МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Московский Авиационный Институт

(Национальный исследовательский университет)

Институт №8

«Компьютерные науки и прикладная математика»

Кафедра 806

«Вычислительная математика и программирование»

Курсовой проект по дисциплине «Фундаментальные алгоритмы»

Тема: «Разработка алгоритмов системы хранения и управления данными на основе динамических структур данных»

Студент: Жураев Д.С.

Группа: М8О-213Б-21

Преподаватель: Ирбитский И.С.

Оценка:

Дата:

Москва 2023

**Содержание**

[Введение 4](#_Toc148961421)

[Аллокаторы 6](#_Toc148961422)

[Освобождение в рассортированном списке 6](#_Toc148961423)

[Освобождение с дескрипторами границ 7](#_Toc148961424)

[Система двойников 7](#_Toc148961425)

[Деревья 8](#_Toc148961426)

[Двоичное дерево поиска 8](#_Toc148961427)

[Операции двоичного дерева поиска 8](#_Toc148961428)

[Splay tree 9](#_Toc148961429)

[Операции косого дерева 10](#_Toc148961430)

[Splay 10](#_Toc148961431)

[Merge 12](#_Toc148961432)

[Поиск 13](#_Toc148961433)

[Вставка 13](#_Toc148961434)

[Удаление 13](#_Toc148961435)

[АВЛ-дерево 13](#_Toc148961436)

[Операции АВЛ-дерева 13](#_Toc148961437)

[Balance 13](#_Toc148961438)

[Поиск 14](#_Toc148961439)

[Вставка 14](#_Toc148961440)

[Удаление 14](#_Toc148961441)

[Интерактивный диалог с пользователем 14](#_Toc148961442)

[Механизм сохранения состояния системы хранения данных в файловую систему и восстановление состояния системы хранения данных из файловой системы 14](#_Toc148961443)

[Хранение объектов строк с помощью паттерна Flyweight 15](#_Toc148961444)

[Клиент-серверная реализация приложения 16](#_Toc148961445)

[Кастомизация средств IPC 18](#_Toc148961446)

[Shared memory 18](#_Toc148961447)

[File mapping 20](#_Toc148961448)

[Список использованных источников 23](#_Toc148961449)

# Введение

Приложение, реализованное в рамках курсового проекта на языке программирования C++, позволяет выполнять операции над коллекциями данных заданного типа и контекстами их хранения.

Коллекция данных описывается набором строковых параметров (набор параметров однозначно идентифицирует коллекцию данных):

* название пула схем данных, хранящего схемы данных;
* название схемы данных, хранящей коллекции данных;
* название коллекции данных

Коллекция данных представляет собой ассоциативный контейнер (конкретная реализация определяется вариантом), в котором каждый объект данных соответствует некоторому уникальному ключу. Для ассоциативного контейнера необходимо вынести интерфейсную часть (в виде абстрактного класса C++) и реализовать этот интерфейс. Взаимодействие с коллекцией объектов происходит посредством выполнения одной из операций над ней:

* добавление новой записи по ключу;
* чтение записи по её ключу;
* чтение набора записей с ключами из диапазона [𝑚𝑖𝑛𝑏𝑜𝑢𝑛𝑑... 𝑚𝑎𝑥𝑏𝑜𝑢𝑛𝑑];
* обновление данных для записи по ключу;
* удаление существующей записи по ключу. Во время работы приложения возможно выполнение также следующих операций:
* добавление/удаление пулов данных;
* добавление/удаление схем данных для заданного пула данных;
* добавление/удаление коллекций данных для заданной схемы данных заданного пула данных.

Поток команд, выполняемых в рамках работы приложения, поступает из файла, путь к которому подаётся в качестве аргумента командной строки. Формат команд в файле определите самостоятельно.

Также были реализованы дополнительные задания.

Была реализована возможность кастомизации (для заданного пула схем) аллокаторов для размещения объектов данных:

* освобождение в рассортированном списке;
* освобождение с дескрипторами границ;
* система двойников.

Ассоциативный контейнер в рамках реализации данного варианта представляет собой косое дерево.

Также была реализована кастомизация реализаций ассоциативных контейнеров, репрезентирующих коллекции данных: косое дерево, АВЛ-дерево.

