

Chapter 2

Market Analysis Background and Related Work

Abstract In this chapter some fundamental concepts, necessary to understand the developed work, are addressed, particularly the domain relative to financial markets and time series analysis. Furthermore several methodologies applied to market investment and especially to pattern detection are presented. Finally an introduction to the SAX representation method will be presented and previous works using this methodology will be discussed.

Keywords Technical Analysis • Fundamental Analysis • Perceptually Important Points (PIP) • Symbolic Aggregate Approximation (SAX) • Pattern Recognition • Pattern Discovery

2.1 Market Analysis

In order to understand the objective of this work, it is necessary to have some basic knowledge of market analysis. Basically, there are two market analysis techniques, the *Fundamental Analysis* and the *Technical Analysis*, these are presented next. For the present work and almost for all automatic investment algorithms the preferred analysis method is the technical, since it is based on some measurable numeric indicators, easily calculated from stock market time series. Another reason for using of this type analysis is the fact that data is easily obtained from the internet, most of the time for free, and the large of number of data available to train and test the algorithms.

2.1.1 Fundamental Analysis

The fundamental analysis (FA) is based on several economic and financial indicators, which attempt to evaluate the intrinsic value of a company. FA studies everything that can affect the company value, including macroeconomic and business specific factors. So, the fundamental analysts will try to determine the future price of a company and based on the current value takes the investment decision.

Some of the most important economic indicators are released by central banks or other institutions, and include indicators like:

- Unemployment
- Consumer Price Index (CPI)
- Consumer Confidence Index
- Gross Domestic Product (GDP)
- New Home Sales

Other indicators could be found in [1], where an important and detail list of indicators are presented.

The fundamental analysts also use indicators that are related to company analysis. These are gathered from the financial reports issued by companies. Some of the most important ones are presented in Table 2.1.

Finally, industry reports are the source to another important set of indicators. With these industry indicators the analysts try to evaluate how healthy is some particular sector, and how the companies are positioned inside it. One important factor to every investor is how to minimize the risk; one way to achieve this goal is to investment on different industries to avoid possible down cycles of some sectors. Some important indicators to look out are:

- Industry Growth
- Competition
- Costumers
- Suppliers

Since each industry has some business specificity is more difficult to make a detail list of indicators.

2.1.2 Technical Analysis

Technical analysis [2] is based on the stock prices and volumes movements, the technical analyst believe that changes on price and volume already incorporates all the fundamentals factors. In this technique, the stock price and volume is all that matters to describe the market condition and to try to predict future market movements. Based on this premises, the analyst builds a set of indicators that allows him to study, in an easier and deeper way, the stock movements in order to profit from future trends.

Table 2.1 Several important company fundamental indicators

Indicator	Measure	Description
Earning per share (EPS)	$\frac{\text{Net Income} - \text{Dividends on preferred Stock}}{\text{Average Outstanding Shares}}$	The portion of a company's profit allocated to each outstanding share of common stock
Price earning ratio (PER)	$\frac{\text{Market Value per Share}}{\text{EPS}}$	A valuation ratio of a company's current share price compared to its per-share earnings
Price cash flow (PCF)	$\frac{\text{Cash Flow per Share}}{\text{Share Price}}$	A measure of the market's expectations of a firm's future financial health
Price sales ratio (PSR)	$\frac{\text{Share Price}}{\text{Revenue per Share}}$	A ratio for valuing a stock relative to its own past performance
Payout Ratio (POR)	$\frac{\text{Dividends per Share}}{\text{EPS}}$	The amount of earnings paid out in dividends to shareholders
Dividend yield (DY)	$\frac{\text{Annual Dividends per Share}}{\text{Price per Share}}$	A financial ratio that shows how much a company pays out in dividends each year relative to its share price
Price book value (PBV)	$\frac{\text{Total Assets} - \text{Intangible Assets} - \text{Liabilities}}{\text{Stock Price}}$	A ratio used to compare a stock's market value to its book value
Return on equity (ROE)	$\frac{\text{Net Income}}{\text{Shareholder's Equity}}$	The amount of net income returned as a percentage of shareholders equity
Return on assets (ROA)	$\frac{\text{Net Income}}{\text{Total Assets}}$	An indicator of how profitable a company is relative to its total assets
Debt equity ratio (DER)	$\frac{\text{Total Liabilities}}{\text{Shareholders Equity}}$	Measure of financial leverage
Quick ratio (QR)	$\frac{\text{Current Assets} - \text{Inventories}}{\text{Current Liabilities}}$	Short-term liquidity, ability to meet short-term obligations
Market capitalization (MC)	$\text{Company's Shares} \times \text{Market Price}$	The total market value of all of a company's outstanding shares

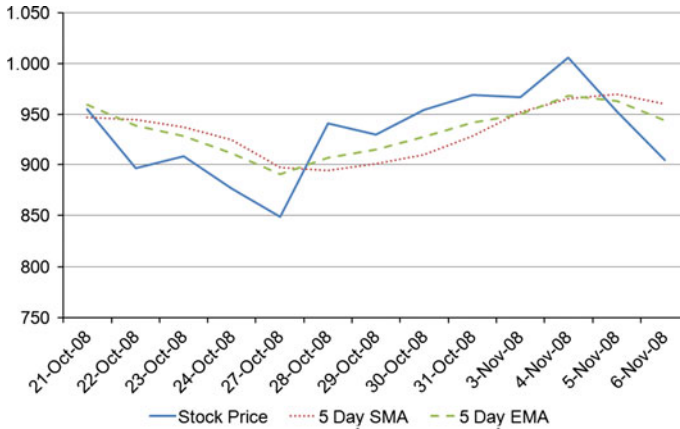


Fig. 2.1 S&P 500 daily chart with a SMA and an EMA

In addition to the technical indicators, that are calculated using the price and transaction volumes, the technical analyst studies pattern formations [3, 4] in the stock price and sometimes in some indicators.

Next, some technical indicators and relevant known pattern formations are presented.

2.1.2.1 Technical Indicators

A technical indicator is a metrics whose value is calculated from the price/volume of an asset. The objective is that the indicator value helps predicting future price, or simply indicates a general price trend. Some popular technical indicators are presented next, for more information on other indicators see [5].

- **Moving Averages**

This is one of oldest indicator used and is calculated by finding the mean value of the price over a certain amount of time. Two type of moving average are applied, the first one is known as Simple Moving Average (SMA or just MA) and is calculated by averaging the price of the last days, or any other time measure. The second average is the Exponential Moving Average (EMA) and in this case the price for recent days has a higher weight on the average.

In Fig. 2.1 the S&P500 index is presented, in this figure two averages are shown, the SMA for 5 days period and the EMA for the same period. It is clear the difference between them, the EMA follows better the fast changes of price indicating that the present time has more weight.

This indicator could be used to signal the investor to buy when the moving average is rising or crosses down the price line and sell when is descendant or crosses up the price line. Sometimes the moving averages are used together to

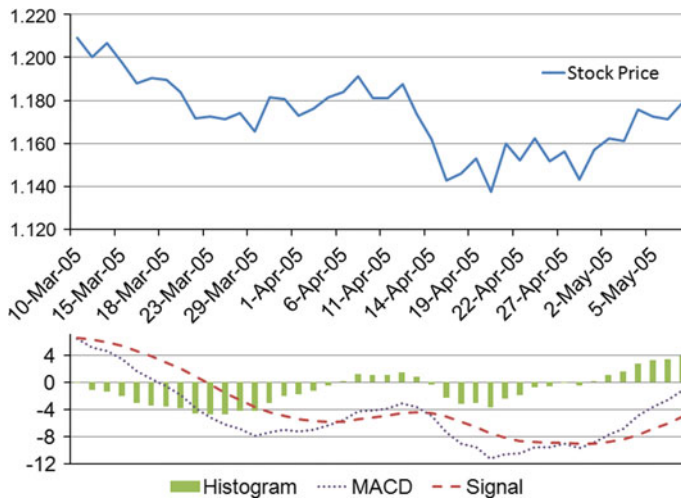


Fig. 2.2 S&P 500 daily chart and MACD indicator, histogram and signal line

generate these signals, the intersection between different moving average is used to generate buy or sell orders.

- **MACD (Moving Average Convergence Divergence)**

One example of the combined used of averages is this indicator, which is calculated from the difference of two EMA, usually the 12 days EMA and the 26 days. An additional line is added to this indicator, is known as the Signal line and is the 9 days moving average of the MACD itself. Added to this indicator usually appears a histogram, which is the difference between the MACD and the Signal, Fig. 2.2. Sometimes the MACD line does not appear, instead just the signal line and the histogram are presented, this is because usually the cross point of the two lines is the point to look for and that point could be identified by the zero crossing of the histogram.

The application rules to the MACD indicator are discussed next:

- The crossing between MACD and the Signal when the MACD is rising is indication to buy in a down movement is to sell.
- The zero cross by the MACD is signal to buy, since when MACD is above zero the market tends to be bullish.
- Positive or negative divergence is also a signal. If the price is on uptrend and the MACD does not, then a negative divergence is present and is time to sell. If the price in downtrend and the MACD begins an uptrend then is a buy signal.

