При **закрытой карточной сортировке** участники получают карточки с названиями пунктов меню или материалов и готовый список основных групп, в которые нужно поместить выданные карточки. Этот метод часто используется, когда на сайте появилась новая информация и ее надо вписать в существующую структуру.

В результате сортировки мы получим иерархическую организационную схему, которая показывает, как наши пользователи представляют картину мира: как называют группы материалов на сайте, как группируют информацию: по процессу, по предмету или как-то еще.

Полезно попросить участников сортировки в конце теста найти информацию в той структуре, которую они сами создали: так можно проверить, внимательно ли они отнеслись к заданию.

Также можно попросить участников выделить те материалы на карточках, к которым они бы обращались чаще всего при использовании такого сайта. Так вы узнаете, к какой информации необходимо предоставить доступ с главной страницы.

Самый главный плюс этого метода — объективно удобная архитектура сайта, которую спроектировали сами пользователи, а не менеджер проекта или дизайнер на основании своих догадок.

Главный недостаток метода в том, что использовать его для больших сложных сайтов специфической тематики достаточно трудно. Этот способ оптимален для однородных по структуре ресурсов, например, для интернет-магазинов.

## **Практическая часть**

Применим упрощенный метод экспертной оценки, то есть при сравнении будем учитывать только базовый показатель Q0 = 1 (идеальная система). Определим следующие критерии оценки при эксплуатации информационной системы и их весовые коэффициенты (в соответствии с ISO 9241-11):

* Результативность – степень реализации запланированной деятельности и достижения запланированных результатов. a1 = 0,5;
* Удовлетворенность – Отсутствие дискомфорта и положительное отношение к использованию продукции. a2 = 0,2;
* Эффективность – связь между достигнутым результатом и использованными ресурсами. a3 = 0,2;
* Условия использования – Пользователи, задачи, оборудование, физическая и социальная среда, в которых используют продукцию. a4 = 0,1.

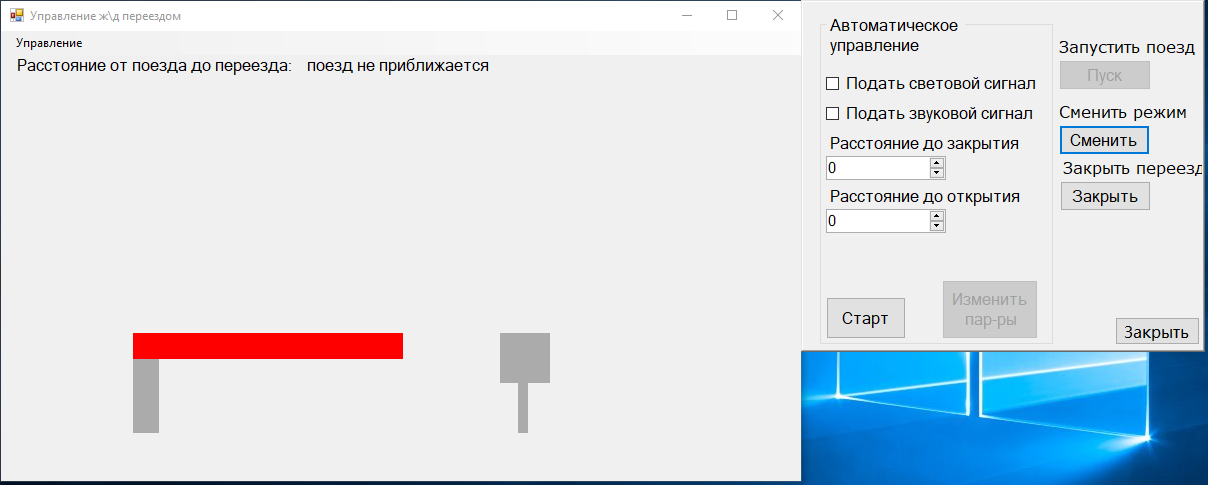


Рисунок 1 - скриншот разработанной ИС

Проанализировав ИС (рисунок 1), мы получили следующие показатели по шкале от 0 (совершенно неудовлетворительно) до 10 (максимальный уровень удовлетворенности):

* Результативность – x1 = 8;
* Удовлетворенность – x2 = 8;
* Эффективность – x3 = 10;
* Условия использования– x4 = 8.