Также были реализованы следующие дополнительные задания:

* интерактивный диалог с пользователем;
* хранение объектов строк, размещенных в объектах данных, на основе структурного паттерна проектирования «Приспособленец»;
* механизм сохранения состояния системы хранения данных в файловую систему и восстановление состояния системы хранения данных из файловой системы;
* функционал приложения в виде сервера, запросы на который поступают от клиентских приложений. При этом взаимодействие между сервером и клиентом происходит через средства межпроцессного взаимодействия (IPC);
* кастомизация средств IPC: Unix message queues, Unix shared memory + Unix semaphores, Unix file mapping.

Коллекция данных представляет собой данные о встречах в рабочем календаре, содержащие следующие поля:

* id встречи (данное поле формирует уникальный ключ объекта данных);
* вид встречи (ежедневная встреча/встреча по результатам отчётного периода/собеседование/корпоратив);
* формат проведения (очный/дистанционный);
* описание встречи;
* ссылка на встречу в дистанционном формате;
* ФИО создателя встречи (раздельные поля);
* дата встречи;
* время начала встречи;
* продолжительность встречи (в минутах);
* список приглашённых на встречу (в виде строки).

# Аллокаторы

Аллокатор представляет собой объект, который управляет выделением памяти под данные, хранящиеся в коллекциях. Сам же аллокатор содержит блок памяти с сервисной информацией и блоки, которые выделяются по внешнему запросу.[3]

Реализованы были 4 вида аллокаторов:

* на глобальной куче;
* с освобождением на рассортированном списке;
* с освобождением с дескрипторами границ;
* система двойников.

Также были реализованы методы поиска подходящего для выделения блока:

* метод первого подходящего: выбирается первый по списку свободный участок подходящего размера;
* метод лучшего подходящего: выбирается наименьший по размеру участок из всех подходящих;
* метод худшего подходящего: выбирается наибольший по размеру участок из всех подходящих.

## Освобождение в рассортированном списке

Идея данного аллокатора в том, что каждый свободный блок хранит указатель на следующий блок и свой размер, а в занятых блоках памяти хранится только размер этого блока без указателя на следующий занятый блок. Таким образом, получается рассортированный односвязный список из свободных блоков памяти.[1]

При выделении памяти с помощью этого аллокатора возникают два случая:

* Размер блока равен размеру, запрашиваемому пользователем. В таком случае указатель предыдущего блока переставляется с текущего на следующий;
* Размер блока больше размера, запрашиваемого пользователем. В таком случае свободная часть блока отделяется от занятой, размер блока уменьшается до размера выделяемой памяти и указатель с предыдущего блока переставляется с текущего на оставшийся свободный участок.

Выделенный блок помечается как занятый и результат возвращается пользователю. Также если данный блок был первым в списке, то указатель в сервисной части аллокатора на первый свободный блок переставляется на следующий за этим.

При освобождении памяти надо найти адреса блоков, между которыми стоит наш блок. Как только мы нашли блок, адрес которого больше адреса нашего участка, то значит мы нашли освобождаемый блок. Тут также может быть два случая:

* Освобождаемый участок стоит перед следующим свободным. В таком случае мы склеиваем эти два блока, то есть увеличиваем размер, прибавляя размер этого свободного блока, и переставляем указатель на следующий свободный участок.
* Между освобождаемым и найденным участками есть еще участки. В таком случае у освобождаемого блока указатель ставится на найденный

Сложность операций выделения и освобождения памяти: O(n).

## Освобождение с дескрипторами границ

Идея данного аллокатора заключается в образовании двусвязного списка из занятых участков памяти, благодаря чему освобождение происходит достаточно быстро.

Чтобы найти необходимый свободный блок при выделении, нужно вычислить разницу адресов занятых блоков. После того, как мы нашли блок, необходимо указатели у соседних блоков переставить на наш, а у нашего блока установить указатели на соседние занятые. После проверок можно вернуть данный блок пользователю.[4]

Освобождение в данном аллокаторе достаточно простое. Необходимо найти соседние занятые блоки для нашего и переставить их указатели друг на друга. На этом освобождение заканчивается.

Сложность операции выделения: O(n).

Сложность операции освобождения: O(1).

## Система двойников

В данном аллокаторе размер кучи равен кратен целой степени двойки, то есть .

