

Практика программирования. Заметки.

Семён Григорьев

4 декабря 2020 г.

Содержание

1	Лекция 1: Введение	3
2	Домашняя работа 1	3
3	Лекция 2	4
3.1	Первое знакомство с F#	4
3.2	Тестирование программ	4
3.3	Отладка кода	5
4	Домашняя работа 2	5
5	Лекция 3	5
6	Домашняя работа 3	6
7	Лекция 4	6
8	Домашняя работа 4	6
9	Лекция 5	7
10	Домашняя работа 5	7
11	Лекция 6	7
12	Лекция 7	8
13	Домашняя работа 6	9
14	Лекция 8	9
15	Домашняя работа 7	9
16	Лекция 9	10
17	Домашняя работа 8	11
18	Лекция 10	11
19	Лекция 11	12
20	Домашняя работа 9	12

1 Лекция 1: Введение

Программирование — не только написание кода. Документация, сборка, тестирование, версионирование, обработка отзывов пользователей.

Инфраструктура проекта, рабочее окружение, система контроля версий, непрерывная сборка.

Соответствующие решения на примере инфраструктуры вокруг GitHub. GithubActions, внешние сервисы для CI (<https://travis-ci.org/>, <https://www.appveyor.com/>, <https://circleci.com/>).

Практика развёртывания соответствующей инфраструктуры.

1. Для начала, завести аккаунт на GitHub (<https://github.com/>). Важно, чтобы имя аккаунта (логин) было NameSurname или Name_Surname.
2. Создаём репозиторий для проекта (для всех домашних работ). Название должно отражать содержимое репозитория. Не забываем добавить описание. Лицензию, readme и gitignore лучше не добавлять.
3. Устанавливаем git (<https://git-scm.com/>) и графическую оболочку для работы с ним (если кому нужно).
4. Теперь пора приступить к созданию проекта. Так как дальше мы будем пользоваться F#, то в качестве шаблона предлагается использовать <https://github.com/TheAngryByrd/MiniScaffold>.
 - (a) Установить .NET Core: <https://dotnet.microsoft.com/download>
 - (b) Прочитать инструкции (<https://github.com/TheAngryByrd/MiniScaffold#install-the-dotnet-template-from-nuget>) и выполнить соответствующие шаги. Нам нужно создать консольное приложение. Это может занять некоторое время.
5. Устанавливаем связь только что созданного локального репозитория с удалённым репозиторием: <https://docs.github.com/en/github/importing-your-projects-to-github/adding-an-existing-project-to-github-using-the-command-line>

С этого момента домашние работы только через GitHub с налаженной сборкой.

2 Домашняя работа 1

Задачи:

1. **(1 балл)** Инициализировать рабочее окружение: репозиторий на GitHub, CI, readme, лицензия. Добавить преподавателя в совладельцы. Оформить тестовый pull request: например, оформленное readme (поправить описание проекта, удалить лишнее, что досталось от шаблона, поправить ссылку на статус сборки). Запросить ревью у преподавателя.

3 Лекция 2

3.1 Первое знакомство с F#

Общие сведения о платформе .NET, общие представления об архитектуре платформы.

F# — это один из языков платформы. Некоторые ресурсы для того, чтобы начать изучать F#.

- Главный сайт сообщества: <https://fsharp.org/>. Там много чего интересного, от библиотек и инструментов до ссылок на другие ресурсы.
- Неплохая книга по основам языка: https://en.wikibooks.org/wiki/F_Sharp_Programming
- Подборка обучающих материалов от сообщества: <https://fsharp.org/learn/>
- Официальная документация от Microsoft: <https://docs.microsoft.com/ru-ru/dotnet/fsharp/>

Структура проекта вообще и на языке F# в частности. Разбиение кода на модули, файлы, библиотеки. Переиспользование кода. Зависимости между модулями, библиотеками. Пространства имён.

Основные особенности F#, примеры кода, базовые языковые конструкции и типы.

Примитивные типы: логические, числовые, символы, строки. Массивы, основные функции работы с ними: инициализация, взятие элемента, запись элемента.

Базовые конструкции управления: ветвления (if), циклы (различные варианты for, while). Двумерный синтаксис.

