New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume II: Applications

Editors

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny

SRI PAS **IBS PAN**

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume II: Applications

Systems Research Institute
Polish Academy of Sciences

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics
Volume II: Applications

Editors

Krassimir T. Atanassov
Władysław Homenda
Olgierd Hryniewicz
Janusz Kacprzyk
Maciej Krawczak
Zbigniew Nahorski
Eulalia Szmidt
Sławomir Zadrożny

IBS PAN SRI PAS

Systems Research Institute
Polish Academy of Sciences
Newelska 6, 01-447 Warsaw, Poland
www.ibspan.waw.pl

ISBN 83-894-7541-3

Dedicated to Professor Beloslav Riečan on his 75th anniversary

Fuzzy geometric protoforms for price patterns recognition and stock trading

Piotr Ładyżyński[1] **and Przemysław Grzegorzewski**[1,2]

[1]Faculty of Mathematics and Computer Science,
Warsaw University of Technology
Plac Politechniki 1, 00-661 Warsaw, Poland
[2]Systems Research Institute, Polish Academy of Sciences,
Newelska 6, 01-447 Warsaw, Poland
p.ladyzynski@mini.pw.edu.pl, pgrzeg@ibspan.waw.pl

Abstract

A novel approach for detecting patterns in price time series was suggested. The system for the consolidation phase recognition was proposed. It utilizes the concept of so-called fuzzy geometric protoform and classification trees.

Keywords: fuzzy protoform, patterns recognition, fuzzy sets, classification trees, consolidation phase, time series.

1 Introduction

While observing price charts of financial instruments (shares, futures contracts, commodities) we can realize that prices tend to form some unique patters before extraordinary market events. Investors try to identify such patterns to predict significant price movements, market volatility changes, bull or bear market commencement. Savin [5] proposed an interesting method assessing the predictive power of price patterns.

New Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications *(K.T. Atanassow, W. Homenda, O. Hryniewicz, J. Kacprzyk, M. Krawczak, Z. Nahorski, E. Szmidt, S. Zadrożny, Eds.), IBS PAN - SRI PAS, Warsaw, 2012.*

Figure 1: Well formed price channel and an upstroke in late November.

In spite of the importance of the problem there are not too many research papers dealing with the price patterns recognition. Most of publications try to solve the problem using neural networks ([1],[4]) being trained with preprocessed price time series (usually smoothing averages and technical analysis indicators are calculated). Other systems extract patterns with manually established rules comparing some indicators calculated for the fixed time windows. Unfortunately, such systems cannot adapt to changing markets.

In this paper we propose a novel approach for detecting patterns in price time series. Our system utilizes the new concept of so-called fuzzy geometric protoform which seems to be quite effective for modeling patterns.

2 Patterns in technical analysis

Investors use information about historical movements in price and trading volume summarized in the form of charts to forecast future price trends. In Fig. 1 prices form a price channel while oscillating between two parallel lines called **resistance** (upper line) and **support** (lower line). An upstroke in late November (we say that there is an upstroke when prices break resistance line) is an indicator of approaching uptrend. On the other hand, when prices came back to the old resistance line in the first week of December, was the strong negation of previous signal.

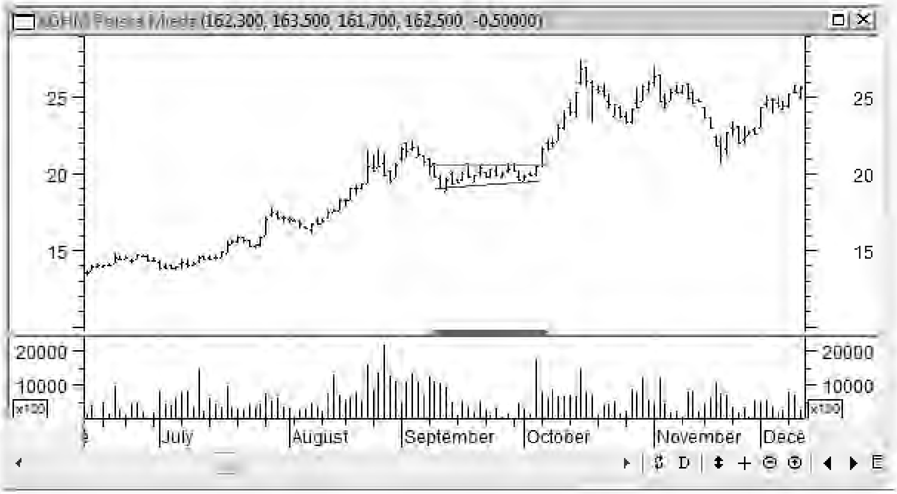

Figure 2: Consolidation phase and an upstroke confirmed by strong uptrend.