- **RSI (Relative Strength Index)**

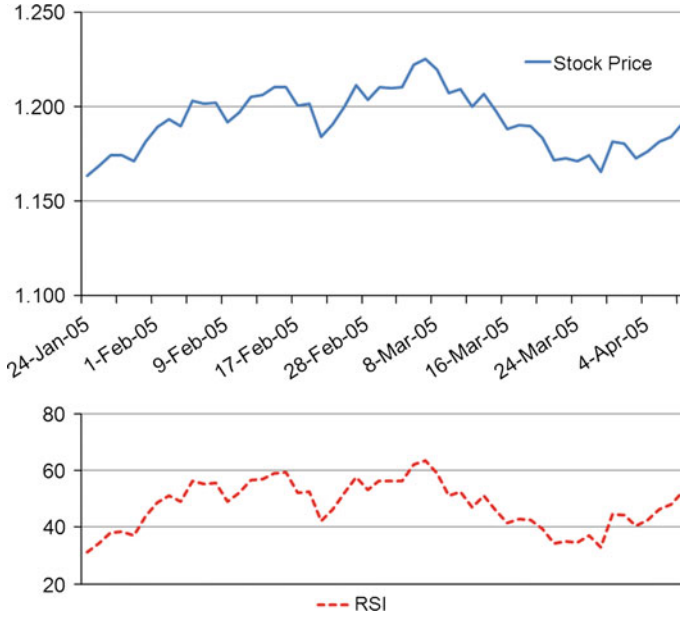


Fig. 2.3 RSI indicator for the S&P 500 daily chart

This indicator is one of the most popular in the momentum family, which compares the magnitude of recent gains to recent losses to identify overbought and oversold conditions of an asset. This indicator representation oscillates between 0 and 100 and usually is calculated for 14 days period. In Eq. (2.1) the formula for calculating this indicator is presented (Fig. 2.3).

$$RSI = 100 - \frac{100}{1 + \frac{\bar{G}}{\bar{L}}} \quad (2.1)$$

where,

\bar{G} Average gain of the x last periods of time;

\bar{L} Average loss of the x last periods of time.

In the use of this indicator it is also possible to detect graphical pattern formations, or by applying the following rules:

- Crossing the 50 value in an uptrend is a buy signal, in case of a downtrend indicates a sell signal.
- If the RSI is above 70 is an overbought signal and the asset should be sold. Below 30 the asset is oversell and a buy signal is issued.
- Like in MACD, this indicator presents positive and negative divergence. If the price is in a downtrend and the RSI does not follows then a buy signal must be

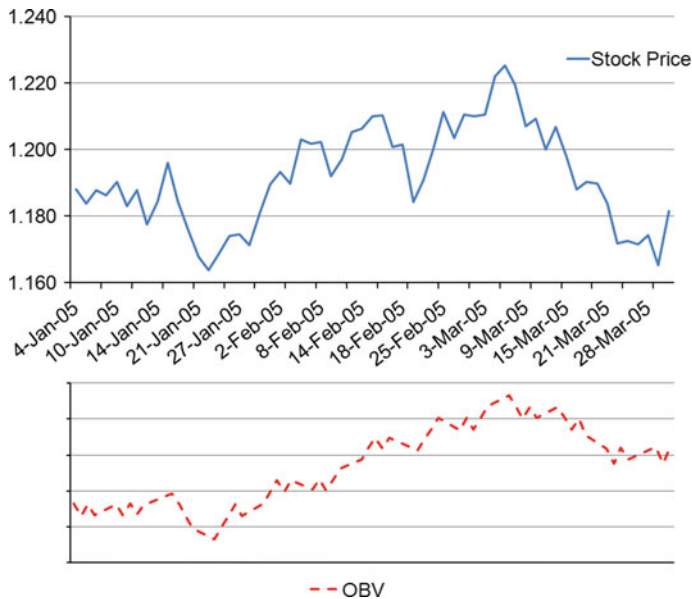


Fig. 2.4 OBV indicator for the EUR/USD Forex pair

issued, this is a positive divergence. If the price is at maximums and the RSI is not, then a negative divergence is present and the asset should be sold.

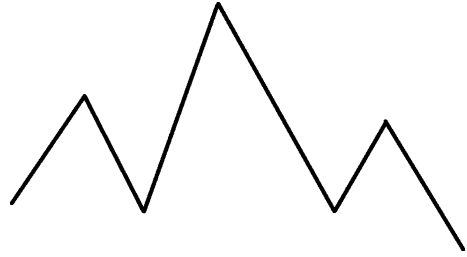
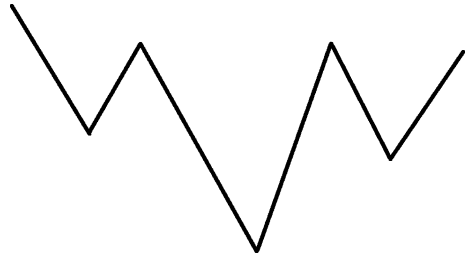
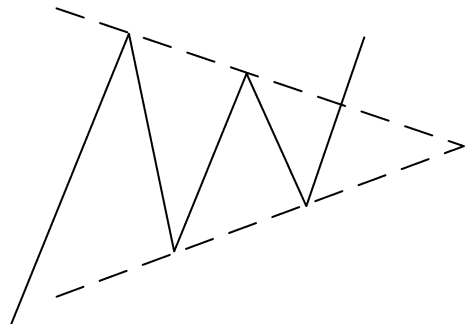
- **OBV (On Balance Volume)**

This indicator, as the name states, uses the volume for its calculation and the price as well. It measures buying and selling pressures. It supports the idea that volume precedes price, the value itself is not important what is important is characteristics of the OBV line. The OBV is calculated by Eq. (2.2) (Fig. 2.4).

$$OBV(t) = \begin{cases} OBV(t-1) + Volume(t), & \text{if } Price_t > Price_{t-1} \\ OBV(t-1) - Volume(t), & \text{if } Price_t < Price_{t-1} \\ OBV(t-1), & \text{Otherwise} \end{cases} \quad (2.2)$$

Several signals could be perceived from this indicator:

- A bullish market is in formation when OBV moves up or forms a higher low, even if the prices move down or forms a lower low, in this case a buying signal is issued. A bearish market forms when OBV moves down or forms a lower low, even when the prices move up or a higher high is created, a sell signal is generated for this case.
- In case the OBV slope is positive and the price is in an uptrend this is a confirmation of the trend. The same applies for the opposite case.

Fig. 2.5 Head and shoulders**Fig. 2.6** Inverse head and shoulders**Fig. 2.7** Symmetrical triangle in a uptrend

2.1.2.2 Chart Patterns

When looking at a graphic representation of a financial time series, it is possible to identify some similar graphic formation along the time. These formations are caused by the repeated actions of the investors when presented to similar market conditions. So, by looking for these similar graphic formations or chart patterns, it will be possible with some degree of confidence to deduce what will happen next. This is the base idea behind the technical analysts that hunt for chart patterns and are known as Chartists.

Next, a set of important chart patterns is presented, the associated investment strategy is easily guessed by the outcome of the patterns. The presented patterns are shown due to its relevance and in order to possibly identify some of the patterns found by the SAX-GA approach. Many more patterns exist, for more information on this graphic formations see [3], and for additional trading tips using patterns [4] and more specific to the Foreign Exchange Market (Forex) [6].

Fig. 2.8 Symmetrical triangle in a downtrend

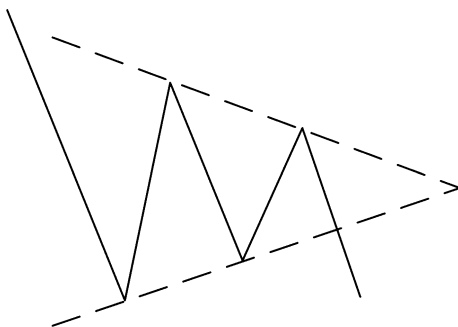


Fig. 2.9 Ascending triangle

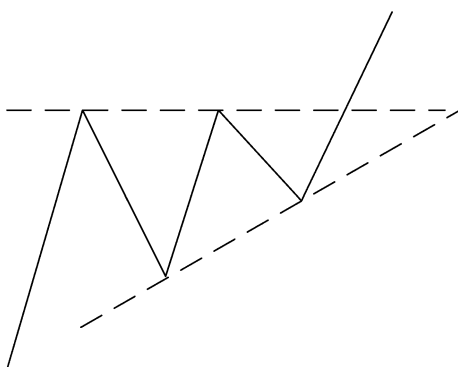
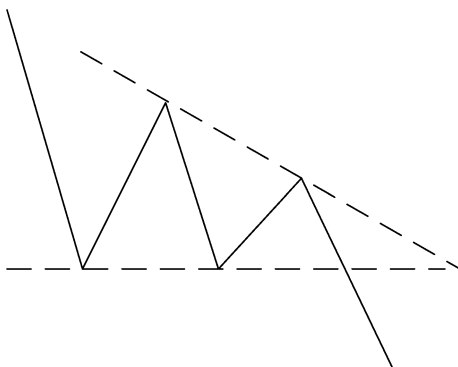


Fig. 2.10 Descending triangle



- Head and Shoulders

This is one of the most famous chart formations. This is a reversal type that indicates a change on trend. The two possible reversal trend formations are presented in Fig. 2.5 and Fig. 2.6.

