

Алгоритмы и структуры данных–2
2024–2025 учебный год

SET 8. Домашняя работа

Строки–1. Поиск вхождений шаблонов.
Редакционные расстояния

апрель						
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				
пн	вт	ср	чт	пт	сб	вс

Немного инструкций

Домашняя работа SET 8 содержит *Блок Р* «Задания на разработку» — задачи, связанные с реализацией и применением алгоритмов поиска вхождений строки-шаблона в тексте, а также с вычислением редакционного расстояния между строками.

Решения заданий Блока Р загружаются в систему CODEFORCES и проходят автоматизированное тестирование. Для загрузки нужно перейти на <https://dsahse.contest.codeforces.com> и выбрать соответствующее соревнование. Доступ к соревнованию предоставлен по тем же учетным данным, что и к системе Яндекс.Контест.

Домашняя работа SET 8 содержит 7 обязательных задач. Баллы, которые можно получить за их решение, распределены следующим образом:

Блок Р							
P1	P2	P3	P4	P5	P6	P7	P8b
6	6	7	7	9	8	11	10

Задачи, помеченные “b” не являются обязательными — баллы за их решение относятся к *бонусным*. Подтверждение решений бонусных задач сопровождается обязательной *устной защитой*.

Важные даты

1. Домашняя работа SET 8 открыта с **14:30 14 апреля 2025 г.**
2. Прием решений завершается в **02:00 28 апреля 2025 г.**
3. Защита решения бонусной задачи P8b принимается до **16 мая 2025 г.**

Содержание

Задача P1.	Секрет посылки из Простоквашино	1
Задача P2.	Котенок Гав и бесконечный моток ниток	2
Задача P3.	Винни Пух и уточненные грани	3
Задача P4.	Ожерелья и бусы	5
Задача P5.	Расстояние Левенштейна	7
Задача P6.	Чебурашка и загадка зеркальных слов	9
Задача P7.	Ну, Погоди! Ищи слова с Ахо-Корасик	10
Задача P8b.	Доктор Ливси и лекарства	12

Успехов!

Задача Р1. Секрет посылки из Простоквашино

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1 секунда
Ограничение по памяти: 32 мегабайта

Кот Матроскин и Шарик получили от почтальона Печкина таинственную посылку с надписью:

*Чтобы открыть коробку, найдите волшебные числа-ключи для строки-шифра "RYBARYBARY".
Каждое число покажет, сколько раз узоры на замке повторяются в начале и конце!*

Мелким шрифтом Печкин приписал: «Это как в вашем любимом сыре — чем длиннее совпадение краёв, тем вкуснее результат!»



Для этой посылки Матроскин и Шарик смогли сами вычислить числа-ключи, однако Печкин стал шифровать замки от всех посылок таким способом, поэтому нашим героям потребовалась ваша помощь. Они догадались, что им необходимо вычислить значения префикс-функции. Помогите им и разработайте универсальный алгоритм, который вычисляет значения префикс-функции для любого шифра посылки.

Формат входных данных

Одна строка s ($1 \leq |s| \leq 10^6$), состоящая из букв латинского алфавита. Будем считать, что элементы строки нумеруются от 0 до $|s| - 1$.

Формат выходных данных

Выведите $|s|$ чисел — значений префикс-функции, разделённых пробелами.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$ s \leq 100$	0	первая ошибка
2	1	$ s \leq 1000$	0–1	первая ошибка
3	2	$ s \leq 10^5$	0–2	первая ошибка
4	2	$ s \leq 10^6$	0–3	первая ошибка

Примеры

стандартный ввод	стандартный вывод
RYBARYBARY	0 0 0 0 1 2 3 4 5 6
abracadabra	0 0 0 1 0 1 0 1 2 3 4

Задача Р2. Котенок Гав и бесконечный моток ниток

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1 секунда
Ограничение по памяти: 32 мегабайта

Котёнок Гав и щенок Шарик нашли в углу двора загадочный моток ниток с узором из постоянно повторяющейся строки s . Шарик настойчиво крутит клубок, но узор повторяется снова и снова! Старый Уличный Фонарь подмигнул и сказал:

— Чтобы узнать, какой длины был исходный кусочек нитки, найдите минимальную длину узора, который, повторяясь, создаёт этот бесконечный моток!



Помогите им разобраться в устройстве бесконечного мотка. Разработайте алгоритм, который по начальному отрезку нитки определяет минимальную длину s .

Формат входных данных

Строка p ($1 \leq |p| \leq 10^6$), состоящая из букв латинского алфавита.

