

Mining information from time series in the form of natural language expressions

Vilém Novák

University of Ostrava, Institute for Research and Applications of Fuzzy Modeling
NSC IT4Innovations, 30. dubna 22, 701 03 Ostrava 1, Czech Republic

Abstract

In this paper, we discuss three tasks of mining information from time series, namely finding intervals with monotonous trend, evaluation of it in sentences of natural language, and summarizing information using intermediate quantifiers. The mined information is presented in simple sentences of natural language. For estimation of the trend of time series, we apply the theory of first-degree F-transform. For generating sentences of natural language we apply the formal theory of fuzzy natural logic. Then the former are obtained by interpretation of special formulas.

Keywords: Fuzzy Natural Logic, Time series data mining, F-transform, Intermediate quantifiers

1. Introduction

This paper is focused on specific tasks of mining information from time series and presenting them in the form of natural language expressions. To solve them, we apply techniques based on the theory of *fuzzy (F-)transform* that makes it possible to analyze and elaborate time series and *fuzzy natural logic* (FNL) that provides methods for automatic generation of comments in natural language.

There are many tasks in the field of time series data mining (cf. [1]). In this paper, we will focus to three problems: finding areas (i.e., time intervals) with definite character of trend, its characterization using expressions of natural language and summarization of the characteristics of time series using the, so called, intermediate quantifiers (part of the theory of FNL).

Linguistic characterization of local trend of time series using our methods was elaborated in [2, 3]. In these papers and also in [4] we outlined and gave examples of application of the theory of intermediate quantifiers to generation of summarizing comments to the structure of time series but we did not elaborate the details. In this paper, we give more details how the mined summarizing linguistic expressions can be obtained. In fact, we first form special logical formulas. The linguistic expressions are then obtained by interpretation of them.

2. Elements of fuzzy natural logic

The *fuzzy natural logic* is a formal mathematical theory that consists of the following special theories:

- (a) A formal theory of evaluative linguistic expressions explained in detail in [5] (for less formal explanation see also [6]).
- (b) A formal theory of fuzzy IF-THEN rules and approximate reasoning presented (see [7, 8]).
- (c) A formal theory of intermediate and generalized fuzzy quantifiers (see [9, 10, 11, 12]).

FNL is a mathematical theory developed using formalism of the *fuzzy type theory* (FTT) that was in detail elaborated in [13, 14]. Its formal language is lambda calculus. Because of limited space, we cannot make precise distinction between formula (or variable) and its interpretation in a model. Therefore, we will sometimes write, e.g., $z \in Form_{o\alpha}$ as a variable*) of type $o\alpha$ and then we say “a fuzzy set z ” having in mind interpretation of z in some model that may not, however, be precisely specified in the text.

2.1. Evaluative linguistic expressions

The central role in all these theories and also in this paper is played by the theory of *evaluative linguistic expressions*. Since we presented this theory in many papers, we will only very briefly remind the main concepts.

Evaluative linguistic expressions have the following general form

$$\langle \text{linguistic modifier} \rangle \langle \text{TE-adjective} \rangle \quad (1)$$

where $\langle \text{TE-adjective} \rangle^\dagger$ is one of the adjectives “small, medium, big” (and possibly other specific adjectives, especially the so called gradable or evaluative ones), or “zero”, or an arbitrary symmetric fuzzy number. The $\langle \text{linguistic modifier} \rangle$ is a special expression that belongs to a wider linguistic phenomenon called *hedging* and that specifies more closely the topic of utterance. In our case, the linguistic modifier makes the meaning of the

*)In FTT, variables are also taken as formulas. $Form_\alpha$ denotes a set of all formulas of the given type α .

†)The “TE” is a short for “trichotomic evaluative”

⟨TE-adjective⟩ more specific. Quite often it is represented by an intensifying adverb such as “very, roughly, approximately, significantly”, etc.

