While we believe our paper is self contained, this presentation contains:

1. Augmented and larger scale versions of experiments shown in the paper.
2. Additional experiments that we could not fit in the paper.
3. Comparisons to other techniques (deliberately omitted from the paper for reasons discussed therein).

Database: MIT-BIH ST Change Database (stdb) Select database

Record: 308

Annotator: atr (reference beat annotations)

Start time: 0

Chart width: small medium large

Show chart E-mail chart to:

Convert signals to text

Convert annotations to text

Annotation key

Chart-O-Matic Help

Record stdb/308

Download a high-resolution PostScript version of this chart

We have changed the original screen shot only by adding a red circle to highlight the anomaly

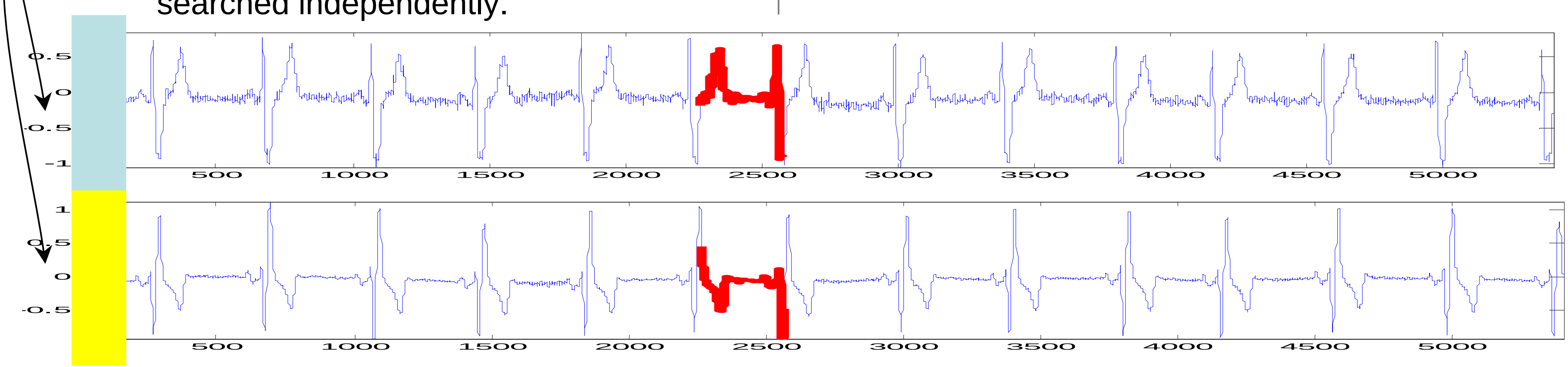
In all the examples below, we have included screen dumps of the MIT ECG server in order to allow people to retrieve the original data independent of us.

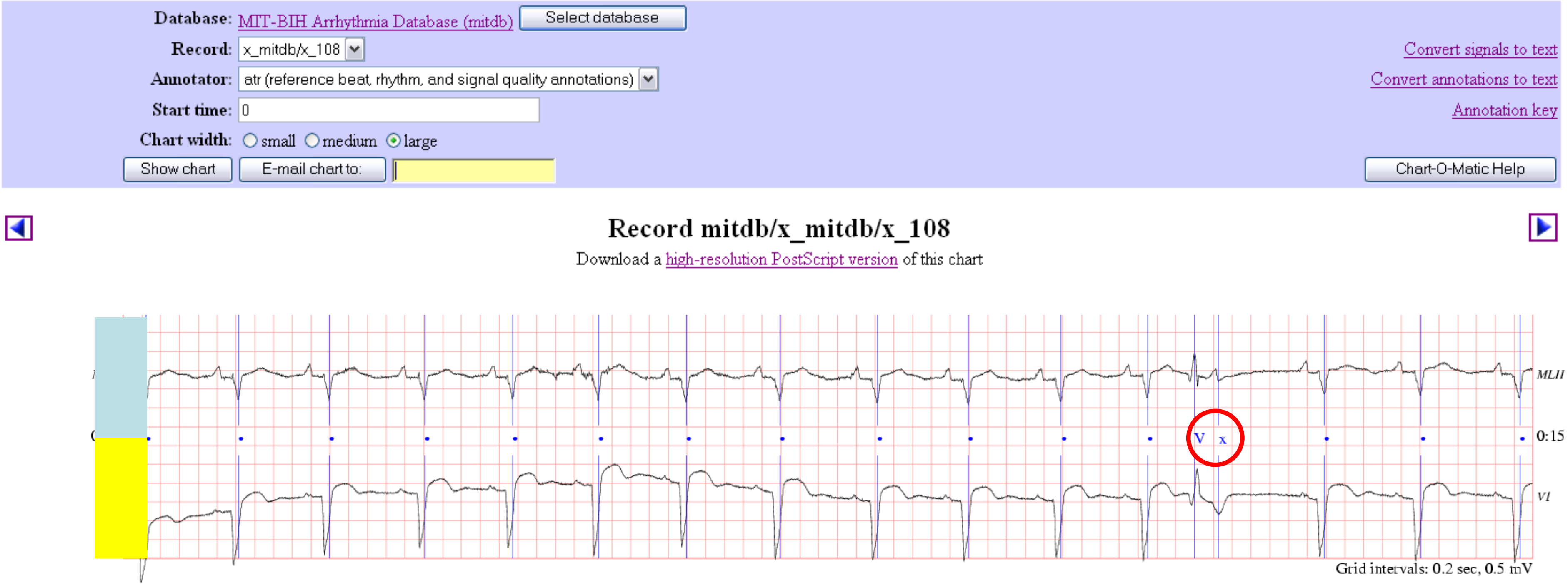
However, all data is also available from us in a convenient zip file.

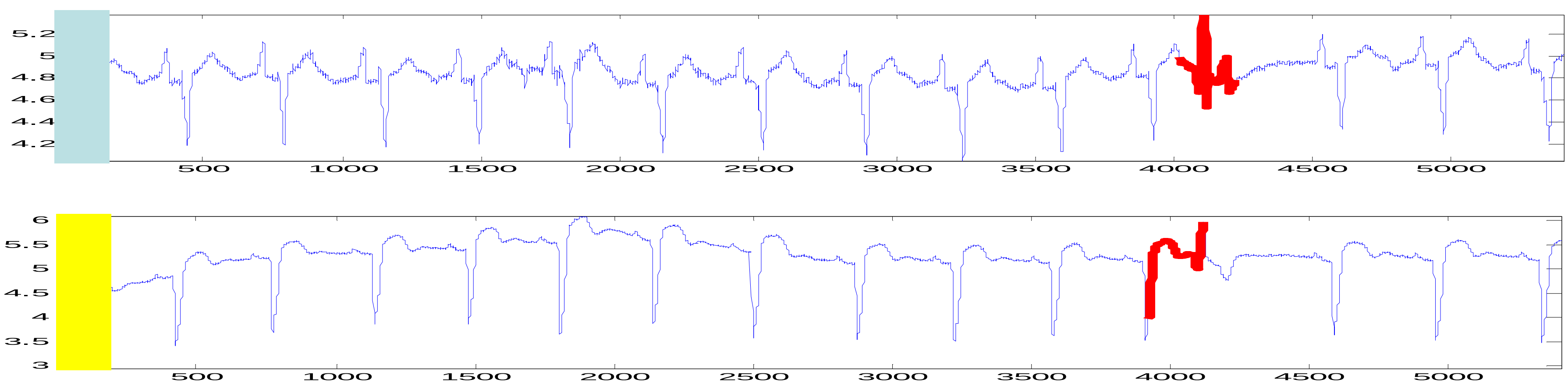
This is KEY only, the next 8 slides show examples in this format

The annotated ECG from PhysioBank (two signals)

Anomalies (marked by red lines) found by the discord discovery algorithm. Each of the two traces were searched independently.







Each of the two traces were searched independently.

Each of the two traces were searched independently.

Database: BIDMC Congestive Heart Failure Database (chfdb) | Select database

Record: chf01

Annotator: ecg (unaudited beat annotations from an automated detector)