Формат выходных данных

Выведите одно число — ответ на основной вопрос задачи.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$ p \leq 100$	0	первая ошибка
2	1	$ p \leq 10^4$	0–1	первая ошибка
3	2	$ p \leq 10^5$	0–2	первая ошибка
4	2	$ p \leq 10^6$	0–3	первая ошибка

Примеры

стандартный ввод	стандартный вывод
xxx	1
bcabcab	3

Задача Р3. Винни Пух и уточненные грани

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1 секунда
Ограничение по памяти: 128 мегабайт



У ослика Иа день рождения, поэтому Винни Пух и Пятачок собираются подарить ему подарок. Иа записал все желаемые подарки в строку T , а особенно желаемые подарки продублировал несколько раз. Винни Пух и Пятачок просят Вас, используя алгоритм Кнута-Морриса-Пратта с помощью уточненных граней строки, определить, сколько раз подарок P был записан в строку T .

Формат входных данных

Первая строка содержит P ($0 \leq |P| \leq 1000$) — строку-шаблон, вхождения которой ищем.
Вторая строка содержит T ($0 \leq |T| \leq 10^7$) — текст, в котором ищутся вхождения строки.
Гарантируется, что строки состоят только из букв латинского алфавита.

Формат выходных данных

На первой строке выходного файла выведите количество вхождений строки в текст. Гарантируется, что количество вхождений строки в текст не превышает 10^6 .

На каждой следующей строке выходного файла номера позиций (индексирование ведётся с 0), с которых начинаются очередные вхождения строки. За последним номером также следует перенос строки.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$ P, T \leq 100$	0	первая ошибка
2	1	$ P, T \leq 10^3$	0–1	первая ошибка
3	4	$ P \leq 10^3, T \leq 10^6$	0–2	первая ошибка
4	2	$ P \leq 10^3, T \leq 10^7$	0–3	первая ошибка

Примеры

стандартный ввод	стандартный вывод
tree	2
chesstreetrainchristmastreealgorithm	5
	23
honey	4
hohoneyhoneyhoneyhoneymorehohoho	2
	7
	12
	17

Задача Р4. Ожерелья и бусы

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1 секунда
Ограничение по памяти: 128 мегабайт

Представьте: у вас есть длинная-длинная веревка бус (текст T), где каждая бусина – отдельный символ. Мартышка и Попугай дали Вам срочное поручение: определить, сколько раз ожерелье (текст P) встречается в длинной веревке бус. Так как ожерелье может быть очень длинным, то персонажи предлагают Вам использовать алгоритм Бойера-Мура-Хорспула.



Формат входных данных

Первая строка содержит P ($0 \leq |P| \leq 1000$) — строка-шаблон, вхождения которой ищем.
Вторая строка содержит T ($0 \leq |T| \leq 10^7$) — текст, в котором ищутся вхождения строки.
Гарантируется, что строки состоят только из строчных букв латинского алфавита.

Формат выходных данных

На первой строке выходного файла выведите количество вхождений строки в текст.
Гарантируется, что количество вхождений строки в текст не превышает 10^6 .

На каждой следующей строке выходного файла номера позиций (индексирование ведётся с 0), с которых начинаются очередные вхождения строки. За последним номером также следует перенос строки.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	–	тесты из условия	–	полная
1	1	$ P, T \leq 100$	0	первая ошибка
2	1	$ P, T \leq 10^3$	0–1	первая ошибка
3	4	$ P \leq 10^3, T \leq 10^6$	0–2	первая ошибка
4	2	$ P \leq 10^3, T \leq 10^7$	0–3	первая ошибка

