МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «Крымский федеральный университет имени В.И. Вернадского»

Физико-технический институт (наименование структурного подразделения (филиала)

ОТЧЕТ

по научно-исследовательской практике

обучающегося      Шкурко Виктора Владимировича

Физико-технический институт

Кафедра компьютерной инженерии и моделирования

Направление подготовки/специальность 09.04.01 «Информатика и вычислительная техника»

Направленность(профиль)/специализация\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Курс 2 Группа ИВТ-м-о-212 Форма обучения очная

Руководитель практики от Университета \_\_\_\_\_\_\_\_\_\_\_ профессор,Сухов А.М. (подпись) (должность, ФИО)

Руководитель практики от \_\_\_\_\_\_\_\_\_\_\_ доцент, \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

организации (подпись) (должность, ФИО)

Отчет защищен с оценкой\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ВВЕДЕНИЕ

Научно-исследовательская практика проводиться для формирования понимания и умения у обучающегося самостоятельно проводить научно-исследовательскую работу.

Данная работа содержит в себе теоретические сведенья о объекте и цели выбранной работы для подготовки к научной деятельности в сфере, связанной с выбранной темой и рассматриваемой проблемой.

Основные задачи научно-исследовательской работы:

* Формирования навыков постановки задач и структуры ВКР.
* Обучение качественной оценки сложности выполняемых задач и эффективности их решения.
* Рассмотрение различных научных подходов к решению заданной задачи.
* Формирование навыков проведения научных экспериментов, оценивания результатов исследований;

Поставленные задачи для прохождения практики:

1. Выбор темы научного исследования
2. Составление плана работы
3. Описание объекта и проблемы выпускной квалификационной работы.
4. Создание теоретической части для выпускной и квалификационной работы с описанием подходов решения основной задачи.
5. Оформление реферата

Реферат

Объект исследования – изучение алгоритмов поиска кратчайшего покрытия в неполном автомате.

Предмет исследования – методы поиска кратчайшего покрытия в ЧДКА.

Цель работы – исследование методов эффективного поиска кратчайшего покрытия в неполном автомате. Рассмотрения проблемы доказуемости нахождения самого оптимального решения при применении минимизации неполного автомата.

**Список сокращений и условных обозначений использующихся в ВКР**

КА – конечный автомат

ДКА – Детерминированный конечный автомат

НКА – Недетерминированный конечный автомат

ЧДКА – Частичный детерминированный конечный автомат

ГРАММАТИКА, КОНЕЧНЫЙ АВТОМАТ, ЧАСТИЧНЫЙ АВТОМАТ ДЕТЕРМИНИРОВАННЫЙ АВТОМАТ, ГРАФ, СОТСТОЯНИЕ, СИМВОЛ, МИНИМИЗАЦИЯ АВТОМАТОВ, КРАТЧАЙШЕЕ ПОКРЫТИЕ, ЭКВИВАЛЕНТНОСТЬ АВТОМАТОВ, АВТОМАТ МИЛИ

**Введение**

В наше время практически все сферы деятельности вовлечены в ИТ-инфраструктуру и человек значительную часть времени взаимодействует с различными программными продуктами. Для решения какой-либо специализированной вычислительной задачи создают специализированное ПО.

Часто в качестве моделей как систем, так и самого ПО используют автоматы. Это достаточно изученный и широко применяемый класс моделей дискретной математики. Автоматность обработки информации подразумевает шаговость вычислительной процедуры, причем новое состояние системы обычно определяется текущим состоянием и каким-либо воздействием (к примеру, результатом предшествующих вычислений).

Целью данной работы является исследование методов эффективного поиска кратчайшего покрытия в неполном автомате. Неполным (частичным) принято называть автомат, у которого имеется хотя бы одна неопределенность значения функции перехода или функции реакции. Для вычислительного процесса первая неопределенность обычно означает окончание обработки (финализация процесса вычислений), вторая – несущественность результата на промежуточном шаге.