**Вывод:** разработанная система имеет неплохие показатели и имеет потенциал развития.

Далее посчитаем, сколько действий необходимо совершить для достижения следующих целей:

1. В ручном режиме управления открыть переезд и шлагбаум
2. Включить автоматический режим с автоматической подачей светового сигнала
3. В ручном режиме управления открыть переезд и шлагбаум
   1. Нажать кнопку меню «Управление»

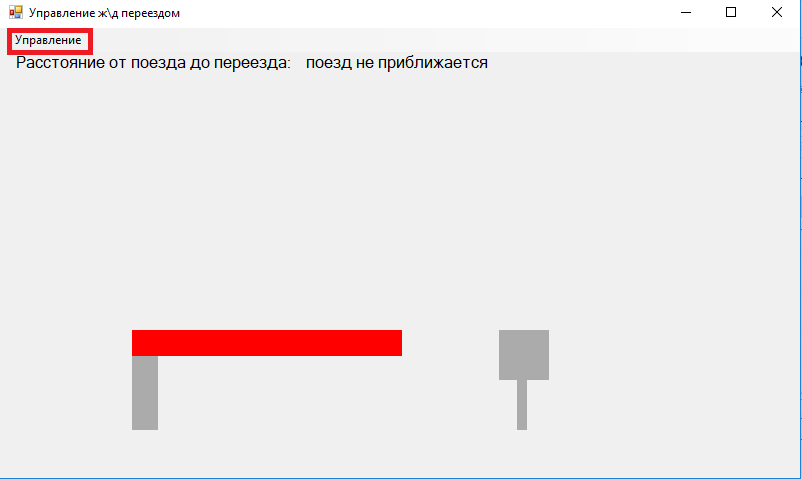


Рисунок 2 - скриншот работы программы

* 1. Нажать кнопку «Открыть» (переезд)

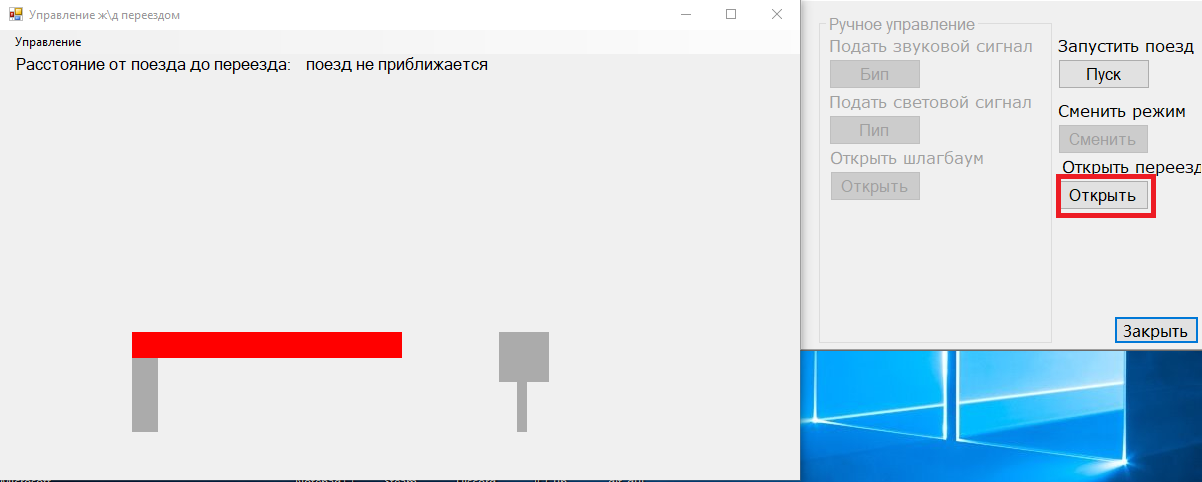


Рисунок 3 - скриншот работы программы

Получается, что данная цель достигается в 2 действия и занимает около 2-х секунд.