In Fig. 2 prices form a consolidation phase. After resistance line break prices elevate in significant uptrend. Well formed consolidation phase is a rather rare market phenomena. Moreover consolidation phase breakup is one of the strongest technical analysis indicators. It provides us with an opportunity to open a position with high probability of exceeding return and low risk. There are two most known trading rules connected with the consolidation phase:

- **Upper breakup** – uptrend is expected. We open a long position and place STOP LOSS order in the middle of breakup day price candle. If prices follow uptrend we move our STOP LOSS up to protect the profit.
- **Lower breakup** – downtrend is expected. We open a short position and place STOP LOSS order in the middle of breakup day price candle. If prices follow downtrend we move our STOP LOSS down to protect the profit.

The support and resistance lines in the consolidation phase price pattern form a semi-rectangle of unknown length.

3 The support and resistance lines

Let $\{X_t\}_{t=1}^{n}$ denote a time series of prices for a given financial instrument (or a part of the longer time series $\{X_t\}_{t\in T}$). To examine patterns that may appear in this time series the following sets will be helpful.

Definition 1 *Let d, $d \in \{1, \dots n\}$, denote the length of a subwindow. Then a set $\mathcal{SB}_d \subset \mathbb{R}^2$ given by*

$$\mathcal{SB}_d = \{(t, Y_t) : Y_t = \min\{X_i, X_{i+1}, \dots, X_{i+d}\}, i = 1, \dots, n-d\} \quad (1)$$

is called the d-span **support base** *of the time series $\{X_t\}_{t=1}^n$. Similarly, a set $\mathcal{SC}_d \subset \mathbb{R}^2$ given by*

$$\mathcal{SC}_d = \{(t, Y_t) : Y_t = \max\{X_i, X_{i+1}, \dots, X_{i+d}\}, i = 1, \dots, n-d\} \quad (2)$$

is called the d-span **resistance base** *of the time series $\{X_t\}_{t=1}^n$.*

We also define the truncations of these sets:

Definition 2 *Lets t_p, t_0 denote time moments such that $t_p < t_0$. Then*

$$\mathcal{SB}_{d|[t_p,t_0]} = \{(t, Y) \in \mathcal{SB}_d : t \in [t_p, t_0]\} \quad (3)$$

$$\mathcal{SC}_{d|[t_p,t_0]} = \{(t, Y) \in \mathcal{SC}_d : t \in [t_p, t_0]\} \quad (4)$$

We'll utilize the support base and the resistance base to define a price-quadrangle for our time series $\{X_t\}_{t=1}^n$. First of all let us define the support line and the resistance line for the given time interval $[t_p, t_0]$. Please, note, that all concepts discussed above are obtained for a fixed subwindow d.

Definition 3 *Let $y = \alpha_b^{(t_p,t_0)} t + \beta_b^{(t_p,t_0)}$ denote a line obtained using the least square method for points $\mathcal{SB}_{d|[t_p,t_0]} \subset \mathbb{R}^2$. Then*

$$\mathcal{B}_{d|[t_p,t_0]} = \left\{(t, y) : t_p \le t \le t_0, y = \alpha_b^{(t_p,t_0)} t + \beta_b^{(t_p,t_0)}\right\} \quad (5)$$

is called the d-span **support** *of the time series $\{X_t\}$ for the interval $[t_p, t_0]$.*

Similarly, let $y = \alpha_c^{(t_p,t_0)} t + \beta_c^{(t_p,t_0)}$ denote a line obtained using the least square method for points $\mathcal{SC}_{d|[t_p,t_0]} \subset \mathbb{R}^2$. Then

$$\mathcal{C}_{d|[t_p,t_0]} = \left\{(t, y) : t_p \le t \le t_0, y = \alpha_c^{(t_p,t_0)} t + \beta_c^{(t_p,t_0)}\right\} \quad (6)$$

is called the d-span **resistance** *of the time series $\{X_t\}$ for the interval $[t_p, t_0]$.*

Further on, unless it is declared differently, we call the d-span support base (resistance) just the support base (resistance) and the d-span support support (resistance) as the the support (resistance). Empirical studies shows that d-span of 10 days works quite well.