В данной работе рассматриваются в основном частичные детерминированные конечные автоматы и вычисление их более оптимальных вариантов для ускорения выполнения анализа или преобразования. Такие автоматы применяются в основном в лексических анализаторах и компиляторах.

Целью работы является рассмотрения автоматизации построения кратчайшего покрытия в частичном автомате, что позволяет выполнить его минимизацию.

# **ОСНОВНЫЕ СВЕДЕНЬЯ О ТЕОРИИ АВТОМАТОВ. ЧАСТИЧНО ДЕТЕРПМИНИРОВАННЫЙ АВТОМАТ. МИНИМИЗАЦИЯ И ПОИСК МИНИМАЛЬНОГО ПОКРЫТИЯ.**

## **Основные сведенья о теории автоматов.**

Теория автоматов является разделом дискретной математики, в котором изучают математические и структурные модели преобразователей дискретной информации, называемые автоматами.

Теория автоматов возникла в середине XX в связи с попытками описать на математическом языке функционирование дискретных устройств, вычислительных машин, нервных систем и других реальных и гипотетических дискретных объектов, и процессов. Такие попытки были впервые предприняты зарубежными учеными У. Мак-Каллоком, У. Питсом (1943 г.), С.К. Клини (1951 г.), А. Берксом и Райтом (1954 г.) и др.[4]

Характерной особенностью автоматного описания является дискретность функционирования во времени [4]. Тогда при описании таких моделей внешние воздействия, реакции и состояния автомата рассматриваются как отдельные (дискретные) буквы трех алфавитов, которые соответственно называют входным, выходным алфавитом и алфавитом состояний. Работу автомата можно представить как преобразование по некоторыми законам букв входного алфавита в буквы выходного алфавита.

Для формализации таких законов преобразования задаются две функции – функция переходов и выходов (реакций), аргументами которых являются пары «состояние, входной символ» [4]. Таким образом, в каждый момент дискретного времени автомат, находящийся в определенном состоянии, принимает входной сигнал – символ входного алфавита; выдает выходной сигнал – символ выходного алфавита, определяемую функцией выходов, и переходит в следующее состояние, определяемое функцией переходов. [4]

В теории автоматов рассматриваются различные виды и модификации автоматов классифицируются по уровню абстракции и сложности вычислений, которые они могут выполнить.

* Бесконечные автоматы – к данной группе относят автоматы, у которых хотя бы один из алфавитов (входной, выходной, состояний) бесконечны.
* Детерминированные конечные автоматы – для которых переходы и реакции определены на любых комбинациях «состояние, воздействие»;
* Недетерминированные автоматы, которые реализуют одновременно множественные переходы;
* Вероятностные автоматы – к данной группе относят автоматы, у которых допускаются отношения или случайные функции переходов и выходов.
* Частичные, или неполные, автоматы, у которых значение функции перехода или функции реакции определено не на всех парах «состояние, воздействие»
* Нечеткие автоматы – автоматы, принадлежащие сразу нескольким группам.

В данной работе будет рассматриваться конечные автоматы.

## **Конечные автоматы.**

**Конечный автомат** – это абстрактный автомат, у которого множества входных символов, состояний и выходов конечны.

**Абстрактный автомат** – это абстрактный дискретный исполнитель, преобразующий последовательность входных символов в последовательность выходных символов так, что значение символа на выходе исполнителя зависит не только от значения входного символа, но и от предыдущей последовательности входных символов (предыстории).

Конечный автомат задается пятеркой - кортежем (Q, Σ, δ, q0, F), где:

* Q - конечное множество состояний
* Σ - конечный входной алфавит
* δ - функция переходов, которая отображает Q × Σ в Q
* q0 - начальное состояние, которое является элементом Q
* F - множество конечных состояний, которое является подмножеством Q

Автомат начинает работу в начальном состоянии **q0**. Затем он обрабатывает последовательность входных символов из алфавита **Σ**, на каждом шаге переходя из текущего состояния в какое-то состояние, определяемое функцией переходов **δ**. Если последовательность входных символов приводит автомат в одно из конечных состояний из множества **F**, то автомат принимает данную последовательность входных символов. Если автомат никогда не достигает конечного состояния, то последовательность входных символов не принимается (отвергается) автоматом.

