

## Struktury Baz Danych

# Sprawozdanie z zadania 1.

*Tomasz Piwowski, indeks 180171*

### 1. Zastosowana metoda.

Zastosowałem scalanie naturalne wg schematu „2+1”. Wykorzystałem do tego celu 3 taśmy i dysk. Zarówno dysk jak i taśmy występują pod postacią plików dyskowych, tzn. ich implementacja zakłada zapisywanie i odczytywanie do pliku za każdym razem, gdy zapisywany/odczytywany jest nowy rekord. W moim przypadku, rekord zawiera 5 liczb. Sortowanie jest wykonywana na podstawie średniej arytmetycznej z tych liczb.

### 2. Specyfikacja formatu pliku tekstowego.

Dane można wprowadzać na 3 sposoby: losowo, z klawiatury lub z pliku. Przy uruchomieniu programu należy wybrać, którą opcję wybieramy. System poprosi o wprowadzenie 10 rekordów w postaci kolejnych po sobie liczb. W przypadku wczytywania z pliku wystarczy podać ścieżkę do niego. Zawartość pliku to kolejne liczby, każde 5 kolejny tworzy 1 rekord.

Dane w wewnętrznych plikach dyskowych symulujących dysk i taśmy (w tym: opis kolejnych kroków sortowania) są zapisane w formacie ciągłym, gdzie każdy rekord zaczyna się znakiem '{' i kończy znakiem '}' ; pomiędzy nimi znajduje się 5 liczb; liczby i znaki zakończenia/rozpoczęcia rekordu oddzielają pojedyncze spacje.

### 3. Sposób prezentacji wyników działania programu.

Uruchomienie programu i wybranie opcji „metoda generacji rekordów: losowo”. Wygenerują się rekordy, które zapisane zostaną do reprezentacji dysku. Następnie nastąpi sortowanie metodą „natural merge”. Na koniec wyświetlone zostaną: liczba zapisów i odczytów na dysku i taśmach oraz liczba faz. W pliku o nazwie „natural\_merge.log” można podejrzeć przebieg całego procesu z podziałem na poszczególne fazy i podglądem dla zawartości każdej z taśm.

Analogicznie uruchomienie dla metody generacji „z klawiatury” i „z pliku”.

Przedstawienie wykresów i obliczeń wykonanych na próbie 1000 wywołań programu.

### 4. Eksperyment.

Wg teorii na każdy rekord w schemacie „2+1” w każdej fazie przypadają 2 odczyty i 2 zapisy. Pokrywa się to w 100% dla 1000 próbek.

Ponadto dla 10 rekordów na dysku następowała różna ilość faz. Wykres przypomina rozkład normalny z największą wartością w punkcie 5.

