

## ADON-EX/高速ソート拡張機能ライブラリ

ADON-EX/高速ソート拡張機能ライブラリは、ADON/高速ソート・プログラムの導入された環境に、追加導入する事で次の機能を ADON/高速ソート・プログラムから使用できるようになります。

### 【 主な機能 】

#### 出力形式の指定

ソートの出力時に各フィールドの順番の入れ換え/削除、定数の挿入などを定義して、レコード形式を変更して出力ファイルに書き出すことができます。

#### 重複レコードの処理

キーが同じレコードについて、その範囲の最初あるいは最後のレコードだけを出力させることができます。

#### サマリー・ソート

キーが同じレコードについて、指定したフィールドの値の総合計を計算します。計算された合計値は、キーが同じ範囲の最初あるいは最後のフィールドと置き換えられて出力されます。合計値を計算する対象のフィールド・タイプは 2 バイト又は 4 バイトの符号付き二進整数あるいは符号無し二進整数、または数値フィールドです。合計値の計算中にオーバーフローが発生した場合は、オーバーフローが発生する直前のレコードの合計対象フィールドを中間値と置き換えて出力し、後続レコードの合計値を改めて計算して処理を継続します。

#### 統計情報出力

指定されたファイルに次の情報を書きだします。

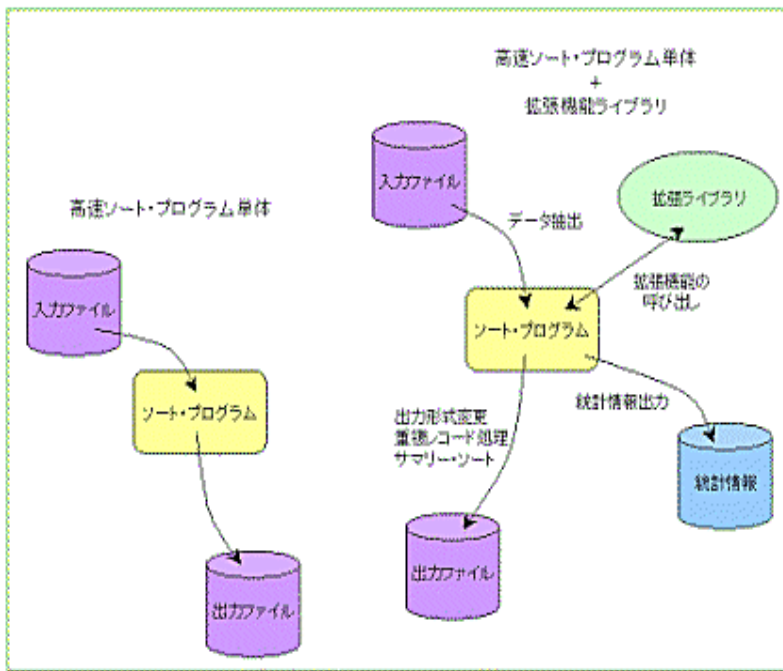
- 各キー・フィールドの開始位置、長さ、種類、昇順/降順
- 同じキー毎のレコード件数とキーの値
- 同じキーの最大/最少レコード件数と、それぞれのキーの値
- キーの数
- キーあたりの平均レコード数

#### データ抽出

条件指定により入力レコードを選択してソートを行う事ができます。

#### 可変長レコードの処理

任意のフィールド区切り文字を指定する事で、CSV の様な可変長フィールドのレコードを処理することができます。



## 【 指定できるキー・タイプの種類 】

キーのタイプは次のものが指定できます。

- キャラクター
- EBCDIC 順キャラクター
- 二進整数(2 バイト又は 4 バイト)
- パック十進数
- 数値フィールド

EBCDIC 順のソートでは、ASCII-EBCDIC の変換テーブルを指定できます。この変換テーブルでは、全 256 文字について変換するコードを指定できるので、単なる ASCII-EBCDIC の変換以外に、独自の重み付をした並び替えルールとして利用する事も可能です。例えば、辞書順(AaBbCc...YyZz)や電話番号順(ABC...YZabc...yz)というルールでソートする事ができます。

数値フィールドにより、ASCII 文字で表現された数字を、文字ではなく数値として大小関係を比較する事ができます。数値フィールドは、n(x,y)の形式で定義され、全体で x 桁、少数点以下 y 桁を意味します。有効な桁数は、符号と少数点を含めて最大 20 桁、そのうち数字部分は 18 桁です。

## 【 条件指定による入力レコードの抽出 】

条件式を与えて、入力レコードのうち条件に一致するものを抽出又は除外してソートすることができます。条件式は

<フィールド定義>,<比較演算子>,<フィールド定義又は定数>

の形式で指定でき、複数の条件式を「AND(論理積)」「OR(論理和)」で結合する事も可能です。また、括弧による条件式の優先順位も指定できます。括弧のネスト・レベルにプログラム上の制限はありません。

フィールド定義でサポートされるフィールド・タイプは次のものです。

- 文字
- 符号付き二進整数(2 バイト又は 4 バイト)
- 符号無し二進整数(2 バイト又は 4 バイト)
- バック十進数
- EBCDIC 順の文字
- バイト・ストリーム
- 数値フィールド