Fig. 2.11 Falling Wedge in an uptrend

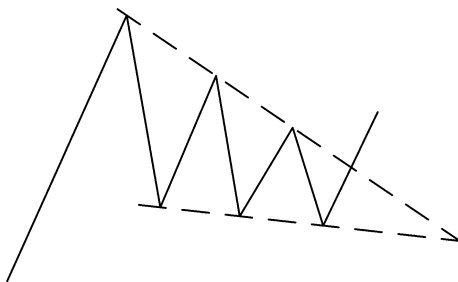


Fig. 2.12 Falling wedge in a downtrend

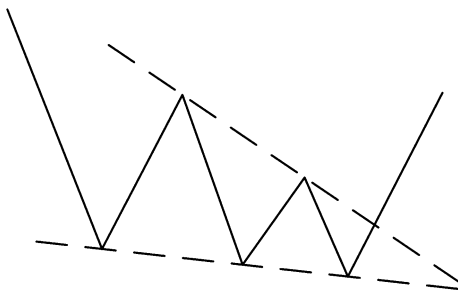


Fig. 2.13 Rising wedge in an uptrend

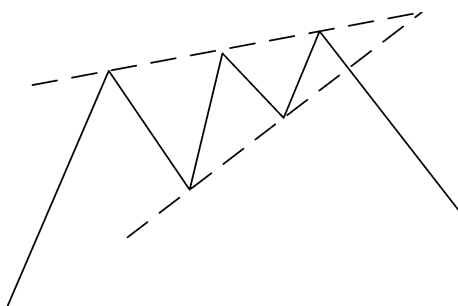
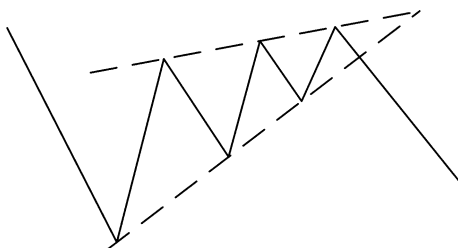


Fig. 2.14 Rising Wedge in a downtrend



- Symmetrical Triangles

This formation is usually defined as neutral, the fact is that some researchers found that most of the time breaks in the previous market direction, could be considered as a trend confirmation. This is illustrated in Fig. 2.7 and Fig. 2.8.

Fig. 2.15 Bull flag in an uptrend

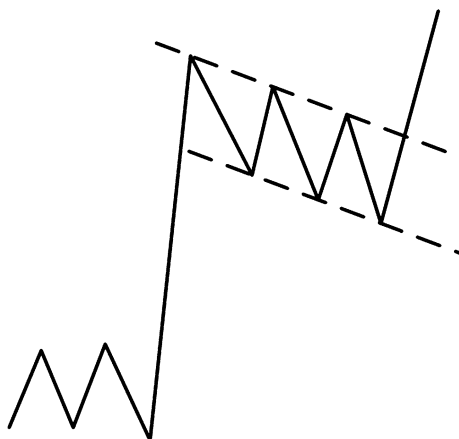
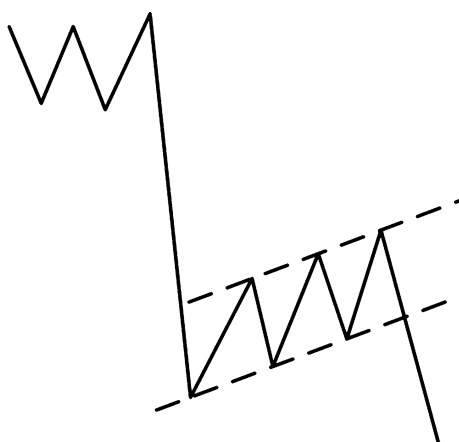


Fig. 2.16 Bear flag in a downtrend



- Ascending Triangle/Descending Triangle

Ascending triangles are considered an uptrend confirmation and naturally are most reliable when found in bullish market conditions, Fig. 2.9.

The descending formation, Fig. 2.10, is the reverse of the previous one, is also a trend confirmation, and naturally is more reliable in a bearish market.

- Falling Wedges

These formations, Figs. 2.11 and 2.12, are quite similar to the triangles, but they are associated to a bullish movement, breaking in an uptrend.

- Rising Wedges

Rising Wedges are the inverse of the previous chart formation and are associated to bearish movements, Fig. 2.13, and appear more frequently in a downtrend market, Fig. 2.14.

Fig. 2.17 Pennant in an uptrend

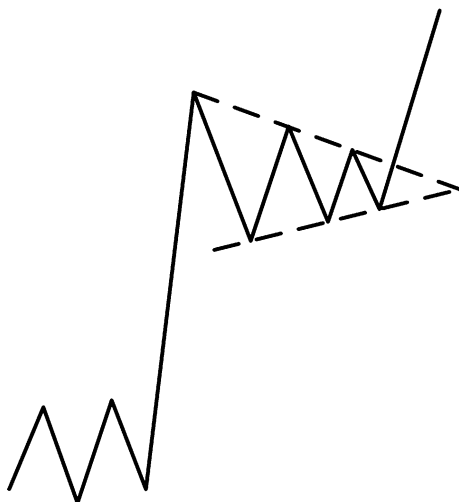


Fig. 2.18 Pennant in a downtrend

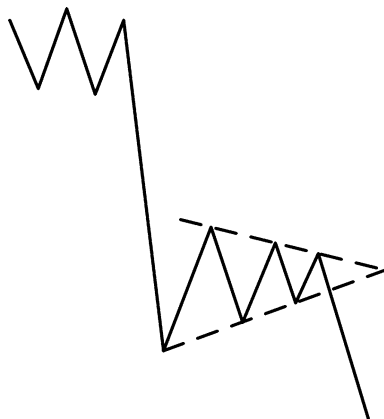
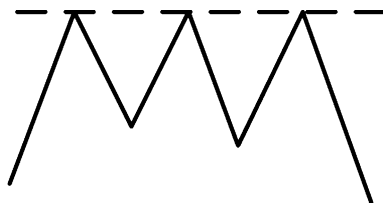
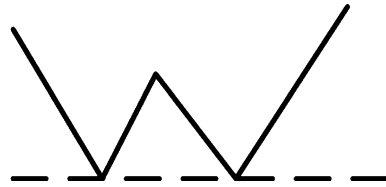


Fig. 2.19 Triple top



- Flags

These charts formations are generally seen as trend confirmations, usually appear after big movements on the graphic, as in Fig. 2.15 and 2.16

Fig. 2.20 Double bottom

- Pennants

These are similar to Symmetrical Triangles, but usually are smaller in amplitude and shorter in time, Figs. 2.17 and 2.18. Like the Flags, the Pennants are generally trend confirmations and represent small pauses in the current trend.

- Tops and Bottoms

These formations are usually trend reversals, they represent a price resistance or a support level, Figs. 2.19 and 2.20 respectively.

2.2 Existing Solutions

In this section several state-of-the-art time series representation and dimensional reduction methods, applied to the financial sector, are presented. Also, their application and use on chart pattern detection methods will be discussed, and at last the SAX method will be briefly presented.

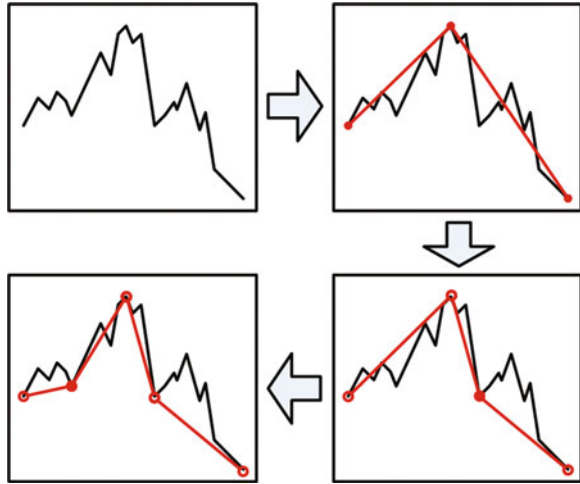
2.2.1 Pattern Detection

From the beginning of stock market history, investors tried to predict market movements. The analysis has always been difficult, and one of the harder tasks is to analyze financial charts, in order to detect patterns on those graphics. The appearance of computers and their large calculus capabilities offers new possibilities of prediction.

First of all a distinction between pattern recognition and pattern discovery should be made. Recognition is identifying some patterns that are known on the time series, this case is a supervised approach, where a library of patterns [3], is created and is made a search on the data market trying to identify them [7]. In pattern discovery, the quest is to find patterns that occur in the time series and that are unknown, in this case typically some data segments or windows are compared with others and this case is associated to unsupervised approach, the case presented on this book.

