User manual EFC

Written by: Bart P.P. van Houte (bvhoute@few.vu.nl)
Version: October 15, 2009

1. Installation

Download the files 'EFC.cc' and 'Makefile' and save them together in the same directory. Compile the source code using the 'make' command in UNIX/LINUX. This generates an executable named 'EFC'. In case you are not working on a UNIX/LINUX platform, use a C/CC++ compiler.

2. Data format

The data file must be a tab-delimited text file, in which each row corresponds to the profile of one sample. The formatting must be as follows:

```
# Header lines that start with a '#' are allowed
# and will be ignored.
<sample ID><tab><value1><tab><value2><tab><value3> …
<sample ID><tab><value1><tab><value2><tab><value3> …
<sample ID><tab><value1><tab><value2><tab><value3> …
<sample ID><tab><value1><tab><value2><tab><value3> …
```

3. Running the program

The program is run using the following command line:

```
./EFC <distance measure> <nr clusters> <file>
```

As distance measure, you can choose from:

- *e* (Euclidean distance)
- *p* (Pearson distance)
- *s* (Spearman distance)
- *a* (agreement distance)
- *c* (concordance distance)

4. Example

Suppose one wants to cluster the synthetic normalized data into four clusters using the Pearson distance, the following command must be used:

```
./EFC p 4 synthetic_normalized.txt
```