При поиске участка для выделения нужно найти блок минимального подходящего размера . Если блок найден, то используем его, в ином же случае находим ближайший блок большего размера и делим его до тех пор, пока не получим необходимый участок. Два блока, образовавшихся в результате деления называются двойниками.[2]

Также если известен адрес и размер участка, то известен и адрес его двойника. Этот адрес можно получить с помощью операции «исключающее или».

В каждом блоке имеется дополнительный бит занятости и два указателя, указывающие на следующий и предыдущий свободные блок.

При освобождении блока памяти также может возникнуть два случая:

* Двойник свободен. В этом случае объединяются двойники до тех пор, пока это возможно.
* Двойник занят. В этом случае ищем соседние от освобождаемого блоки, затем переставляем указатели у всех блоков.

Сложность операции выделения: O(n).

Сложность операции освобождения: O().

# Деревья

## Двоичное дерево поиска

Двоичное дерево - это иерархическая структура данных, в которой каждый узел имеет значение и ссылки на левого и правого потомка. Узел, находящийся на самом верхнем уровне называется корнем. Узлы, не имеющие потомков, называются листьями.

Бинарное дерево поиска отличает наличие некоторых свойств: значение левого потомка меньше значения родителя, а значение правого потомка больше значения потомка родителя. Данные в бинарном дереве поиска хранятся в отсортированном виде. При каждой операции вставки нового или удаления существующего узла отсортированный порядок дерева сохраняется. При поиске элемента сравнивается искомое значение с корнем. Если искомое больше корня, то поиск продолжается в правом потомке корня, если меньше, то в левом, если равно, то значение найдено и поиск прекращается.[5]

### Операции двоичного дерева поиска

Поиск элемента.

Поиск элемента в двоичном дереве поиска достаточно прост, так как данные хранятся в отсортированном виде. Если искомое значение меньше значения узла, то идем в левое поддерево, в ином случае идем в правое поддерево, и так пока не найдем искомый узел.

Сложность операции поиска:

* В худшем случае: O(n);
* В среднем: O().

Вставка элемента.

Как и с поиском элемента, мы спускаемся по дереву, сравнивая значение вставляемого узла со значениями элементов дерева, и как только мы достигаем узла, у которого поддерево имеет значение NULL, вставляем туда наш узел.

Сложность операции вставки:

* В худшем случае: O(n);
* В среднем: O().

Удаление элемента.

При удалении элемента возникает три случая:

* Удаляемый элемент – лист. В этом случае просто удаляем лист из дерева.
* У удаляемого узла есть потомок. В этом случае меняем удаляемый узел на потомка, а затем удаляем наш узел.
* У удаляемого узла есть два потомка. Если в первом узле правого поддерева нет левого узла, то используем его, в другом случае находим самый левый элемент правого поддерева. После нахождения этого элемента заменяем удаляемый узел на найденный, затем удаляем наш узел.

Сложность операции удаления:

* В худшем случае: O(n);
* В среднем: O().

## Splay tree

В косом дереве, как и в АВЛ и красно-черном деревьях, поддерживается свойство сбалансированности. В splay-дереве обеспечивается выполнение операций поиска, вставки и удаления за логарифмическое время. Идея косого дерев в том, что вставка элемента, к которому недавно осуществлялся доступ, происходит в корень дерева, что позволяет получить этот элемент за O(1).[6]

### Операции косого дерева

#### Splay

Основная операция данного дерева, которая заключается в перемещении узла в корень с помощью выполнения следующих операций:

* Zig;
* Zig-Zig;
* Zig-Zag.

Обозначим вершину, которую надо поставить на место корня, за . Родителя обозначим за , а уже его родителя за .

Операция Zig:

Также эта операция называется малый правый поворот.

Данная операция выполняется в случае, когда является корнем. Выполняется поворот по ребру между вершинами и . Эта операция нужна только для разбора крайнего случая и выполняется один раз в конце, когда изначальная глубина нечетна.