The methods that directly compete with the new SAX-GA approach, are techniques that try to detect chart patterns from the financial time series. Basically, are alternative representations of data series and then apply some decision or

Fig. 2.21 PIP identification process



classifier method to identify the patterns. In this area two methods were identified, the first one uses a Perceptually Important Points (PIP) representation and the second uses a matrix representation of data. The matrix data representation and PIP when are used in a pattern detection methods usually are applied in a template based pattern recognition strategy. Where a set of known patterns, are converted to the same representation structure and a matching process, between template and time series, is made. Generally, this matching process is based on the distance measure between the data representation and the template, and if the distance is lower than some threshold, it is considered that the pattern is present in the evaluated time series. A detailed survey comparing several time series data mining techniques is presented in [8].

2.2.1.1 Perceptually Important Points Pattern Recognition

As the name of this alternative data representation and dimensional reduction indicates, this method reduces the time series to a set of points that are considered important. The decision to identify the important points is quite similar to the one a person makes when looking at a graphic, where the points that more stand out from each other are the ones that attract the eye attention. This pattern recognition method was first present by Chung et al. [9] and it first use was in financial applications.

The process to identify PIP's in a time series is quite simple and could be seen in Fig. 2.21. It starts by defining that the first two important points are the first and the last of the time series, then draws a line between those points and calculates for the rest of the points in the time series which is further apart from the line, the more distant point will now be an important point. After, draws other line between the first point and this new one, and from the new to the last point, and for each of

Fig. 2.22 PIP evaluation with Euclidian distance (PIP-ED)

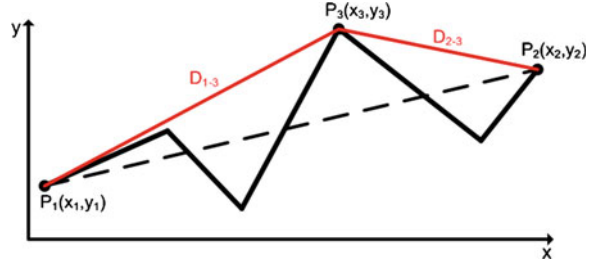
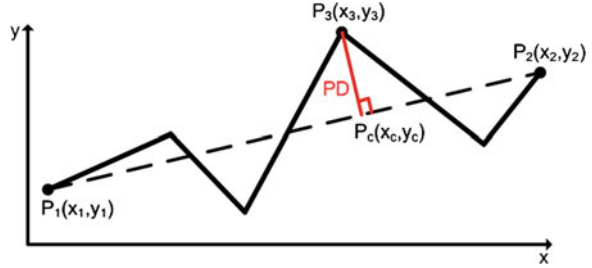


Fig. 2.23 PIP evaluation using perpendicular distance (PIP-PD)



the segments, again finds the point further apart, and in a recursive way it will detect the important points [10]. If this process is taken to the limit then all points from the time series will be important points. So, when to stop? Well, there is no straight answer to this question since is data dependable, if the time series is very irregular with lots of oscillations then more PIP's have to be found to preserve the data structure. Usually is defined a desired compression ratio, Eq. (2.3), or an acceptable level of error between time series and the PIP representation.

$$C_r = \frac{\text{Number of points in time series}}{\text{Number of PIP's to represent time series}} \quad (2.3)$$

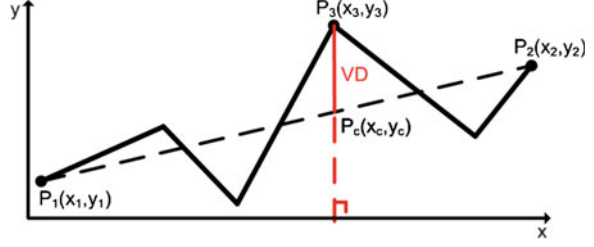
In the method description, was said that the distant point from the line will be considered an important one. So it is important to define how to measure the distance, in [10] three types of measures were considered and are presented next:

- Euclidian distance (ED)—Calculates the sum of distances, Eq. (2.4), between the test point p_3 and the adjacent PIP's p_1 and p_2 —Fig. 2.22

$$ED(p_1, p_2, p_3) = \sqrt{(x_2 - x_3)^2 + (y_2 - y_3)^2} + \sqrt{(x_1 - x_3)^2 + (y_1 - y_3)^2} \quad (2.4)$$

- Perpendicular distance (PD)—Calculates the PD, Eqs. (2.5)–(2.8) between the test point p_3 and the line that connects the two adjacent important points p_1 and p_2 —Fig. 2.23

Fig. 2.24 PIP evaluation using vertical distance (PIP-VD)



$$s = \text{Slope}(p_1, p_2) = \frac{y_2 - y_1}{x_2 - x_1} \quad (2.5)$$

$$x_c = \frac{x_3 + sy_3 + sy_2 - s^2x_2}{1 + s^2} \quad (2.6)$$

$$y_c = sx_c - sx_2 + y_2 \quad (2.7)$$

$$PD(p_3, p_c) = \sqrt{(x_c - x_3)^2 + (y_c - y_3)^2} \quad (2.8)$$

- Vertical distance (VD)—Calculates the VD, Eqs. (2.9)–(2.10), between the test point p_3 and the line that connects the two adjacent important points p_1 and p_2 —Fig. 2.24

$$y_c = y_1 + (y_2 - y_1) \frac{x_3 - x_1}{x_2 - x_1} \quad (2.9)$$

$$VD(p_3, p_c) = |y_c - y_3| \quad (2.10)$$

From tests made using the Hang Seng Index (HSI), the distance method that proves best results is the vertical distance, which was able to capture the essence of the HSI graphic [10].

Now, to complete the pattern recognition method description is necessary to compare the converted time series with the templates, for instance Fig. 2.25.

So, if P important point's exists in the sequence, converted from a time series, and a query template sequence Q is defined, is possible to calculate the distance point-to-point from each other, Eq. (2.11).

$$AD(P, Q) = \sqrt{\frac{1}{n} \sum_{k=1}^n (p_k - q_k)^2} \quad (2.11)$$

where,

p and q Points of the sequence;

n Number of points in the sequence.

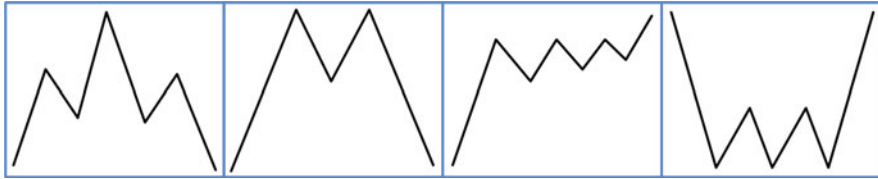
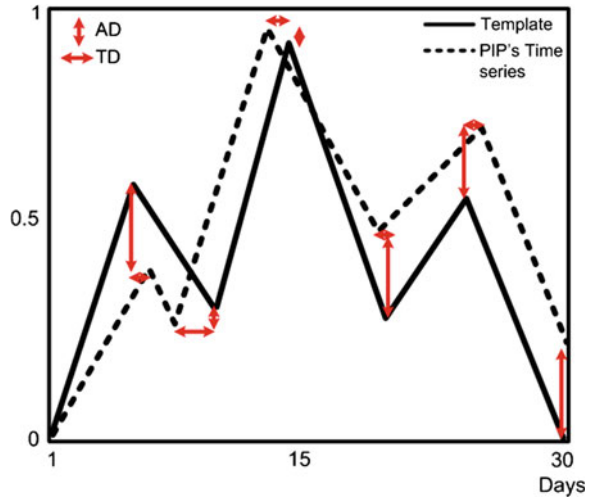


Fig. 2.25 Four PIP technical analysis patterns templates

Fig. 2.26 Matching between template and time series



This last distance measure, Eq. (2.11), makes possible to compare the sequences according to the amplitude and in this manner identify the amplitude similitude of the sequences, but it is also necessary to detect the similarity in the horizontal temporal axes, it is also important to consider time distortions between time series and template, so it is necessary to calculate a temporal distance Eq. (2.12).

$$TD(P, Q) = \sqrt{\frac{1}{n-1} \sum_{k=2}^n (p_k - q_k)^2} \quad (2.12)$$

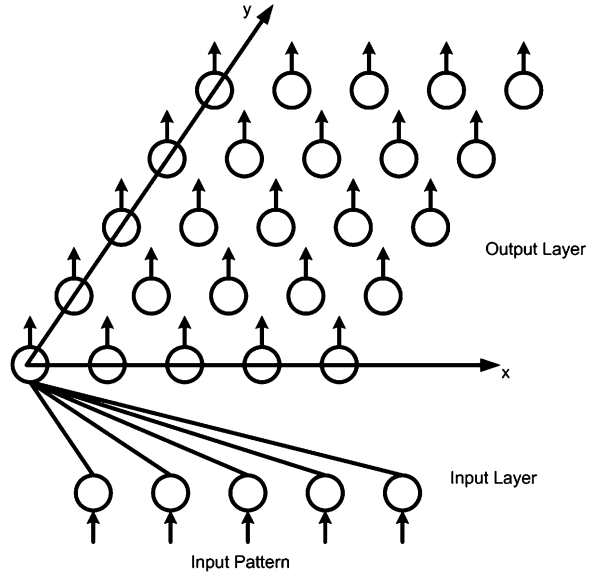
In this measure, Eq. (2.12), the sum begins at 2 because the first points of the two sequences are aligned, so the distance is zero. The p and q are the time coordinates of the sequences.

