

Практика 4. Синхронизация доступа к динамическим спискам

Задача: реализовать многопоточную обработку текстовых файлов с целью вычисления статистики (подсчёт встречаемости слов, символов и др. в соответствии с вариантом, см. ниже). Для экспериментальной оценки производительности алгоритмов использовать количество файлов не менее 20. Требуется реализовать последовательный, три параллельных алгоритма и метод проверки идентичности результатов, полученных разными алгоритмами.

Параллельные алгоритмы:

1. Декомпозиция по файлам с локальными буферами. Все потоки участвуют в чтении файлов. Каждый поток читает определенное подмножество файлов и вычисляет *локальные результаты* по группе файлов. После завершения работы вычисляется глобальная статистика.
2. Декомпозиция по файлам с глобальным буфером. Все потоки участвуют в чтении файлов и работают с *глобальным буфером*. В качестве глобального буфера использовать или конкурентные коллекции, или неконкурентные коллекции с синхронизацией (в соответствии с вариантом).
3. Декомпозиция по задачам. Различаются *потоки-читатели*, выполняющие чтение файлов без какой-либо обработки, и *потоки-статистики*, выполняющие обработку текста и подсчёт статистики. В первой группе использовать 1-2 потока, которые наполняют общий буфер (тип буфера определяется вариантом); другие потоки обрабатывают общий буфер с целью подсчета статистики.

Вариант: $V = (c1 + c2) \% 8$, где $c1$ – первая буква имени в верхнем регистре на английском языке, $c2$ – первая буква фамилии в верхнем регистре на английском языке

Вариант	Вычислительная задача	Глобальный буфер в алгоритме №2	Буфер для взаимодействия читателей и статистиков
0	1	Конкурентный	BlockingCollection
1	2	Неконкурентный	ConcurrentQueue
2	3	Конкурентный	ConcurrentStack
3	4	Неконкурентный	ConcurrentBag
4	4	Неконкурентный	BlockingCollection
5	3	Конкурентный	ConcurrentQueue
6	2	Неконкурентный	ConcurrentStack
7	1	Конкурентный	ConcurrentBag

Вычислительные задачи:

1 – подсчёт частот встречаемости слов (результат: словарь «слово – частота»);

2 – подсчёт частот встречаемости символов (результат: словарь «символ – частота»);

3 – подсчёт частот типов предложений: вопросительные, восклицательные, утвердительные (результат: словарь «тип предложения – частота»);

4 – подсчёт частот длин слов (результат: словарь «длина – кол-во слов»).