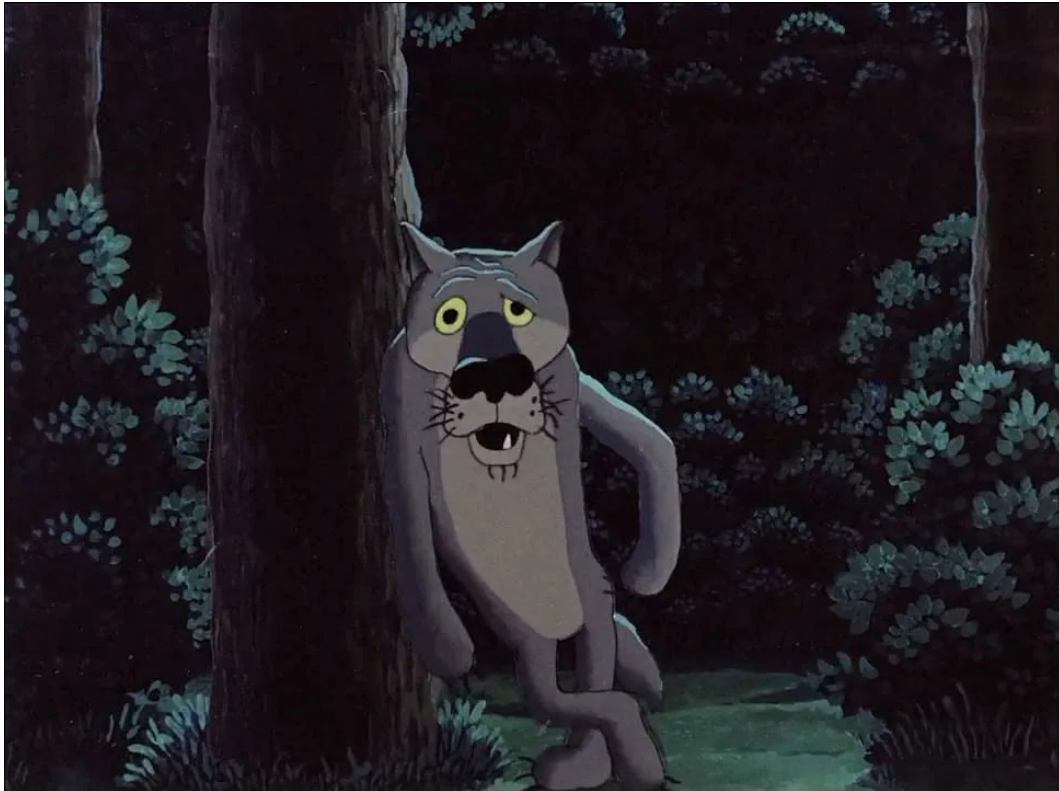
Примеры

стандартный ввод	стандартный вывод
bc	3
dcbbabcaababccscbba	5
	11
	15
dad	0
dcbdbdcabbcdcdcacaa	

Задача Р5. Расстояние Левенштейна

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1.25 секунд
Ограничение по памяти: 64 мегабайта

"Шо, опять?! – именно так сказал волк, когда его в очередной раз попросили оценить, насколько одна строка отлична от другой. Волк не прогуливал курс по алгоритмам и структурам данных и поэтому знает, что такое расстояние Левенштейна, но все же попросил Вас помочь ему в этом деле.



Формат входных данных

В первой строке задается число n ($1 \leq n \leq 10^6$) — количество пар строк, которые надо сопоставить и вычислить расстояния Левенштейна.

Затем идут $2 \cdot n$ строк s_i ($0 \leq |s_i| \leq 4000$), которые и надо сравнить попарно. Смотрите примеры.

Формат выходных данных

Выведите одну строку, в которой n чисел, разделенных пробелами. Каждое число — это расстояние Левенштейна для пары строк.

Выведенные расстояния должны соответствовать порядку пар строк для сравнения во входных данных (1-е число — расстояние для 1-й пары, 2-е число — расстояние для 2-й пары и т. д.).

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$n \leq 100, s_i \leq 10$	0	первая ошибка
2	1	$n \leq 1000, s_i \leq 50$	0–1	первая ошибка
3	1	$n \leq 10^4, s_i \leq 10$	0–2	первая ошибка
4	2	$n \leq 10^6, s_i \leq 14$	0–3	первая ошибка
5	2	$n \leq 20, s_i \leq 1000$	0–4	первая ошибка
6	2	$n = 1, s_i \leq 4000$	0–5	первая ошибка

Примеры

стандартный ввод	стандартный вывод
2 shoopyat agaopyat sunday sudnay	3 2
3 sunday saturday cat cats cats cat	3 1 1

Задача Р6. Чебурашка и загадка зеркальных слов

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 2 секунды
Ограничение по памяти: 64 мегабайта

Чебурашка и Гена нашли в старом чемодане Шапокляк, который она все время оставляет где ни попадя, таинственную записку:

Ха-ха! Чтобы узнать, где я спрятала все апельсины, найдите число волшебных слов-перевертышей в строке "SHALASH"! Считайте все подстроки, которые читаются одинаково слева направо и справа налево!



Помогите Чебурашке и Гене вернуть апельсины — разработайте универсальный алгоритм для подсчета таких перевертышей в любой строке. Вдруг Шапокляк зашифровала еще что-нибудь важное в других записках?

Формат входных данных

Вводится одна строка s , состоящая из латинских букв. Длина строки не превышает 100000 символов.

