



Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ «Робототехника и комплексная автоматизация»

КАФЕДРА «Системы автоматизированного проектирования (РК-6)»

РАСЧЁТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

к выпускной квалификационной работе

на тему

«Разработка компонентов графоориентированного
программного каркаса для реализации сложных
вычислительных методов»

Студент РК6-81Б
 группа

подпись, дата

Тришин И.В.
 ФИО

Руководитель ВКР

подпись, дата

Соколов А.П.
 ФИО

Консультант

подпись, дата

Першин А.Ю.
 ФИО

Нормоконтролёр

подпись, дата

Грошев С.В.
 ФИО

Москва, 2022

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Московский государственный технический университет имени
Н.Э. Баумана
(национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)

УТВЕРЖДАЮ

Заведующий кафедрой РК-6
индекс

_____ А.П. Карпенко

« _____ » _____ 2022 г.

ЗАДАНИЕ

на выполнение выпускной квалификационной работы

Студент группы: РК6-81Б

Тришин Илья Вадимович

(фамилия, имя, отчество)

Тема выпускной квалификационной работы: Разработка компонентов графоориентированного программного каркаса для реализации сложных вычислительных методов

Источник тематики (кафедра, предприятие, НИР): кафедра

Тема выпускной квалификационной работы утверждена распоряжением по факультету РК № _____ от « _____ » _____ 2022 г.

Техническое задание

Часть 1. Аналитический обзор литературы.

В рамках аналитического обзора должны быть рассмотрены различные подходы, направленные на упрощение реализации сложных вычислительных методов. Должно быть проведено сравнение разрабатываемой системы с некоторой аналогичной ей (на усмотрение студента)

Часть 2. Разработка архитектуры программной реализации, программная реализация.

Должны быть спроектированы программные средства для описания и

выполнения обхода графовых моделей сложных вычислительных методов, созданных по методологии GBSE

Часть 3. Проведение вычислительных экспериментов, тестирование.

Более подробная формулировка задания. Должна быть представлена некоторая конкретизация: какие вычислительные эксперименты требовалось реализовать, какие тесты требовалось провести для проверки работоспособности разработанных программных решений. Формулировка задания должна включать некоторую конкретику, например: какими средствами требовалось пользоваться для проведения расчетов и/или вычислительных экспериментов. Например: «Вычислительные эксперименты должны быть проведены с использованием разработанного в рамках ВКР программного обеспечения».

Оформление выпускной квалификационной работы:

Расчетно-пояснительная записка на 42 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.):

количество: 12 рис., 0 табл., 13 источн.
--

/здесь следует ввести количество чертежей, плакатов/
--

Дата выдачи задания «08» февраля 2022 г.

Студент

подпись, дата

Тришин И.В.
ФИО

Руководитель

выпускной

подпись, дата

Соколов А.П.
ФИО

квалификационной работы

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

Министерство науки и высшего образования Российской Федерации
 федеральное государственное бюджетное образовательное
 учреждение высшего профессионального образования
 «Московский государственный технический университет имени
 Н.Э. Баумана
 (национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)

РК
 ФАКУЛЬТЕТ
 КАФЕДРА РК-6
 ГРУППА РК6-81Б

УТВЕРЖДАЮ
 Заведующий кафедрой РК-6
индекс
 _____ *А.П. Карпенко*
 « ____ » _____ 2022 г.

КАЛЕНДАРНЫЙ ПЛАН выполнения выпускной квалификационной работы

Студент группы: РК6-81Б

Тришин Илья Вадимович

(фамилия, имя, отчество)

Тема выпускной квалификационной работы: Разработка компонентов графоориентированного программного каркаса для реализации сложных вычислительных методов

№ п/п	Наименование этапов выпускной квалификационной работы	Сроки выполнения этапов		Отметка о выполнении	
		план	факт	Должность	ФИО, подпись
1.	Задание на выполнение работы. Формулировка проблемы, цели и задач работы	18.02.2022	18.02.2022	Руководитель ВКР	Соколов А.П.
2.	1 часть: аналитический обзор литературы	18.02.2022	31.03.2022	Руководитель ВКР	Соколов А.П.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана (национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

**НАПРАВЛЕНИЕ
НА ЗАЩИТУ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ**

Председателю
Государственной Экзаменационной Комиссии № _____

факультета «Робототехника и комплексная автоматизация» МГТУ им. Н.Э. Баумана

Направляется студент Фамилия Имя Отчество группы РК6-81Б

на защиту выпускной квалификационной работы Тема

Декан факультета подпись декана «11» мая 2020 г.

Справка об успеваемости

Студент Фамилия Имя Отчество за время пребывания в МГТУ имени Н.Э. Баумана
с 2017 г. по 2020 г. полностью выполнил учебный план со следующими оценками:
отлично – [процент] %, хорошо – [процент] %, удовлетворительно – [процент] %.

Инспектор деканата от руки

Отзыв руководителя выпускной квалификационной работы

Студент Фамилия И.О. в процессе выполнения ВКР проявил себя как ... Результаты, полученные в процессе реализации задания, позволили сделать вывод о ... целесообразности/нецелесообразности выбранных путей решения поставленной задачи, ... невозможности применения ... Работа выполнена автором самостоятельно, в полном объеме, в полном соответствии с заданием и календарным планом. Несмотря на сделанные замечания студент достоин «отличной» оценки... и присвоения звания бакалавр техники и технологий по направлению «Информатика и вычислительная техника».

Руководитель ВКР _____ «____» _____ 2020 г.
(подпись) А.П. Соколов (ФИО) (дата)
Студент _____ «____» _____ 2020 г.
(подпись) И.О. Фамилия (ФИО) (дата)

РЕФЕРАТ

выпускная квалификационная работа: 42 с., 12 рис., 0 табл., 13 источн.

CASE-СИСТЕМЫ, РАСПРЕДЕЛЁННЫЕ ВЫЧИСЛЕНИЯ, ГРАФООРИЕНТИРОВАННЫЙ ПОДХОД, СЛОЖНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ, ОПИСАНИЕ БИЗНЕС-ЛОГИКИ.

Данная работа посвящена разработке программного инструментария, позволяющего описывать и реализовать логику решения различных задач, требующих большого количества трудоёмких вычислений. При описании применяется т.н. графоориентированный подход, который позволяет пользователю задавать действия алгоритма или вычислительного метода в виде переходов между состояниями данных. Формируемое описание затем интерпретируется и выполняется с применением стандартных или пользовательских реализаций каждого из переходов.

Реализованные программные средства позволяют структурировать и ускорить разработку наукоёмкого программного обеспечения, применяемого при анализе больших объёмов данных и научно-технических исследованиях.

Тип работы: выпускная квалификационная работа.

Тема работы: *«Разработка компонентов графоориентированного программного каркаса для реализации сложных вычислительных методов».*

Объект исследования: подходы к описанию бизнес-логики в системах автоматизированной разработки программного обеспечения.

Основная задача, на решение которой направлена работа:
@Основная задача, на решение которой направлена работа@.

Цели работы: @Цель выполнения работы@

В результате выполнения работы: 1) предложено ...; 2) создано ...; 3) разработано ...; 4) проведены вычислительные эксперименты ...

СОКРАЩЕНИЯ

API – прикладной программный интерфейс (Application Programming Interface).

DFD – диаграмма потоков данных (Data Flow Diagram).

GBSE – графоориентированный подход к разработке программного обеспечения (graph based software engineering).