Now that a set of measures that provides a way to compare the time series and the template have been defined, Fig. 2.26, it is possible to combine the two distances in one expression to measure the similarity between the two Eq. (2.13), and to be able to decide if the pattern is present in the tested data series or not.

$$DM(P, Q) = wAD(P, Q) + (1 + w)TD(P, Q) \quad (2.13)$$

where,

Fig. 2.27 Self-organizing map architecture



w Weight factor that allows to specify which measure is more important, according to tests made in [11], 0.4 is a reasonable value for this factor.

The time series representation using PIP's, has the advantage of preserving some of the important features of financial time series. Many of the preserved points are important indicators of trend inversion. One of the problems when using this method to recognize patterns, is that it is necessary to convert the time series to the same number of PIP's present in the template, in order to evaluate the degree of similitude between time series and template. So, if the application has to find a head-and-shoulders pattern, it is necessary to convert the time series to a seven important points representation. After, if template changes to a different pattern with a different number of important points, it is necessary to reprocess the time series to match the same number of points. The authors tried to solve this problem by loading the PIP's to a binary tree like structure, where the most important points are stored in the higher levels of the tree and the less important are near the leaves [10]. This way it is easy to prune several branches of the tree to have the needed number of PIP's representation.

In [12] the authors also feel the need of pattern discovery, of creating a system where the data series itself define the patterns. To solve this problem the authors used a Self-Organizing Map (SOM) which is a neural network based unsupervised learning algorithm that allows similar time series data windows to be clustered together. SOM is based on competitive learning where the only one of the output nodes is activated. The winner node will then see their weights adjusted this will cause that this node becomes specialized on the kind of patterns that cause the activation. The SOM structure used is presented in Fig. 2.27.

Fig. 2.28 Bull flag matrix pattern template

0.5	0	-1	-1	-1	-1	-1	-1	-1	0
1	0.5	0	-0.5	-1	-1	-1	-1	0.5	0
1	1	0.5	0	-0.5	-0.5	-0.5	-0.5	0	0.5
0.5	1	1	0.5	0	-0.5	-0.5	-0.5	0	1
0	0.5	1	1	0.5	0	0	0	0.5	1
0	0	0.5	1	1	0.5	0	0	1	1
-0.5	0	0	0.5	1	1	0.5	0.5	1	1
-0.5	-1	0	0	0.5	1	1	1	1	0
-1	-1	-1	-0.5	0	0.5	1	1	0	-2
-1	-1	-1	-1	-1	0	0.5	0.5	-2	-2.5

The authors identified two major problems; the first was the efficiency of the discovery process, the increase of the number of data points in the pattern lead to an exponential increase of the pattern discovery. The second was the multi-resolution problem, where the patterns can appear with different lengths, causing to reprocess the time series with different SOM architectures. In the paper, the authors solve the problems by using the PIP representation and converting a set of different lengths data window time series, to the same number of PIP's. This will limit the number of points in the patterns and also convert the data to the same representation, allowing to use the same SOM architecture. This approach will cause, that in some cases, the compression of data will be rather large and some important features will be lost, also the training process is long, needing many iterations trough the training set.

From the previous descriptions of this approach, it is clear that the PIP method has lots of potentials, from the representation point of view and pattern recognition, but creating an algorithm to discover new pattern formations using this method, will result on a time consuming approach. The multi-resolution problem can affect the pattern discovery process and increasing the compression of data is not the solution, since will cause the loss of important information. To solve this problem, the algorithm will have to save the several dimensional representations of the time series windows and then search between the windows with equal number of PIP's, causing a large and complex process of pattern search.

2.2.1.2 Matrix Template Pattern Recognition

This approach is based on the works of Leigh et al. [13–15]. As the name implies this method will recognize patterns based on a template pattern approach and the templates are in a matrix format, like Fig. 2.28. Is visible by the matrix values, that

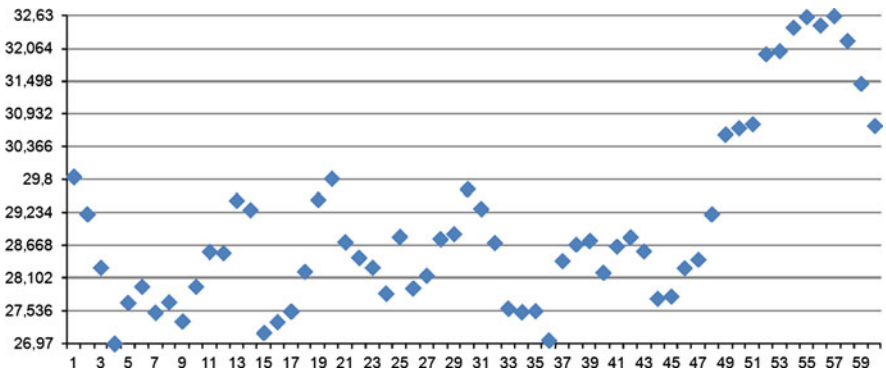


Fig. 2.29 60 points time series

Fig. 2.30 Matrix representation of the time series

0	0	0	0	0	0	0	0	0	0.17	0.67
0	0	0	0	0	0	0	0	0	0.33	0
0	0	0	0	0	0	0	0	0	0	0.17
0	0	0	0	0	0	0	0	0	0.5	0.17
0.17	0	0	0.17	0	0	0	0	0	0	0
0	0	0.33	0.17	0	0.17	0	0	0	0	0
0.17	0	0	0.17	0.5	0.17	0.5	0.17	0	0	0
0.17	0.33	0.17	0.33	0.17	0	0.5	0.5	0	0	0
0.33	0.33	0	0.17	0.17	0.33	0	0.33	0	0	0
0.17	0.33	0.5	0	0	0.33	0	0	0	0	0

the region where the pattern is present is populated by the maximum value 1, other regions of the matrix farther away from the pattern, have negative values. So, the method used to match the time series with the template, is based on the conversion of the time series to an identical size matrix, like the template, and the result of the product between the two matrixes, will indicate the level of similarity between time series and pattern.

The conversion of the time series to a matrix format, consists on dividing the time series or a window of the time series, in a grid of 10×10 , for the present example. To convert the time series like the example in Fig. 2.29, the vertical axes will be divided in 10 levels and on the horizontal temporal axes the 60 points will be divided into 10 groups of 6 points each. Accordingly to the number of points that lay on each cell of the grid a percentage value is calculated, with the constraint that the sum of these values will be 1 for each column, Fig. 2.30.

So, after converting the times series to a matrix format, a match between template and data must be calculated, and depending on the similarity result an investment decision should be made.

The authors have tested an artificial neural networks (ANN) and genetic algorithm (GA) to implement an investment decision model [16, 17]. The ANN will have as inputs the sum of the product between template and price/volume matrix by column, so from these operation the ANN will have 20 inputs, an additional 2 input values will be considered, those values correspond to the window height of price and volume, these parameters correspond to the relation of the difference between the lowest price/volume and the highest price/volume with the price/volume window size. The output of the ANN will be a price forecasting in one of the case studies and in the other, the two outputs contain a confidence factor and based on a threshold mechanism tries to predict the market. The GA was used by one of these approaches [17] to reduce the input number of variables to the ANN by determining a sub set of the 22 inputs that optimize the R2 correlation coefficient between the neural estimated price increase and the actual. The results are shown in Table 2.3 at the final of the chapter.

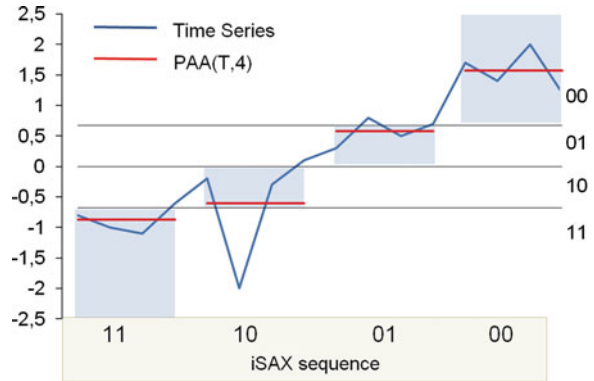
This method has several limitations, when the patterns to be recognized are more complex, for instance a head-and-shoulders, since the level of detail that the matrix has to offer is not enough to represent those kind of complex patterns. It is possible to increase the matrix size, but then the time series window size will also have to increase. This approach is entirely design to work as pattern recognition method with simple patterns. Although, the use of the GA proves to get good results, this algorithm was able to improve the performance of the ANN, by selecting a set of important parameters to feed the ANN classifier. In [7] the GA was also used, with good results, but in this case to select the degree of similitude between template and data in order to create trading rules.