Now we are able to define one of the fundamental concepts of our contribution called a price-quadrangle.

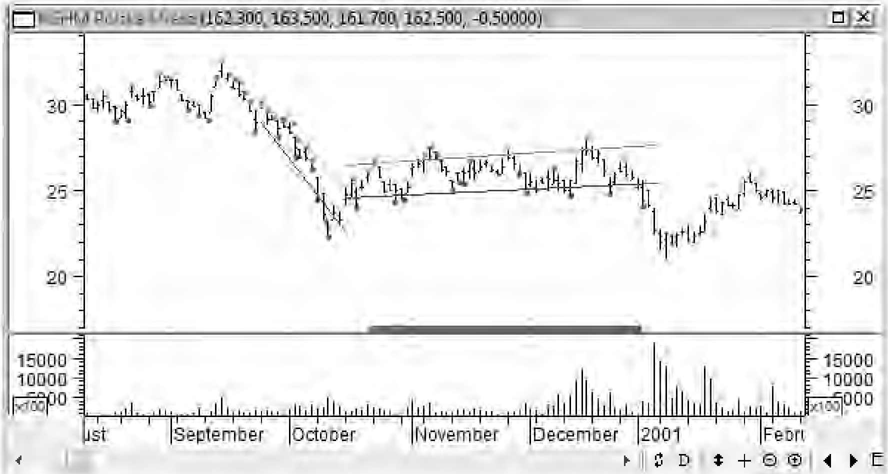

Figure 3: Dark grey points represents minimum prices from last ten trading days and light grey points represents maximum prices (visualistion of sets $\mathcal{SB}^d$ and $\mathcal{SC}^d$ – see def. 1).

Definition 4 *A* **price-quadrangle** *based on the time series* $\{X_t\}_{t=1}^n$ *for the given interval* $[t_p, t_0]$ *is a set*

$$Q_{d|[t_p,t_0]} = \left\{(t, y) : t_p \le t \le t_0, \mathcal{B}_{d|[t_p,t_0]} \le y \le \mathcal{C}_{d|[t_p,t_0]}\right\}. \tag{7}$$

Different examples of the price-quadrangles are given in Fig. 1, Fig. 2 and Fig. 3. In particular, the price-quadrangle in Fig. 1 form the so-called increasing price channel with parallel support and resistance. In Fig. 2 we have shown the price-quadrangle for the consolidation phase, while in Fig. 3 we find two price-quadrangles: one corresponding to the decreasing price channel and the second illustrating the consolidation phase.

These few examples that the shape of the price-quadrangle may be applicable for detecting different patterns in the time series under study. In particular, if both the support and the resistance are increasing (decreasing) then we get an increasing (decreasing) price channel, while the price-quadrangle made by the support and resistance parallel to the time axis seems to be characteristic for the consolidation phase. In other words, the desired shape of the price-quadrangle indicating the consolidation phase is a rectangle parallel (perpendicular) to the time (price) axis.

Of course, to get the ideal rectangle might be impracticable for the real data. However, we should expect that the price-quadrangle corresponding to the consolidation phase would be at least semi-rectangular, i.e. the slope of the support α_b and the slope of the resistance α_c should be “nearly”equal and “close”to 0.

Thus a natural question arises, how to model such expressions like “nearly”, “close”, etc. It seems that fuzzy set theory might be quite useful for it.

Besides modeling vague statements we have to be aware that using sample data we should also consider the random error that may occur. Since each price-quadrangle is formed by two straight lines obtained by means of the least square method, we define the error corresponding to a price-quadrangle as

$$\sigma_Q = \sqrt{SSE_b^2 + SSE_c^2}, \tag{8}$$

where SSE_b^2 and SSE_c^2 denote the sum of squared error for the support and resistance, respectively.

