Introducing the Annotation Vector

Takeaway Messages
The Matrix Profile (introduced a year ago) is a domain independent approach to make most time series tasks trivial.
It allows motif discovery, density estimation, anomaly detection, rule discovery, joins, segmentation, clustering, visualization etc.
However, sometimes you may wish to impose domain dependent constraints from your particular domain. The Annotation Vector is a simple principled way to do so.

Matrix Profile and Time Series Motifs Discovery
If you know the nearest neighbors of every subsequence in a time series, you can solve many time series data mining tasks trivially. For example, to find time series motifs, you locate the subsequence pairs with the smallest mutual distance.
The Matrix Profile is a data structure that gives you that information. The Matrix Profile can be computed fast thanks to [1].

Classic Motif Search: What you see may not be what you want
The top motif pair are the two normal heartbeats
A cardiologist is, however, more interested in the two repeated ventricular contractions
A snippet of ECG data from BIDMC Congestive Heart Failure Database

Annotation Vector Framework
Main idea: combines the Matrix Profile (MP) with the Annotation Vector to produce a new Matrix Profile, which incorporates the contextual bias for the problem at hand.
The Annotation Vector is a time series consisting of real-valued numbers between [0 – 1].
A lower value indicates the subsequence starting at that index is less desirable, and therefore should be biased against.

The Math behind
\[ CMP_i = MP_i + (1 - AV_i) \cdot \max(MP) \]
where
- \( CMP_i \) is the corrected Matrix Profile value at index \( i \)
- \( MP_i \) is the original Matrix Profile value at index \( i \)
- \( AV_i \) is the Annotation Vector value at index \( i \)
- \( \max(MP) \) is the maximum value of the original Matrix Profile

Case Study: Suppressing Hard-limited Artifacts
A snippet of a left-eye EOG sampled at 50 Hz from an individual with sleep disorder

References
To learn more about the Matrix Profile and the wonderful things it has to offer:
To learn more about the Annotation Vector and download code and datasets:
http://www.cs.ucr.edu/~hdau001/guided_motif_search/