

Задачи, предлагаемые для летней стажировки в Яндексе

Содержание

Стажировка в поисковом направлении в группе работы с пользовательскими данными (C++, Москва).....	1
Стажировка в отделе разработки геоинформационных сервисов (C++, Москва, Санкт-Петербург)...	1
Стажировка в отделе разработки геоинформационных сервисов, служба разработки интерфейсов (JavaScript, Москва)	2
Стажировка в службе разработке Маркета (Java, Москва).....	3
Стажировка в отделе компьютерного зрения (C++, Москва)	4
Стажировка в команде разработки браузера (C++, Новосибирск)	4

Стажировка в поисковом направлении в группе работы с пользовательскими данными (C++, Москва)

Вам предстоит заниматься разработкой/поддержкой парсеров пользовательских данных.

Тестовое задание:

Необходимо разработать консольную программу, решающую следующую задачу.

На вход программе даем путь к файлу (книга в текстовом формате). После запуска программа принимает от пользователя запросы. Каждый запрос - несколько слов. Окончание ввода запроса - Enter. По запросу программа выводит на экран все предложения, в которые входят введенные пользователем слова. Без учета регистра, без учета порядка слов. Лемматизация/стемминг - не требуется, но приветствуется.

Язык программирования:

C++03

Кроме STL, можно использовать boost.

Решение должно включать:

- * файлы с кодом решения (*.h, *.cpp)
- * [не обязательно] make-файл для сборки утилиты

Стажировка в отделе разработки геоинформационных сервисов (C++, Москва, Санкт-Петербург)

Вам придётся решать нетривиальные алгоритмические задачи, связанные с обработкой большого количества пространственных данных, параллельными вычислениями, алгоритмами на графах, статистическим анализом данных, вычислительной геометрией; познакомиться со всеми этапами

создания промышленного C++ кода и научиться писать хороший код; участвовать в разработке серверных компонент для Яндекс.Пробок, автомобильной маршрутизации, панорам улиц, геопоиска, Народной.Карты и других геоинформационных сервисов; научиться проектировать высоконагруженные отказоустойчивые системы.

Тестовое задание:

Нужно написать консольную программу вычисления угла между стрелками часов (на C++). Программа принимает на вход момент времени, формат, в котором нужно вывести результат, и тип часов. Выводит посчитанный угол на стандартный выход. В `argv[1]` передаётся время, которое может быть указано в 12-часовом (HH:MM AM/PM, где HH может принимать значения от 0 до 11) или 624-часовом (HH:MM) формате. В `argv[2]` передаётся один из трёх форматов вывода угла: радианы (*rad*), градусы (*deg*) или градусы, минуты и секунды (*dms*). В `argv[3]` передается тип часов: механические, у которых положение часовой стрелки зависит от положения минутной и кварцевые, у которых положение часовой стрелки не зависит от положения минутной.

Как будет оцениваться результат: (a) Программа компилируется; (b) Правильно завершается, даже при неправильных данных и недостаточных ресурсах, а не просто падает в случайных местах; (c) Корректно считает угол и выводит его в требуемом формате; (d) Код можно довольно быстро прочитать и понять и легко поддерживать (пишите код так, как Вы писали бы его у нас на работе); (e) Программа должна демонстрировать Ваши представления об объектноориентированном дизайне; (f) Желательно наличие юнит-тестов.

Решение должно представлять собой один `cpp`-файл, который компилируется `g++`

4.x и(или) `ms visual studio c++ compiler` версии ≥ 7.1 . Желательно в реше-

нии не использовать `copy` и `paste` из библиотек, `STL` и `boost` использовать можно.

Примеры входных и выходных данных:

вход: "03:00"deg; выход: 90

вход: "15:00"rad; выход: 1.5708

вход: "09:00 PM"dms; выход: 90.00'00"

Стажировка в отделе разработки геоинформационных сервисов, служба разработки интерфейсов (JavaScript, Москва)

Тестовое задание.

Вам дана стопка посадочных карточек на различные виды транспорта, которые доставят вас из точки А в точку В. Карточки перепутаны, и вы не знаете, где начинается и где заканчивается ваше путешествие. Каждая карточка содержит информацию о том, откуда и куда вы едете на данном отрезке маршрута, а также о типе транспорта (номер рейса, номер места и прочее).