1. Включить автоматический режим с автоматической подачей светового сигнала
   1. Нажать кнопку меню «Управление»

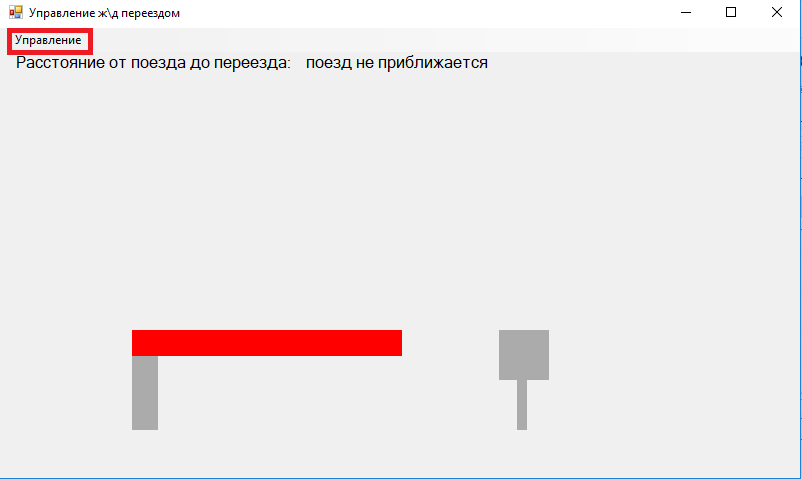


Рисунок 4 - скриншот работы программы

* 1. Нажать кнопку «Открыть» (переезд)

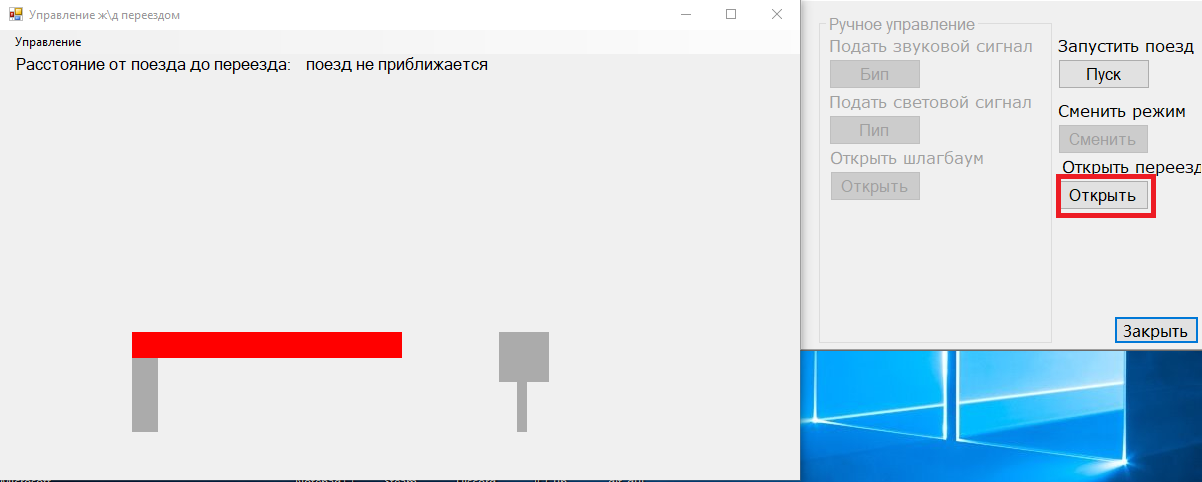


Рисунок 5 - скриншот работы программы

* 1. Нажать кнопку «Сменить» (режим)

