

MapReduce w Sparku: Mnożenie macierzy

21 stycznia 2018

Opis pliku z zadaniami

Wszystkie zadania na zajęciach będą przekazywane w postaci plików `.pdf`, sformatowanych podobnie do tego dokumentu. Zadania będą różnego rodzaju. Za każdym razem będą one odpowiednio oznaczone:

- Zadania do wykonania na zajęciach oznaczone są symbolem \triangle – nie są one punktowane, ale należy je wykonać w czasie zajęć.
- Punktowane zadania do wykonania na zajęciach oznaczone są symbolem \diamond – należy je wykonać na zajęciach i zaprezentować prowadzącemu.
- Zadania do wykonania w domu oznaczone są symbolem \star – są one punktowane, należy je dostarczyć w sposób podany przez prowadzącego i w wyznaczonym terminie (zwykle przed kolejnymi zajęciami).

1 Mnożenie macierzy

25p.◇

Treść

Napisz program, który pozwoli na mnożenie dwóch macierzy przez siebie. Zakładamy, że żadna z macierzy nie mieści się w pamięci i obie należy wczytać jako dane wejściowe. Plik `result_mm.txt` zawiera oczekiwany wynik mnożenia macierzy z pliku `M.txt` przez macierz zawartą w pliku `N.txt`.

Format plików wejściowych: `wiersz kolumna wartość`

Podpowiedź:

- Zajrzyj do wykładu :)
- Wczytaj macierze podobnie jak na poprzednich zajęciach,
- Odpowiednio przemapuj wczytane dane (`map`),
- Następnie wykorzystaj następujące polecenia: `join`, `groupByKey`, `flatMap`, dwukrotnie `map`, `reduceByKey`, `sortByKey`, `foreach`, `println`.