Start time: 275

Chart width: small | medium | large

Show chart | E-mail chart to: | Chart-O-Matic Help

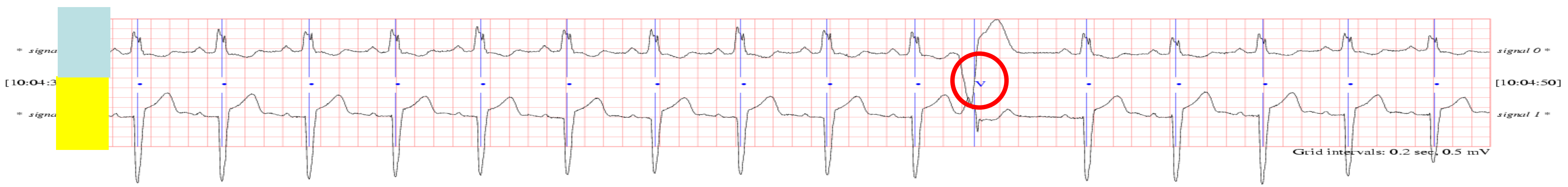
Convert signals to text

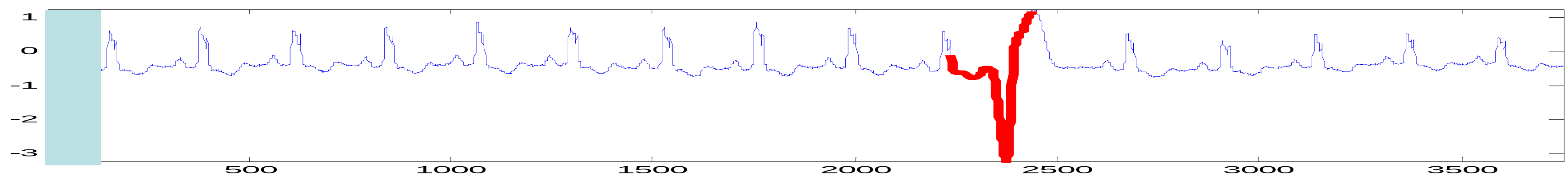
Convert annotations to text

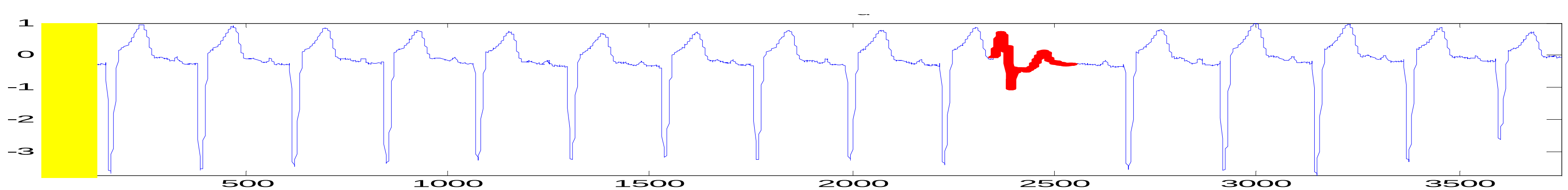
Annotation key

Record chfdb/chf01

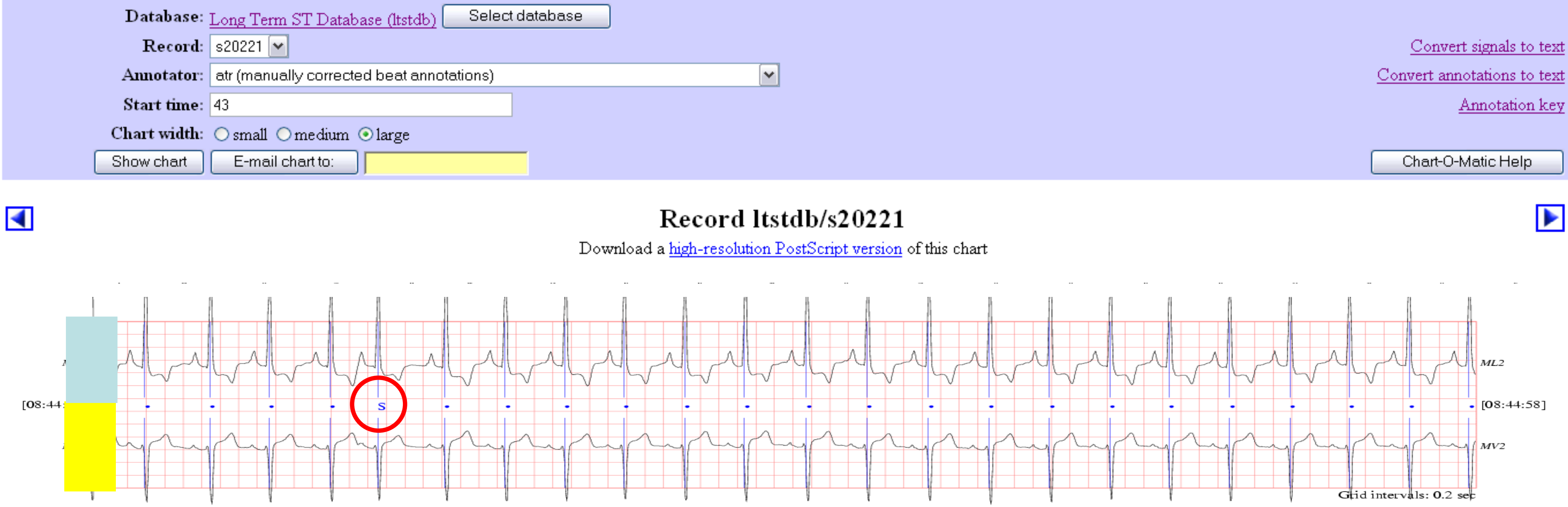
Download a high-resolution PostScript version of this chart

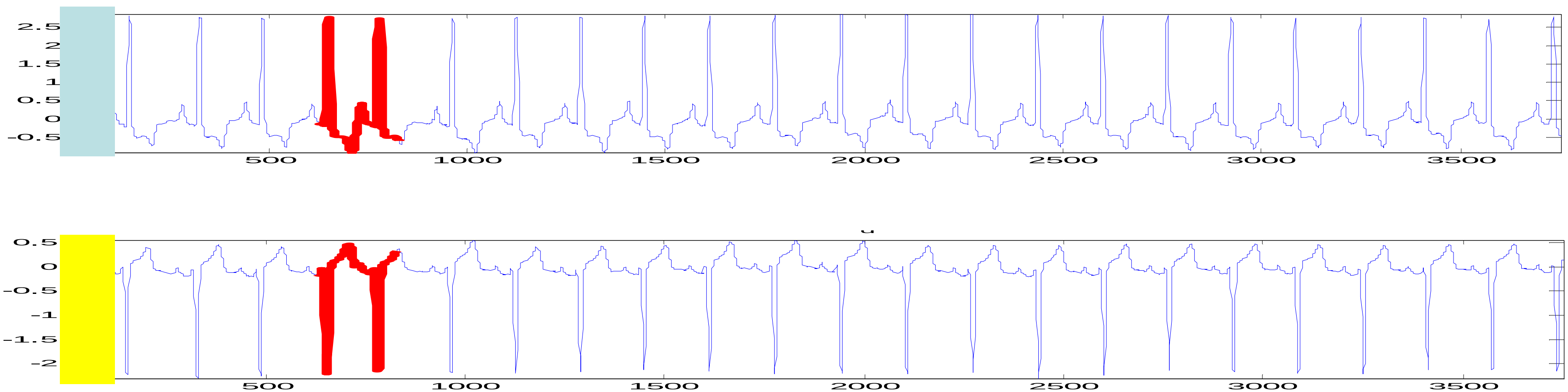






Each of the two traces were searched independently.





Each of the two traces were searched independently.

Database: Long Term ST Database (ltstdb) Select database

Record: s20321

Annotator: atr (manually corrected beat annotations)

Start time: 240

Chart width: small medium large

Show chart E-mail chart to:

Convert signals to text

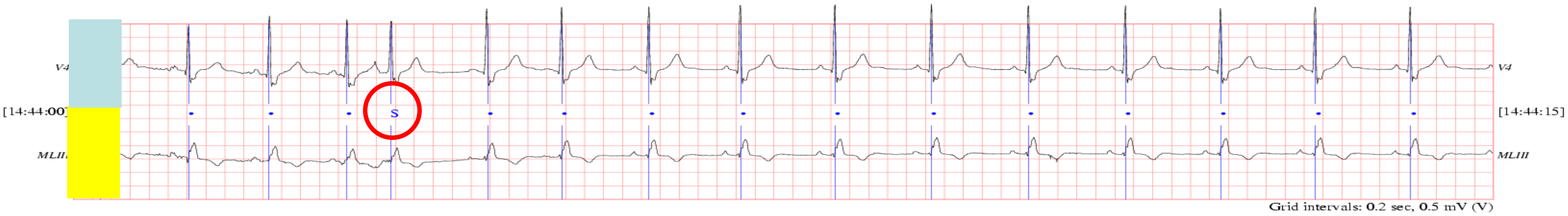
Convert annotations to text

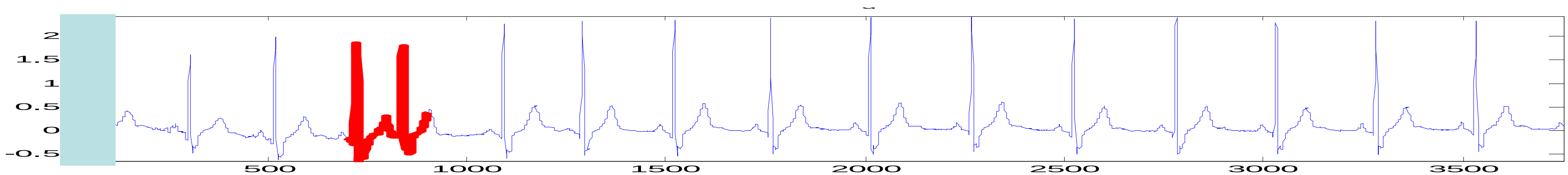
Annotation key

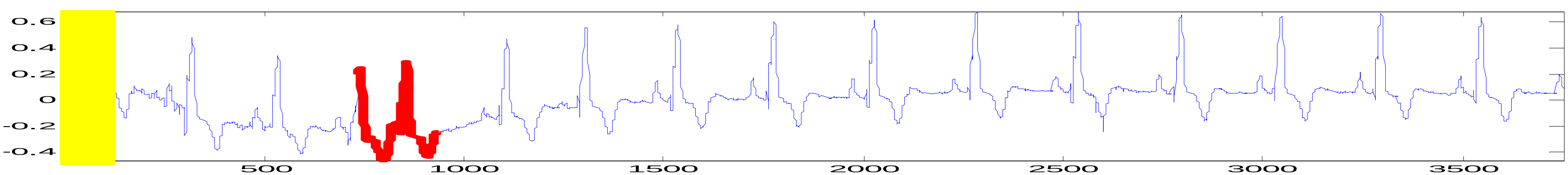
Chart-O-Matic Help

Record ltstdb/s20321

Download a high-resolution PostScript version of this chart



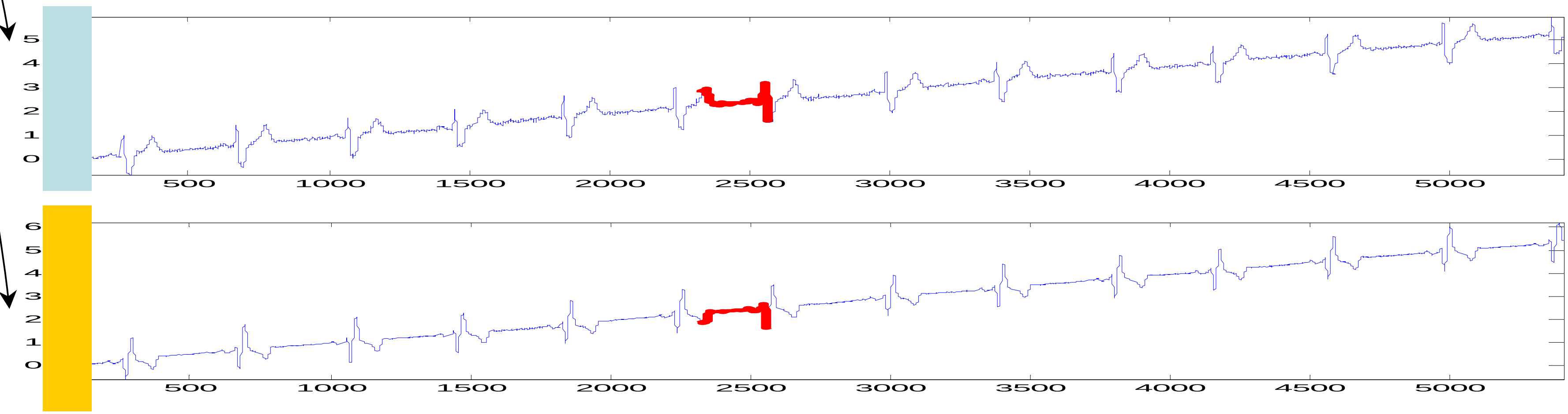




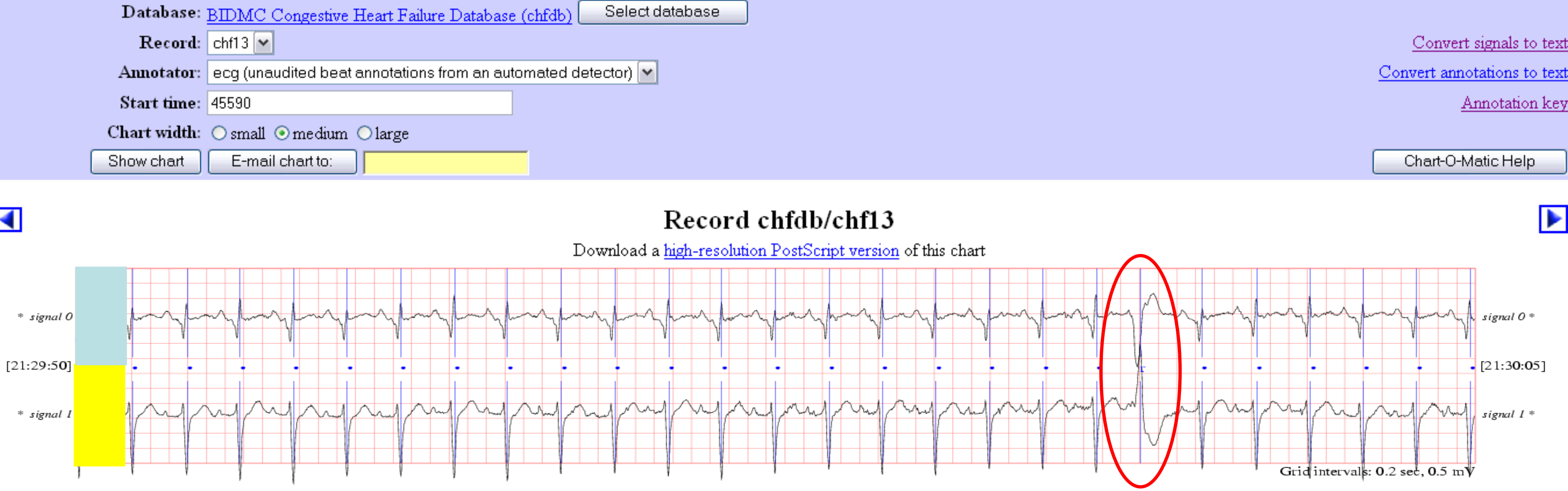
Each of the two traces were searched independently.

This is a dataset shown in a previous example

To demonstrate that the discord algorithm can find anomalies even with the presence of linear trends, we added linear trend to the ECG data on the top. The new data and the anomalies found are shown below. This is important in ECGs because of the *wandering baseline* effect, see Figure 11 in the paper.

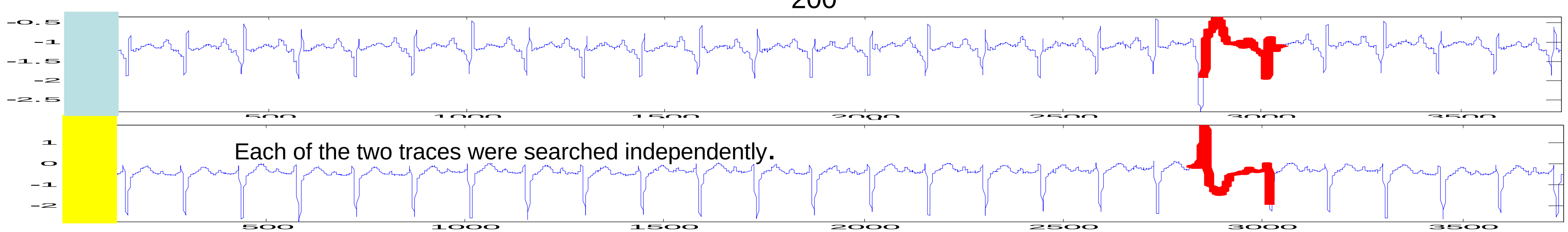


Each of the two traces were searched independently.

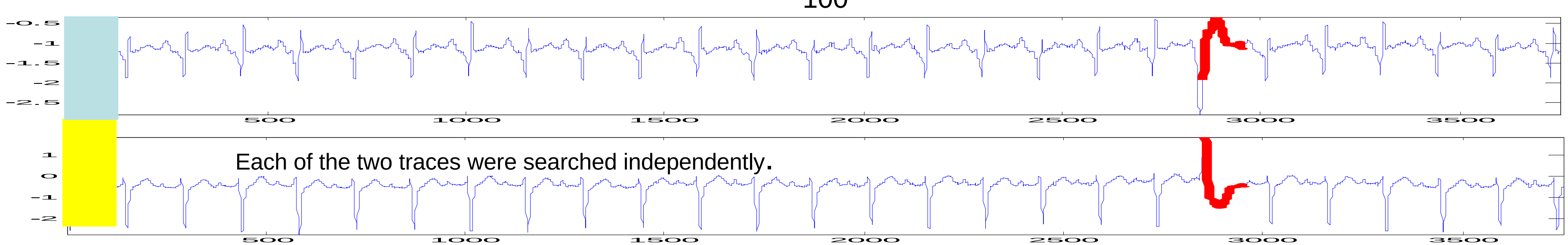


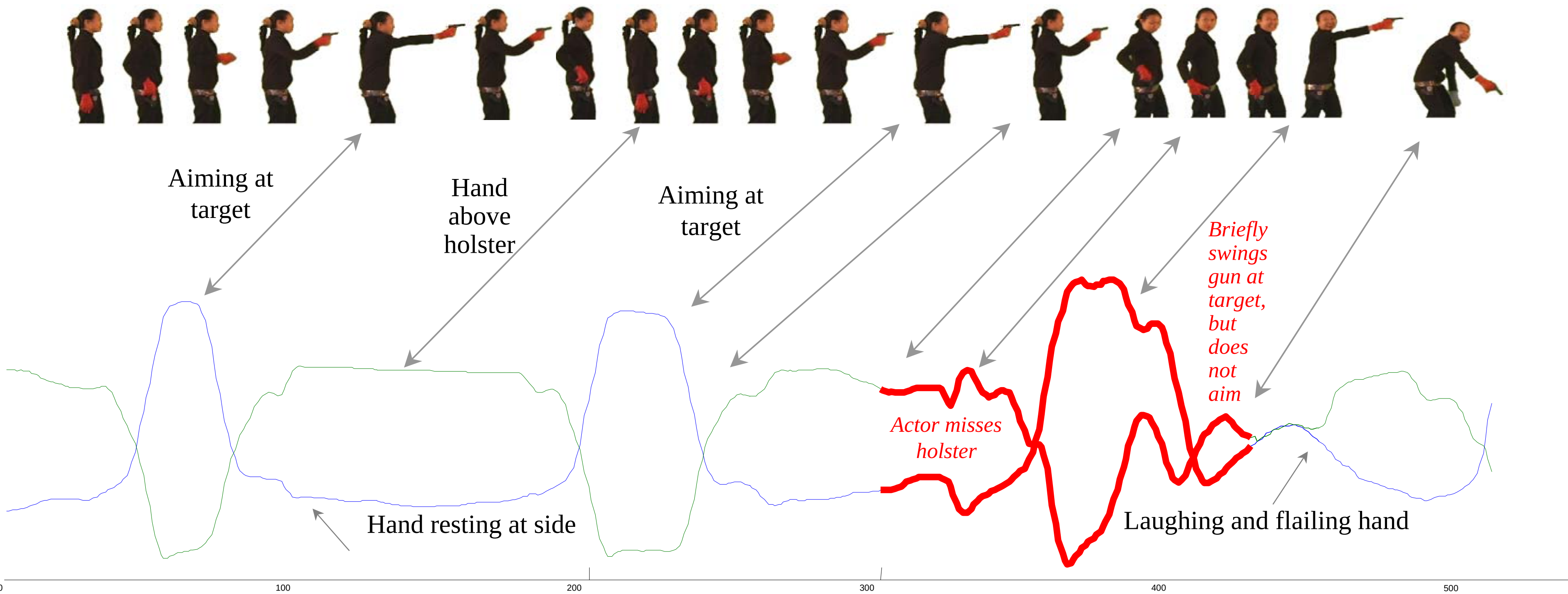
This example shows that the discord algorithm is not sensitive to the window size. In fact on all 8 problems above, we can double or half the discord length and still find the anomalies. Below is just one example for clarity.

discord$_{200}$



discord$_{100}$

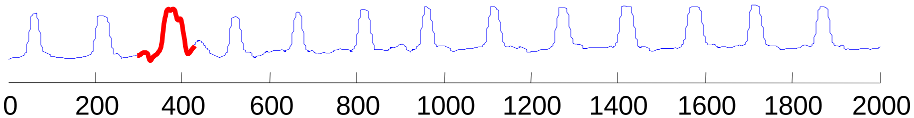




The 2D time series was extracted from a video of an actor performing various actions with and without a replica gun. The film strip above illustrates a typical sequence. The two time series measure the X and Y coordinates of the actors right hand. The actor draws a replica gun from a hip mounted holster, aims it at a target, and returns it to the holster. Watching the video we discovered that at about ten seconds into the shoot, the actor misses the holster when returning the gun. An off-camera (inaudible) remark is made, the actor looks toward the video technician, and convulses with laughter. At one point (frame 450), she is literally bent double with laughter.

This is dataset ann_gun_CentroidA

Here is a longer subsection for context. Later in this presentation we test two other algorithms on this dataset.

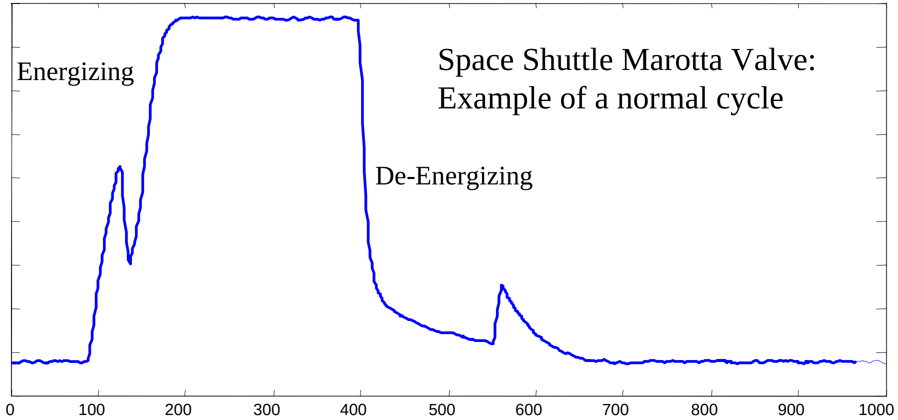


The next few slides demonstrate utility of discords in finding anomalies in Space Shuttle Marotta Valve time series.

In every case, there are five examples of an Energize/De-Energize cycle. Exactly *one* cycle has been annotated by a domain expert as been abnormal.

Each cycle is of length 1000, and we know in advance that an anomaly can be just a part of a cycle, so we set the length of the discord to be an fraction of this (in particular, 128) for all experiments shown here. (We note that we get correct results for all experiments here if we double or half this arbitrary choice.)

