POSTER P5311 | CONTACT NAME Christopher Painter: Chris.Painter@zigzag.co.uk

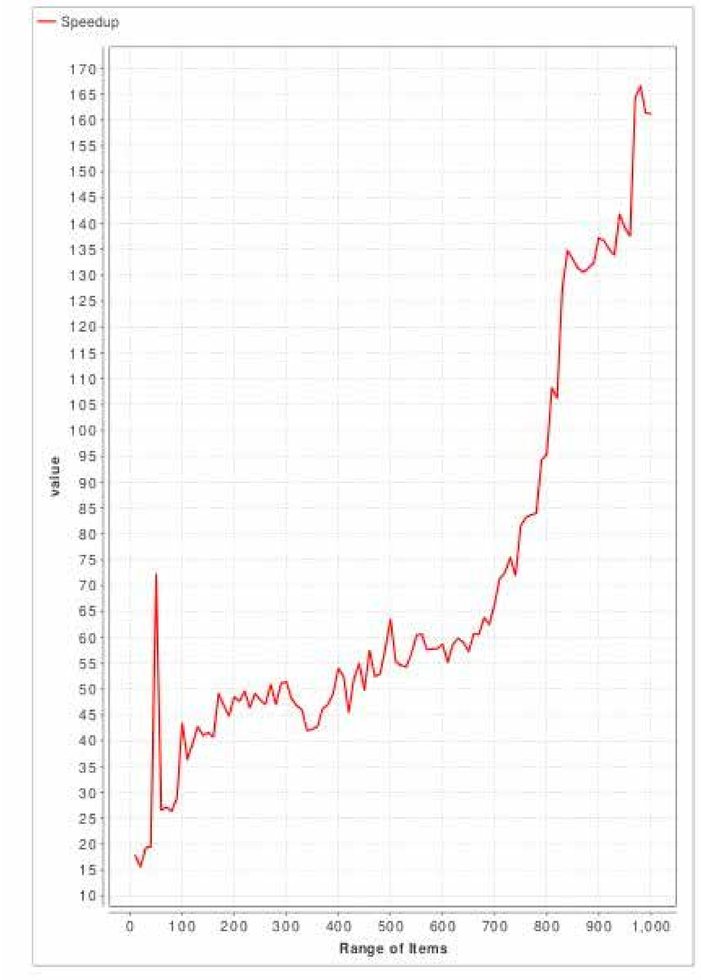
Counting Coincidences Makes Sense.

How to be objective

The Associator generates a graded list of Association Rules from binominal data, using a parallel implementation of the A Priori algorithm to generate Frequent Itemsets, and Iterative Proportional Fitting to standardise the contingency tables that the Itemsets contain, and from which rules are extracted.

This standardisation results in all measures of interest ranking the contingency tables in the same order. The implementation extends the work that appears in Tan, Pang-Ning; Kumar, Vipin; and Srivastava, Jaideep; "*Selecting the right objective measure for association analysis.*"

How to be fast enough



Frequent Itemset mining: early GPU based implementations using CUDA+Thrust accelerated a RapidMiner application by more than 100 times using modest onboard graphics cards.

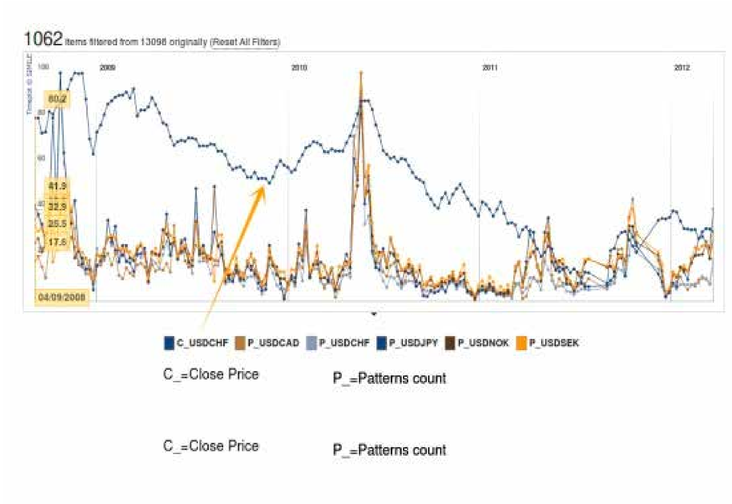
Crucially, the acceleration is proportional to the dimensions of the problem.