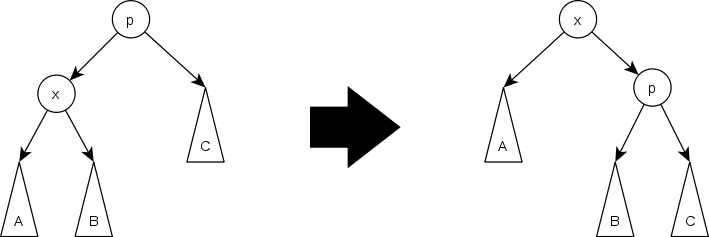


Рисунок 1. Zig (малый правый поворот)

Листинг 1. Операция Zig.

void right\_rotation(node \*&subtree\_root) const  
{  
 node \*tmp = subtree\_root;  
 subtree\_root = subtree\_root->left\_subtree\_address;  
 tmp->left\_subtree\_address = subtree\_root->right\_subtree\_address;  
 subtree\_root->right\_subtree\_address = tmp;  
}

Операция Zig-Zig:

Данная операция выполняется в случае, когда обе вершины и являются левыми сыновьями. Выполняется поворот по ребру между вершинами и , затем по ребру между вершинами и .

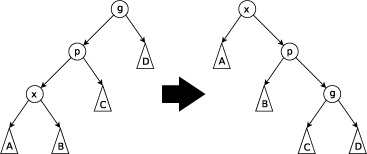


Рисунок 2. Операция Zig-Zig (большой правый поворот)

Аналогично, когда вершины и являются правыми сыновьями. В этом случае выполняется операция Zag-Zag, то есть большой левый поворот.

Операция Zig-Zag:

Данная операция выполняется в случае, когда вершина является правым потомком, а вершина – левым. Выполняется поворот по ребру между узлами и , затем по ребру между узлами и .

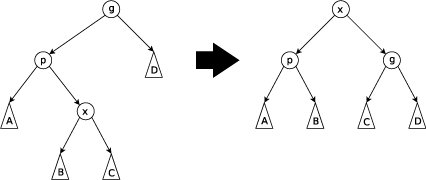


Рисунок 3. Операция Zig-Zag (малый правый поворот и малый левый поворот)

Аналогично, когда – левый потомок, а – правый. В этом случае выполняется операция Zag-Zig, то есть малый левый поворот и малый правый поворот.

Листинг 2. Операция Zag.

void left\_rotation(node \*&subtree\_root) const  
{  
 node \* tmp = subtree\_root;  
 subtree\_root = subtree\_root->right\_subtree\_address;  
 tmp->right\_subtree\_address = subtree\_root->left\_subtree\_address;  
 subtree\_root->left\_subtree\_address = tmp;  
}

#### Merge

Данная операция вызывается, когда требуется удалить узел из дерева. Эта операция нужна для слияния двух поддеревьев удаляемого узла. Сначала находим максимальный элемент поддерева, значения узлов которого больше, чем у другого поддерева, то есть у левого. Затем вызываем операцию Splay для этого узла. После этого приклеиваем правого потомка удаляемого узла к получившемуся поддереву. Таким образом мы получаем целое дерево без нарушения сбалансированности.[6]

Листинг 3. Операция Merge.

node \* splay\_tree<tkey, tvalue, tkey\_comparer>::merge(node \*&left\_subtree, node \*&right\_subtree)  
{  
 if (left\_subtree == nullptr && right\_subtree == nullptr)  
 {  
 return nullptr;  
 }  
  
 if (left\_subtree == nullptr && right\_subtree != nullptr)  
 {  
 return right\_subtree;  
 }  
 else if (left\_subtree != nullptr && right\_subtree == nullptr)  
 {  
 return left\_subtree;  
 }  
 else  
 {  
 std::stack<node \*\*> path\_to\_subtree\_root\_exclusive;  
  
 node \* successor = left\_subtree;  
 node \*\* previous\_node\_address = &left\_subtree;  
  
 while (successor->right\_subtree\_address != nullptr)  
 {  
 path\_to\_subtree\_root\_exclusive.push(previous\_node\_address);  
 previous\_node\_address = &(successor->right\_subtree\_address);  
 successor = successor->right\_subtree\_address;  
 }  
  
 splay(successor, path\_to\_subtree\_root\_exclusive);  
  
 left\_subtree->right\_subtree\_address = right\_subtree;  
  
 return left\_subtree;  
 }  
}

#### Поиск

Поиск в косом дереве выполняется так же, как и в обычном бинарном дереве поиска, однако после завершения поиска выполняется операция Splay для найденного узла.

#### Вставка

Вставка в косом дереве похожа на вставку элемента в бинарное дерево поиска, как и операция поиска. Как и в поиске, после вставки элемента выполняется операция Splay для вставленного узла, вследствие чего узел становится корневым.