Структура программы. Точка входа, модули, функции. Модуль и пространство имён.

Основы работы с консолью, библиотека Argv.

3.2 Тестирование программ

Тестирование программ: ручное, автоматизированное, автоматическое.

Доказательство корректности vs тестирование. Тестирование не есть доказательство корректности. А можно ли всё таки доказать корректность? Да (иногда).

- Формальная верификация используя внешние инструменты.
- Использование специальных языков программирования, таких как Coq (на самом деле это целая система, так называемый proof assistant), Agda, Idris, F*,

Типы тестов и особенности их применения: модульные, интеграционные, unit и т.д. Автоматизация создания тестов. Intellitest — пример инструмента для автоматической генерации тестов. Примеры инструментов для тестирования: Expecto, FsUnit, NUnit, FsCheck, Canopy.

С этого момента все домашние работы должны быть снабжены автоматически запускаемыми при сборке тестами.

3.3 Отладка кода

Отладка кода. Некоторые методы отладки: отладочная печать, логгирование, использование пошаговых отладчиков. Некоторые шаги отладки: формулировка гипотезы и её проверка, локализация ошибки, работа с тестами. Практика по использованию отладчика.

4 Домашняя работа 2

В задачах, связанных с обработкой массивов на вход необходимо принимать длину массива и затем создавать случайный массив соответствующей длины. Для всех задач необходимо реализовать чтение входных данных из консоли и вывод результата в консоль.

Задачи:

1. (1 балл) Реализовать функцию, вычисляющую значение выражения $x^4 + x^3 + x^2 + x + 1$ “наивным” способом.
2. (1 балл) Реализовать функцию, вычисляющую значение выражения $x^4 + x^3 + x^2 + x + 1$, применив минимальное число умножений и сложений.
3. (1 балл) Вычислить индексы элементов массива, не больших, чем заданное число.
4. (1 балл) Вычислить индексы элементов массива, лежащих вне диапазона, заданного двумя числами.
5. (1 балл) Дан массив длины 2. Поменять местами нулевой и первый элементы, не используя дополнительной памяти/переменных.
6. (1 балл) Поменять местами i -й и j -й элементы массива не используя дополнительной памяти/переменных.

5 Лекция 3

Ещё раз произнесения: в реквесте должно быть только то, что непосредственно относится к сдаваемой домашке.

Ещё раз про функции, про то, как выделять и разделять функциональность, не надо записывать всё в одну функцию. Про то, где должны быть проверки.

Про консоль.

Про обработку крайних случаев. Про исключения.

Про тесты и ошибки: нашёл ошибку — создал тест.

Про стиль кодирования: про пробелы вокруг скобок и операций, про отступы и переводы строк. Про соглашения о наименовании. `camelCase` `CamelCase`

Про единицы измерения.

Базовые структуры данных, алгоритмы и их выражение в F#. Функция. Рекурсия и итерация. Базовые типы и основы работы с ними: матрицы, массивы, списки, структуры.

Числа Фибоначчи.

6 Домашняя работа 3

Для всех задач обеспечить чтение n из консоли и печать результата в консоль.

1. (1 балл) Реализовать вычисление n -ого числа Фибоначчи рекурсивным методом.
2. (1 балл) Реализовать вычисление n -ого числа Фибоначчи итеративным методом.
3. (1 балл) Реализовать вычисление n -ого числа Фибоначчи используя хвостовую рекурсию (не используя `mutable` и других изменяемых структур). Подсказка: нужно использовать рекурсию с аккумулятором.
4. (2 балла) Реализовать вычисление n -ого числа Фибоначчи через перемножение матриц “наивным” методом. Функции построения единичной матрицы, умножения и возведения в степень должны быть реализованы в общем виде.
5. (2 балла) Реализовать вычисление n -ого числа Фибоначчи через перемножение матриц за логарифм.
6. (1 балл) Реализовать вычисление всех чисел Фибоначчи до n -ого включительно.

7 Лекция 4

Рассказать про то, что выделенные значения — это плохо. Немного про исключительные ситуации. Про создание проектов. Про версии пакетов и вообще версии артефактов. Про то, что заимствование кода не поощряется. Тем более неправомерное заимствование. Про классный пример форматирования $x*x*x + x*x + x + 1$.