JSON – файловый формат для хранения структур данных (Javascript Object Notation).

LCPD – платформы малокодовой разработки (low-code development platforms).

ПО – программное обеспечение.

СОДЕРЖАНИЕ

СОКРАЩЕНИЯ	8
ВВЕДЕНИЕ	10
1 Постановка задачи	24
2 Архитектура программной реализации	25
2.1 Требования к алгоритму обхода графовых моделей	25
2.2 Функциональные структуры данных	28
2.3 Информационные структуры данных	30
2.4 Управляющие структуры данных	31
3 Описание реализованных средств	33
3.1 Используемые технологии	33
3.2 Описание реализованных алгоритмов	33
3.3 Тестирование	35
4 Анализ результатов	36
4.1	36
ЗАКЛЮЧЕНИЕ	37
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	38
ПРИЛОЖЕНИЯ	41
А	41

ВВЕДЕНИЕ

Современные научно-технические исследования зачастую включают в себя задачи, при решении которых требуется большое количество вычислений, для которых задействуются большие вычислительные мощности. К таким задачам относятся, например, задачи анализа, определения характеристик материалов или технических объектов, моделирования сложных динамических процессов. Как правило, для решения подобных задач применяется или разрабатывается специализированное программное обеспечение (далее – ПО).

Среди прочих применяются программные продукты, предоставляющие пользователю формальный язык описания математических выражений и его интерпретатор, выполняющий необходимые вычисления на машине пользователя. К таким системам относятся, например, Mathcad. Также стоит отметить системы специализирующиеся на символьной алгебре, такие, как Maple[1] и Wolfram Mathematica. В настоящее время данные программные комплексы поддерживают решение задач из различных областей математики, включающих в себя теорию графов, теорию множеств и т.д, предоставляют инструменты визуализации и анализа результатов. Все они позволяют выполнять математическое моделирование, в том числе, сложных технических объектов. При всех их преимуществах необходимость формулировать математические постановки решаемых задач (т.е. формировать математические модели, составлять системы уравнений и т.д.) остаётся за пользователем. Зачастую требуется решать множество задач с схожей постановкой, но с различными входными параметрами. Такая необходимость, например, возникает при решении задач оптимизации, где критерием является некоторая характеристика, получаемая в результате решения задачи анализа. Следовательно, целесообразны автоматизированные средства решения типовых задач анализа и моделирования.

Данные средства относятся к специализированному ПО, а потому при их разработке требуются глубокие познания в предметной области. Кроме того, важно, чтобы создаваемая кодовая база была рассчитана на дальнейшую поддержку, что предъявляет соответствующие требования к

структуре исходного кода и документации. Таким образом целесообразно применение некоторых средств, позволяющих организовать разработку программного обеспечения для решения задач моделирования и анализа и повысить его поддерживаемость.

В наши дни популярность приобретает применение т.н. научных систем управления потоком задач (англ. scientific workflow systems). Они предоставляют средства организации этапов решения вычислительной задачи и управления вычислительными ресурсами. Процесс работы с подобными системами состоит из 4 основных этапов:

- 1) составление описания операций обработки данных и зависимостей между ними;
- 2) распределение процессов обработки данных по вычислительным ресурсам;
- 3) выполнение обработки данных;
- 4) сбор и анализ результатов и статистики [2].

Примерами подобных систем могут служить Pegasus[3], Kepler[4] и pSeven[5]. Помимо инструментов загрузки пользовательских реализаций этапов решения задачи они, как правило, представляют библиотеку типовых действий и преобразований, таких, как считывание данных и их сохранение в файлы одного из поддерживаемых форматов, операции со строками, работы с базами данных, и т.д. Кроме того, некоторые из них имеют средства интеграции с другими системами моделирования и анализа, что позволяет задействовать их при расчётах. На рисунке В.1 изображён пример описания некоторого процесса в системе Kepler.

Кроме того, для облегчения процесса разработки трудоёмкого ПО существуют т.н. платформы малокодовой разработки (англ. low-code development platforms, LCPD)[6]. В них, подобно системам управления потоком задач, логика разрабатываемого программного продукта описывается при помощи некоторого формального языка или с использованием графического редактора. От системы к системе подход к описаниям варьируется. Может применяться структурный подход, описывающий шаги алгоритма, или предметно-ориентированный, при

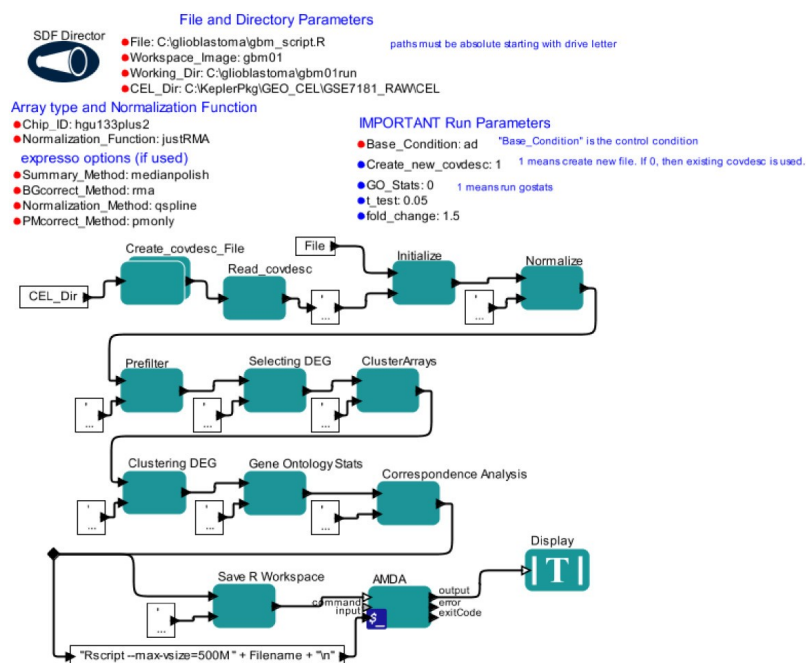


Рисунок В.1. Описание процесса обработки данных в системе Kepler

котором описываются взаимодействующие сущности. Некоторые системы позволяют по созданному описанию генерировать готовые компоненты будущего программного продукта. Так платформа Codebots реализует предметно-ориентированный подход и по составленным UML-диаграммам взаимодействующих сущностей позволяет генерировать API, JSON-схемы данных и документацию[6]. Тем не менее, при реализации сложных вычислительных методов целесообразнее использовать структурный подход.

Одной из ключевых особенностей описанных технологических решений является выделение операций обработки данных в отдельные программные модули (функции, подпрограммы, скрипты). Как правило, при создании описаний алгоритмов в них используется следующий подход. Поскольку известно, что выходные данные одного программного модуля могут являться входными для одного или нескольких других модулей, можно сказать, что между ними формируются зависимости по входным и выходным данным. Тогда возможно составить такой ориентированный граф, описывающий общую логику алгоритма, в котором узлами являются операции обработки данных, а рёбрами – пути данных. Такой подход получил название “диаграммы потоков данных” (англ. Dataflow Diagram, DFD). При известных входных и выходных данных каждого модуля становится возможной их независимая разработка[7]. Таким образом,

уменьшается объём работы по написанию исходных кодов, приходящийся на одного исследователя. Это в свою очередь облегчает отладку и написание документации, что положительно сказывается на общем качестве реализуемого ПО.

```
!!! ----- WARNING ! MISSING PART ----- !!!  
!!! Здесь нужен какой-то переход к тому, зачем может потребоваться  
вводить абстракцию над обрабатываемыми данными  
!!! ----- !!!
```

Таким образом, в некоторых случаях может быть целесообразен такой подход к построению описания логики реализуемого решения, что в нём не указываются конкретные обрабатываемые данные. Последовательность выполнения отдельных этапов в таком случае должна задаваться явно. В предпринимательстве и управлении проектами подобный подход широко распространён и реализован в сетевых графиках. Сетевой график представляет собой ориентированный граф, в котором вершины – это события или состояния проекта, а рёбра – это работы. В работе [8] рассматривается применение идеи переходов между состояниями при описании логики вычислительных алгоритмов. Описанный подход получил название *graph-based software engineering (GBSE)*. Кроме того в указанной работе описана реализация GBSE в библиотеке *comsdk* для языка C++.