#### Удаление

При выполнении данной операции сначала находим узел, который необходимо удалить, затем вызываем операцию Splay для этого узла. После этого вызываем операцию Merge для потомков удаляемого узла.

## АВЛ-дерево

АВЛ-дерево – это прежде всего двоичное дерево поиска, ключи которого удовлетворяют стандартному свойству: ключ любого узла дерева не меньше любого ключа в левом поддереве данного узла и не больше любого ключа в правом поддереве этого узла.

Особенностью АВЛ-дерева является то, что оно является сбалансированным в следующем смысле: для любого узла дерева высота его правого поддерева отличается от высоты левого поддерева не более чем на единицу. Также узел АВЛ-дерева хранит поле height – высоту поддерева с корнем в данном узле.[7]

### Операции АВЛ-дерева

#### Balance

Данная операция выполняет балансировку дерева. Она может понадобиться, когда разница высот некоторых узлов оказывается равными 2 или -2, т.е. возникает разбалансировка поддерева. Такая ситуация возникает в процессе добавления и удаления узла.[7]

Балансировка происходит посредством поворотов, рассмотренных в разделе про Splay-дерево.

В реализации АВЛ-дерева доопределяется лишь один шаг – это операция, которая выполняется после удаления и вставки. Операция поиска не требует никаких модификаций.

#### Поиск

В реализации АВЛ-дерева операция поиска никак не модифицируется, то есть она аналогична операции поиска в обычном бинарном дереве поиска.

#### Вставка

После вставки для каждого узла, который находится в стеке, содержащий путь к вставленному узлу, выполняется функция балансировки.

#### Удаление

В операции удалении, как и в операции вставки для АВЛ-дерева, вызывается операция балансировки для каждого узла, находящегося в стеке пути.

# Интерактивный диалог с пользователем

Диалог дает возможность пользователю выбирать формат ввода команд, то есть можно вводить команды из консоли и можно передать путь к файлу, из которого считываются команды. Также есть возможность сохранить состояние базы данных в файл и восстановить базу данных из файла.

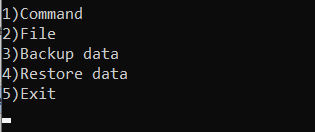


Рисунок 4. Меню выбора операции над базой данных

# Механизм сохранения состояния системы хранения данных в файловую систему и восстановление состояния системы хранения данных из файловой системы

Данный механизм реализован в виде отдельного класса, в котором реализованы методы backup (сохранение) и restore (восстановление). Также в нем реализованы вспомогательные методы, которые вызываются в методах backup и restore. Данное задание реализовано по принципу журналирования.

При вызове метода backup нынешнее состояние базы данных сохраняется в файл. При этом файл можно создать, либо сохранить в существующий. Существует служебный файл broker.txt, в который записываются все команды, которые передавал пользователь. Когда пользователь выбирает backup, команды из служебного файла записываются в файл для сохранения.

При вызове метода restore база данных восстанавливается из файла. Файл выбирается из предоставляемого списка существующих файлов. При этом если ни одного файла не существует, то выводится информация о том, что файлов нет, и операция restore не выполняется. Данный метод возвращает вектор строк, в котором лежат команды, которые отправляются на сервер для выполнения.

Также существует служебный файл existing\_recover\_files.txt, в котором хранятся пути к файлам, в которых сохранены некоторые состояния базы данных, сохраненные пользователем.

Также при старте работы программы есть опция удаления файла из системы. При этом после удаления файла стартовое меню не пропадает, и пользователь может дальше удалять файлы или восстановить базу данных.

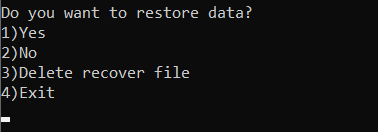


Рисунок 5. Стартовое меню программы

# Хранение объектов строк с помощью паттерна Flyweight

Flyweight (Легковес, Приспособленец) – это структурный паттерн проектирования, который позволяет вместить большее количество объектов в отведённую оперативную память. Легковес экономит память, разделяя общее состояние объектов между собой, вместо хранения одинаковых данных в каждом объекте.