Работа с файлами.

Сортировки: пузырьком, вставкой, Хоара. Различные сценарии использования: поддержание отсортированного набора, сортировка всего набора целиком. Некоторые особенности реализации: наивная функциональная реализация Хоара, реализация на массиве.

Основы машинного представления данных. Представления чисел. Представление чисел с плавающей точкой. Проблемы переполнения. Битовые операции. Строки, кодировки.

8 Домашняя работа 4

Во всех задачах на сортировку необходимо реализовать чтение массива из файла и печать результата в файл. Функции чтения и записи необходимо переиспользовать.

В задачах на битовые операции продолжаем работать с консолью: чтение данных с консоли и печать результата туда же.

Для данной домашней работы необходимо создать отдельный проект.

При создании тестов необходимо, в задачах на сортировку, убедиться, что, во-первых, сортировки ведут себя одинаково на одинаковых данных, во-вторых, что они ведут себя так же, как системные сортировки для соответствующих коллекций. Для задачи о запаковке и распаковке надо проверить, что реализованные функции являются взаимно обратными. `FsCheck (testProperty)` в помощь.

1. (1 балл) Реализовать сортировку пузырьком массива.
2. (1 балл) Реализовать сортировку пузырьком списка.
3. (1 балл) Реализовать быструю сортировку для списка.
4. (1 балл) Реализовать быструю сортировку для массива.
5. (1 балл) Реализовать запаковку двух 32-битных чисел в одно 64-битное и распаковку обратно.
6. (1 балл) Реализовать запаковку четырёх 16-битных чисел в одно 64-битное и распаковку обратно.

9 Лекция 5

Основы анализа алгоритмов. Модель вычислителя. Понятие элементарной операции. Асимптотика, “O”-символика.

Постановка эксперимента и оформление результатов. Эксперимент по сравнению и анализу производительности. Точность проведения замеров. “Масштабы времени”, цель эксперимента и точность измерений, инструменты измерений. Базовая статистическая обработка данных: медиана и среднее, выбросы, распределение. Проверка гипотез.

Технические средства. О системе вёрски L^AT_EX. TexLive, Overleaf. Способы визуализации результатов. Python + Matplotlib, другие библиотеки. GnuPlot.

Отдельно про скрипты сборки. Теперь мы руками попробуем сделать свой маленький скрипт. Shell, cmd (PowerShell).

10 Домашняя работа 5

В данной задаче предполагается, что для обработки сырых данных и отрисовки графиков используется Python и Matplotlib, однако другие сравнимые по выразительности и качеству результата пакеты разрешены. Использование офисных пакетов (типа LibreOffice, MS Office и т.д.) запрещено. Шрифты должны быть векторными. Внимательно следите за тем, что в репозиторий должны быть только исходники, сгенерированные файлы в репозиторий попасть не должны.

1. (5 баллов) Провести сравнительное исследование реализованных в предыдущей домашней работе сортировок и стандартных реализаций сортировок соответствующих коллекций. Оформить отчёт: введение, детали реализации, постановка эксперимента, результаты экспериментов, анализ результатов. Отчёт оформляется в L^AT_EX, исходники выкладываются так же как и обычный код, снабжаются скриптом сборки (Shell), который генерирует графики и собирает pdf-документ.

11 Лекция 6

11. Контрольная работа.

1. **[3 балла]** Что такое хвостовая рекурсия? Чем она отличается от “обычной” рекурсии? Приведите пример кода на F# для следующих функций, реализованных с использованием хвостовой рекурсии

- (a) сумма элементов списка,
- (b) минимальный элемент списка,
- (c) максимальный элемент списка,
- (d) факториал,
- (e) длина списка

2. **[3 балла]**

- (a) Расскажите про этапы жизни программного продукта (проекта). Опишите каждый этап.
- (b) Для каких целей может проводиться экспериментальное исследование реализации алгоритма. В чём отличие между теоретическим и экспериментальным исследованием?
- (c) Что такое отсортированная коллекция? Каким условиям должны удовлетворять элементы коллекции, чтобы коллекцию из этих элементов можно было отсортировать? Какие алгоритмы сортировки вы знаете? Чем они отличаются?
- (d) Перечислите компоненты инфраструктуры проекта. Какую роль они выполняют? Приведите примеры конкретных реализаций компонент.
- (e) Что такое отладка? Что такое тестирование? Чем отладка отличается от тестирования? Какие средства отладки вы знаете? Какие средства тестирования вы знаете?