4 Fuzzy geometric protoform

The concept of a protoform (prototypical form), that is recently vividly advocated by Zadeh (see [6]), is defined as an abstract prototype. Zadeh points out the relevance of protoforms in the formalization of human consistent reasoning in various problems related to broadly perceived information technology and knowledge engineering. For example, Kacprzyk and Zadrożny [3] proposed the implementation of protoforms for generating families of linguistic database summaries and fuzzy queries. In their application protoforms are so powerful conceptual tools because one can formulate many different types of linguistic summaries in a uniform way and then devise a uniform and universal way to handle different linguistic summaries.

Below we define a quite abstract concept of a fuzzy geometric protoform and then we propose how to apply this idea for modeling and validating the consolidation phase in a time series.

Definition 5 *A **fuzzy geometric protoform (FGP)** is an ordered triple* $(\mathcal{P}, \mathcal{F}, \mathcal{M})$*, where:*

- $\mathcal{P}$ *- is a family of parameters describing the geometric structure of the object under study,*
- $\mathcal{F}$ *- is a (fuzzy) domain for parameters in* $\mathcal{P}$*,*
- $\mathcal{M}$ *- is a valuation function defined on* $\mathcal{F}$*.*

Here we do not discuss fuzzy geometric protoforms in general but we show that this concept might be useful for the consolidation phase detection.

Definition 6 *Let (X_t) be a given price time series. A* ***fuzzy geometric protoform of consolidation phase*** *is is an ordered triple $(\mathcal{P}, \mathcal{F}, \mathcal{M})$ such that*

- $\mathcal{P} = (\alpha_b^{(t_p,t_0)}, \alpha_c^{(t_p,t_0)}, T)$*, where T is a time window, $t_p, t_0 \in T$ are certain time moments such that $t_p < t_0$, while $\alpha_b^{(t_p,t_0)}$ and $\alpha_c^{(t_p,t_0)}$ denote the slope of the support and resistance calculated for time interval $[t_p, t_0]$, respectively;*

- $\mathcal{F} = (A_c, A_{par}, A_T)$*, where A_c, A_{par} and A_T denote fuzzy subsets of the real line $\mathbb{R}$ which are domains of the successive parameters in $\mathcal{P}$, respectively;*

- $\mathcal{M} = \mathcal{S}(\mathcal{T}(\mu_c(\alpha_c^{(t_p,t_0)}), \mu_{par}(\alpha_c^{(t_p,t_0)} - \alpha_b^{(t_p,t_0)}), \mu_T(t_0)))$*, where $\mathcal{S}$ and $\mathcal{T}$ denote an S-conorm and T-norm, respectively, while μ_b, μ_c and μ_T stand for the membership functions of successive fuzzy sets A_b, A_c, A_T that appear in $\mathcal{F}$.*

Please note, that classical S-conorm and T-norm are defined on a unite square $[0,1]^2$, while function considered above are specified for more than two arguments. Hence they are actually S-multiconorm and T-multinorm. However, such functions are easily obtained from the classical ones due to associativity of S-conorms and T-norms.

A natural question arises: How to define fuzzy domains of the parameters enclosed in $\mathcal{P}$ and a valuations function $\mathcal{M}$. Of course, each particular fuzzy geometric protoform of consolidation phase should depend on the problem under study and hence all objects which characterize such fuzzy protoform ought to be strongly consistent with the real-life data.

A fuzzy set A_c corresponds to possible values of the resistance. However, since our intention is to detect the consolidation phase, the most desired values of the slope of the resistance are close to 0 while values far from 0 are unwanted for the consolidation. Thus a natural membership function for A_c is symmetrical about 0. The same is true for A_b. However, in our study we have decided to substitute the fuzzy set A_b corresponding to the support by a fuzzy A_{par} describing parallelism. Actually, instead of checking whether the slope of the resistance and support are both close to 0 we may examine only one of these coefficients and then test if the support and resistance are parallel, i.e. if the distance between their slopes is zero. Thus the membership function of A_{par} should be also symmetrical about 0. We refer the reader to Fig. 4 for the graphs of μ_c and μ_{par}.