Был проведён сравнительный анализ программного каркаса *comsdk* с аналогичным программным комплексом, в котором реализованы диаграммы потоков данных. В качестве такой реализации был рассмотрен программный комплекс *pSeven*, разработанный отечественной компанией *DATADVANCE*. Он направлен в первую очередь на решение конструкторских, оптимизационных задач и, помимо этого, задач анализа данных, что в первом приближении делает его аналогом *comsdk* по предметному назначению.

В терминах *pSeven*: графовое описание процесса решения задачи называется *расчетной схемой* (англ. *workflow*); узлам оргграфа поставлены в соответствие процессы обработки данных (используется термин *блоки*), а рёбра определяют *связи* между блоками и направления передачи данных

между процессами [5]. При работе с pSeven используются следующие понятия:

- расчётная схема – формальное описание процесса решения некоторой задачи в виде ориентированного графа;
- блок – программный контейнер для некоторого процесса обработки данных, входные и выходные данные для которого задаются через порты (см. ниже);
- порт – переменная конкретного¹ типа, определённая в блоке и имеющая уникальное имя в его пределах;
- связь – направленное соединение типа “один к одному” между выходным и входным портами разных блоков.

С учётом данных понятий можно описать используемую методологию диаграмм потоков данных следующим образом. Расчётная схема содержит в себе набор процессов обработки данных (блоков), каждый из которых имеет (возможно, пустой) набор именованных входов и выходов (портов). Данные передаются через связи. Для избежания т.н. гонок данных (англ. data races) множественные связи с одним и тем же входным портом не поддерживаются. Для начала выполнения каждому блоку требуются данные на всех входных портах. Все данные на выходных портах формируются по завершении исполнения блока [5].

Сравнение реализаций двух подходов проводилось по следующим критериям:

- особенности реализуемого подхода,
- особенности программной реализации,
- особенности взаимодействия пользователя с реализованной системой.

С учётом этого были выделены конкретные признаки для сравнения:

- предметное назначение,
- значение вершины графа, описывающего алгоритм,
- значение ребра графа, описывающего алгоритм,
- топология графа, описывающего решение,

¹Динамическая типизация не поддерживается.

- поддержка иерархических графовых описаний, когда одно графовое описание является частью (ребром или вершиной) другого
- принцип передачи данных между отдельными этапами описываемого алгоритма,
- необходимость указывать входные и выходные данные каждого шага алгоритма
- язык программной реализации,
- файловый формат графовых описаний,
- файловая структура проекта реализуемого алгоритма,
- принцип ввода входных данных для алгоритма и его параметров,
- принцип вывода результатов работы алгоритма,
- поддержка параллельного выполнения независимых шагов алгоритма,
- поддержка распределённого выполнения отдельных этапов алгоритма на вычислительном кластере,
- наличие графического редактора графовых описаний
- средства визуализации результатов работы алгоритма,
- поддержка алгоритмов, требующих принятия решения от пользователя,
- возможность дополнения набора входных данных во время работы алгоритма.

Результаты проведённого сравнения представлены в таблице В.2.

Таблица В.2. Сравнительная таблица

№	Признак	pSeven	comsdk
1	Предметное назначение	Задачи оптимизации, анализ данных	Задачи автоматизированного проектирования, алгоритмизация сложных вычислительных методов, анализ данных
2	Значение вершины графа, описывающего алгоритм	Блок (процесс обработки данных)	состояние данных
3	Значение ребра графа, описывающего алгоритм	Связь (направление передачи данных)	переход между состояниями с указанием функций, осуществляющих переход

4	Топология графа, описывающего решение	По умолчанию ациклические графы без ветвлений. Поддерживаемая топология расширяется за счёт специальных управляющих блоков, которые отслеживают выполнение условий: для ветвления используется блок "Условие"(англ. condition), который перенаправляет данные на один из выходных портов в зависимости от выполнения описанного условия (подробнее см. [9]); Для реализации циклов в общем случае используются блоки "Цикл"(англ. loop)[10], но для некоторых задач существуют специализированные блоки, организующие логику работы цикла (например, блок "Оптимизатор"(англ. optimizer))	Любая
5	Поддержка иерархических графовых описаний	Присутствует	

6	Принцип передачи данных между отдельными этапами описываемого алгоритма	Данные между узлами передаются согласно определенным связям, которые на уровне выполнения создают пространство в памяти для ввода и вывода данных для выполняемых в отдельных процессах блоков. Транзитная передача данных, которые не изменяются в данном блоке, на выход невозможна.	Поскольку узлами графа являются состояния данных, существует возможность задействовать в расчётах только часть данных, оставляя их другую часть неизменной. Фактической передачи данных не производится.
7	Необходимость указывать входные и выходные данные каждого шага алгоритма	Присутствует	Отсутствует
8	Язык программной реализации	C++, Python	

9	Файловый формат графовых описаний	Расчетная схема (в форме орграфа) сохраняется в двоичный файл закрытого формата с расширением .p7wf.	Графовая модель (определяет алгоритм проведения комплексных вычислений в форме орграфа) сохраняется в текстовом файле открытого формата, подготовленного на языке aDOT[11], являющегося “сужением” (частным случаем) известного формата DOT (Graphviz).
10	Файловая структура проекта реализуемого алгоритма	Проект состоит из непосредственно файла проекта, в котором хранятся ссылки на созданные расчётные схемы и локальную базу данных, сами расчётные схемы, файлы с их входными данными, файлы отчётов, где сохраняются выходные данные последних расчётов и результаты их анализа.	Проект состоит из .aDOT файла с описанием графа, .aINI-файлов с описанием форматов входных данных, библиотек функций-обработчиков, функций-предикатов и функций-селекторов, файлов, куда записываются выходные данные.
11	Принцип ввода входных данных для алгоритма и его параметров	Входные данные должны быть указаны при настройках внешних входных портов расчётной схемы.	Входные данные хранятся в файле в формате aINI[12], откуда считываются при запуске обхода графа [13].