Приспособленец – это совместно используемый объект, который можно задействовать одновременно в нескольких контекстах. В каждом контексте он выглядит как независимый объект, то есть неотличим от экземпляра, который не используется совместно. Приспособленцы не могут делать предположений о контексте, в котором работают. Ключевая идея здесь – различие между внутренним и внешним состояниями. Внутреннее состояние хранится в самом приспособленце и состоит из информации, не зависящей от его контекста. Именно поэтому он может использоваться совместно. Внешнее состояние зависит от контекста и изменяется вместе с ним, поэтому совместно не используется. Объекты-клиенты отвечают за передачу внешнего состояния приспособленцу, когда в этом возникает необходимость.[2]

Приспособленцами управляет класс flyweight\_factory, содержащий пул легковесов.

Листинг 4. Класс flyweight\_string

class flyweight\_string  
{  
private:  
  
 std::string \_value;  
  
public:  
  
 const std::string &get\_value() const;  
  
 void set\_value(const std::string &value);  
  
};

Чтобы избежать хранения в памяти одинаковых строк, создается приспособленец и управляющий им объект. Значению поля строки присваивается указатель на легковес, который находится в пуле. Если такой строки нет, то создается новый легковес. Таким образом, избегается хранение одинаковых строк, что оптимизирует программу по памяти.

Листинг 5. Класс flyweight\_factory

class flyweight\_factory  
{  
private:  
  
 std::unordered\_map<std::string, std::shared\_ptr<flyweight\_string>> \_flyweight\_string;  
  
public:  
  
 static flyweight\_factory &get\_instance();  
  
 std::shared\_ptr<flyweight\_string> get\_flyweight\_string(const std::string &value);  
  
};

# Клиент-серверная реализация приложения

Данное задание реализовано при помощи использования средств межпроцессного взаимодействия (IPC).

На сервере хранится база данных, которая заполняется по запросу клиента.

Клиент выбирает один из трех видов межпроцессного взаимодействия, с помощью которого в дальнейшем отправляются команды для выполнения на сервер.

При запуске сервера и клиента создается очередь сообщений Unix, которая связывает клиента с сервером. Через эту очередь передается сообщение клиента о выборе типа межпроцессного взаимодействия, после чего, в зависимости от выбора, клиент и сервер связываются либо очередью сообщений, либо разделяемой памятью, либо с помощью отображения файла в памяти. При этом все средства межпроцессного взаимодействия реализуются на платформе Unix.

Листинг 6. Создание очереди сообщений в клиенте для начальной связи

if((my\_qid = msgget(IPC\_PRIVATE, 0660)) == -1)  
{  
 perror("msgget: main, my\_qid error");  
 exit(1);  
}  
if((server\_queue\_key = ftok(SERVER\_KEY\_PATHNAME, PROJECT\_ID)) == -1)  
{  
 perror("ftok: main, server\_qid error");  
 exit(1);  
}  
if((server\_qid = msgget(server\_queue\_key, 0)) == -1)  
{  
 perror("msgget: main server\_qid error");  
 exit(1);  
}

Листинг 7. Создание очереди сообщений на сервере для начальной связи

if((msg\_queue\_key = ftok(SERVER\_KEY\_PATHNAME, PROJECTD\_ID)) == -1)  
{  
 perror("ftok error");  
 exit(1);  
}  
if((qid = msgget(msg\_queue\_key, IPC\_CREAT | 0660)) == -1)  
{  
 perror("msgget error");  
 exit(1);  
}

Функция ftok использует файл с именем SERVER\_KEY\_PATHNAME, которое должно указывать на существующий файл к которому есть доступ, и младшие 8 бит PROJECT\_ID, который должен быть отличен от нуля, для создания ключа с типом key\_t, используемого в System V IPC для работы с [msgget](https://www.opennet.ru/cgi-bin/opennet/man.cgi?topic=msgget&category=2)(), [semget](https://www.opennet.ru/cgi-bin/opennet/man.cgi?topic=semget&category=2)(), и [shmget](https://www.opennet.ru/cgi-bin/opennet/man.cgi?topic=shmget&category=2)().