3. **[4 балла]**

- (a) Сколько операций сложения и умножения чисел требуется для перемножения двух матриц? Почему?
- (b) Сколько операций сравнения потребуется для сортировки списка длины n сортировкой Хоара? Почему?
- (c) Сколько операций сравнения потребуется для сортировки списка длины n сортировкой пузырьком? Почему?
- (d) Сколько операций сложения потребуется для вычисления n -ого числа Фибоначчи наивным рекурсивным методом? Почему?
- (e) Сколько операций сложения и умножения чисел потребуется для вычисления n -ого числа Фибоначчи через умножение матриц “умным” (не наивным) способом? Почему?

12. Про анализ результатов экспериментов. Бокс-плот. Придумать что-то про ЛІТ?

12 Лекция 7

Пример рефакторинга разбора аргументов командной строки. И то же самое с тестами.

Понятие типа данных. Системы типов: статические, динамические, строгие, нестрогие. Примеры языков с разными системами типов. Понятие о разной выразительности (“мощности”) систем. Пример кода на F^* .

Приведение типов: автоматическое, ручное.

Обобщённые типы данных. Понятие о полиморфизме.

Алгебраические типы данных: кортежи, Discriminated Unions (рамеченные объединения). Примеры на $F\#$. Единицы измерения.

13 Домашняя работа 6

1. (1 балл) Предположим, что мы храним булевы матрицы в виде списка координат ячеек, значение которых `true`. Необходимо реализовать соответствующие типы: единицы измерения для строк и столбцов, пара “строка-столбец”, список пар “строка-столбец”.
2. (2 балла) Реализовать функцию, перемножающую две матрицы, заданных в формате, описанном в предыдущей задаче. Не забыть проверку корректности входных данных. Реализовать подгрузку матриц из файла и запись результата в файл. Файлы с данными и результатом указываются через консоль. Формат хранения матрицы из m строк и n столбцов: в файле m строк, каждая строка состоит из n символов 0 или 1.

14 Лекция 8

Про то, что теперь не указываем типы там, где это не нужно. А точнее, указываем только там, где это необходимо. Про думать головой над постановками задачи, а не просто кодить.

Ещё раз про полиморфизм, Ad-hoc полиморфизм и бинарные операции. Типовые параметры и ограничения на них в $F\#$ (<https://docs.microsoft.com/en-us/dotnet/fsharp/language-reference/generics/constraints>).

Списки, деревья: как формальные объекты, структуры данных и как примеры алгебраических обобщённых типов. Реализация списка и дерева. Обходы списков и деревьев.

Про `fold`, `map`, `iter`.

Длинная арифметика. Практика работы со списками. Ещё раз о проблеме переполнения. Целочисленная арифметика на списках.

15 Домашняя работа 7

1. (1 балл) Реализовать самостоятельно полиморфный непустой список (далее будем называть этот тип `MyList`). Реализовать для него функции сортировки, вычисления длины, конкатенации, `map`, `iter`. Реализовать преобразование из стандартного списка в `MyList`.
2. (1 балл) На основе `MyList` реализовать тип `MyString`, представляющий строку как список символов. Реализовать преобразование стандартной строки в `MyString` и конкатенацию строк для `MyString`.

3. (1 балл) Реализовать тип дерева с произвольным количеством потомков в каждом узле (использовать `MyList`) `MyTree`. Каждый узел должен хранить данные произвольного типа.
4. (1 балл) Пусть есть `MyTree`, хранящий в узлах целые числа. Реализовать функции, которые находят максимальный хранимый элемент, среднее значение всех хранимых элементов.

16 Лекция 9

Ещё раз про то, как условный оператор форматировать.

Граф как формальный объект и как структура данных. Понятие о бинарном отношении и его свойствах: транзитивность, рефлексивность, симметричность. (Не)Ориентированные, (не)помеченные графы. Способы представления графов: список рёбер, список смежности, матрица смежности.