12	Принцип вывода результатов работы алгоритма	Данные с выходных портов схемы сохраняются в локальной базе данных. Для их записи в файлы для обработки/анализа вне pSeven необходимо воспользоваться специально предназначенными для этого блоками.	Для записи выходных/промежуточных данных в файлы или базы данных необходимо добавить соответствующие функции-обработчики. Формат выходных данных не регламентирован.
13	Поддержка параллельного выполнения независимых шагов алгоритма	Присутствует. Блоки, входящие в состав различных ветвлений схемы могут быть выполнены параллельно, поскольку они не зависят друг от друга по используемым данным.	Присутствует. Существует возможность обойти различные ветвления графа одновременно.
14	Поддержка распределённого выполнения отдельных этапов алгоритма на вычислительном кластере	Присутствует	В текущей версии отсутствует

15	Наличие графического редактора графовых описаний	Да	Да ²
16	Средства визуализации результатов работы алгоритма	Реализованы как часть системы формирования отчётов (см. выше)	В текущей версии отсутствуют
17	Поддержка алгоритмов, требующих принятие решения от пользователя	По умолчанию отсутствует. Требуется реализация дополнительных скриптов на языке Python, отвечающих за взаимодействие с пользователем	Частично присутствует засчёт средства генерации форм ввода[13]

²В виде отдельного веб-приложения

18	Возможность дополнения набора входных данных во время работы алгоритма	Отсутствует	Частично реализована при помощи функций-обработчиков специального типа, создающих формы ввода
----	--	-------------	---

Таким образом, на данный момент comsdk обладает сравнительно меньшим числом функциональных возможностей, чем современные научные системы управления потоком задач, подобные pSeven, но предоставляет потенциально больше средств для взаимодействия реализуемых алгоритмов с пользователем. В условиях существующей на сегодняшний день потребности в отечественном программном обеспечении для реализации сложных численных методов, актуально развитие данного программного каркаса.

1 Постановка задачи

Объектом разработки является программный каркас для реализации сложных вычислительных методов comsdk. Целью разработки является создание новых программных средств для описания и представления графовых моделей сложных вычислительных методов и их обхода. Для достижения цели были поставлены следующие задачи:

1. провести аналитический обзор источников по теме “Системы для реализации сложных вычислительных методов”;
2. провести сравнение рассматриваемой разработки с аналогом;
3. спроектировать структуры данных для описания и представления графовых моделей в каркасе comsdk;
4. разработать алгоритм обхода графовых моделей, поддерживающий циклы и возможности условного обхода ветвей графовой модели;
5. реализовать разработанные алгоритмы и структуры данных на языке C++

Разработка должна вестись с использованием системы контроля версий кафедры РК6.

2 Архитектура программной реализация

Была проведена реорганизация исходного кода программного каркаса, в результате которой было выделено 3 модуля. Текущая структура каталогов проекта представлена на рисунке 2. Были выделены отдельные каталоги

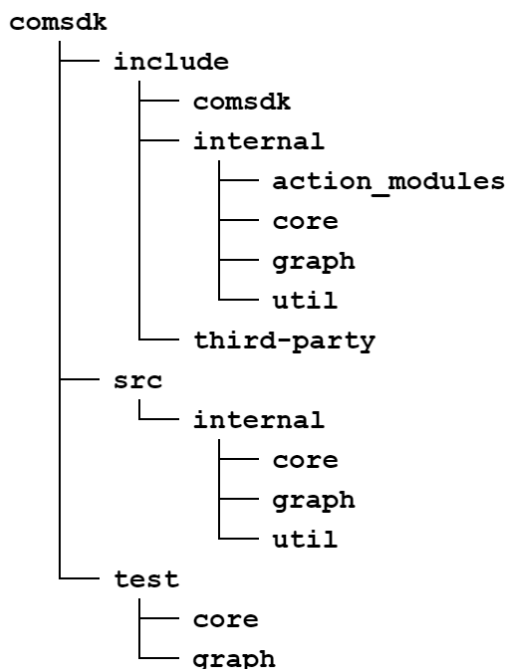


Рисунок 2. Структура каталогов проекта

для заголовочных файлов, где объявлены интерфейсы разработанных структур данных, и для файлов с их реализациями (include/internal и src/internal соответственно). Кроме того, создан отдельный каталог для публичных заголовочных файлов библиотеки include/comsdk, которые будут использоваться конечным пользователем при разработке программных реализаций различных алгоритмов.

2.1 Требования к алгоритму обхода графовых моделей

Методология GBSE подразумевает параллельное независимых ветвей графа. На рисунке 3 после выполнения рёбер F_{12} и F_{13} будет получено два независимых состояния данных S_2 и S_3 соответственно. Далее возникает задача правильным образом перевести данные из этих состояний в общее состояние S_4 .

Данный подход значительно увеличивает эффективность использования ресурсов вычислительной системы и ускоряет процесс решения, однако

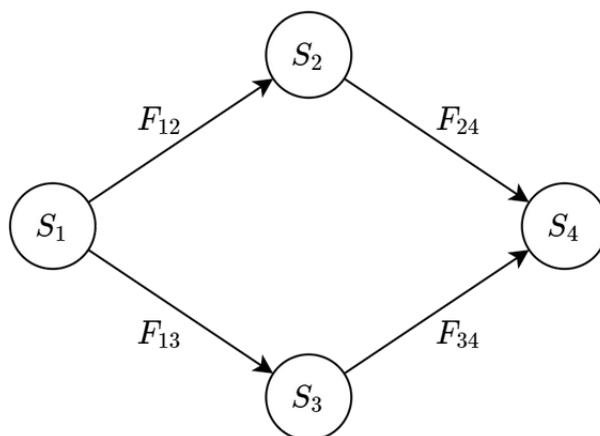


Рисунок 3. Пример графовой модели, требующей параллельного исполнения

добавляет некоторые второстепенные задачи при разработке. Так в примере на рисунке 3 рёбра F_{12} и F_{13} выполнялись параллельно, а значит полученные в результате их выполнения данные существуют в самом общем случае в различных адресных пространствах оперативной памяти (возможно даже на двух разных вычислительных машинах). В момент разветвления графа должно происходить корректное предоставление вычислительным ресурсам (потокам, процессам, узлам кластера и т.п.) доступа к обрабатываемым данным. Помимо этого алгоритм обхода графовой модели должен корректно отрабатывать слияние ветвей графа и в частности при необходимости выполнять сбор данных.

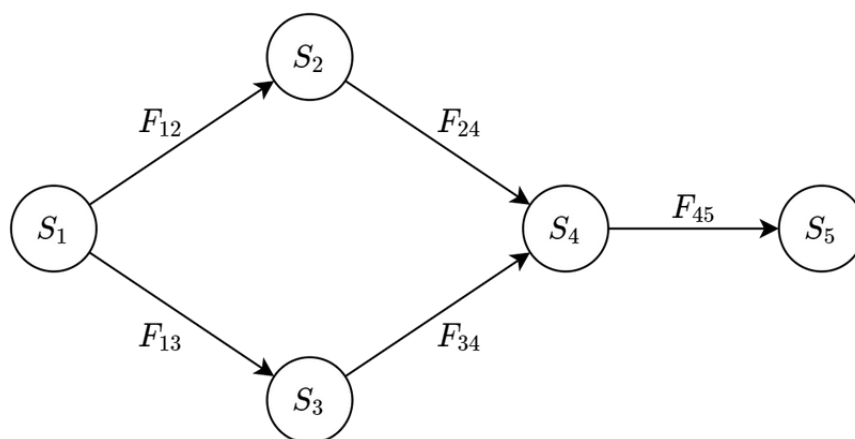


Рисунок 4. Пример графовой модели с совмещением ветвей

На рисунке 4 ветви $S_1 \rightarrow S_2 \rightarrow S_4$ и $S_1 \rightarrow S_3 \rightarrow S_4$ выполняются с использованием различных вычислительных ресурсов, но ребро F_{45} выполняется в пределах одной «общей» ветви графа $S_4 \rightarrow S_5$, и в момент его выполнения ресурсы, выделенные на выполнение двух параллельных ветвей уже не требуются. Таким образом, целесообразно разработать

управляющую структуру, которая бы отвечала за выделение и освобождение вычислительных ресурсов во время работы с несколькими параллельными ветвями графа.