Функция msgget(key\_t key, int msgflag) возвращает идентификатор очереди сообщений, связанный со значением параметра key . Она также создает новую очередь сообщений, если key равен IPC\_PRIVATE; в случае если key не равен IPC\_PRIVATE, то с параметром key не существует ни одной очереди сообщений и в поле msgflg включен флаг IPC\_CREAT.[8]

# Кастомизация средств IPC

Данное задание реализовано с помощью очереди сообщений, то есть сначала сервер и клиент связываются очередью сообщений Unix, затем клиент выбирает тип межпроцессного взаимодействия, и сообщение отправляется на сервер. На сервере уже, в зависимости от выбора клиента, создается связь с помощью того метода IPC, которое было выбрано клиентом.

## Shared memory

Разделяемая память (Shared memory) является самым быстрым средством обмена данными между процессами.

В других средствах межпроцессового взаимодействия (IPC) обмен информацией между процессами проходит через ядро, что приводит к переключению контекста между процессом и ядром, т.е. к потерям производительности.

Листинг 7. Создание разделяемой памяти и семафора в клиенте

int shm\_id = shmget(SHARED\_MEMORY\_KEY, SHARED\_MEMORY\_SIZE, IPC\_CREAT | 0666);  
if(shm\_id == -1)  
{  
 perror("shmget error");  
 exit(1);  
}  
  
auto \*shared\_data = (SharedData\*) shmat(shm\_id, nullptr, 0);  
if(shared\_data == (SharedData\*)(-1))  
{  
 perror("shmat error");  
 exit(1);  
}  
  
int sem\_id = semget(SEMAPHORE\_KEY, SEMAPHORE\_COUNT, IPC\_CREAT | 0666);  
if(sem\_id == -1)  
{  
 perror("semget error");  
 exit(1);  
}

Листинг 8. Создание разделяемой памяти и семафора на сервере

int shm\_id = shmget(SHARED\_MEMORY\_KEY, SHARED\_MEMORY\_SIZE, 0666);

if(shm\_id == -1)

{

perror("shmget error");

exit(1);

}

auto \*shared\_data = (SharedData\*) shmat(shm\_id, nullptr, 0);

if(shared\_data == (SharedData\*)(-1))

{

perror("shmat error");

exit(1);

}

int sem\_id = semget(SEMAPHORE\_KEY, SEMAPHORE\_COUNT, 0666);

if(sem\_id == -1)

{

perror("semget error");

exit(1);

}

Функция shmget(key\_t key, int size, int shmflag) возвращает идентификатор разделяемому сегменту памяти, соответствующий значению аргумента key. Создается новый разделяемый сегмент памяти с размером size( округленным до размера, кратного PAGE\_SIZE), если значение key равно IPC\_PRIVATE или если значение key не равно IPC\_PRIVATE и нет идентификатора, соответствующего key; причем, выражение shmflg&IPC\_CREAT истинно.

Функция shmat подстыковывает сегмент разделяемой памяти к адресному пространству вызывающего процесса. При ошибке функция возвращает -1, а переменной errno присваивается номер ошибки. При удачном выполнении shmat возвращает адрес подстыкованного сегмента памяти.

## File mapping

Отображение файла в память (File mapping) – это способ работы с файлами в некоторых операционных системах, при котором всему файлу или некоторой непрерывной его части ставится в соответствие определённый участок памяти.

Файл, отображаемый в памяти, содержит содержимое файла в виртуальной памяти. Такое сопоставление файла и пространства памяти позволяет приложению, включая несколько процессов, изменять файл путем чтения и записи непосредственно в память.

Листинг 9. Создание файла, отображенного в памяти, и семафора в клиенте

int fd = open(FILE\_MAPPING\_PATHNAME, O\_CREAT | O\_RDWR);  
if(fd == -1)  
{  
 perror("file error");  
 exit(1);  
}  
  
char \*addr = (char\*) mmap(NULL, file\_size, PROT\_READ | PROT\_WRITE, MAP\_SHARED, fd, 0);  
if(addr == MAP\_FAILED)  
{  
 perror("file mapping error");  
 close(fd);  
 exit(1);  
}  
close(fd);  
  
int sem\_id = semget(SEMAPHORE\_KEY, SEMAPHORE\_COUNT, IPC\_CREAT | 0666);  
if(sem\_id == -1)  
{  
 perror("semget error");  
 exit(1);  
}

