File and folder info

1. Datasets.xlsx: Excel file indicating datasets used along with dataset parameters
2. test-out: contains queries and their results (for best match operation)
3. {Dataset}_Out_Other.csv: results for other CORAL operations

test-out folder has a folder for each dataset. Inside each dataset folder, **Inside(Outside)** corresponds to inside(outside) queries folder. Each respective folder has query files and results files for each system. Each dataset folder should have 50 inside and 50 outside queries.

Definitions

1. **Inside Queries**: subsequences present in dataset i.e. they have a correlation of 1 with (exact) best match
2. **Outside Queries**: subsequences not present in datset. For more info about their source, look in experimental section of CORAL paper.

Best match result file

{Dataset}_ {System} _{IQ/OQ}.csv IQ/OQ stand for inside/outside query. Following is the column information (left to right) for each system's results file

Naive

Correlation between query and each possible subsequence of dataset is computed. Subsequences with highest correlation is taken as best-match and is considered to be ground truth. Results for Naive method are with CORAL results file

CORAL

- index: query no.
- Query ID: query has been taken from this time series(0 based indexing) in dataset (only applies to inside queries).
- Query start: query starts at this offset(0 based indexing) in dataset (only applies to inside queries)
- Query Length: lenght of the query
- CORAL TSID: time series of subsequence found as best match by CORAL
- CORAL TStart: offset of subsequence found as best match by CORAL
- CORAL Corr: correlation between query and subsequence found as best match by CORAL
- CORAL Time (s): time taken by CORAL online processing (seconds)
- Error: error between correlation with CORAL-best-match and Naive-best-match
- CORAL no. Candidate Clusters: UNDEFINED
- Naive TSID: time series of subsequence found as best match by Naive
- Naive TStart: offset of subsequence found as best match by Naive

- Naive TEnd: end of subsequence found as best match by Naive
- Naive Corr: correlation between query and subsequence found as best match by Naive (should be 1 for inside)
- Naive Time (s): time taken by Naive online processing (seconds)

PAA

- index: query no.
- TSID: time series of subsequence found as best match
- TStart: offset of subsequence found as best match
- TEnd: end of subsequence found as best match
- Corr: correlation between query and subsequence found as best match
- Time(S): time taken by Naive online processing (seconds)

JOCOR

- FileID: query number (same as *index* above except it is 1 based)
- mi: 0 (see below)
- mj: offset of best match (see below)
- mlen: length of best match subsequence
- maxC: correlation between query and subsequence found as best match by JOCOR
- FFT+Search: time taken by FFT and JOCOR algorithm
- FFT time: time taken by FFT in seconds

NOTE: JOCOR considers all dataset as 1 long time series. So *mi* is fixed to zero and mj is the offset when all time series are flattened into one long time series.

JOCOR-LA

Same as JOCOR

NOTE: *JOCOR* has been represented by the string *JACOR*

Other operation result file

- Index: query no.
- Length: length of query
- GCS-CorrThres: correlation threshold used for Group of Correlated Sequence (GCS) operation
- GCS-NaiveTime: time taken by Naive method for GCS operation (seconds)
- GCS-CORALTime: time taken by CORAL for GCS operation (seconds)
- TS: time series used in Self Correlation (SC) operation
- SC-Naive: time taken by Naive method for GCS operation (seconds)
- SC-CORAL: time taken by CORAL for SC operation (seconds)