2.2.1.3 Computational Intelligence

This area of computational science and specially when connect to the finance sector had suffer a rapidly expansion. Evolutionary computation (EC) like the GA has been largely applied, like in the examples of the previous section. Other works, instead of dealing with patterns, have used GA to find sets of technical indicators and tune their parameters [18] in order to define profitable trading rules, in this work was also addressed the problem of overfitting, which this kind of algorithms tend to suffer. A similar work [19], defines trading rules based on a set of MA chosen by the GA, in order to define an investment decision support system.

Other machine learning methodologies had been applied to the financial sector. ANN has been used with success in classifying market condition and forecast future prices of assets [20]. Other method largely used is the Support Vector Machine (SVM), which has been used as a classifier and as an estimator regression [21, 22] in order to predict future market conditions.

Fig. 2.31 Four symbol iSAX time series representation



The fact is that, in several works, these last methods always have been used as classifiers and optimization methods to predict markets, based on technical indicators or to recognize patterns and never to detect new pattern formations. The use of clusters to discover patterns, like in [12], prove that this kind of solutions did not offer the flexibility that allows the algorithm to adapt to the data and patterns characteristics. The ability that GA reveals, by the easy way how a problem could be coded in their genes, in order to optimize parameters and find solutions, also the possibility of creating chromosome structures that could change dynamically, makes the GA the ideal candidate to discover new pattern formations.

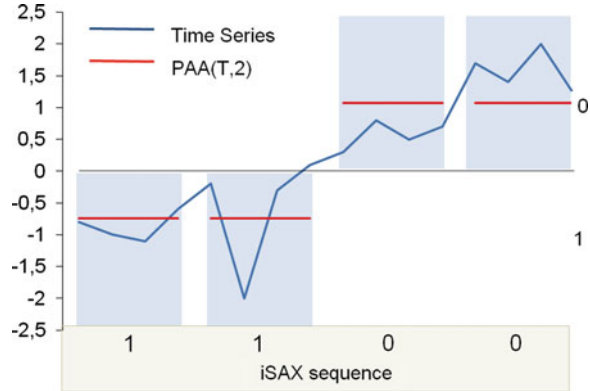
2.2.1.4 Symbolic Aggregate Approximation

In these last few years the appearance of algorithms for efficient string manipulation and the bioinformatics applications, for instance the Human Genome Project (HGP), turn the scientific community attention to the use of symbolic representation of data. In 2003 this symbolic method was first presented by Lin et al. [23, 24].

Traditionally time series representation and data dimensionality reduction use numeric methods, like Discrete Fourier Transform (DFT) [25] or Singular Value Decomposition (SVD) [26], those methods allow to define a similarity metrics between data representation that relates to the real distance between the raw series. The use of symbolic representation of time series had always suffer from the fact that the distance between sequences have low correlations to the distance defined between the original time series. The SAX method solves the problem of distance between representation and real data, since a lower bounding approximation from the Euclidian distance could be obtained [27].

Symbolic aggregate approximation (SAX) is based on the Piecewise Aggregate Approximation (PAA) [26], which basically consists in dividing the time series in equal size segments and then calculates the mean of the points in each segment, this new value will represent that segment. With this new representation a dimensional reduction of the time series is possible. To finally get to SAX, the

Fig. 2.32 Two symbol iSAX time series representation



PAA method suffers an additional step at the end, where the mean value for representing the time series section is discretized to a symbol, this process will be presented on [Chap. 3](#).

One of the advantages of using a symbolic representation, when using a database management system (DBMS) to hold the time series representation, has to do with the fact that the index structure of the DBMS can easily index sequences of strings and allows to query data to search for patterns. To improve time series data mining and indexing, when dealing with large amounts of data, an extension of the standard SAX was implemented [28], this new method is “indexable Symbolic Aggregate approXimation” or iSAX. This new approach allows for a fast time series data mining and reduced index timing, in [29], using iSAX 2.0, was possible to reduce the index building time of a data series by 72 % when comparing to standard iSAX, for this tests were used 1,000,000,000 (one billion) time series of length 256 and took less than 400 h. This new iSAX representation replaces the alphabetic symbols by binary sequences, Fig. 2.31.

The special codification of binary sequences allows creating less detail representations for the same data, with less number of symbols. For instance in Fig. 2.32 is the same time series of the previous Fig. 2.31, where it is possible to verify that thanks to the careful choice of the symbolic binary representation the symbols below zero all begin with the number one and above with zero, on both figures. So if the system codifies the time series with greater detail it is possible to go to a rough representation by removing trailing bits.

This characteristic allows the creation of an index tree structure that allows to fast search for a specific representation [28]. The present work did not use this kind of representation since the SAX-GA approach tries to discover new meaningful patterns in a time series and not search known patterns in time series database.

Another important variation of SAX for financial applications is eSAX, which stands for extended SAX [30]. This new approach to the SAX method tries to avoid the loss of some important characteristics of the financial time series by adding additional information to the representation. As was identified by studies using PIP, some important points exists in financial data, those points could

Fig. 2.33 PAA/SAX representation misses important points

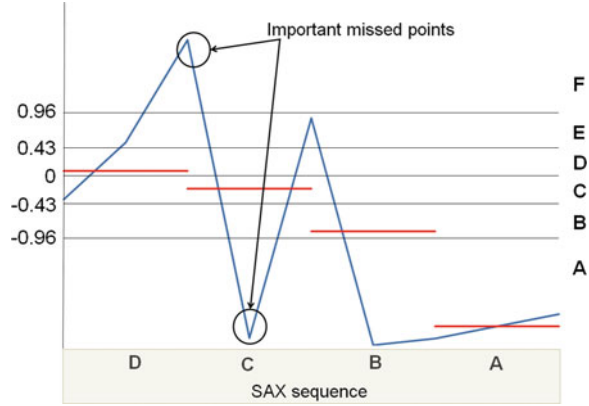
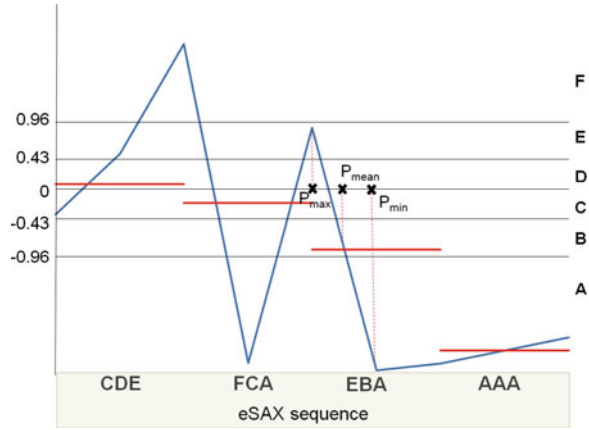


Fig. 2.34 eSAX time series representation



indicate a reverse trend on the market and the standard SAX tends to smooth them since is based on PAA, which is a method based on mean values, Fig. 2.33.

As can be seen on Fig. 2.33 some important points that could increase volatility are missed by the SAX representation, to solve this problem eSAX uses three symbols per division of the time series indicating the maximum, mean and minimum points in each division, so instead of one symbol per segment this method uses three symbols according to Eq. (2.14).

$$\langle S_1, S_2, S_3 \rangle = \begin{cases} \langle S_{\max}, S_{\text{mean}}, S_{\min} \rangle, & \text{if } P_{\max} < P_{\text{mean}} < P_{\min} \\ \langle S_{\min}, S_{\text{mean}}, S_{\max} \rangle, & \text{if } P_{\min} < P_{\text{mean}} < P_{\max} \\ \langle S_{\min}, S_{\max}, S_{\text{mean}} \rangle, & \text{if } P_{\min} < P_{\max} < P_{\text{mean}} \\ \langle S_{\max}, S_{\min}, S_{\text{mean}} \rangle, & \text{if } P_{\max} < P_{\min} < P_{\text{mean}} \\ \langle S_{\text{mean}}, S_{\max}, S_{\min} \rangle, & \text{if } P_{\text{mean}} < P_{\max} < P_{\min} \\ \langle S_{\text{mean}}, S_{\min}, S_{\max} \rangle, & \text{otherwise} \end{cases} \quad (2.14)$$

Table 2.2 Comparison between pattern detections methods

Method	Advantages	Disadvantages
Matrix template pattern	Gives good results when identifying trends or trend inversions, because the noise removal associated with the method that smooth the data allows to clearly identify these market movements	With large datasets and low dimensional reduction can be less effective, since has to convert to matrixes the financial time series and them multiply two matrixes to take any decision. Used only on pattern recognition
Perceptually important points (PIP)	Even when applying high levels of dimensional reduction to the data is possible to preserve the important information in the financial time series. This method was created to deal with financial data	Hard to use as pattern discover method, is mainly used to pattern recognition. The data has to be converted to several representations corresponding to the same dimensions of the existing templates that the method is trying to identify
Symbolic aggregate approXimation (SAX)	Converts data to a symbolic sequence, allowing to easily identifying patterns by comparing strings. Easy to implement and fast converting the data to the symbolic representation	Tends to lose important information of the financial time series if the method conversion parameters are not well chosen

where,
S The SAX symbol;
P Position in the horizontal axes where the maximum, minimum and mean points occur.

