

На входе есть большой текстовый файл, где каждая строка имеет вид Number. String

Например:

415. Apple

30432. Something something something

1. Apple

32. Cherry is the best

2. Banana is yellow

Обе части могут в пределах файла повторяться. Необходимо получить на выходе другой файл, где все строки отсортированы. Критерий сортировки: сначала сравнивается часть String, если она совпадает, тогда Number.

Т.е. в примере выше должно получиться

1. Apple

415. Apple

2. Banana is yellow

32. Cherry is the best

30432. Something something something

Требуется написать две программы:

1. Утилита для создания тестового файла заданного размера. Результатом работы должен быть текстовый файл описанного выше вида. Должно быть какое-то количество строк с одинаковой частью String.

2. Собственно сортировщик. Важный момент, файл может быть очень большой. Для тестирования будет использоваться размер ~100Gb.

При оценке выполненного задания мы будем в первую очередь смотреть на результат (корректность генерации/сортировки и время работы), во вторую на то, как кандидат пишет код. Язык программирования на усмотрение кандидата любой из C#/C++/Delphi. С нашей точки зрения, выбранный язык ни на что не влияет.