Use case - Intermarket Dynamics

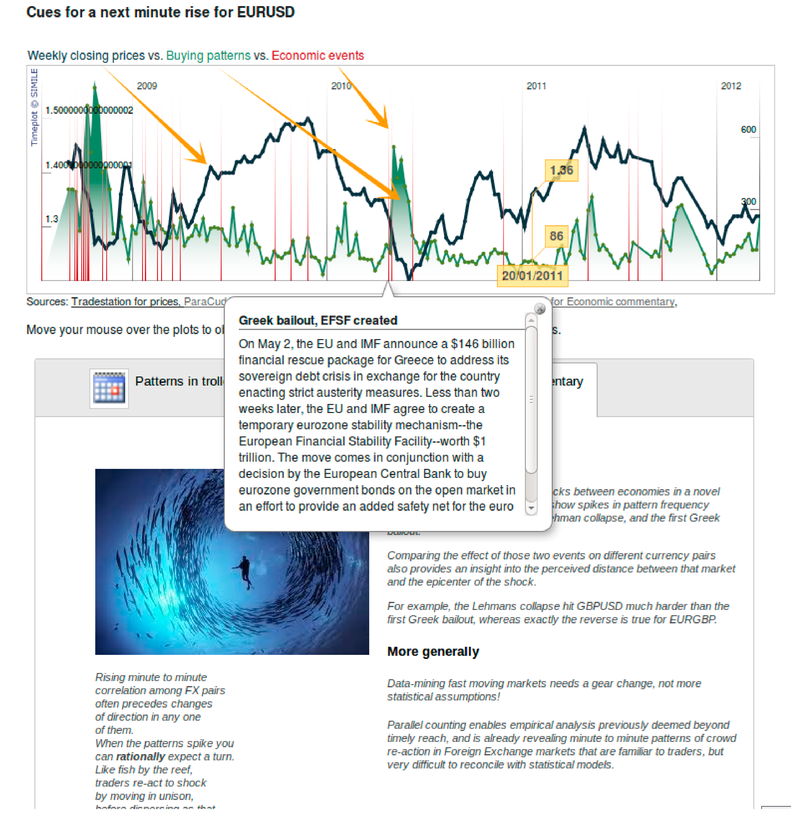
A comprehensive survey of the billions of combinations of 37 major currency pairs, minute to minute, over the post crash years reveals clear spikes in identifiable patterns.

The analysis matches rises for a particular currency pair in one minute against rises in **any** pairs in the **previous** minute; any match that recurs sufficiently frequently is then counted as a pattern.

Example: Pattern spikes that signal change in the outlook for the Dollar against the Swiss Franc



Example: pattern spikes that signal change on EUR-USD, and tie in with real-world events.



Use case – Key Theme Identification

Putting key words into a search engine will return a collection of relevant documents.

But what if you already have a collection of documents, and need to identify the key phrases that they contain?

The News Digest at meme-machines.com looks for exactly such keyword patterns 24/7

Example: Hundreds of news stories, four big themes, fourteen words.

Association Rules can be combined and accumulated into usable rulebases. So the meme-machines.com News Desk also records the 'ifs' and 'thens' for later research.

Example: What words connect to 'Ebola', and how?

```
?- lookup('Ebola',If-Then).

If = ['Africa', 'Chimerix', 'City', 'Contact', 'Doctor', 'Health', 'Hospital', 'Liberia',
'Nurse', 'Outbreak', 'Patient', 'Quarantine', 'US', 'Virus', 'West', 'York'],
Then = ['Africa', 'City', 'Leone', 'Nurse', 'Quarantine', 'Sierra', 'West', 'York'].
```