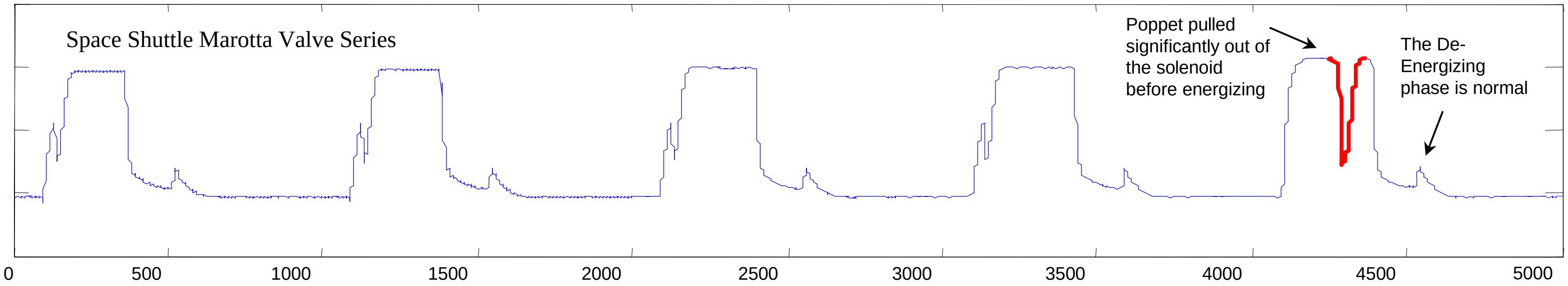
We tested on 3 different challenges, of increasing difficulty. Note that all the annotations shown are those of the domain expert.



This is Figure 5 in the paper.

Test 1: A simple problem

This is dataset TEK16.txt

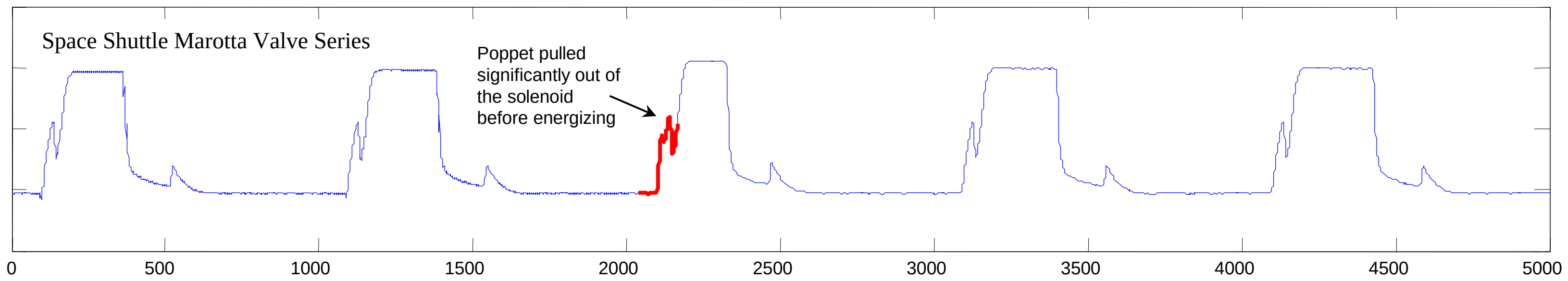


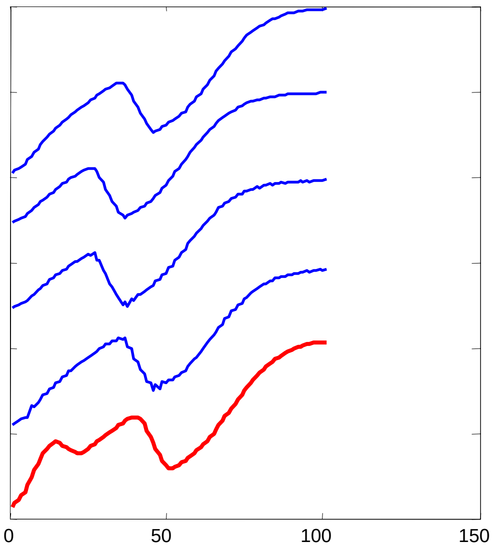
In this case the anomaly is very obvious, and the discord (marked in red) easily finds it.

This is Figure 6 in the paper.

Test 2: A more subtle problem

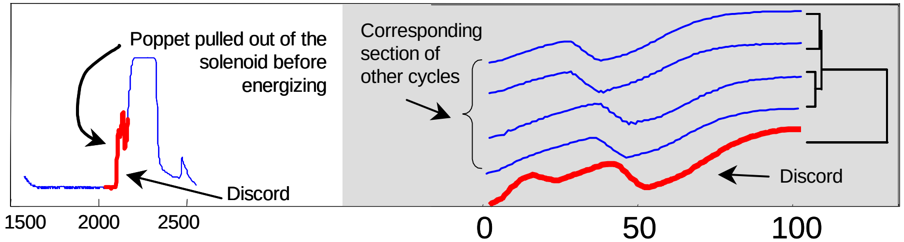
This is dataset TEK17.txt



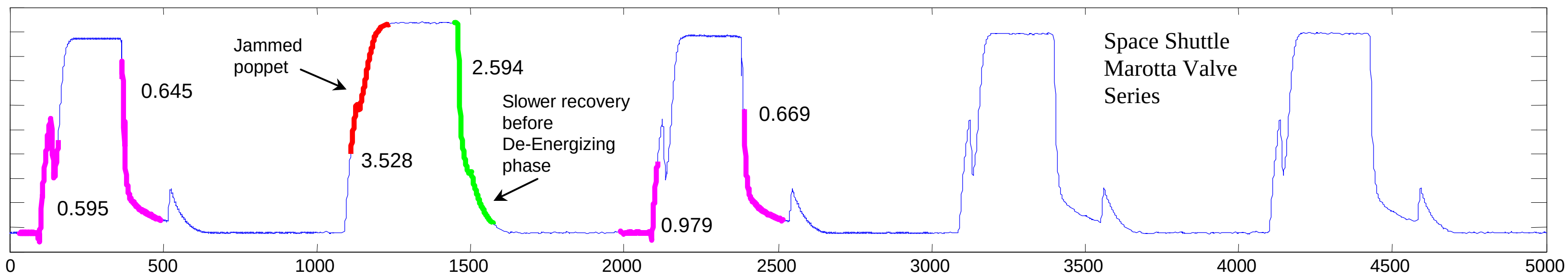


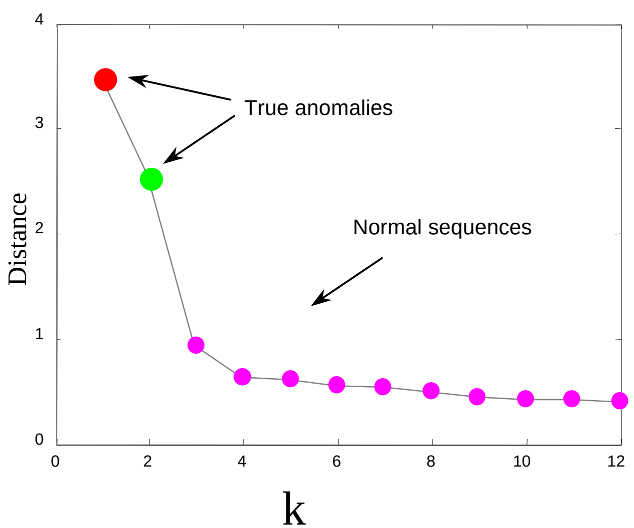
Here the discord (marked in red) easily finds the anomaly marked by the domain expert, but it is not obvious (at this scale) what the anomaly was.

A *zoom-in* of the anomaly, and the 4 corresponding segments from the normal cycle (*left*), explains what the discord discovered. Only the anomalous cycle has a "*double hump*".
These are from Figures 7 and 8 in the paper.



Test 3: Finding multiple discords





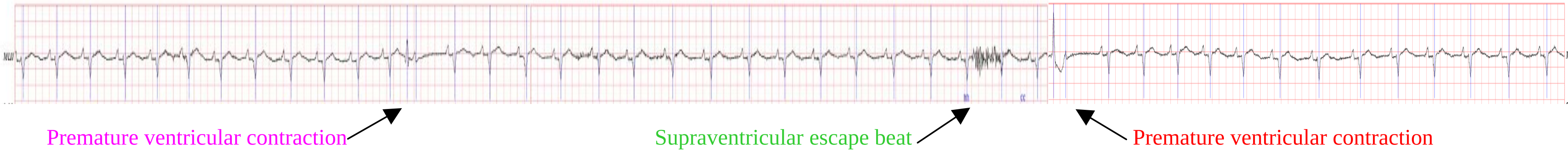
In this example we consider the problem of knowing when an discord is significant.

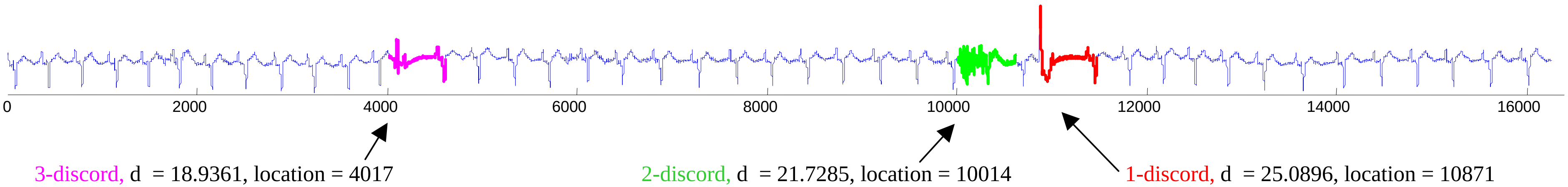
We found the top 12 discords (only 6 are shown above for clarity). The top 2 correspond to true anomalies, in red we see a missing small peak before the large plateau, and in green we see a slower recovery before the de-energizing phase. The next 4 discords are shown in pink.

If we plot the discords scores against K (*left*) we can see that we could potentially assess the significance of an discord with some kind of “knee finding” algorithm.

This example was *not* in the paper.

The next 6 slides are larger versions of the examples shown in the paper.

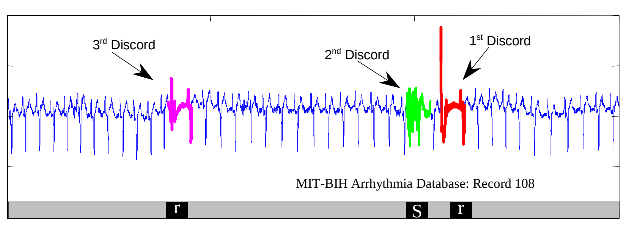




The time series is record mitdb/x_mitdb/x_108 from the PhysioNet Web Server (The local copy in the UCR archive is called mitdbx_mitdbx_108.txt). It is a two feature time series, here we are looking at just the MLII column.

Cardiologists from MIT have annotated the time series, here we have added colored makers to draw attention to those annotations.

