

Name: Koch Benjamin  
Matrikelnummer: 2110458043

Aufwand in h: 6h  
Punkte: /100

---

## Aufgabe 1

### Lösungsidee

Es wird eine Datei mit Zahlenwerten eingelesen (diese kann auch generiert werden).

Anschließend geschieht eine Partitionierung der Datei in zwei kleinere Dateien, um diese anschließend für sich mittels Merge Sort zu sortieren. Der Sinn hinter der Aufteilung ist, dass die komplette Datei eventuell zu groß für den Arbeitsspeicher ist und daher halbiert werden muss.

Das Prinzip von Merge Sort ist, ein rekursiver Algorithmus, der den zu sortierenden Vektoren (in unserem Fall Vektoren mit Integer) kontinuierlich in zwei Hälften teilt, bis er nicht mehr weiter geteilt werden kann.

Das bedeutet, dass die Teilung aufhört, sobald der lower Index gleich der Mitte ist, bzw. der higher Index gleich dem Ende des Vektors ist. Dies ist der Anker, um die Rekursion zu beenden.

Wenn der Vektor mehrere Elemente hat, wird der Vektor in Hälften geteilt und rekursiv die Zusammenführungssortierung für jede der Hälften aufgerufen.

Wenn beide Hälften sortiert sind, wird schließlich die Merge-Operation durchgeführt. Bei dieser Merge-Operation werden zwei kleinere sortierte Vektoren zu einem größeren Array kombiniert.

Ist dies bis zum Rekursionsanfang geschehen, liegen die zwei partitionierten Dateien in einem sortierten Zustand vor. Sie müssen anschließend nur noch geordnet zusammengefügt werden, was mithilfe einer repartition (merge) Funktion geschieht.

Annahmen: In dieser Version ist nur eine Halbierung der Datei möglich. Müsste man die Datei vierteln oder achteln, wäre die repartition Funktion am Ende, die die Dateien wieder zusammenfügt rekursiv aufzurufen und müsste jeweils eine Datei nach der anderen mergen.

## Testfälle

Testfälle siehe testing();

Testfälle, die zum Programmabbruch, bzw. Inkonsistenz führen:

- Zufällig Strings können zwar generiert aber nicht sortiert werden.  
Auch Sonderzeichen und sonstige Datentypen führen zu einer Exception beim Sortieren.
- Das Einfügen von mehr als zwei zu partitionierenden Daten führt zu Datenverlusten, da nur von der Halbierung der Datei ausgegangen wird.
- Dateinamen müssen korrekt eingegeben werden
- Der Stringcontainer dst muss mit zwei Dateinamen gefüllt werden, in die Partitioniert werden soll.