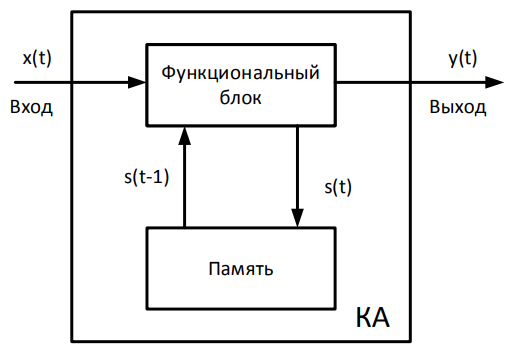


Рисунок 1 - Структурная схема абстрактного конечного автомата.

Конечные автоматы могут делиться на детерминированные и недетерминированные.

### **Способы задания конечных автоматов**

Как уже упоминалось ранее автомат есть модель, исполняющая пошагово некий набор действий переходов исходя из полученных входных сигналов. И для задания такой модели используют автоматные языки. В автоматных языках поведение автомата задается следующими стандартными способами:

* Таблично (таблицы переходов и выходов) – представляет собой таблицу с двойным входом, строки которой нумерованы входными буквами, а столбцы − состояниями. На пересечении строк и столбцов указывается состояние, в которое переходит автомат или выходной сигнал, выдаваемый им.

Таблица Таблица переходов

|  |  |  |
| --- | --- | --- |
|  | 0 | 1 |
| S1 | S2 | S1 |
| S2 | S1 | S2 |

* С помощью графа автомата. Графом автомата называют ориентированный граф, в котором в качестве узла представлены состояния, а дуги представляют переходы между ними.

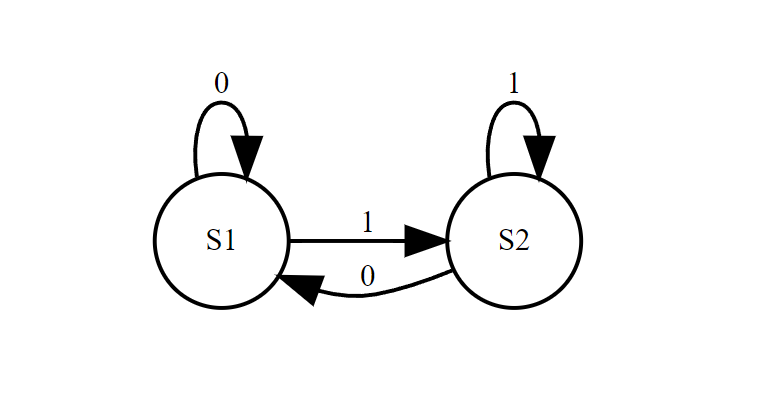


Рисунок Граф автомата

### **Детерминированный и недетерминированный конечный автомат**

Детерминированный автомат (ДКА) — это конечный автомат, в котором каждому состоянию и символу входного алфавита соответствует ровно один переход в следующее состояние [2]. Другими словами, для любого состояния Qi и входного символа Σ k переход δ(Qi, Σ k) = Qj является единственным.

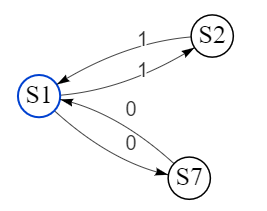


Рисунок 3 - ДКА определяющий язык, слова которого содержат только нули и единицы при чем только четное количество.

**Недетерминированный автомат** – автомат, для которого допускаются многозначные (множественные) значения функции переходов, т.е. для некоторых состояний и символов входного алфавита автомат может переходить в совокупность состояний. Другими словами, функция переходов δ не определена однозначно для всех пар (q, a) из Q × Σ. Вместо этого для некоторых пар (q, a) существует множество состояний, в которые автомат может перейти при чтении входного символа. Для диаграмм состояний это соответствует нескольким альтернативным переходам из одного состояния в несколько состояний по одному и тому же входному символу.