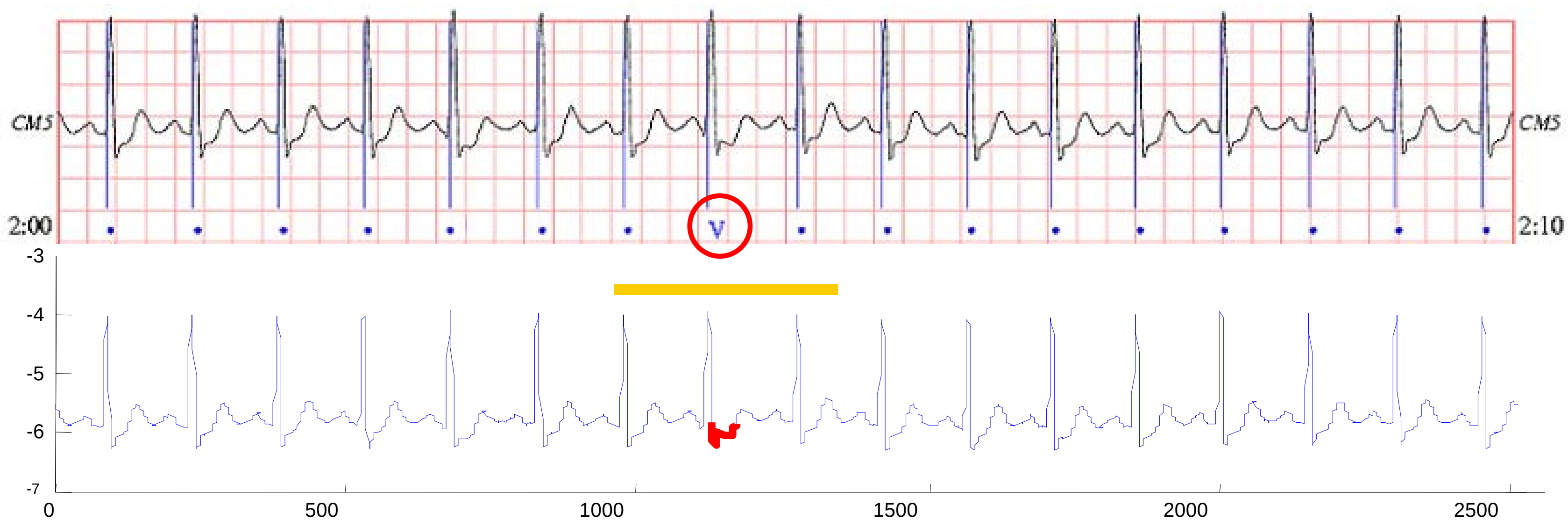
Here we show the results of finding the top 3 discords on this dataset. We chose a length of 600, because this a little longer than the average length of a single heartbeat.

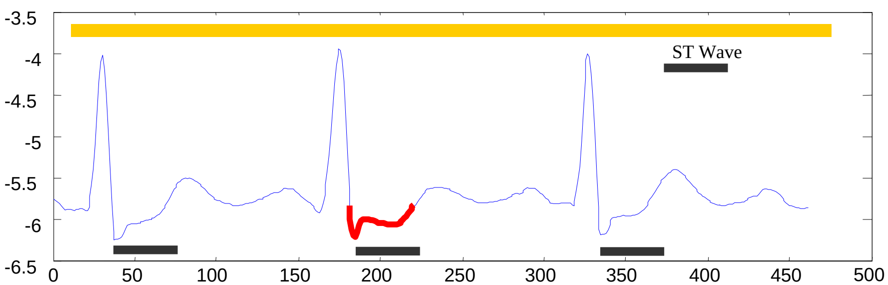


This is Figure 12 in the paper.

This is part of Record qtdbsele0606from the PhysioBank QT Database (qtdb) (the local copy in the UCR archive is called qtdbsele0606.txt)

A cardiologist noted subtle anomalies in this dataset. Let us see if the discord algorithm can find them.

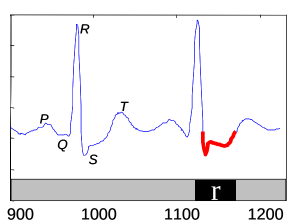


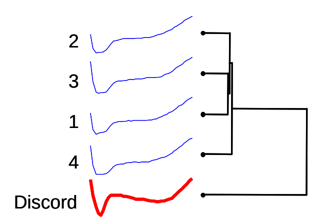


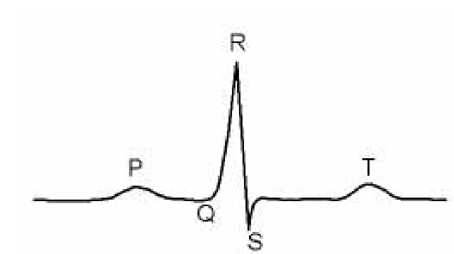
The time taken was 0.3% of the time for brute force

This is figures 12/13 in the paper.

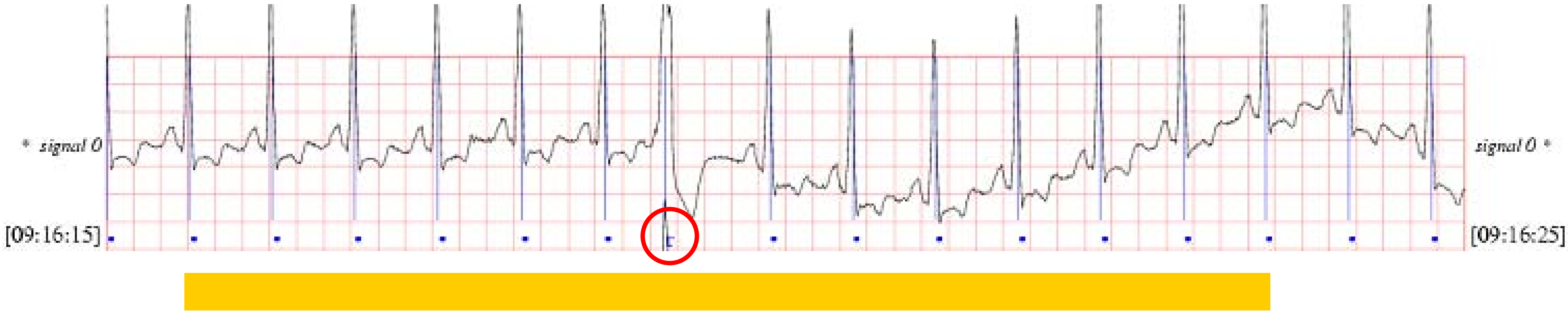
How was the discord able to find this very subtle Premature ventricular contraction? Note that in the normal heartbeats, the ST wave increases monotonically, it is only in the Premature ventricular contractions that there is an inflection.NB, this is not necessary true for all ECGS

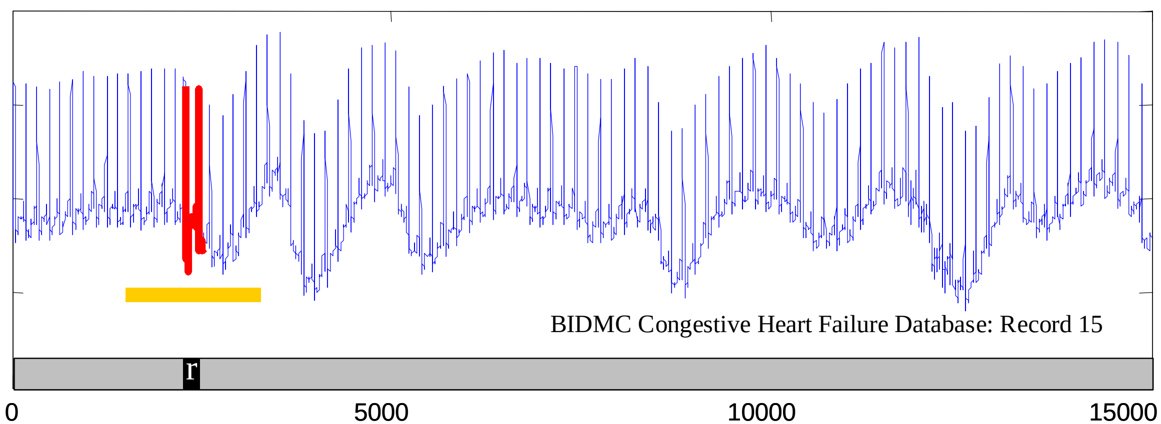




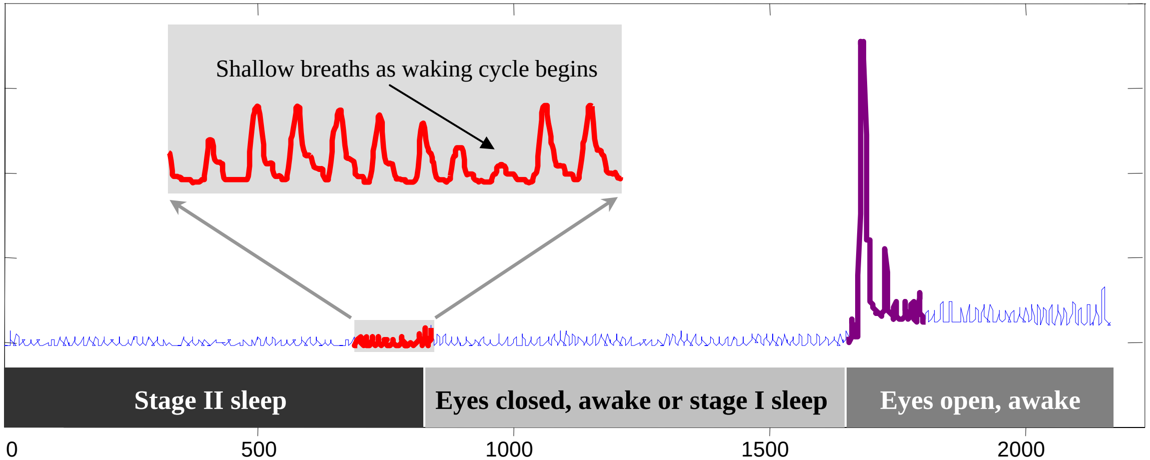


The is file chfdbchf15.txt





The time taken was 0.047% of the time for brute force
This is Figure 11 in the paper.



A time series showing a patients respiration (measured by thorax extension), as they wake up. A medical expert, Dr. J. Rittweger, manually segmented the data. The 1-discord is a very obvious deep breath taken as the patient opened their eyes. The 2-discord is much more subtle and impossible to see at this scale. A zoom-in suggests that Dr. J. Rittweger noticed a few shallow breaths that indicated the transition of sleeping stages.

Institute for Physiology. Free University of Berlin.

Data shows respiration (thorax extension), sampling rate 10 Hz.