GraphViz

Базовые алгоритмы на графах. Обходы в глубину и ширину, построение транзитивного замыкания, поиск кратчайшего пути.

Линейная алгебра. Основы: матрица, вектор, полукольцо, кольцо, моноид, полугруппа. Сведение некоторых задач к операциям линейной алгебры (транзитивное замыкание, кратчайшие пути, пересечение автоматов). Особенности практического использования такого подхода: разреженные структуры данных, абстрактность, композициональность.

Множество S , с заданными на нем бинарными операциями $+$ и \cdot , называется полукольцом, если для любых элементов a, b, c верно следующее:

1. $\langle S, + \rangle$ — коммутативный моноид. То есть имеют место свойства:
 - (a) Коммутативности: $a + b = b + a$
 - (b) Ассоциативности: $(a + b) + c = a + (b + c)$
 - (c) Существования нейтрального элемента (нуля): $a + 0 = 0 + a = a$
2. $\langle S, \cdot \rangle$ — полугруппа. Необходимо свойство ассоциативности: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
3. Умножение дистрибутивно относительно сложения:
 - (a) Левая дистрибутивность: $a \cdot (b + c) = a \cdot b + a \cdot c$
 - (b) Правая дистрибутивность: $(a + b) \cdot c = a \cdot c + b \cdot c$
4. Мультипликативное свойство нуля: $a \cdot 0 = 0 \cdot a = 0$

Моноид — это полугруппа с нейтральным элементом. Кольцо, в отличие от полукольца, по сложению образует коммутативную группу (содержит обратные по сложению).

В отечественной культуре в полукольце нет нейтрального по умножению, зато есть полукольцо с единицей — полукольцо с нейтральным по умножению. Фактически, умножение начинает задавать моноид. Однако часто определение полукольца включает требование наличие нейтрального по умножению.

Полукольцо называют коммутативным, если операция умножения в нём коммутативна.

Полукольцо называют идемпотентным, если для любого $s \in S, s + s = s$

Дерево квадрантов.

Примеры матриц смежности и тензорное произведение можно посмотреть в этом документе. Соответственно, разделы 2.1 “Основные определения” и 7.2 “Тензорное произведение”.

17 Домашняя работа 8

При создании новых структур данных необходимо расширять библиотеку, созданную в предыдущей домашней работе.

1. **(5 балла)** Используя тип `MyList` из предыдущей домашней работы, реализовать целочисленную длинную арифметику: операции сложения, умножения, вычитания, целочисленного деления.
2. **(2 балл)** Реализовать представление разреженных матриц в виде дерева квадрантов. Реализовать функцию поэлементного сложения двух матриц в таком формате.
3. **(3 балл)** Реализовать функцию умножения двух матриц в формате дерева квадрантов.
4. **(3 балл)** Реализовать функцию тензорного умножения двух матриц в формате дерева квадрантов.
5. **(5 балла)** Реализовать построение транзитивного замыкания ориентированного графа через произведение матриц. Использовать представление матриц из первой задачи. Визуализировать результат с помощью `GraphViz`: исходный граф и выделенные рёбра, появившиеся в результате транзитивного замыкания. Для задания графа использовать формат из задачи 2 6-й домашней работы.
6. **(5 балла)** Реализовать вычисление кратчайших путей между всеми парами вершин в ориентированном графе. Использовать представление матриц из первой задачи. Визуализировать результат с помощью `GraphViz`: исходный граф и выделенные рёбра со значением кратчайшего пути между соответствующей парой вершин. Для задания графа использовать формат, аналогичный формату из задачи 2 6-й домашней работы.

18 Лекция 10

Рассказать про иерархию типов и про коллекции (`Seq`)

Регулярные выражения и конечные автоматы. Определения. Построение автомата по регулярному выражению. Применения регулярных выражений (поиск, анализ текста, моделирование систем, анализ программного кода).

Устройство языков программирования: лексика, синтаксис, семантика. Определения, примеры. Шаги обработки кода: лексический и синтаксический анализы, “семантический” анализ.