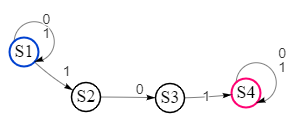


Рисунок 4 - НКА определяющий палиндром нулей и единиц.

Разница между ДКА и НКА заключается в том, что ДКА строго следует одной траектории переходов на каждом шаге обработки входной последовательности, в то время как НКА может выбирать нескольких путей переходов для каждой входной последовательности.

Существует доказательство [3], что вычислительная мощность ДКА и НКА одинакова, т.е. для любого языка, представляемого ДКА, есть эквивалентный НКА, и наоборот.

В данной работе будут фигурировать главным образом ДКА и частичные детерминированные конечные автоматы.

### **Автомат Мили**

Автомат Мили — это конечный автомат, который определяет выходное значение на основе текущего состояния и входного символа. Он отличается от автомата Мура, который определяет выходное значение только на основе текущего состояния.

В автомате Мили выходное значение изменяется в момент перехода в новое состояние, в зависимости от текущего состояния и входного символа. Каждый переход в автомате Мили имеет выходное значение, которое может быть связано с действием, которое должно быть выполнено в ответ на этот переход.

Формально, автомат Мили определяется как пятерка элементов:

* Конечное множество состояний (Q)
* Конечное множество входных символов (Σ)
* Конечное множество выходных символов (Γ)
* Функция перехода (δ: Q × Σ → Q)
* Функция выхода (γ: Q × Σ → Γ)

Функция перехода определяет, в какое состояние перейти при получении определенного входного символа в текущем состоянии. Функция выхода определяет, какое выходное значение должно быть сгенерировано для каждого перехода.

Автомат может быть представлен в виде ориентированного графа, где вершины соответствуют состояниям [4], ребра - переходам между состояниями, и метки ребер - паре (входной символ, выходной символ), связанной с переходом.

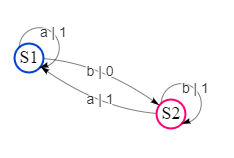


Рисунок 5 - Автомат Мили

## **Частичный детерминированный конечный автомат**

Частичный детерминированный конечный автомат (ЧДКА) — это автомат, у которого некоторые переходы или некоторые реакции не определены. Неполная определенность системы переходов из некоторых состояний позволяет рассматривать неполный автомат как частный случай недетерминированного автомата; детерминированный автомат также можно рассматривать как частным случай неполного (частичного) автомата.

Частичный детерминированный конечный автомат (ЧДКА) формально определяется как пятерка A = (Q, Σ, Γ, δ, q0, F), где:

* Q - конечное множество состояний автомата
* Σ - конечный входной алфавит
* Γ - конечный магазинный алфавит
* δ - отображение из Q × (Σ ∪ {ε}) × Γ в конечные подмножества Q × Γ\*
* q0 - начальное состояние автомата
* F - множество принимающих состояний автомата

Отличительной особенностью ЧДКА является то, что значения функции переходов и реакций определены не на всем множестве пар «состояние, воздействие». То есть у такого автомата могут быть неопределенны некоторые переходы или состояния.

### **Применимость**

ЧДКА могут в общем случае применятся в тех же областях что и полностью определенные ДКА, но под их специфику попадают и области, в которых присутствует факт неполного определения чего-либо.

ЧДКА нашли свое применение в различных областях, например:

* В языковой обработке для распознавания и генерации языковых структур, которые могут быть частично определены или недетерминированы.
* В автоматическом управлении для моделирования и анализа процессов с неполной информацией о состоянии системы.
* В обработке изображений и видео, например, для распознавания жестов, объектов и т.д.
* В биоинформатике, для анализа генетических последовательностей и поиска паттернов в белковых структурах.
* В других областях, требующих моделирования сложных систем, в которых информация может быть частично определена или недоступна.