In Fig. 2.34 is an example of this method representation, where the use of the extra symbols allows to preserve the information about the oscillations in each of the segments.

A problem detected in the SAX representation method when using financial time series is the loss of some important points, because smoothes the data since is based on a mean method. To overcome this problem it could be suggested the use of eSAX, but the additional symbols correspond to a lower dimensional reduction of the time series, this would mean more data to be handled. It is clear that the dimensional reduction that can be achieved and the discretization of data in amplitude by the basic SAX method are characteristics that could be used to create an efficient algorithm to pattern discover. To solve the problem of the smooth data it is possible to try to choose the method parameters in a more careful way, for instance considering smaller segments when converting the time series.

Table 2.3 Market forecasting methods and results

References	Years	Method	Used data	Financial market	Period	Algorithm performance
[15]	2008	Bull flag pattern w/matrix template	Stock price	NYSE composite. index	1967–2003	4.59 % (Transaction average over the period)
[31]	2007	Bull flag pattern w/matrix template	Stock price	NASDAQ	1985/04/03–2004/03/20	4.38 % (Transaction average over the period)
[16]	2002	Hybrid neural network w/ pattern detection	Stock price and vol.	NYSE composite Index	1984/07/24–1998/06/11	66 % (Days market goes up after buying order)
[32]	2006	Template-based	Stock price	Several	N/A	96 % (Hits on pattern identification)
[32]	2006	Rule-based	Stock price	Several	N/A	38 % (Hits on pattern identification)
[33]	2009	ENN + PLR	Stock price	Google.com	2007/05/01–2007/10/19	95 % (Average rate of return)
[34]	2003	AANN	Stock price	Kospi 200 futures	2001	31 % (year profit)
[35]	2009	ANN	Stock price	Stocks from Bovespa	2008/05/02–2008/12/02	Worst return 23 % Best return 130 %
[36]	2008	SOM GA-BPN	Several	BSE-30 index	2007/11/12–2008/01/01	30 % more return than the index
[37]	2008	GACS-M	Stock price	Stocks from TAIEX	2000/09/21–2007/09/21	21.42 % (Average rate return)
[18]	2010	GA	Several	Nikkei 225	Jan.1999–Dec.2009	57.4 % (Profit rate)
[38]	2010	GA-ANN	Stock price	Shenzhen	N/A	0.7176 (Correlation between prediction and actual value)

(continued)

Table 2.3 (continued)

References	Years	Method	Used data	Financial market	Period	Algorithm performance
[39]	2011	MCS RBFNN	Stock price	Hang Seng index	10 years	167 (Average earning)
[40]	2011	NN	Price	Forex	2010/09/ 20–2011/ 01/21	0.666186e-3–0.101144e-2 (Mean square error)
[20]	2011	HLP-ANN	Stock price	Shanghai index	1991/11/ 18–2009/ 02/10	7.16 % –9.65 % (Prediction error)
[41]	2010	WMM + Kalman filter	Daily trading volume	Bowin Technology & Denghai Seed industry	2008/02/ 13–2009/ 02/13	0.13–0.1879 (SNR—Prediction)
[42]	2010	GRA + GNP	Stock price	Several	2005/01/ 08–2007/ 01/04	4 % (Average monthly return rate)

2.2.2 Why Choosing GA and SAX

In Sect. 2.2.1.3 the GA was chosen as the solution with the ability to discover patterns. The easy way of coding new patterns on the chromosome genes and then evolve those patterns to match important patterns existent on the financial data, makes this optimization tool the ideal to complete this task with success. Now it is necessary a way to represent the time series and patterns efficiently in order to use this form of representation in the GA. In Table 2.2 is presented a comparison between the previous time series representation methods, PIP, Matrix and SAX.

From Table 2.2 and as was referred in Sect. 2.2.1.2, the matrix representation has large limitations, it is only capable of supporting simple graphic formations, so for the discover process is not an interesting choice and is left out. The PIP representation has the advantage of being born to represent financial data, but the problem of having to keep several dimensional representations of the same data accordingly to the complexity of the patterns, makes this method a more complex to adapt to the GA. Finally SAX, which uses a symbolic representation of the time series, proves to be ideal to be coded in the gene format. The symbolic discretization, which applies in the vertical axes of the data, creates a symbol that is possible to save on chromosome gene and easily manipulated in the crossover and mutation process. Also the discretization factor will help to reduce the overfitting, since the pattern terms will be chosen from a limited set of values rather than a real number interval. The problem identified in this representation, the loss of some important information in the financial time series, could be overcome by the inclusion of the adjustable SAX representation factors as parameters to be optimized by the genetic algorithm.

2.3 Conclusions

As was previously referred, chart pattern detection is an important method of trying to forecast financial markets. Due to the difficulty of this task several methods, using technical or fundamental indicators or even the price of financial data, were used to forecast markets, some of those methods results are referred on Table 2.3. The lack of solutions that tries to predict market behavior based in graphic pattern discovery is notorious, so it is reasonable to think that a solution on this area could bring a natural advantage on financial trading. The methods presented in Sect. 2.2.1, based on pattern recognition, did not offer the ability of discovering new pattern formations and in the present market conditions to be able to discover new patterns is an important characteristic, because allows the forecast method to adapt and learn new graphic pattern formations, which could be relevant to the trading decision support system. Even all the works about the SAX methodology used this method as a pattern recognition tool. In areas other than

financial, SAX has been used to discover patterns, working with K-means or K-motifs as are called in [24].

In the discovery pattern area one of the goals is to find patterns with algorithms more efficient. One tool largely used to find solutions in an effective way is the GA. So, now what is needed is a data time series representation that could work together with the GA in an efficient way and could be represented in a chromosome format. From the several data representation studied the first choice was SAX, since the symbolic representation of data could be easily inserted on genes of a chromosome structure and assists the GA in finding new and important graphic pattern formations.

References

1. B. Baumohl, *The Secrets of Economic Indicators: Hidden Clues to Future Economic Trends and Investment Opportunities*, 2nd edn. (Wharton School Publishing, Pennsylvania, 2007)
2. S.B. Achelis, *Technical Analysis From A-To-Z*. (Vision Books, New Delhi, 2000)
3. T.N. Bulkowski, *Encyclopedia of Chart Patterns*, 2nd edn. (Wiley, New Jersey, 2005)
4. R. Fischer, J. Fischer, *Candlesticks, Fibonacci, and Chart Pattern Trading Tools: A Synergistic Strategy to Enhance Profits and Reduce Risk*. (Wiley, New Jersey, 2003)
5. R.W. Colby, *The Encyclopedia of Technical Market Indicators*, 2nd edn. (McGraw-Hill, New York, 2003)
6. E. Ponsi, *Forex Patterns & Probabilities—Trading Strategies for Trending and Range-Bound Markets*. (Wiley Trading, Wiley, New Jersey, 2007)
7. P. Parracho, R. Neves, N. Horta, Trading with optimized uptrend and downtrend pattern templates using a genetic algorithm kernel. *IEEE Congr. Evol. Comput.* 1895–1901 (2011). doi:[10.1109/CEC.2011.5949846](https://doi.org/10.1109/CEC.2011.5949846)
8. T.-C. Fu, A review on time series data mining. *Int. J. Eng. Appl. Artif. Intell.* **24**(3), 164–181 (2011). ISSN 0952-1976. doi:[10.1016/j.engappai.2010.09.007](https://doi.org/10.1016/j.engappai.2010.09.007)
9. F.-L. Chung, T.-C. Fu, R. Luk, V. Ng, Flexible time series pattern matching based on perceptually important points. *Int. Jt. Conf. Artif. Intell. Workshop on Learn from Temporal and Spatial Data*, 1–7 (2001)
10. T.-C. Fu, F.-L. Chung, R. Luk, C.M. Ng, Representing financial time series based on data point importance. *Eng. Appl. Artif. Intell.* **21**(2):277–300 (2008). ISSN 0952-1976. doi:[10.1016/j.engappai.2007.04.009](https://doi.org/10.1016/j.engappai.2007.04.009)
11. F.-L. Chung, T.-C. Fu, V. Ng, R.W.P. Luk, An evolutionary approach to pattern-based time series segmentation. *IEEE Trans. Evol. Comput.* **8**(5), 471–489 (2004). doi:[10.1109/TEVC.2004.832863](https://doi.org/10.1109/TEVC.2004.832863)
12. T.-C. Fu, F.-L. Chung, R. Luk, V. Ng, Pattern discovery from stock time series using self-organizing maps. *The 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2001)
13. W. Leigh, N. Modani, R. Purvis, T. Roberts, Stock market trading rule discovery using technical charting heuristics. *Expert Sys. Appl.* **23**(2), 155–159 (2002). ISSN 0957-4174. doi:[10.1016/S0957-4174\(02\)00034-9](https://doi.org/10.1016/S0957-4174(02)00034-9)
14. W. Leigh, N. Paz, R. Purvis, Market timing: a test of a charting heuristic. *Econ. Lett.* **77**(1), 55–63 (2002). ISSN 0165-1765. doi:[10.1016/S0165-1765\(02\)00110-6](https://doi.org/10.1016/S0165-1765(02)00110-6)
15. W. Leigh, C.J. Frohlich, S. Hornik, R.L. Purvis, T.L. Roberts, Trading with a stock chart heuristic. *IEEE Trans. Part A: Sys. Hum. Sys. Man Cybern.* **38**(1), 93–104 (2008). doi:[10.1109/TSMCA.2007.909508](https://doi.org/10.1109/TSMCA.2007.909508)