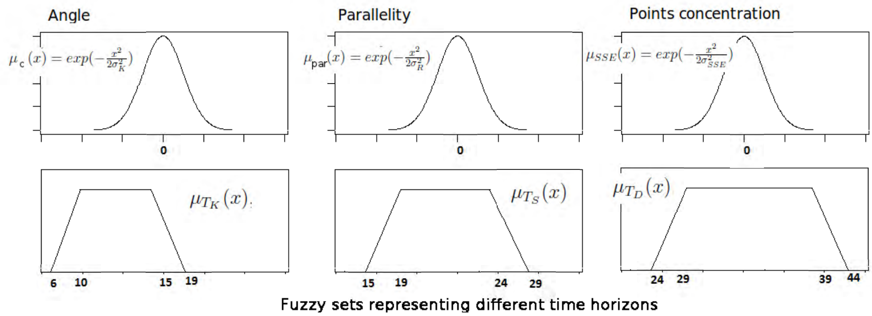

Figure 4: Membership functions of fuzzy sets used for construction of valuation function.

Finally, we need a valuation function $\mathcal{M}$. Actually we are interested in a value of $\mathcal{M}$ for a fixed time moment t_0. Since the price-quadrangle is calculated for given interval $[t_p, t_0]$ from time window T, $\mathcal{M}$ would also strongly depend on $[t_p, t_0]$. Without loss of generality we may assume that $T = t_1, \ldots, t_n$ such that $t_n = t_0$.

In our study (see Sec. 7 for the description and results) we compare price-quadrangles corresponding to short-term, medium-term and long-term investments horizons. Of course, such statements like "short-term", "medium-term"and "long-term"are fuzzy and therefore they are satisfactory modeled by fuzzy subsets A_{T_S}, A_{T_M}, A_{T_L} of the real line $\mathbb{R}$. In our study we have decided to utilize trapezoidal fuzzy numbers - both because of their simplicity and a natural interpretation which is easy to grasp by the users. Membership functions μ_{T_S}, μ_{T_M} and μ_{T_L} of these fuzzy sets are given in Fig. 4. Therefore, we define three valuation functions - each one connected with different investment horizon.

We have decided to apply max operator as S-conorm and the product as T-norm. Moreover, in our opinion the random error connected with the resistance and support should be also taken into consideration. So we propose the following valuation functions

$$\mathcal{M}_\tau = \max_{t_p \in T}(\mu_c(\alpha_c^{(t_p,t_0)}) \cdot \mu_{par}(\alpha_c^{(t_p,t_0)} - \alpha_b^{(t_p,t_0)}) \cdot \mu_{SSE}(\sigma_Q^{(t_p,t_0)}) \cdot \mu_\tau(t_p)), \quad (9)$$

where $\tau \in \{T_S, T_M, T_L\}$ and μ_{SSE} is a membership function of a fuzzy set A_{SSE} which describes the acceptable values of the error $\sigma_Q^{(t_p,t_0)}$ corresponding to the price-quadrangle calculated for the time interval $[t_p, t_0]$. As before, μ_{SSE} is symmetrical about 0. It's graph is also shown in Fig. 4.

Please note that any valuation function takes values in the unit interval $[0, 1]$ and hence the estimated value of $\mathcal{M}_\tau$ may be perceived as a degree of conviction that we are in the consolidation phase. In particular, high values of valuation functions $\mathcal{M}_\tau$ indicate that we are already in the consolidation phase.

5 Trend following indicators

Exponential smoothing of the rate of returns from a certain financial instrument could be used as a trend following indicator. Tests for trend may also be useful for the consolidation phase detection. It is so since in well formed consolidations there is no significant trend.

Now let us recall two classical tools for handling with a time series that will be useful in the next section.

Definition 7 *The* ***exponential moving average*** *for the time series* $\{X_t\}$ *is a time series* $\{F_t\}$ *defined by*

$$F_t = \lambda X_t + (1 - \lambda) F_{t-1}, \tag{10}$$

where $F_0 = X_0$.

Definition 8 *The* ***exponential*** n***-moving average of rate of returns*** $\hat{m}^{exp,(n)}$ *for the price time series* $\{X_t\}$ *is the exponential* n*-moving average calculated for time series* $\{m_t\}$*, where*

$$m_t = \frac{X_t - X_{t-1}}{X_{t-1}} \tag{11}$$

is the rate of return from X_t *in the time interval* $[t-1, t]$.