### **Частичная определенность**

Как уже приводилось ранее ЧДКА отличается от ДКА тем что ЧДКА имеет переходы в неопределенные состояния или при переходе, может быть, не определён выходной символ или и может быть неопределенно и то и другое.

Для ЧДКА можно выделить два вида неопределенности:

* Не определенность реакции на воздействие
* Не определенность перехода в состояние

Так же ДКА может считаться неполным при неопределенном начальном или конечном состоянии.

Таблица Совмещенная таблица для ЧДКА

|  |  |  |  |
| --- | --- | --- | --- |
|  | a | b | c |
| s1 | s2/0 | s3/- | -/- |
| s2 | -/- | s1/1 | s3/0 |
| s3 | s3/0 | s3/1 | s2/0 |

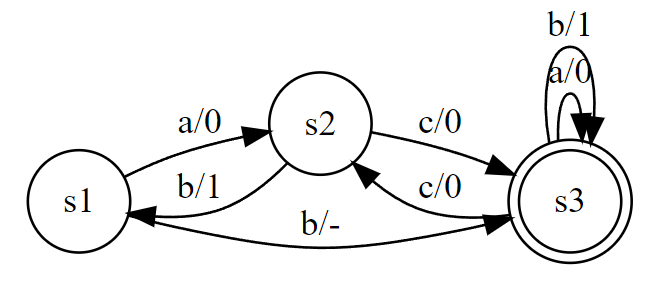


Рисунок Граф ЧДКА

Для таких неполных автоматов можно применить метод доопределения такого автомата чтоб преобразовать его в полностью детерминированный конечный автомат, но в некоторых случаях выгоднее оставить все как есть.

## **Минимизация и поиск кратчайшего покрытия ЧДКА**

Основной задачей данной работы является выбор и реализация алгоритма минимизации КА. Минимизация подразумевает собой получения более простого автомата эквивалентного заданному. Минимизация КА используется для упрощения автоматов и уменьшения затрат на их реализацию и тестирование.

Минимизация конечного автомата (КА) — это процесс получения эквивалентного КА, который имеет минимальное число состояний среди всех эквивалентных ему КА. Формально, если задан КА A = (Q, Σ, δ, q0, F), то минимизация КА заключается в построении эквивалентного КА B = (Q', Σ, δ', q0', F'), такого, что:

1. Количество состояний в автомате B минимально возможно среди всех эквивалентных автоматов.
2. Автоматы A и B распознают одинаковый язык L(B) = L(A), т.е. для любой входной строки w, автомат A и автомат B дадут одинаковый результат: w принадлежит языку или не принадлежит.
3. Каждое состояние автомата B достижимо из начального состояния q0'.

Так как минимизация КА это процесс поиска эквивалентного автомата исходному с меньшим количеством состояний то в случае неполных автоматов задача минимизации сводиться к поиску квазиэквивалентного автомата так как для ЧДКА понятие эквивалентности не применимо.

Квазиэквивалентный автомат – это такой автомат, который по отношению к допустимым для исходного автомата входным последовательностям ведет себя на выходах так же, как и исходный автомат, но имеет меньшее число состояний. Такой автомат называют сокращенной формой исходного.

Отношение квазиэквивалентности рефлексивно и транзитивно, однако не симметрично. Исходя из этого говорят, что сокращенная форма неполного автомата включает исходный неполный автомат.

Так же проблема минимизации конечного автомата заключается в том, что она является NP-полной задачей, то есть в общем случае нельзя построить минимальный автомат за полиномиальное время. Это означает, что при увеличении размера автомата время вычисления минимального автомата растет экспоненциально.

### **Поиск кратчайшего покрытия автомата**

Поиск кратчайшего покрытия (minimum state cover) для конечного автомата (КА) заключается в нахождении минимального подмножества его состояний, которое сохраняет все его функции перехода. Это подмножество состояний эквивалентно исходному КА и содержит минимальное количество состояний, необходимых для его описания. Другими словами, это процесс нахождения эквивалентного минимального детерминированного конечного автомата (ДКА), содержащего наименьшее количество состояний.