Лексический и синтаксический анализ. Введение в формальные языки как способ описания синтаксиса. Контекстно-свободные грамматики как система переписываний. Способы реализации лексических и синтаксических анализаторов. `ANTLR`, `fslex+fsyacc`, `FParsec`.

19 Лекция 11

Продолжение про FsLex+FsYacc, полноценный пример интерпретатора.

Интерпретация и компиляция. Особенности, разновидности интерпретаторов, основные шаги. Особенности, разновидности компиляторов, основные шаги. Примеры, пример реализации простого интерпретатора.

Устройство сред разработки и компиляторов, интерпретаторов: общие шаги, классические возможности, JIT/AOT. Примеры из .NET, F#, JVM, LLVM.

$F\# \rightarrow IL \rightarrow C\# \rightarrow ASM$

module M

```
type Lst<'a> =  
    | Nil  
    | Cons of 'a * Lst<'a>
```

```
let rec map f l =  
    match l with  
    | Nil -> Nil  
    | Cons (x,t1) -> Cons(f x, map f t1)
```

```
Cons(1,Nil)  
|> map ((+)1)  
|> printfn "%A"
```

Этот код на sharplab.io

module M

```
let f x =  
    let h y = if y / 2 = 0 then y * x else y + x  
    if x > 5 then x + 2 else h (x * 6)
```

```
let rec g y = f (y + g 2)
```

Этот код на sharplab.io

20 Домашняя работа 9

В задачах ниже необходимо максимально переиспользовать результаты предыдущих домашних работ. Для реализации синтаксического анализатора можно использовать не только FsYacc, рассматриваемый на паре, но и другие инструменты (например FParsec). Нельзя писать анализатор руками.

1. Разработать библиотеку конечных автоматов, использующую разреженные матрицы из предыдущей работы для представления переходов автомата, и предоставляющую следующие возможности.

- Построение автомата по регулярному выражению. Можно ограничиться заданием регулярного выражения через конструкторы типа.
 - Построение пересечения двух автоматов.
 - Возможность проверять, принимается ли строка автоматом.
2. Релизовать синтаксический анализатор регулярных выражений, позволяющий в предыдущей задаче задавать регулярное выражение как строку. поддерживаемые операции регулярных выражений: конкатенация, альтернатива, звезда Клини. Алфавит регулярных выражений — строчные и прописные латинские символы, цифры, арифметические знаки, знаки припенания. Обеспечить построение автомата по регулярному выражению. Предусмотреть возможность задавать регулярное выражение с консоли, а на выход получать представление автомата в DOT.
 3. Реализовать синтаксический анализатор для арифметических выражений над целыми числами. Поддерживаемые операции: сложение, умножение, вычитание, деление. Также бывают группирующие скобки. Числа могут быть очень большими (но целыми). Реализовать вычисление значения выражения на основе операций длинной арифметики. Предусмотреть возможность задавать выражение с консоли, а на выход получать результат его вычисления (в консоль) и дерево разбора в формате DOT.
 4. Расширить язык регулярных выражения следующими конструкциями.
 - Операцией пересечения, повторения один или более раз, повторения 0 или 1 раз. Все эти операции могут встречаться в произвольном месте выражения.
 - Функцией проверки, что строка принадлежит языку, задаваемому выражением.
 - Функцией поиска всех подстрок, удовлетворяющих заданному регулярному выражению.
 - Функцией печати автомата, задаваемого выражением, в файл в формате DOT.
 - Функцией печати результата в консоль.
 - Переменными. Переменные могут использоваться в правых частях всех выражений.

Реализовать интерпретатор получившегося языка. Предусмотреть возможность его консольного запуска: на входе файл с кодом на нашем языке, на выходе — результат интерпретации.

5. Расширить язык арифметики следующими конструкциями.
 - Операцией возведения в степень, взятия остатка (от целочисленного деления), модуля, добавить унарный минус.
 - Функцией перевода числа в двоичную систему исчисления.
 - Функцией печати результата в консоль.
 - Переменными. Переменные могут использоваться в правых частях всех выражений.

Реализовать интерпретатор получившегося языка. Предусмотреть возможность его консольного запуска: на входе файл с кодом на нашем языке, на выходе — результат интерпретации.