6 Consolidation phase detection system

In previous sections we have shown how to compute measures aggregating information that might be useful for detecting patterns in a price time series. In this section we suggest how to construct a decision rule based system for detecting the consolidation.

As it is known machine learning algorithms could adapt to the problem structure and automatically capture decision rules from data. Classification trees, for instance, seem to work quite well for our problem (see [2]).

The following steps show how to implement a system for consolidation phase recognition:

1. Prepare a training set by selecting days which belong to the consolidation phase. For every trading day we assign a class y_i, such that $y_i = 1$ corresponds to the consolidation phase and $y_i = 0$ indicates the lack of consolidation.

2. For each trading day from a learning set we calculate a vector
$$\mathbf{x}_i = (\mathcal{M}_{T_S}, \mathcal{M}_{T_M}, \mathcal{M}_{T_L}, m_{15}^{(exp)}, m_{25}^{(exp)}, m_{35}^{(exp)}). \quad (12)$$

3. Train the decision tree using the learning set $\{(y_i, \mathbf{x}_i)\}$.

7 Empirical resuls

The system for detecting consolidation phases discussed above was trained on KGHM (the largest copper mining company in Europe) quotations from 1-01-1998 to 1-01-2006. Then the performance of the system was examined using data obtained from 2-01-2006 to 1-09-2011. During this period 29 consolidation phases occurred 27 of which were recognized by the system which gave the total accuracy of 93%.

Dark grey thick lines in Fig. 5 and Fig. 6 mark days classified as the consolidation phase. Looking at Fig. 5 one may conclude that the system has found the consolidation phase lasting from the second week of September to the first week of October. Technical analyst would say that this consolidation phase begins in the third week of August. It is true but we should remember that when the consolidation starts we are unable to verify its existence until it forms a certain pattern.

The existence of the consolidation phase itself is not precisely defined. When constructing the learning set we assume that the consolidation phase starts when prices rebound from one of boundary lines (support, resistance) for the second time without crossing it. Empirical results show that the system is able to adapt to this assumption (see Fig. 5 and Fig. 6).

Despite of promising performance the system has a weak point. In some cases the detected consolidation phases were not coherent which means that inside such consolidation phase single days or small group of days were classified as non-consolidation.

8 Conclusions

In this paper we presented a novel system architecture for detecting patterns in price time series. The system utilizes so-called fuzzy geometric protoforms sug-

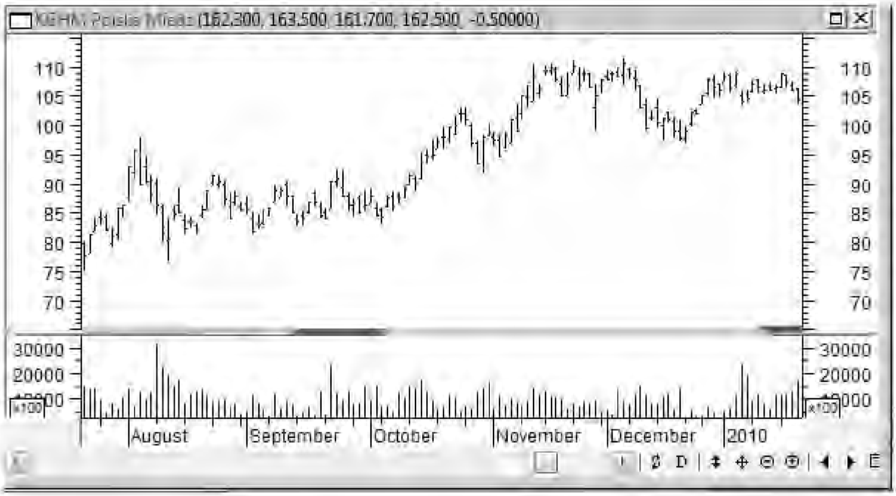

Figure 5: Sample consolidations extracted by the system.