Кроме того, разрабатываемая архитектура должна поддерживать несколько вариантов параллельного исполнения. Среди прочих желательна поддержка:

- поочерёдного выполнения (в первую очередь для отладки) в одном потоке управления;
- выполнения с использованием нескольких процессов операционной системы;
- выполнения с использованием нескольких потоков процессора;
- выполнения на удалённых узлах (через SSH-соединение).

Таким образом целесообразна поддержка единого интерфейса обозначенной управляющей структуры для разных режимов выполнения (последовательный, параллельный, распределённый и пр.).

В случае, когда параллельного выполнения не требуется и предусмотрено условное ветвление, оно должно быть реализовано при помощи специальных функций, привязываемых к узлам графовой модели. В контексте «графоориентированного подхода» такие функции называются *селекторами*. Формально функция-селектор h_i , привязанная к узлу v_i , должна входному набору данных \bar{D} ставить в соответствие множество рёбер E_i , выходящих из v_i , проход по которым нужно совершить. Пример работы функции-селектора демонстрирует рисунок 5.

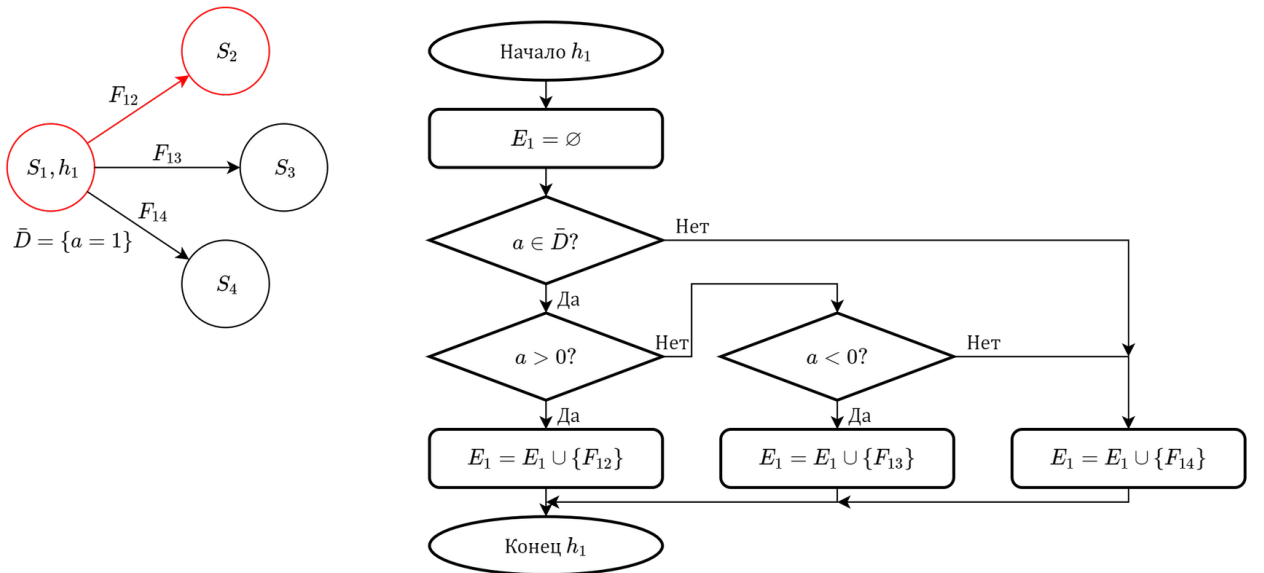


Рисунок 5. Пример фрагмента графовой модели с функцией-селектором

На рисунке 5 красным обозначено ребро, переход по которому будет совершён после вызова функции-селектора.

Алгоритмы обхода графовых моделей, предусматривающие условное ветвление и параллельное выполнение ветвей представлены в разделе 2.1.

2.2 Функциональные структуры данных

2.2.1 Базовый класс

Поскольку методология GBSE позволяет во время работы алгоритма загружать функции-предикаты, функции-обработчики и функции-селекторы из стандартных или пользовательских динамических библиотек, требовалась некоторая абстракция вида «функция, загружаемая из динамической библиотеки». При анализе исходных компонентов библиотеки comsdk, было обнаружено, что такая абстракция уже реализована в классе `ActionItem`. Он представляет собой абстракцию над любой функциональной возможностью реализуемой системы. После выполнения `ActionItem` возвращается код, сообщающий об успехе или ошибке. Возможные значения кода были вынесены в отдельный перечислимый тип `ActionRetCode`. Все значения кодов приведены на рисунке 6. Изначально входными данными для `ActionItem` мог быть только объект класса `Anymap`, поэтому был разработан шаблон класса для поддержки различных типов входных данных. UML-диаграмма разработанных шаблонов классов представлена на рисунке 6.

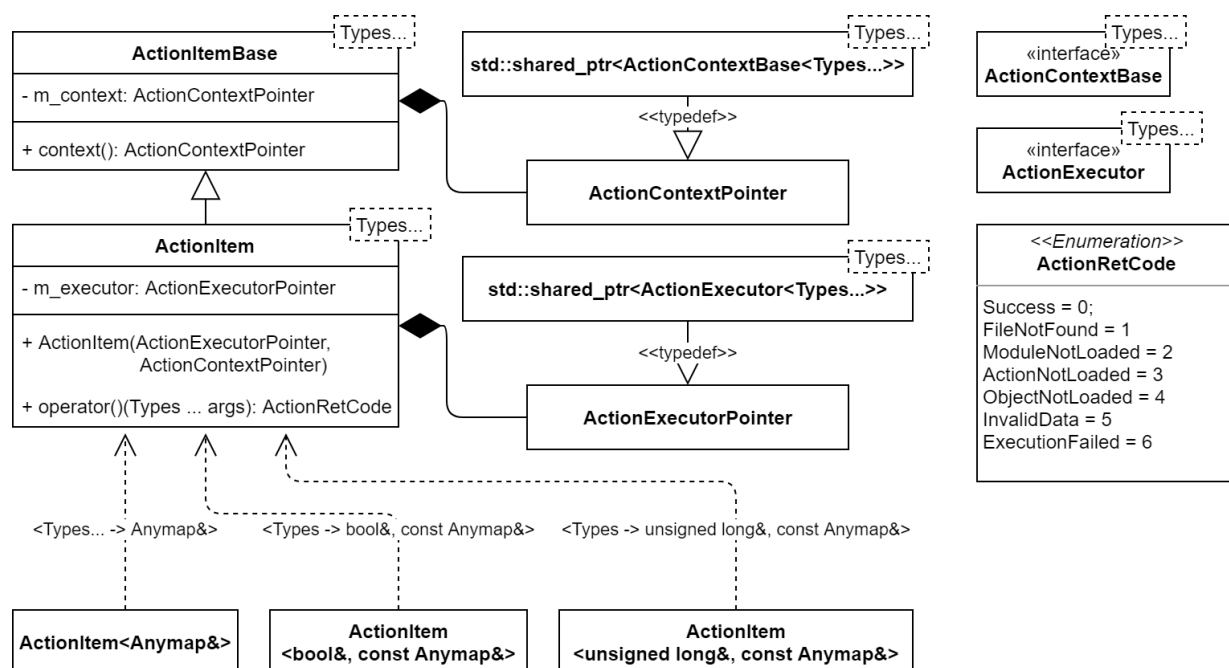


Рисунок 6. Структура классов, описывающих функциональные возможности системы

Запуск функциональной возможность осуществляется посредством перегруженного оператора вызова. Абстрактный класс **ActionContextBase** представляет собой интерфейс для объектов, отвечающих за запуск и хранение мета-информации о выполнении различных **ActionItem**. Базовый класс **ActionExecutor** представляет собой интерфейс для объектов, отвечающих за синхронное выполнение **ActionItem**. В исходной версии **comsdk** уже была реализация этого интерфейса, позволяющая выполнять загружаемые из динамических библиотек функции – класс **SharedLibFuncExecutor**.