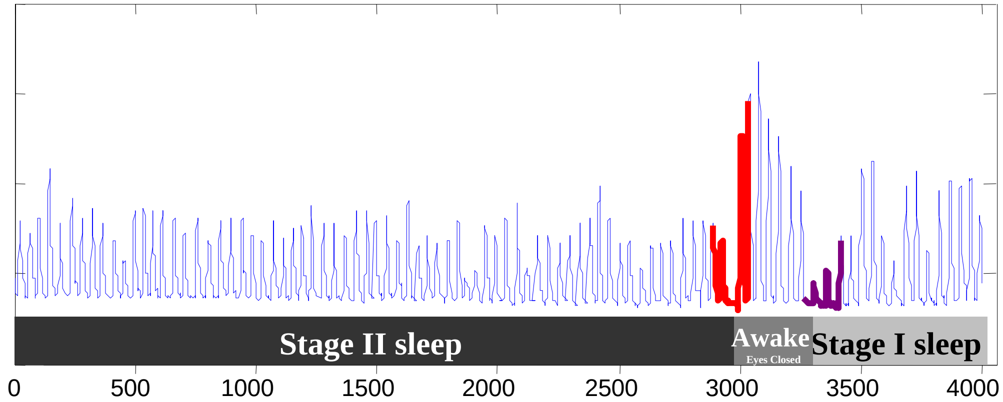
This is Figure 9 in the paper.

This is dataset nprs44

Beginning at 15500

Ending at 22000

The beginning and ending points were chosen for visual clarity (given the small plot size) they do not effect the results



A time series showing a patients respiration (measured by thorax extension), as they wake up. A medical expert, Dr. J. Rittweger, manually segmented the data.

Institute for Physiology.Free University of Berlin.

Data shows respiration (thorax extension), sampling rate 10 Hz.

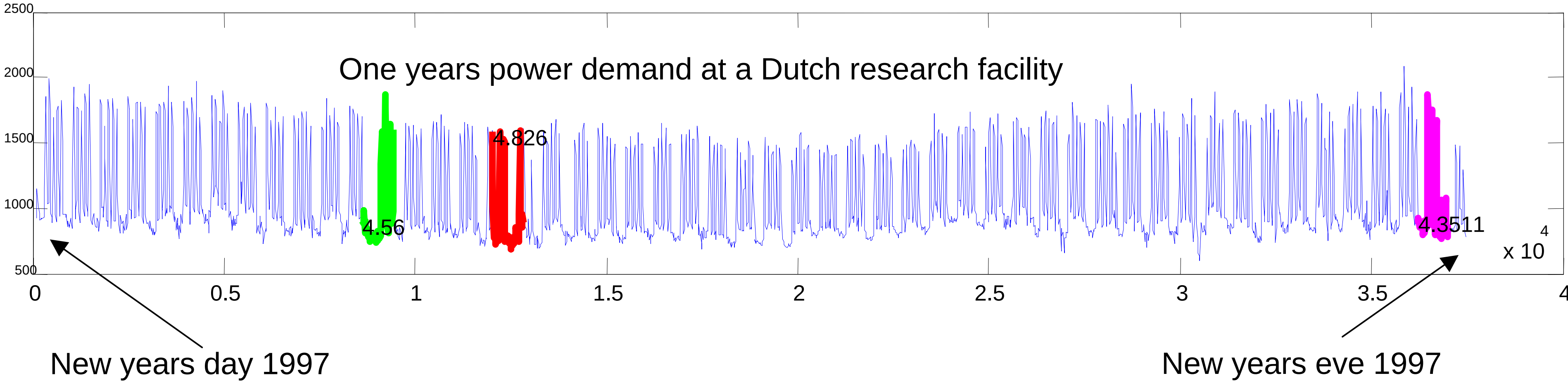
This is Figure 10 in the paper.

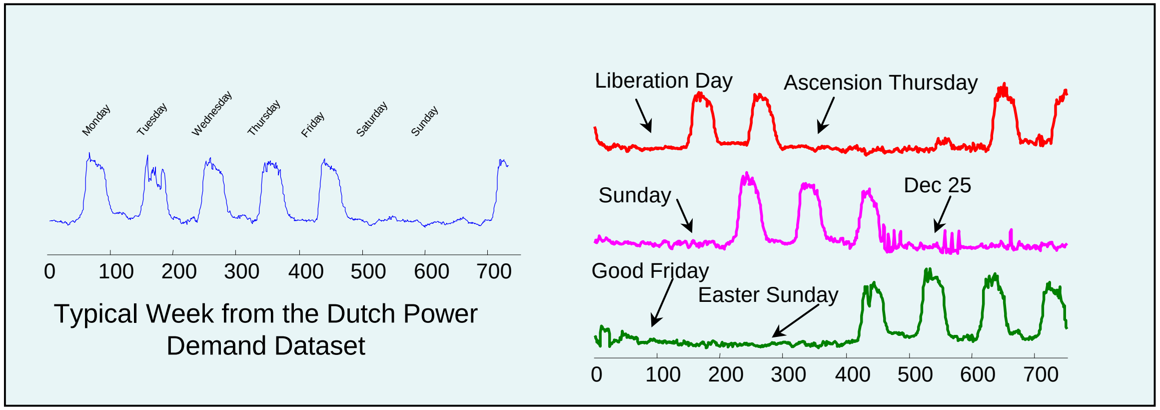
This is dataset nprs43

Beginning at 1

Ending at 4000

The beginning and ending points were chosen for visual clarity (given the small plot size) they do not effect the results





This is dataset power_data

This is Figure 15 from the paper. Many more details about this dataset are available in this paper:

van Wijk J. J. and van Selow E. R. *Cluster and calendarbased visualization of time series data*. In Proc. IEEE Symposium on Information Visualization, pages 4-9, Oct. 25-26, 1999.

Comparisons to other techniques

As we noted in the paper, any comparisons we could do to other techniques are probably unfair to the rival methods.

This is because discords only require a single parameter, and as we have seen above, we can typical double or half this parameter without effecting the results.

In contrast, most other anomaly detection schemes require require 3 to 7 parameters , including some parameters for which we may have poor intuition, such as *Embedding dimension*, *Kernel function, SOM topology* or *number of Parzen windows*.

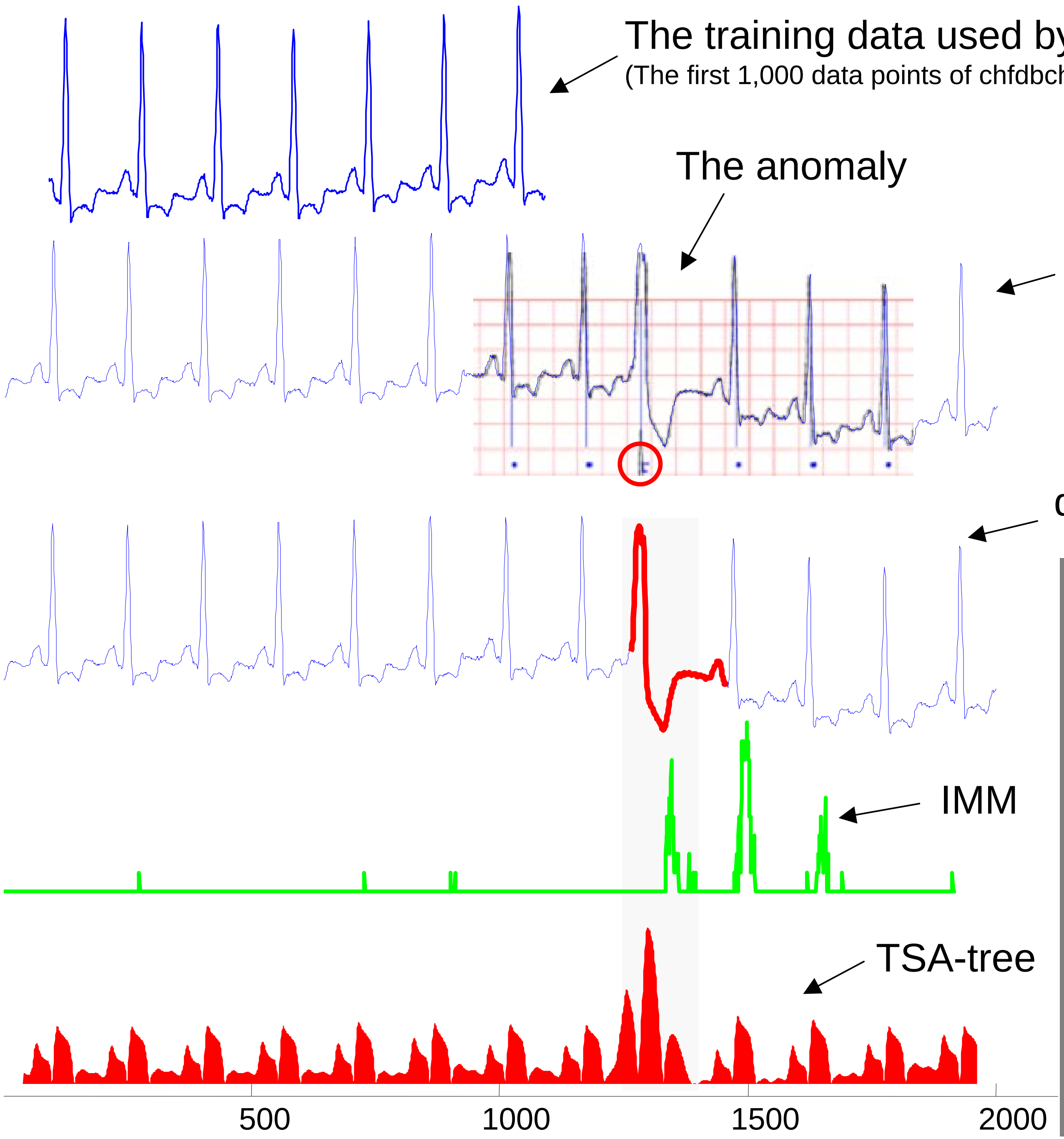
Nevertheless, comparisons to rival methods are a cornerstone of science, and we have shown some of the results below. We selected the experiments where the rival techniques did the best. While we actually did comparisons to the 4 techniques listed in [1], we only show two of these in the experiments below. These two techniques were chosen because they each have at least 20 references and each performed reasonably well on at least one dataset.

The two rival techniques are:

- Immunology (IMM) inspired approach (5 parameters) [2].
- TSA-tree Wavelet based approach (3 parameters) [3].

For each rival technique, for each experiment, we spent one hour of human time searching for the best parameters.