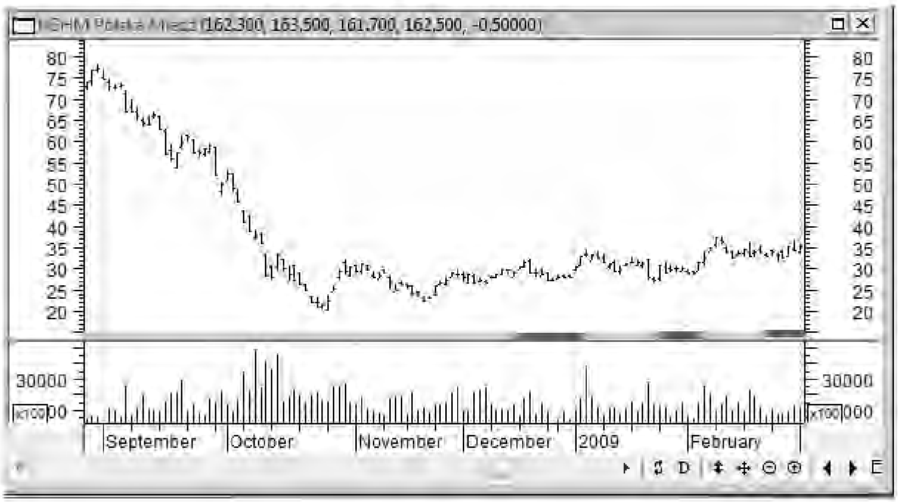

Figure 6: Sample consolidations extracted by the system.

gested also in the paper. Promising results of the empirical study proves that the concept of fuzzy geometric protoform is a useful tool for identifying patterns in graphical visualizations of data.

Although the performance of the system is quite satisfactory, some further improvements would be desirable. Firstly, we should try to upgrade a classifier using boosting technique. Secondly,the problem of single misclassified days inside the consolidation phase should be solved.

References

[1] Guo X., Liang X., Li X. (2007), A Stock Pattern Recognition Algorithm Based on Neural Networks. Natural Computation, 2007. ICNC 2007.

[2] Hastie T., Tibshirani R., Friedman J. (2009), The Elements of Statistical Learning. Springer-Verlag.

[3] Kacprzyk J., Zadrożny S. (2005), Linguistic database summaries and their protoforms: towards natural language based knowledge disresistancey tools, Information Sciences 173, 281304.

[4] Kamijo K., Tanigawa T. (1990), Stock Price Pattern Recognition - A Recurrent Neural Network Approach. Neural Networks, 1990., 1990 IJCNN International Joint Conference.

[5] Savin N.E., Weller P.A., Zvingelis J. (2007), The Predictive Power of Head-and-Shoulders Price Patterns in the U.S. Stock Market. Journal of Financial Econometrics, Vol. 5, Issue 2, pp. 243-265.

[6] Zadeh L. (2002), A prototype-centered approach to adding deduction capabilities to search enginesthe concept of a protoform. BISC Seminar, 2002, University of California, Berkeley.

The papers presented in this Volume 2 constitute a collection of contributions, both of a foundational and applied type, by both well-known experts and young researchers in various fields of broadly perceived intelligent systems.

It may be viewed as a result of fruitful discussions held during the Tenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2011) organized in Warsaw on September 30, 2011 by the Systems Research Institute, Polish Academy of Sciences, in Warsaw, Poland, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences in Sofia, Bulgaria, and WIT - Warsaw School of Information Technology in Warsaw, Poland, and co-organized by: the Matej Bel University, Banska Bystrica, Slovakia, Universidad Publica de Navarra, Pamplona, Spain, Universidade de Tras-Os-Montes e Alto Douro, Vila Real, Portugal, and the University of Westminster, Harrow, UK:

Http://www.ibspan.waw.pl/ifs2011

The consecutive International Workshops on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGNs) have been meant to provide a forum for the presentation of new results and for scientific discussion on new developments in foundations and applications of intuitionistic fuzzy sets and generalized nets pioneered by Professor Krassimir T. Atanassov. Other topics related to broadly perceived representation and processing of uncertain and imprecise information and intelligent systems have also been included. The Tenth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets (IWIFSGN-2011) is a continuation of this undertaking, and provides many new ideas and results in the areas concerned.

We hope that a collection of main contributions presented at the Workshop, completed with many papers by leading experts who have not been able to participate, will provide a source of much needed information on recent trends in the topics considered.

ISBN-13 9788389475411