Для реализации обработчиков, предикатов и селекторов были созданы специализации шаблона **ActionItem**, которые были использованы в качестве базовых классов для реализуемых в дальнейшем функциональных структур данных.

2.2.2 Класс предиката

Функция-предикат должна принимать на вход данные в начальном состоянии и ставить им в соответствие логические 0 или 1. Поскольку результат выполнения **ActionItem** – это числовой код **ActionRetCode**, то к входным данным помимо ссылки на объект класса **Anymap** была добавлена

ссылка на переменную типа `bool`, куда будет сохраняться результат работы предиката. UML-диаграмма класса предиката, представлена на рисунке 7.

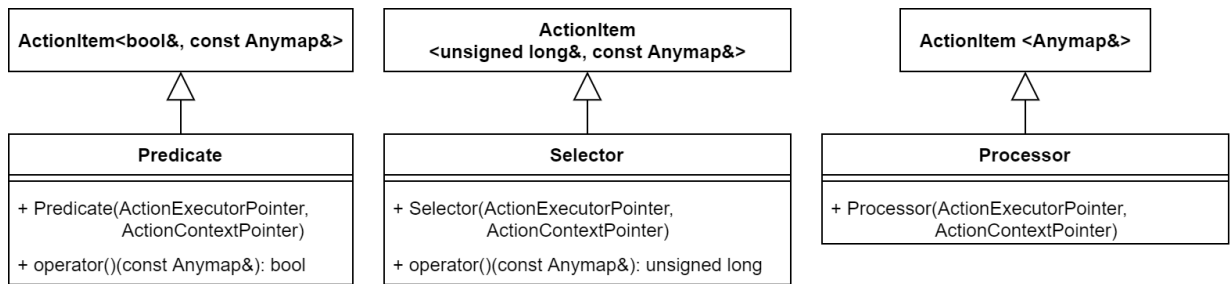


Рисунок 7. UML-диаграмма классов, описывающих функции, выполняемые при обходе графовой модели

Была добавлена обёртка вокруг оператора вызова базового класса, которая возвращает переменную типа `bool`. Это даёт предикату интуитивно понятный интерфейс. В случае, если при выполнении оператора базового класса возникнет непредвиденная ошибка, оператор производного класса вернёт значение `false`.

2.2.3 Класс обработчика

Функция обработчик отвечает за преобразование данных. Ей на вход только требуется ссылка на объект класса `Анумар`, содержащий обрабатываемые данные. Возвращаемым значением такой функции является код ошибки.

2.3 Информационные структуры данных

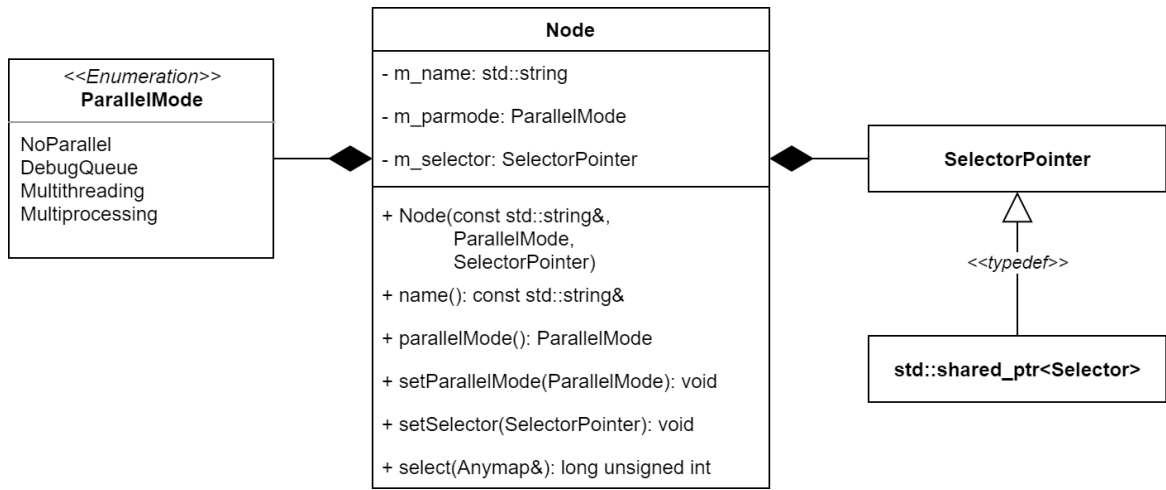


Рисунок 8. UML-описание вершины графовой модели

Структура данных, описывающая вершину графовой модели, хранит в себе имя вершины, режим параллельного обхода нескольких ветвей, выходящих из данной вершины и указатель на функцию-селектор, реализация которых описана в разделе 2.2.

2.4 Управляющие структуры данных

Основной структурой данных, поддерживающей разработанный алгоритм обхода (описан в разделе 3.2) является структура данных «контейнер выполнения». Задача её внутренней логики – контролировать параллельный обход нескольких ветвей графа и отслеживать их слияние. Помимо прочего это подразумевает выделение и освобождение вычислительных ресурсов для параллельного обхода. Поскольку в общем случае могут применяться различные вычислительные ресурсы (потoki ядер процессора, процессы операционной системы, узлы вычислительного кластера и пр.), было принято решение разработать единый интерфейс структуры «контейнер выполнения» без привязки к конкретным ресурсам. Это позволит независимо разрабатывать различные реализации данной структуры для различных режимов параллельного обхода.

Помимо прочего при проектировании интерфейса структуры данных «контейнер выполнения» были задействованы разработанные информационные структуры данных «операция с вершиной» (NodeOp) и «операция с ребром» (EdgeOp).

Интерфейсы спроектированных структур данных были описаны с помощью UML-диаграмм, представленных на рисунке 9. Были учтены особенности описания классов на языке C++.

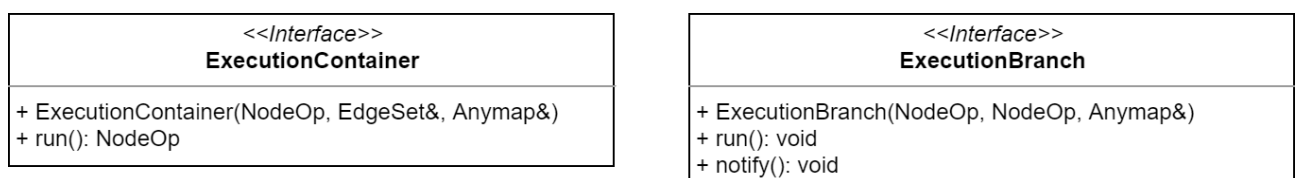


Рисунок 9. UML-диаграммы разработанных структур данных

Конструктор класса **ExecutionContainer** соответствует алгоритму, описанному в разделе 3.2. Метод `run` выполняет алгоритм, описанный блок-схемой на рисунке 11.