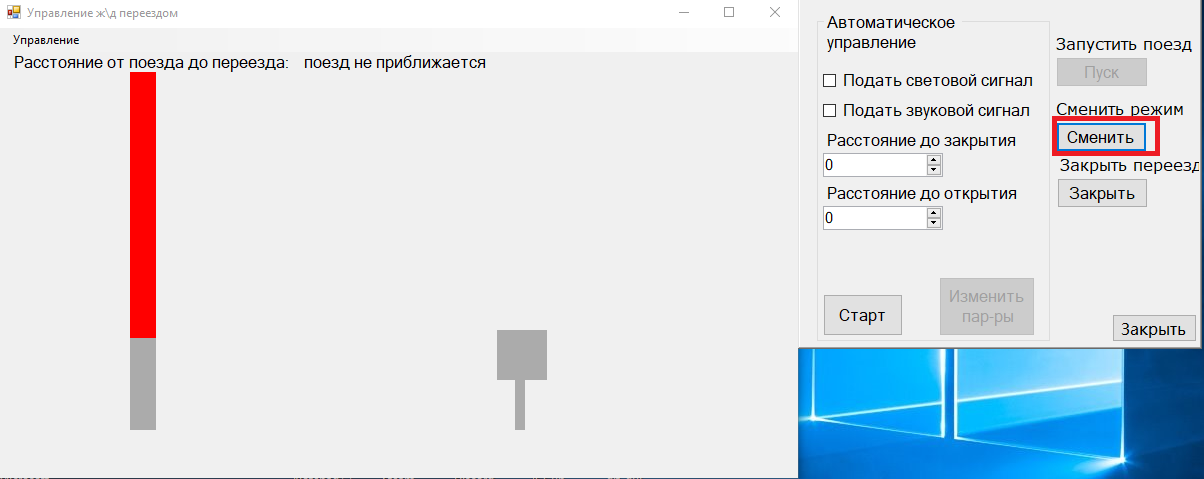


Рисунок 6 - скриншот работы программы

* 1. Нажать флажок «Подать световой сигнал»
  2. Ввести значение в поле «Расстояние до закрытия»
  3. Ввести значение в поле «Расстояние до открытия»
  4. Нажать кнопку «Старт»

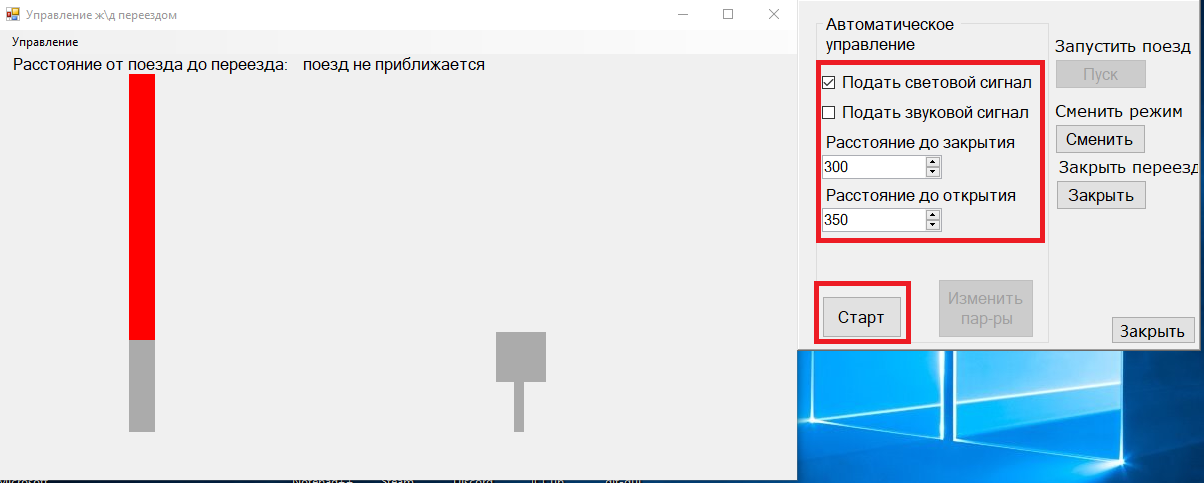


Рисунок 7 - скриншот работы программы

Получается, что данная цель достигается за 7 действий и занимает около 8-ми секунд.

## **Список литературы**

1. Головач В. Дизайн пользовательского Интерфейса: Искусство мыть слона [Электронный ресурс] – Режим доступа: http://uibook2.usethics.ru/ (дата обращения: 17.05. 2019)
2. ГОСТ 28195-89, ОЦЕНКА КАЧЕСТВА ПРОГРАММНЫХ СРЕДСТВ
3. Раскин Д. Интерфейс: новые направления в проектировании компьютерных систем. — СПб: Символ-плюс, 2010.
4. Торрес Р.Дж. Практическое руководство по проектированию и разработке пользовательского интерфейса – М.: Вилиямс, 2002
5. Cooper A., Reimann R. About Face 2.0: The Essentials of Interaction Design– Wiley Publishing Inc., 2003.
6. Michael C. Medlock. Using the RITE method to improve products – Microsoft Corporation, 2002