Предоставьте JavaScript API, который отсортирует такой список карточек и вернет словесное описание, как проделать ваше путешествие. API должен принимать

на вход несортированный список карточек в формате придуманном вами и возвращать, например, такое описание:

7(a) Take train 78A from Madrid to Barcelona. Seat 45B.

(b) Take the airport bus from Barcelona to Gerona Airport. No seat assignment.

(c) From Gerona Airport, take flight SK455 to Stockholm. Gate 45B. Seat 3A.

Baggage drop at ticket counter 344.

(d) From Stockholm, take flight SK22 to New York JFK. Gate 22. Seat 7B. Baggage will be automatically transferred from your last leg.

Требования:

(a) Алгоритм должен работать с любым количеством карточек, если все карточки образуют одну неразрывную цепочку.

(b) Время прибытия и отправления неизвестно и не важно. Подразумевается, что средство передвижения для следующего отрезка дожидается вас.

(c) Структура кода должна быть расширяема для использования любых типов транспорта и информации, которая может быть связана с каждым типом транспорта.

(d) API будет вызываться из других частей JavaScript-кода без необходимости дополнительных запросов между браузером и сервером.

(e) Не используйте библиотеки и фреймворки, напишите все с нуля.

(f) ЗадOCUMENTИРУЙТЕ в коде формат входных и выходных данных.

Что нас интересует:

(a) Какой формат входных данных вы придумаете.

(b) Как вы структурируете свой код, чтобы он был расширяем, легок к пониманию и изменениям другими программистами.

(c) Какой алгоритм сортировки вы придумаете.

Стажировка в службе разработке Маркета (Java, Москва)

Наше направление занимается доработкой и оптимизацией внутренних интерфейсов Маркета. Мы разрабатываем сводные отчеты и мониторинги по различным внутренним процессам Маркета.

Требуемые скиллы: Java Core, SQL, базовые знания HTTP.

Тестовое задание

<http://yadi.sk/d/CDLAEWzcLiJpJ>

Стажировка в отделе компьютерного зрения (C++, Москва)

Мы занимаемся исследовательскими задачами в области компьютерного зрения, разработка и реализация на C++ алгоритмов анализа и классификации визуальных данных, распознавания образов.

Тестовое задание.

Реализовать метод, описанный в статье.

Прислать нам код и сообщить достигнутые характеристики.

Статья:

<http://www.cs.cmu.edu/~har/iaai2005.pdf>

"Boosting Sex Identification Performance", Baluja, Rowley

Данные для обучения: <http://yadi.sk/d/dZUw3bNNLiKVE>

Там три варианта векторов - для размеров патча 16x16, 20x20 и 24x24.

В каждой папке есть превью паттернов в виде bmp, и бинарный файл с векторами (samplesxx.vec).

Формат файлов-векторов:

1 байт - ширина патча (w)

1 байт - высота патча (h)

Далее - w*h байт картинки (первой идёт нижняя строка изображения),

1 байт - признак пола (1=male, 255=female)

Все паттерны в векторах уже нормализованы по уровню среднего (=128) и дисперсии.

Стажировка в команде разработки браузера (C++, Новосибирск)

Команда браузера в Новосибирске занимается разработкой внутренних кроссплатформенных компонент браузера. Среди них - модуль для воспроизведения html5 медиа-контента, постраничный переводчик текстов, сейфбраузинг, защищающий пользователя при просмотре страниц с опасным для компьютера содержанием, подсистема синхронизации данных между браузерами на разных устройствах.

Тестовое задание

Спроектировать и реализовать мини-библиотеку "пул потоков" для использования другими программистами компании. Пул должен принимать задания, выполнять эти задания в рабочих потоках и возвращать результаты выполнения заданий. Нужно иметь возможность ограничить максимальное число рабочих потоков, а также ставить задачи в очередь, если все потоки пула заняты выполнением ранее добавленных задач.

При выполнении задания необходимо придерживаться style guide: <http://google-styleguide.googlecode.com/svn/trunk/cppguide.xml>, который используется в проекте Chromium.