Кроме того, была спроектирован интерфейс структуры «ветви исполнения» (ExecutionBranch на рисунке 9). Её задача – выполнение обхода одной конкретно взятой ветви (метод run()) графовой модели с уведомлением «контейнера исполнения» о завершении каждой функции перехода (метод notify()), как того требует алгоритм.

Таким образом, разработанные классы реализуют в себе логику параллельного обхода ветвей графовой модели и хранят данные в соответствии с разработанной процедурой обхода.

3 Описание реализованных средств

3.1 Используемые технологии

3.2 Описание реализованных алгоритмов

Сначала был разработан упрощённый алгоритм обхода, поддерживающий только условное ветвление. Его блок-схема представлена на рисунке 10.

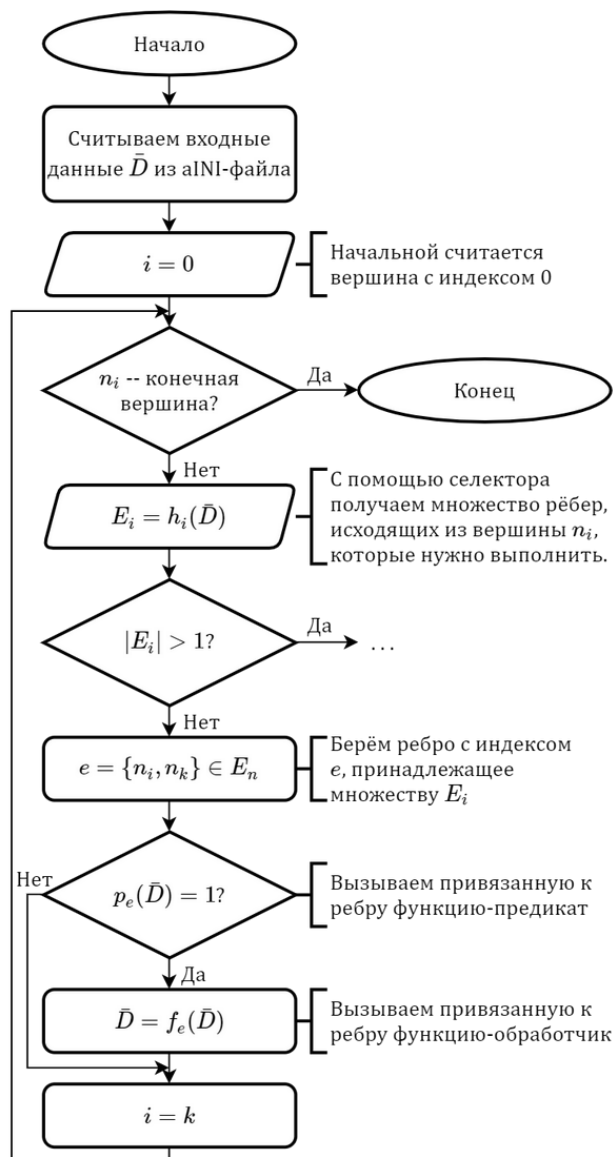


Рисунок 10. Блок-схема алгоритма обхода графовой модели, не предполагающей параллельное исполнение

Далее был отдельно рассмотрен случай, когда необходимо параллельно обойти несколько ветвей. Для контроля за параллельным исполнением функций перехода и выполнения необходимых операций по распределению и сбору данных с задействованных вычислительных ресурсов была разработана управляющая структура «контейнер выполнения». Более

подробно структура описана в разделе 2. Логика работы с данной структурой представлена на рисунке 11.

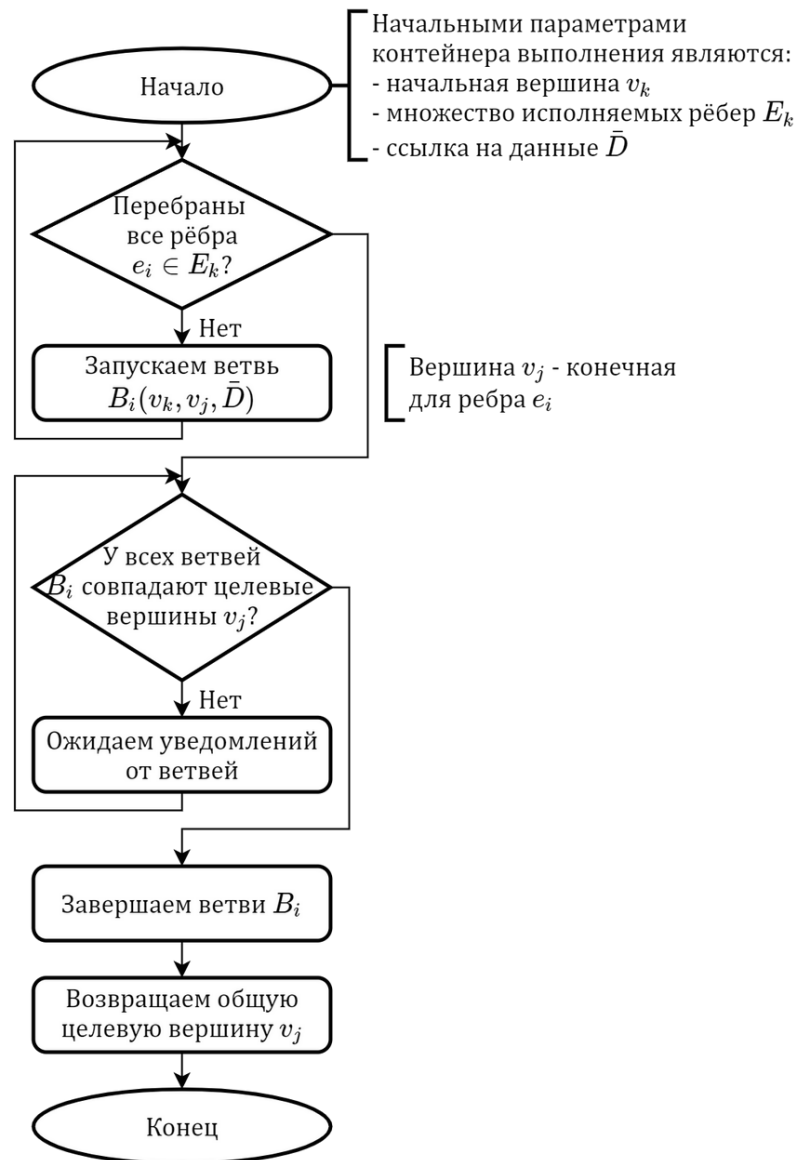


Рисунок 11. Блок-схема алгоритма отслеживания параллельного исполнения ветвей графа

При этом внутри управляющей структуры «контейнер выполнения» подразумевается создание отдельных «ветвей».

Алгоритм обхода одной ветви представлен на рисунке 12.