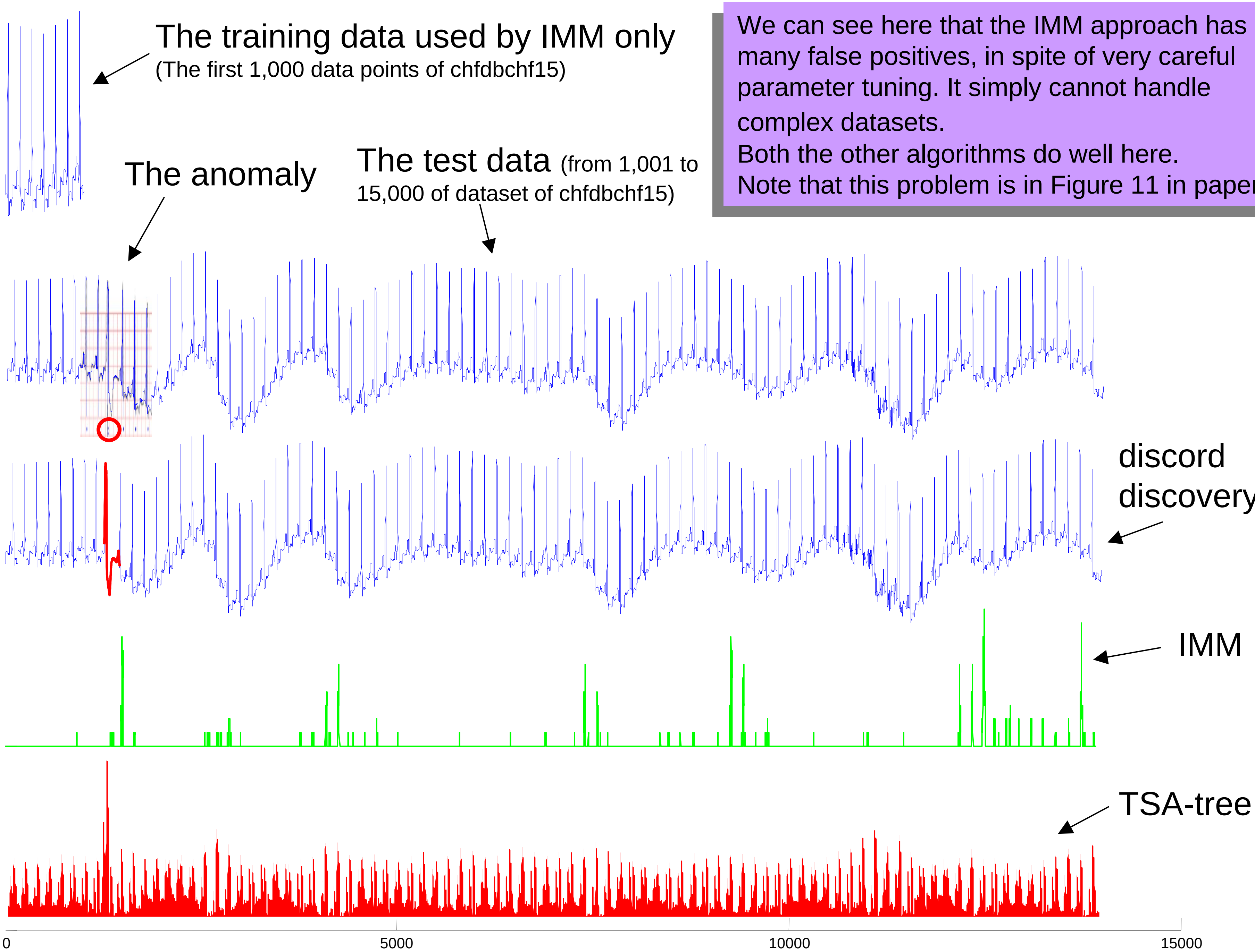
[1] Keogh, E., Lonardi, S. and Ratanamahatana, C. (2004). Towards Parameter-Free Data Mining. In proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, WA, Aug 22-25, 2004.
[2] D. Dasgupta and S. Forrest, Novelty Detection in Time Series Data Using Ideas from Immunology Proceedings *of the 5th International Conference on Intelligent Systems,* Reno, June, 1996
[3] C. Shahabi, X. Tian. & W. Zhao. TSA-tree: a wavelet-based approach to improve the efficiency of multi-level surprise and trend queries. In Proc. of 12th Int'l Conf. on Scientific and Statistical Database Management. pp 55-68, 2000.

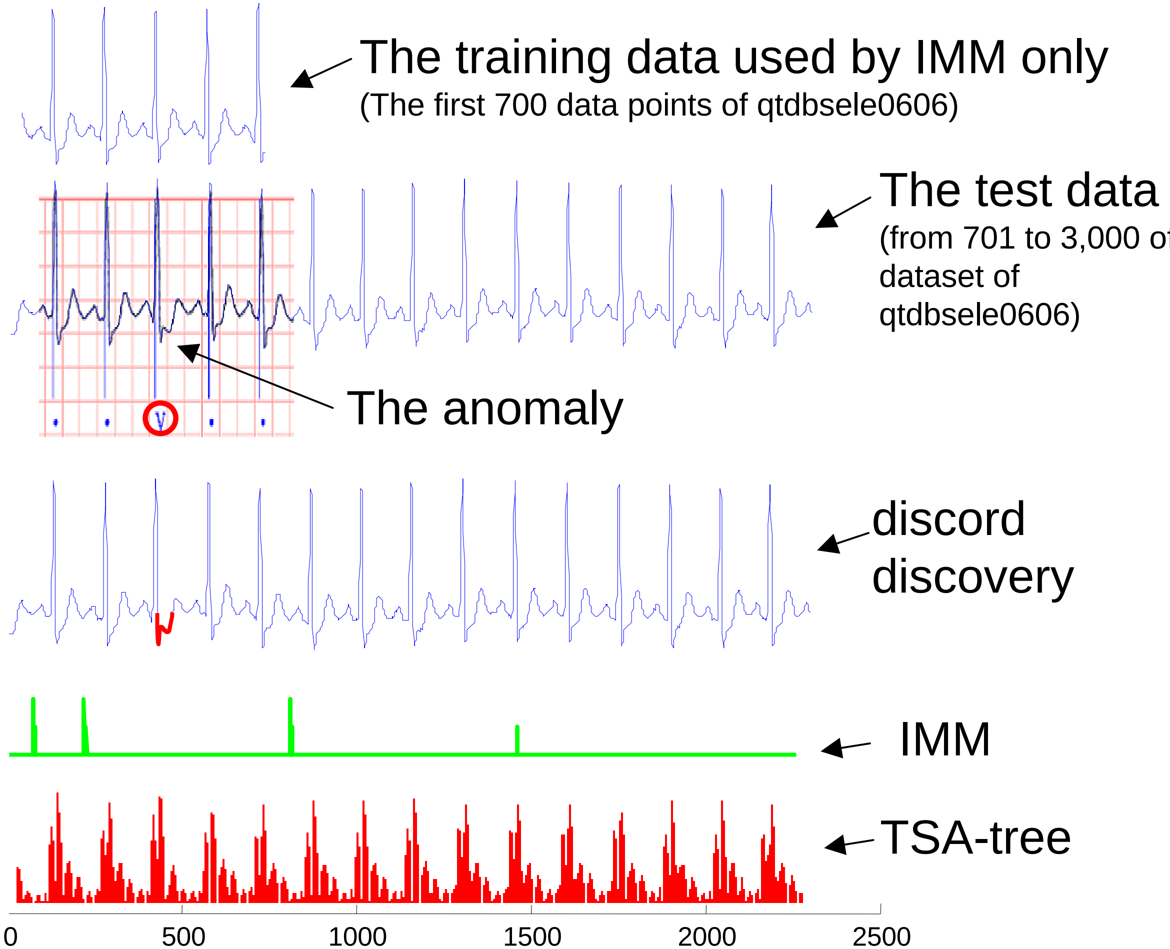


In this experiment, we can say that all the algorithms find the anomaly. The IMM approach has a slightly higher peak value just after the anomaly, but that may simply reflect the slight discretization of the time axis.

In the next slide, we consider more of the time series…

We can see here that the IMM approach has many false positives, in spite of very careful parameter tuning. It simply cannot handle complex datasets.
Both the other algorithms do well here.
Note that this problem is in Figure 11 in paper

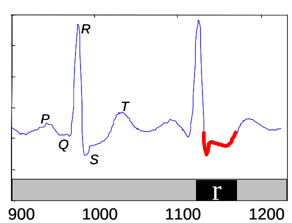


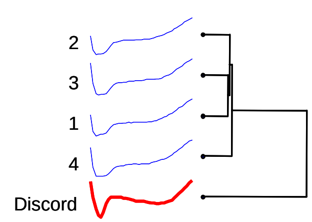


This example is Figure 12/13 in the paper.

Recall that we discussed this example above, it is interesting because the anomaly is extremely subtle.

Here only the discord discovery algorithm can find the anomaly.

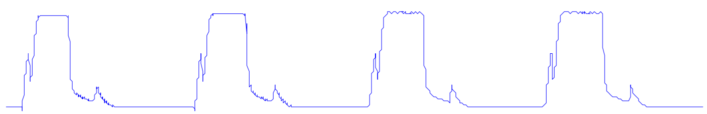




How was the discord able to find this very subtle Premature ventricular contraction? Note that in the normal heartbeats, the ST wave increases monotonically, it is only in the Premature ventricular contractions that there is an inflection.NB, this is not necessary true for all ECGS

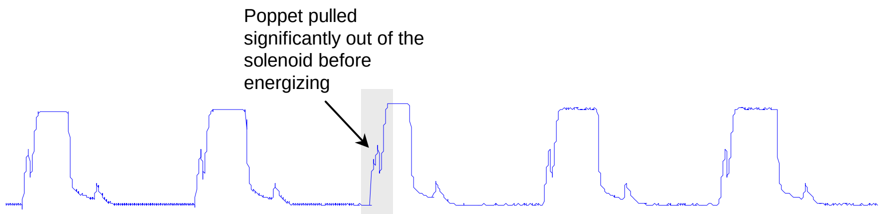
The training data used by IMM only

4 normal cycles of Space Shuttle Marotta Valve Series

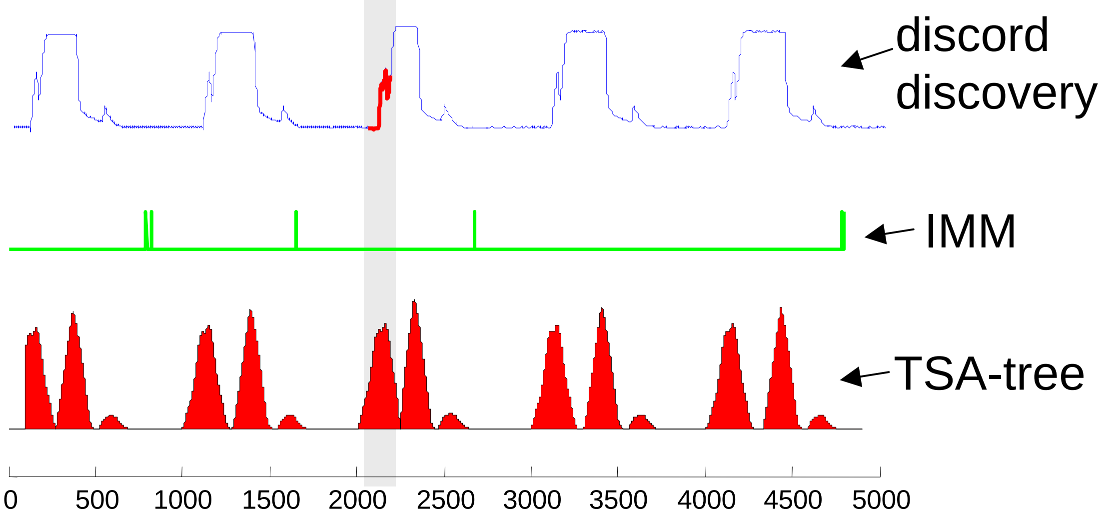


The test data

TEK17.txt



Space Shuttle Marotta Valve Series

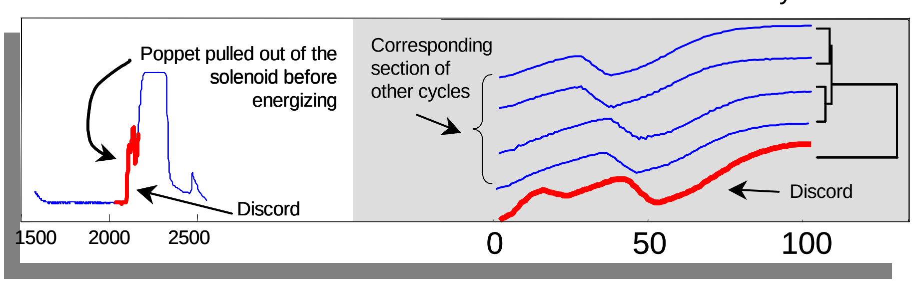


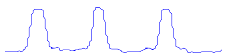
This example is Figure 7/8 in the paper.