Состояние sj автомата реализует (покрывает) состояние si, если:

* любая цепочка, допустимая для si является допустимой для sj
* на одинаковые входные символы в допустимой цепочке для состояний si и sj символы выходной цепочки совпадают там, где для si они определены

Вследствие неполноты задания ЧДКА, покрывающий автомат в общем случае- не единственный, т.е. в общем случае существует множество точных решений. И доказуемость нахождения самого оптимального решения невозможна.

Для поиска кратчайшего покрытия был выбран алгоритм Ангера-Пола так как данный алгоритм может работать с неполными автоматами. На вход принимается неполный детерминированный конечный автомат и возвращает его квазиэквивалентный автомат.

Для поиска кратчайшего покрытия неполного автомата с помощью алгоритма Ангера-Пола нужно выполнить следующие действия:

1. С помощью треугольной таблицы совместимости выделить:
   1. Явно несовместимые состояния
   2. Явно совместимые
   3. Неявно совместимые и совместимые
2. Нормализовать полученную группировку из таблицы совместимости.
3. Проверить правильность и замкнутость полученной максимальной группировки с помощью таблицы переходов и реакций.
4. Найти кратчайшее покрытия с разбором полученной максимальной группировкой с помощью таблицы покрытия и замкнутости.
5. Выбрать одно из полученных решений

## **Постановка задачи исследования**

Задача состоит в выборе алгоритма для эффективного поиска кратчайшего покрытия в неполном конечном автомате. Входными данными является неполный конечный автомат, заданный в виде совмещенной таблицы (автомат Мили) а выходными данными, является таблица и граф эквивалентного заданному автомата с наименьшим количеством состояний.

Необходимо реализовать алгоритм, который находит минимальное количество состояний, которые покрывают все пути в автомате. Для этого необходимо изучить следующие шаги:

1. Поиск всех достижимых состояний в автомате.
2. Конструирование детерминированного эквивалентного автомата по исходному автомату.
3. Поиск минимального покрытия в детерминированном эквивалентном автомате с помощью алгоритма поиска кратчайшего пути в графе.
4. Трансляция найденного покрытия обратно в исходный автомат.

Целью данной задачи является выбор эффективного алгоритма, который способен обрабатывать большие неполные конечные автоматы за приемлемое время. Решение данной задачи может быть применено в различных областях, таких как компиляция, анализ программного кода, распознавание речи и др.

# **ЗАКЛЮЧЕНИЕ**

При выполнении работы были изучены теоретические сведенья о теории автоматов и алгоритмы минимизации ДКА и ЧДКА. Был выбран оптимальный алгоритм эффективного поиска кратчайшего покрытия ЧДКА, а также рассмотрены основные проблемы минимизации КА.

# Список литературы

1. Пентус, А. Е. Математическая теория формальных языков : учебное пособие / А. Е. Пентус, М. Р. Пентус. — 2-е изд. — Москва : ИНТУИТ, 2016. — 218 с. — ISBN 5-9556-0062-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/100633 (дата обращения: 09.03.2021). — Режим доступа: для авториз. пользователей.
2. Миронов, С. В. Формальные языки и грамматики : учебное пособие / С. В. Миронов. — Саратов : СГУ, 2019. — 80 с. — ISBN 978-5-292-04612-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/148854 (дата обращения: 09.03.2021). — Режим доступа: для авториз. пользователей.
3. Жильцова, Л. П. Основы теории автоматов и формальных языков в примерах и задачах : учебно-методическое пособие / Л. П. Жильцова, Т. Г. Смирнова. — Нижний Новгород : ННГУ им. Н. И. Лобачевского, 2017. — 64 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/152819 (дата обращения: 09.03.2021). — Режим доступа: для авториз. пользователей.
4. Акинина Ю.С. Элементы теории автоматов: учеб. пособие / Ю.С. Акинина, С.В. Тюрин. Воронеж: ФГБОУ ВО «Воронежский государственный технический университет», 2017. 184с.