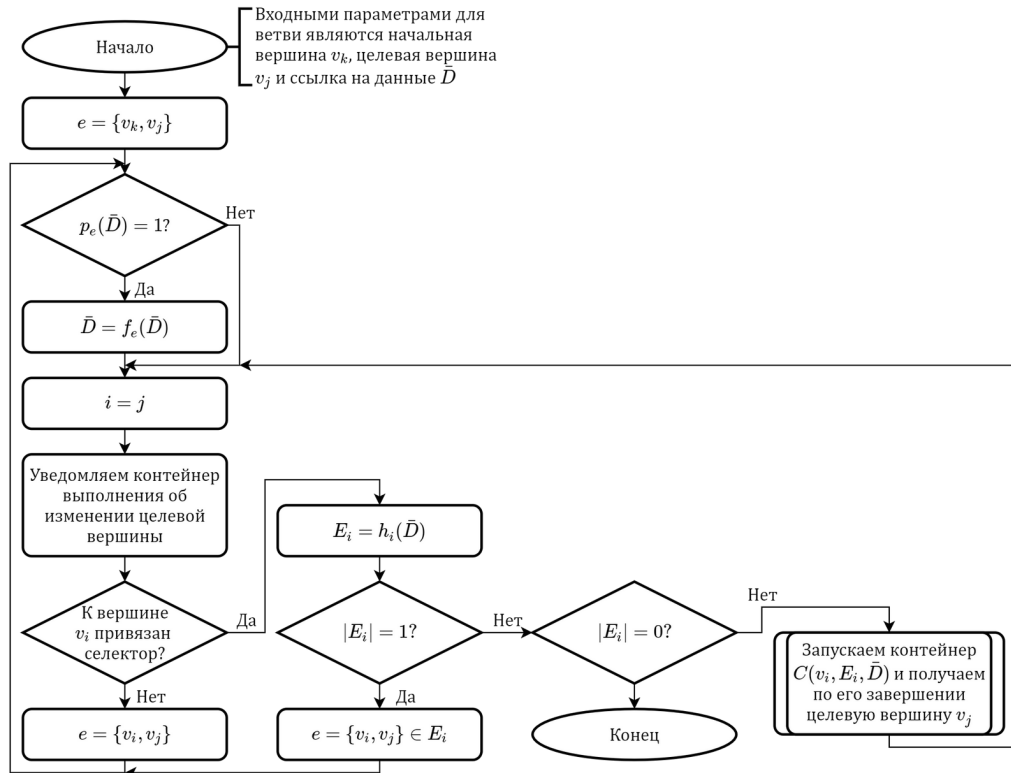


Рисунок 12. Блок-схема алгоритма исполнения одной из параллельных ветвей

Общий алгоритм предполагает завершение выполнения ветви по сигналу от «контейнера выполнения». Ситуация, когда по какой-то причине для текущей вершины v_i и данных \bar{D} не было выбрано ни одного ребра, является исключительной и должна обрабатываться отдельно.

3.3 Тестирование

4 Анализ результатов

4.1 ...

В разделе следует представить анализ полученных результатов, включая указание перспектив развития созданных научно-технических решений.

Обязательность представления: раздел обязателен.

Объём: объём не ограничен, но, как правило, не должен быть меньше 2 страниц.

ЗАКЛЮЧЕНИЕ

В разделе следует представить выводы по работе в целом. Каждый вывод **не должен** быть банальным указанием факта реализации поставленных задач. Каждый вывод должен быть результатом проведенной работы в целом, включая результаты тестирования, вычислительных экспериментов и анализа результатов.

Обязательность представления: раздел обязателен.

Объём: как правило, не должен быть больше 1-2 страниц.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 The design of Maple: A compact, portable, and powerful computer algebra system / Bruce W. Char, Keith O. Geddes, W. Morven Gentleman [и др.] // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 1983. T. 162 LNCS. С. 101 – 115.
- 2 Workflows and e-Science: An overview of workflow system features and capabilities / D. E., G. D., S. M. et al. // Future Generation Computer Systems. 2009. Vol. 25, no. 5. P. 528 – 540.
- 3 Pegasus in the cloud: Science automation through workflow technologies / Deelman E., Vahi K., Rynge M. [и др.] // IEEE Internet Computing. 2016. T. 20, № 1. С. 70 – 76.
- 4 Kepler: An extensible system for design and execution of scientific workflows / Altintas I., Berkley C., Jaeger E. [и др.]. Т. 16. 2004. С. 423 – 424.
- 5 Alexey M. Nazarenko Alexander A. Prokhorov. Hierarchical Dataflow Model with Automated File Management for Engineering and Scientific Applications // Procedia Computer Science. 2015. Т. 66. URL: <https://www.sciencedirect.com/science/article/pii/S1877050915034055?pes=vor>.
- 6 Low-code development and model-driven engineering: Two sides of the same coin? / Davide Di Ruscio, Dimitris Kolovos, Juan de Lara [и др.] // Software and Systems Modeling. 2022. Т. 21, № 2. С. 437 – 446.
- 7 Данилов А.М., Лапшин Э.В., Беликов Г.Г., Лебедев В.Б. Методологические принципы организации многопоточной обработки данных с распараллеливанием вычислительных процессов // Известия вузов. Поволжский регион. Технические науки. 2001. № 4. С. 26–34.
- 8 Соколов А.П. Першин А.Ю. Графоориентированный программный каркас для реализации сложных вычислительных методов // Программирование. 2018. № X.
- 9 Condition - pSeven 6.31.1 User Manual [Электронный ресурс] [Офиц. сайт]. 2022. (дата обращения 07.03.2022). URL:

<https://www.datadvance.net/product/pseven/manual/6.31.1/blocks/Condition.html>.

10 Расчётные схемы - Руководство пользователя pSeven 6.27 [Электронный ресурс] [Официальный сайт]. 2021. Дата обращения: 15.11.2021. URL: <https://www.datadvance.net/product/pseven/manual/ru/6.27/workflow.html#workflow-links>.


11 Соколов А.П. Першин А.Ю. Описание формата данных aDOT (advanced DOT). 2020.

12 Соколов А.П. Описание формата данных aINI (advanced INI) [Электронный ресурс]. Облачный сервис SA2 Systems. [Официальный сайт]. 2020. URL: <https://sa2systems.ru/nextcloud/index.php/f/403527>.

13 Соколов А.П. Першин А.Ю. Программный инструментальный для создания подсистем ввода данных при разработке систем инженерного анализа // Программная инженерия. 2017. Т. 8, № 12. С. 543–555.

Выходные данные

Тришин И.В.. Разработка компонентов графоориентированного программного каркаса для реализации сложных вычислительных методов по дисциплине «Модели и методы анализа проектных решений». [Электронный ресурс] — Москва: 2022. — 42 с. URL: <https://sa2systems.ru>: 88 (система контроля версий кафедры РК6)

Постановка:  канд. физ.-мат. наук, Соколов А.П.

Решение и вёрстка:  студент группы РК6-81Б, Тришин И.В.

2022, весенний семестр

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана (национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

АКТ
проверки выпускной квалификационной работы

Студент группы РК6-81Б

Фамилия Имя Отчество
(Фамилия, имя, отчество)

Тема выпускной квалификационной работы: [Тема]

Выпускная квалификационная работа проверена, размещена в ЭБС «Банк ВКР» в полном объеме и соответствует / не соответствует требованиям, изложенным в Положении о порядке подготовки и защиты ВКР.
ненужное зачеркнуть

Объем заимствования составляет % текста, что с учетом корректного заимствования соответствует / не соответствует требованиям к ВКР
ненужное зачеркнуть

← от руки

↑ бакалавра, специалиста, магистра
от руки

Нормоконтролёр

Согласен:

Студент

Дата:

↑ от руки

↑ подписать
С.В. Грошев
(подпись) (ФИО)

↑ подписать
И.О. Фамилия
(подпись) (ФИО)