Листинг 10. Создание файла, отображенного в памяти, и семафора на сервере

int fd = open(FILE\_MAPPING\_PATHNAME, O\_CREAT | O\_RDWR, S\_IRUSR | S\_IWUSR);  
if(fd == -1)  
{  
 perror("file error");  
 exit(1);  
}  
  
lseek(fd, file\_size - 1, SEEK\_SET);  
write(fd, "", 1);  
  
char \*addr = (char\*) mmap(NULL, file\_size, PROT\_READ | PROT\_WRITE, MAP\_SHARED, fd, 0);  
if(addr == MAP\_FAILED)  
{  
 perror("file mapping error");  
 close(fd);  
 exit(1);  
}  
close(fd);  
  
int sem\_id = semget(SEMAPHORE\_KEY, SEMAPHORE\_COUNT, 0666);  
if(sem\_id == -1)  
{  
 perror("semget error");  
 exit(1);  
}

Чтобы отобразить файл в память, можно воспользоваться функцией mmap(void \*addr, size\_t len, int prot, int flag, int filedes, off\_t off). Она возвращает адрес начала участка отображаемой памяти или MAP\_FAILED в случае неудачи. Первый аргумент – желаемый адрес начала участка отображенной памяти; len – количество байт, которое нужно отобразить в память;  
prot – число, определяющее степень защищённости отображенного участка памяти; flag – описывает атрибуты области; filedes – дескриптор файла, который нужно отобразить;  
off – смещение отображенного участка от начала файла.

В этом методе межпроцессного взаимодействия, как и в разделяемой памяти, используются семафоры.

Семафор  – примитив синхронизации работы [процессов](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%BE%D1%86%D0%B5%D1%81%D1%81_(%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0)) и [потоков](https://ru.wikipedia.org/wiki/%D0%9F%D0%BE%D1%82%D0%BE%D0%BA_%D0%B2%D1%8B%D0%BF%D0%BE%D0%BB%D0%BD%D0%B5%D0%BD%D0%B8%D1%8F), в основе которого лежит счётчик, над которым можно производить две [атомарные операции](https://ru.wikipedia.org/wiki/%D0%90%D1%82%D0%BE%D0%BC%D0%B0%D1%80%D0%BD%D0%B0%D1%8F_%D0%BE%D0%BF%D0%B5%D1%80%D0%B0%D1%86%D0%B8%D1%8F): увеличение и уменьшение значения на единицу, при этом операция уменьшения для нулевого значения счётчика является блокирующейся. Служит для построения более сложных механизмов синхронизации и используется для синхронизации параллельно работающих задач, для защиты передачи данных через [разделяемую память](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D0%B7%D0%B4%D0%B5%D0%BB%D1%8F%D0%B5%D0%BC%D0%B0%D1%8F_%D0%BF%D0%B0%D0%BC%D1%8F%D1%82%D1%8C), для защиты [критических секций](https://ru.wikipedia.org/wiki/%D0%9A%D1%80%D0%B8%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D0%B5%D0%BA%D1%86%D0%B8%D1%8F), а также для управления доступом к аппаратному обеспечению.

Системный вызов semget(key\_t key, int nsems, int semflg) возвращает идентификатор набора семафоров, связанный с аргументом key. Создается новый набор из семафоров nsems, если значение key равно IPC\_PRIVATE или c ключом key не связано ни одного существующего набора семафора, а выражение semflg & IPC\_CREAT истинно.

При удачном завершении возвращаемое значение будет представлять собой идентификатор набора семафоров (целое неотрицательное значение), иначе возвращается -1, а переменной errno присваивается номер ошибки.

# Руководство пользователя

# Список использованных источников

1. Д. Кнут. Искусство программирования том 1.
2. Э.Гамма, Р.Хелм, Р.Джонсон, Дж. Влиссидес. Паттерны объектно-ориентированного программирования.
3. [Электронный ресурс] <https://habr.com/ru/articles/707032/>
4. [Электронный ресурс] <https://habr.com/ru/articles/473294/>
5. [Электронный ресурс] <https://codechick.io/tutorials/dsa/dsa-binary-search-tree>
6. [Электронный ресурс] <https://habr.com/ru/companies/JetBrains-education/articles/210296/>
7. [Электронный ресурс] <https://habr.com/ru/articles/150732/>
8. [Электронный ресурс] <https://www.opennet.ru/man.shtml?topic=msgget&category=2&russian=0>
9. [Электронный ресурс] <https://www.opennet.ru/man.shtml?topic=shmget&category=2&russian=0>