Формат выходных данных

Выведите одно число — количество подстрок данной строки, являющихся палиндромами.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$s \leq 100$	0	первая ошибка
2	1	$s \leq 1000$	0–1	первая ошибка
3	2	$s \leq 10^4$	0–2	первая ошибка
4	4	—	0–3	первая ошибка

Примеры

стандартный ввод	стандартный вывод
SHALASH	8
aaa	6
aba	4

Задача Р7. Ну, погоди! Ищи слова с Ахо-Корасик

Имя входного файла:	стандартный ввод
Имя выходного файла:	стандартный вывод
Ограничение по времени:	2 секунды
Ограничение по памяти:	512 мегабайт

Волк устроил настоящий хаос в городе! Он перепутал все таблички с названиями улиц, и теперь Заяц не может найти дорогу к Большому Осеннему Празднику, где его ждут с главным угощением — гигантским салатом из ингредиентов, заданных в словаре $D = \{P_1, \dots, P_n\}$. Заяц должен найти эти слова в испорченных Волком надписях. Однако времени мало — праздник совсем скоро начнется!



Волк издевается:

— Ха! Без меня тут всё развалится! Ты будешь прыгать по буквам до следующей осени!

Заяц не сдаётся:

— Ну, погоди! Я знаю алгоритм *Ахо-Корасик* — это мой секретный прыжок через заборы! Сначала построю **Огородный лабиринт** (префиксное дерево) из букв ингредиентов, потом добавлю **Волчьи тропки** (суффиксные и выходные ссылки), чтобы не бегать по кругу. А потом — прыг-скок! — найду все слова за один забег!

Помогите Заяцу и реализуйте алгоритм Ахо-Корасик для заданного словаря ингредиентов D .

Формат входных данных

В первой строке задается строка, в которой необходимо осуществлять поиск. В следующей строке содержится число N ($1 \leq N \leq 1\,000\,000$) — количество образцов. В каждой из N следующих строк содержится по одному образцу. Суммарная длина образцов не превосходит $1\,000\,000$. Строка и образцы состоят из маленьких латинских букв.

Формат выходных данных

Выведите N строк. В i -ой строке выведите количество вхождений i -го образца, а затем — индексы начала вхождения этого образца в строку в порядке возрастания. Нумерация индексов с единицы. Суммарное количество вхождений образцов гарантированно не превосходит $1\,000\,000$.

Система оценки

$|D|$ обозначает суммарную длину образцов.

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	3	$ D \leq 1000$	0	первая ошибка
2	3	$ D \leq 5 \cdot 10^4$	0–1	первая ошибка
3	5	—	0–2	первая ошибка

Пример

стандартный ввод	стандартный вывод
abrachkacadabrachka	2 1 12
4	1 9
abrachka	2 1 12
cadabrachka	0
ab	
marazmik	

Задача Р8b. Доктор Ливси и лекарства

Имя входного файла: стандартный ввод
Имя выходного файла: стандартный вывод
Ограничение по времени: 1 секунда
Ограничение по памяти: 256 мегабайт



Как известно, для некоторых слова «ром» и «смерть» означают одно и то же.

Доктор Ливси понял, что на кону жизнь Билли Бонса. Он пытается подобрать лекарства, которые уберут аритмию и спасут печень, селезёнку и все остальные органы пирата. Сейчас доктор проверяет очередной набор лекарств на совместимость.

Каждое лекарство представляет из себя последовательность нулей и единиц. Лекарства совместимы, если существует бесконечная последовательность из нулей и единиц (обозначающая хрупкое здоровье старого пирата), в которую не входит код ни одного из лекарств.

Помогите Доктору и определите для заданных кодов лекарств, есть ли такая последовательность.

Формат входных данных

Первая строка содержит одно целое число n ($1 \leq n \leq 10^6$), равное количеству кодов лекарств. Каждая из следующих n строк содержит слово s_i ($1 \leq |s_i| \leq 10^6$), составленное из символов 0 и 1 — код лекарства. Суммарная длина всех слов не превосходит 10^6 .

Формат выходных данных

Первая и единственная строка выходного файла должна содержать слово:

- **ТАК** — если набор лекарств совместим (т.е., существует требуемая бесконечная последовательность из нулей и единиц);
- **НЕТ** — в противном случае.

Система оценки

Подзадача	Баллы	Дополнительные ограничения	Необходимые подзадачи	Информация о проверке
0	—	тесты из условия	—	полная
1	1	$n \leq 10^3$, $ s_i \leq 100$	0	первая ошибка
2	2	$n \leq 10^6$, $ s_i \leq 10^3$	0–1	первая ошибка
3	3	$n \leq 10^6$, $ s_i \leq 10^4$	0–2	первая ошибка
4	4	$n \leq 10^6$, $ s_i \leq 10^6$	0–3	первая ошибка

Примеры

стандартный ввод	стандартный вывод
3 01 11 00000	NIE
3 011 11 0000	ТАК