Here the anomaly very subtle.

Only the discord discovery algorithm can find the anomaly.

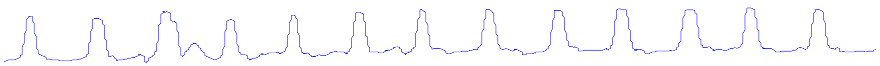
A reminder of the cause of the anomaly





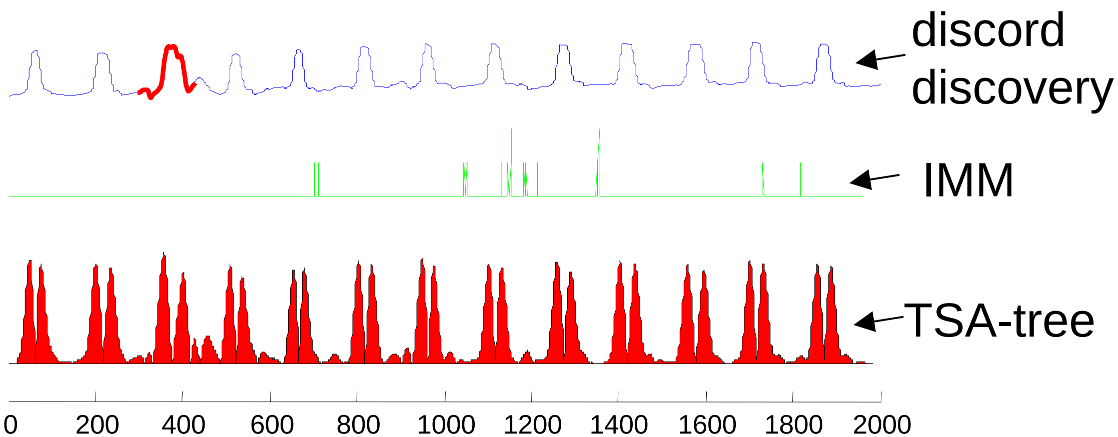
The training data used by IMM only

3 normal cycles of gun draw.



The test data

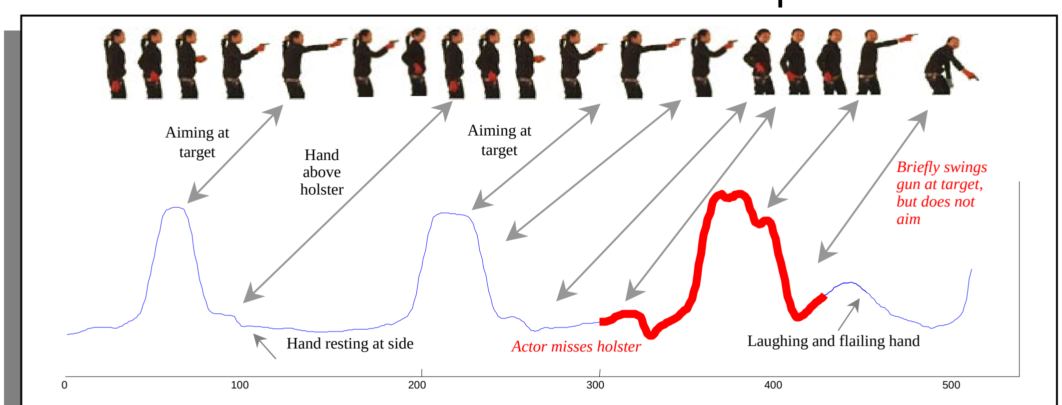
the first 13 cycles of ann_gun_CentroidA



This example is not in the paper, the cause of the “anomaly” is explained above

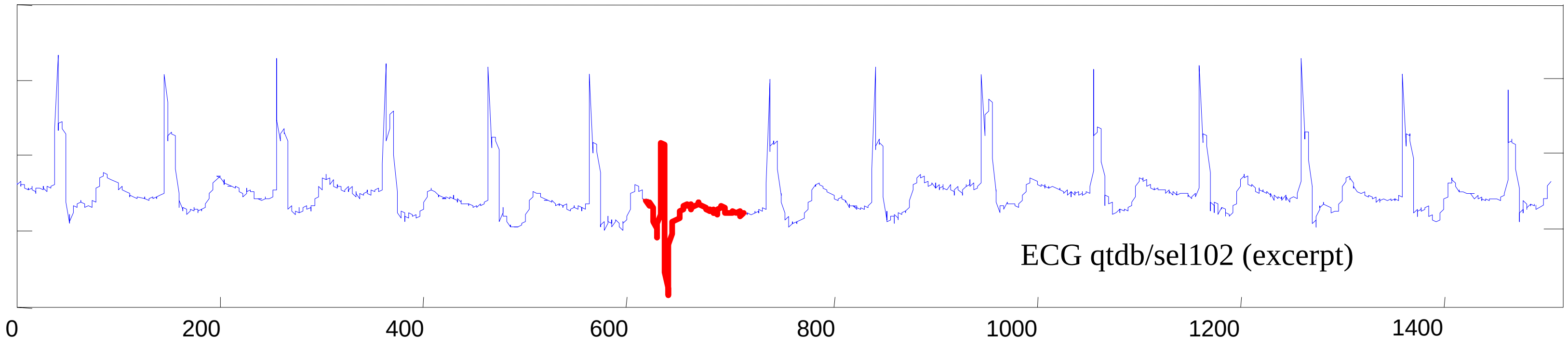
Only the discord discovery algorithm can find the anomaly.

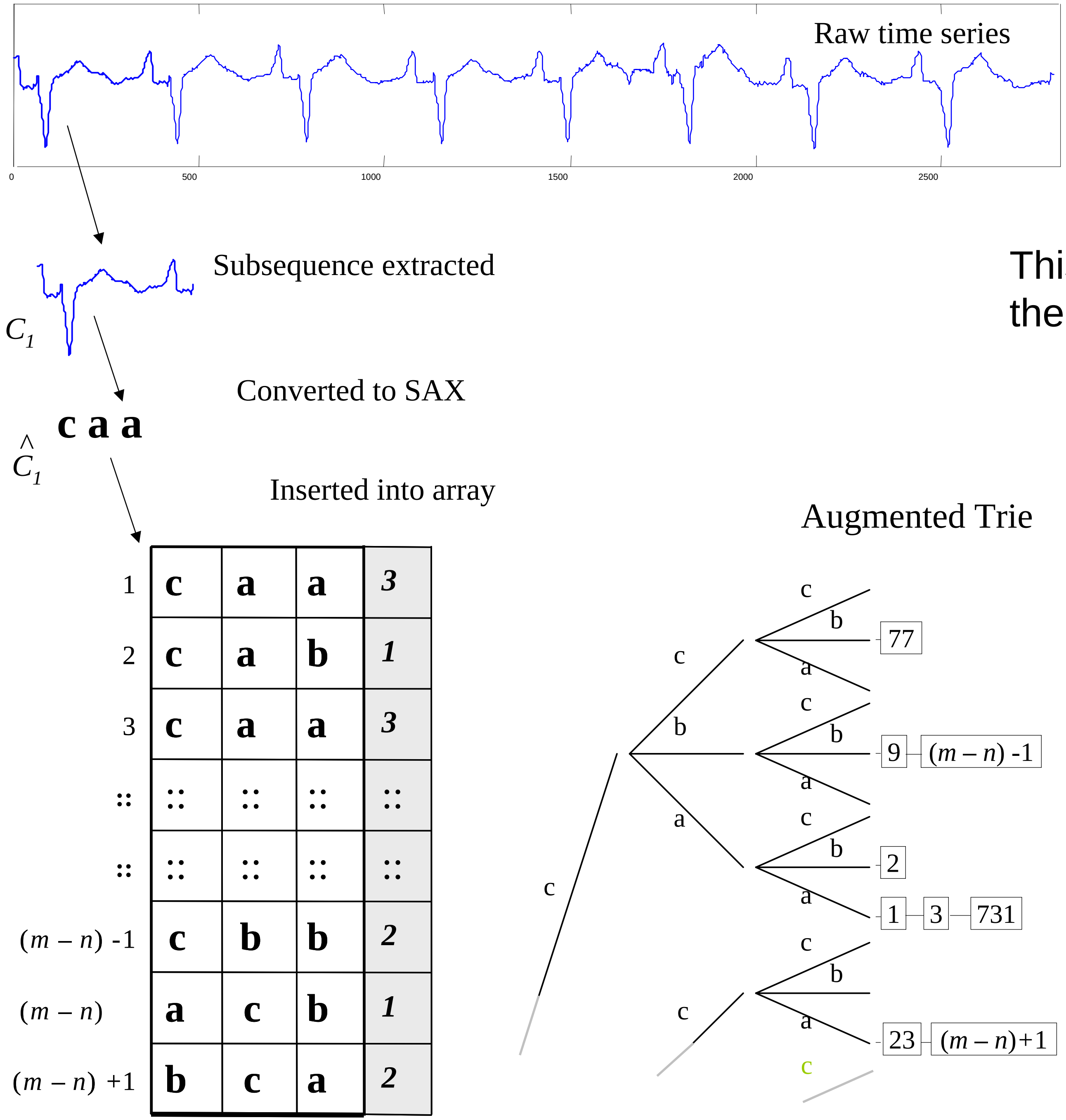
Recall that we have seen this example before…



Below are some larger versions of figures from the paper, for reference.

This is figure 1 from the paper





This is figure 4 from the paper