定数の種類は次のものが使えます。

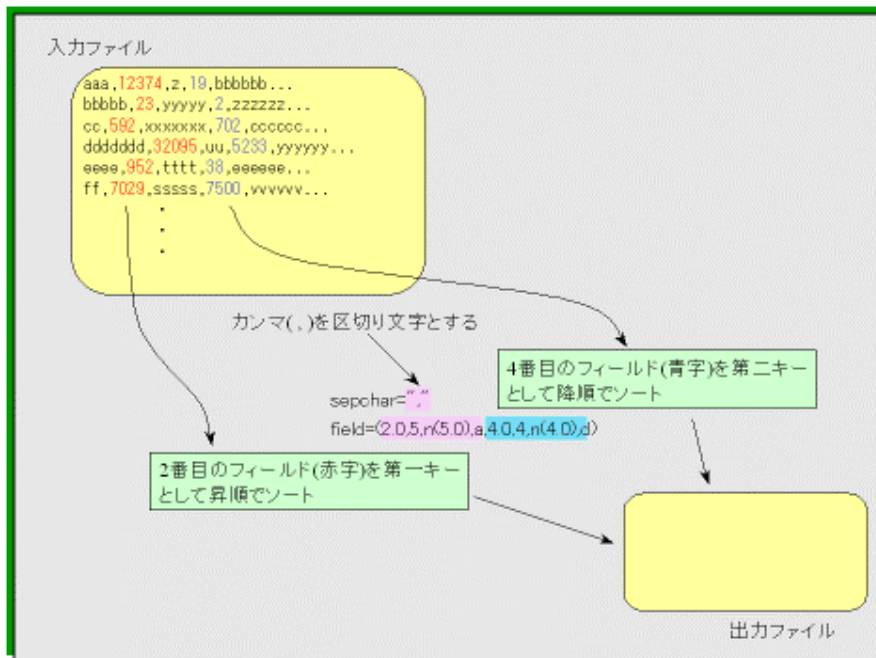
- 文字列
- 16 進数
- 十進数
- 数値

比較演算子は次の種類がサポートされます。

- EQ (等しい)
- NE (等しくない)
- GT (...より大きい)
- GE (...以上)
- LT (...より小さい)
- LE (...以下)

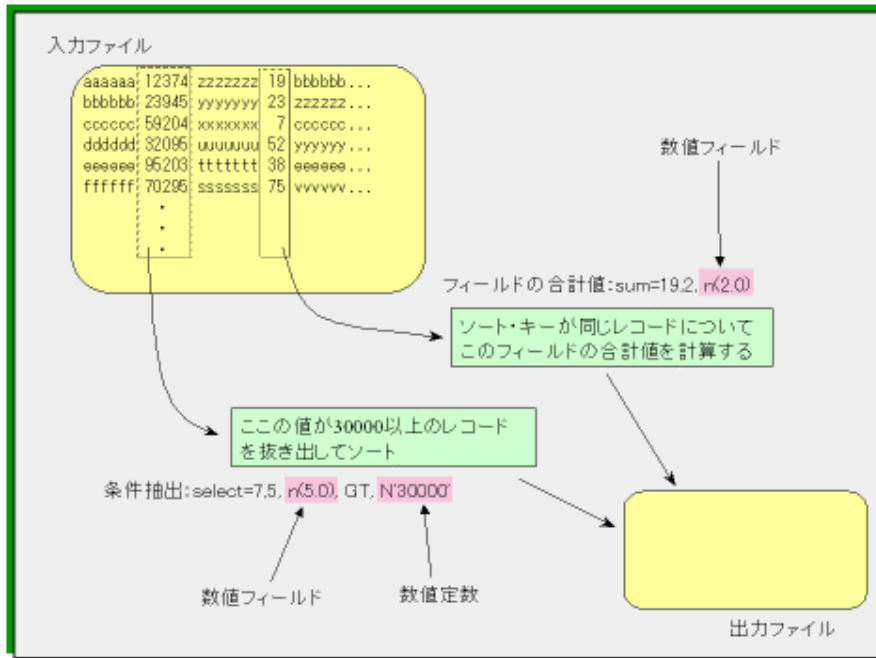
## 【 可変長フィールドのサポート 】

フィールド区切り文字を指定することで、特定の文字で区切られた可変長フィールドを扱うことができるようになります。例えば、各フィールドがカンマ「,」で区切られた CSV ファイルを入力ファイルとして、ソート処理を行うことができます。



## 【 数値フィールドのサポート 】

数値フィールドとは、ASCII 文字で表現された数字を「文字」ではなく「数値」として大小関係を判断するフィールド・タイプで、符号や少数点を含んだ数字に対してサマリー・ソートやデータの抽出などの拡張機能が使えます。



数値フィールドは  $n(x,y)$  の形式で定義され、全体で  $x$  桁、少数点以下  $y$  桁を意味します。数値フィールドは、符号と少数点を含めて最大 20 桁で、そのうち数字部分は最大 18 桁となります。