16. W. Leigh, M. Paz, R. Purvis, An analysis of a hybrid neural network and pattern recognition technique for predicting short-term increases in the NYSE composite index. *Omega*, Elsevier **30**, 69–76 (2002)
17. W. Leigh, R. Purvis, J.M. Ragusa, Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *J. Decis. Support Sys.* **32**(4), 361–377 (2002). doi:[10.1016/S0167-9236\(01\)00121-X](https://doi.org/10.1016/S0167-9236(01)00121-X)
18. K. Matsui, H. Sato, Neighborhood evaluation in acquiring stock trading strategy using genetic algorithms. *The International Conference on Soft Computing and Pattern Recognition (SoCPaR)*, pp. 369–372 (2010). doi:[10.1109/SOCPAR.2010.5686733](https://doi.org/10.1109/SOCPAR.2010.5686733)
19. J. Pinto, R. Neves, N. Horta, *Fitness function evaluation for MA trading strategies based on genetic algorithms*. ed. by N. Krasnogor. *Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO '11)*. ACM, New York, NY, USA, pp. 819–820 (2011). doi:[10.1145/2001858.2002105](https://doi.org/10.1145/2001858.2002105)
20. L. Wang, Q. Wang, Stock market prediction using artificial neural networks based on HLP. *Int. Conf. Intel. Hum. Mach. Sys. Cybern.* **1**, 116–119 (2011). doi:[10.1109/IHMSC.2011.34](https://doi.org/10.1109/IHMSC.2011.34)
21. L.J. Cao, F.E.H. Tay, Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Trans. Neural Netw.* **14**(6), 1506–1518 (2003). doi:[10.1109/TNN.2003.820556](https://doi.org/10.1109/TNN.2003.820556)
22. D. Zhang, H. Song, P. Chen, Stock market forecasting model based on a hybrid ARMA and support vector machines. *Proceedings of the 15th International Conference on Management Science and Engineering*, Long Beach, USA (2008)
23. J. Lin, E. Keogh, S. Lonardi, B. Chiu, A symbolic representation of time series, with implications for streaming algorithms. *Proceedings of the 8th ACM SIGMOD International Conference on Management of Data, Workshop on Res. Issues in Data Mining and Knowledge Discovery*, pp. 2–11 (2003). doi:[10.1145/882082.882086](https://doi.org/10.1145/882082.882086)
24. J. Lin, E. Keogh, S. Lonardi, P. Patel, Finding motifs in time series. *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining . 2nd Workshop on Temporal Data Mining*, pp. 53–68 (2002)
25. C. Faloutsos, M. Ranganathan, Y. Manolopoulos, Fast subsequence matching in time-series databases. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Minneapolis, pp. 419–429 (1994). doi:[10.1145/191839.191925](https://doi.org/10.1145/191839.191925)
26. E. Keogh, K. Chakrabarti, M. Pazzani, S. Mehrotra, Dimensionality reduction for fast similarity search in large time series databases. *J. Knowl. nf. Sys.* (2000). doi:[10.1145/191839.191925](https://doi.org/10.1145/191839.191925)
27. J. Lin, E. Keogh, L. Wei, S. Lonardi, Experiencing SAX: A novel symbolic representation of time series. *Data Min. Knowl. Discov.* **15**(2), 107–144 (2007). doi:[10.1007/s10618-007-0064-z](https://doi.org/10.1007/s10618-007-0064-z)
28. J. Shieh, K. Keogh, iSAX: Indexing and mining terabyte sized time series. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, pp. 623–631 (2008). doi:[10.1145/1401890.1401966](https://doi.org/10.1145/1401890.1401966)
29. A. Camerra, T. Palpanas, J. Shieh, E. Keogh, iSAX 2.0: Indexing and mining one billion time series. *Proceedings of the 2010 IEEE International Conference Data Mining (ICDM '10)*. IEEE Computer Society, Washington, DC, USA, pp. 58–67 (2010). doi:[10.1109/ICDM.2010.124](https://doi.org/10.1109/ICDM.2010.124)
30. B. Lkhagva, Y. Suzuki, K. Kawagoe, New time series data representation ESAX for financial applications. *Proceedings of the 22nd International Conference Data Engineering Workshops (ICDEW '06)*, p. 115 (2006). doi:[10.1109/ICDEW.2006.99](https://doi.org/10.1109/ICDEW.2006.99)
31. J.-L. Wang, S.-H. Chan, Stock market trading rule discovery using pattern recognition and technical analysis. *Expert Sys. Appl.* **33**(2), 304–315 (2007). doi:[10.1016/j.eswa.2006.05.002](https://doi.org/10.1016/j.eswa.2006.05.002)
32. T.-C. Fu, F.-L. Chung, R. Luk, C.M. Ng, Stock time series pattern matching: Template-based vs rule-based approaches. *Eng. Appl. Artif. Intell.* **20**(3), 347–364 (2007)
33. P.-C. Chang, C.-Y. Fan, C.-H. Liu, Y.-W. Wang, J.-J. Lin, Evolving neural network with dynamic time warping and piecewise linear representation system for stock trading decision

- making, 2009 WRI World Congr. Comput. Sci. Inf. Eng. **5**, 303–307 (2009). doi:[10.1109/CSIE.2009.36](https://doi.org/10.1109/CSIE.2009.36)
34. L. Junmyung, C. Sungzoon, B. Jinwoo, Trend detection using auto-associative neural networks: Intraday KOSPI 200 futures. Proceedings of the IEEE International conference on Computational Intelligence for Financial Engineering, pp. 417–420. (2003) doi:[10.1109/CIFER.2003.1196290](https://doi.org/10.1109/CIFER.2003.1196290)
 35. L.C. Martinez, D.N. da Hora, J.R. de M Palotti, W. Meira, G.L. Pappa (2009) From an artificial neural network to a stock market day-trading system: A case study on the BM&F BOVESPA, The International Joint Conference on Neural Networks 2006–2013 (2009). doi:[10.1109/IJCNN.2009.5179050](https://doi.org/10.1109/IJCNN.2009.5179050)
 36. A.U. Khan, T.K. Bandopadhyaya, S. Sharma, Classification and identification of stocks using SOM and genetic algorithm based backpropagation neural network. International Conference on Innovation in Management and Information Technology, pp. 292–296 (2008). doi:[10.1109/INNOVATIONS.2008.4781644](https://doi.org/10.1109/INNOVATIONS.2008.4781644)
 37. P.-C. Ko, P.-C. Lin, C.-S. Shih, Stock valuation and dynamic asset allocation with genetic algorithm and cubic spline. Int. Conf. Mach. Learn. Cybern. **7**, 3997–4000 (2008). doi:[10.1109/ICMLC.2008.4621101](https://doi.org/10.1109/ICMLC.2008.4621101)
 38. H.-N. Hao, Short-term forecasting of stock price based on genetic-neural network. 6th Int. Conf. Nat. Comput. **4**, 1838–1841 (2010). doi:[10.1109/ICNC.2010.5584528](https://doi.org/10.1109/ICNC.2010.5584528)
 39. W.W.Y. Ng, X.-L. Liang, P.P.K. Chan, D.S. Yeung, Stock investment decision support for Hong Kong market using RBFNN based candlestick models. Int. Conf. on Mach. Learn. Cybern. **2**, 538–543 (2011)
 40. H. Tahersima, M. Tahersima, M. Fesharaki, N. Hamed, Forecasting stock exchange movements using neural networks: A case study. Proceedings of the International Conference on Future Computer Science and Applications, pp. 123–126 (2011)
 41. Z. Fang, G. Luo, F. Fei, S. Li, Stock forecast method based on wavelet modulus maxima and kalman filter. Proceedings of the 4th International Conference on Management of e- Commerce and e-Government, pp. 50–53 (2010). doi:[10.1109/ICMeCG.2010.19](https://doi.org/10.1109/ICMeCG.2010.19)
 42. V. Parque, S. Mabu, K. Hirasawa, Enhancing global portfolio optimization using genetic network programming. Proceedings of the SICE Annual Conference, pp. 3078–3083 (2010)

Investment Strategies Optimization based on a SAX-GA
Methodology

Canelas, A.M.L.; Neves, R.F.M.F.; Horta, N.

2013, XII, 81 p. 81 illus., 19 illus. in color., Softcover

ISBN: 978-3-642-33109-1