In this paper, we will use several chosen hedges that can be ordered according to their semantics effect from the most narrowing to the most widening:

$$\begin{aligned} Ex \text{ (extremely)} &\lll Si \text{ (significantly)} \lll Ve \text{ (very)} \\ &\lll \langle \text{empty hedge} \rangle \lll ML \text{ more or less} \lll \\ Ro \text{ (roughly)} &\lll QR \text{ (quite roughly)} \lll \\ &VR \text{ (very roughly)}. \end{aligned} \quad (2)$$

Since there is also the natural ordering Ze (zero) $\lll Sm$ (small) $\lll Me$ (medium) $\lll Bi$ (big), we obtain a linear lexicographic ordering \lll of simple evaluative expressions (1) interpreted as “to be sharper”.

If ⟨linguistic hedge⟩ is missing (expressions such as “weak, large”, etc.) then we understand that an *empty linguistic hedge* is present. Thus, all the simple evaluative expressions have the same form (1). Since they characterize values on an ordered scale, we may consider also scales divided into two parts that are usually interpreted as *positive* and *negative*. Hence, the evaluative expressions may have also a sign, namely “positive” or “negative”.

The evaluative expressions (1) are called *simple*. We may form also *compound* ones using logical connectives (usually “and” and “or”). Note, however, that syntactic and semantic limitations of natural language *prevent the compound evaluative expressions to form a boolean algebra!*

We distinguish abstract evaluative expressions from more specific *evaluative predications*. The latter are expressions of natural language of the form ‘ X is \mathcal{A} ’ where \mathcal{A} is an evaluative expression and X is a variable which stands for objects, for example “degrees of temperature, height, length, speed”, etc. Examples are “temperature is high”, “speed is extremely low”, “quality is very high”, etc. In general, the variable X represents certain features of objects such as “size, volume, force, strength,” etc. and so, its values are often real numbers.

Important notion is that of *linguistic context*^{*}. In our theory it is a triple of (real) numbers $w = \langle v_L, v_S, v_R \rangle$ where v_L is the leftmost typically small value, v_S is typically medium value and v_R is the rightmost typically big value. For example, when speaking about height of people in Europe, we may set $v_L = 140$ cm, $v_S = 170$ cm and $v_R = 220$ cm. In the sequel, we will consider a set of all linguistic contexts

$$W = \{w = \langle v_L, v_S, v_R \rangle \mid v_L, v_S, v_R \in \mathbb{R}, v_L < v_S < v_R\}. \quad (3)$$

The element x belongs to a context $w \in W$ if $x \in [v_L, v_R]$. Then we write $x \in w$.

^{*}In general logic, a more general concept of *possible world* is used instead.

The meaning of an evaluative linguistic expression and predication is represented by its *intension*

$$\text{Int}(X \text{ is } \mathcal{A}) : W \longrightarrow \mathcal{F}(\mathbb{R}) \quad (4)$$

where $\mathcal{F}(\mathbb{R})$ is a set of all fuzzy sets on \mathbb{R} (set of real numbers). The set of intensions of all considered evaluative expressions is denoted by *EvExpr*.

For each context $w \in W$, the *extension* $\text{Ext}_w(X \text{ is } \mathcal{A}) \subseteq \mathbb{R}$ is a specific fuzzy set on \mathbb{R} . Note that $\text{Ext}_w(X \text{ is } \mathcal{A}) = \text{Int}(X \text{ is } \mathcal{A})(w)$.

In formal language, intensions of specific evaluative expressions are denoted by formulas $Ze, Sm \boldsymbol{\nu}, Me \boldsymbol{\nu}, Bi \boldsymbol{\nu} \in \text{Form}_{(o\alpha)\omega}$ where $\boldsymbol{\nu}$ is a linguistic hedge, o is a type of truth value, α arbitrary object and ω is a type of context. For formal reasons, we write hedge on the second place. For example $Sm Ve$ is intension of *very small*. We will usually use a meta-formula $Ev \in \text{Form}_{(o\alpha)\omega}$ to denote intension of an arbitrary evaluative expression.

Important is the function of *local perception* $LPerc : \mathbb{R} \times W \longrightarrow \text{EvExpr}$ defined as follows:

$$LPerc(u, w) = \begin{cases} \text{Int}(X \text{ is } \mathcal{A}) & \text{if } u \in w, \\ \text{undefined} & \text{otherwise.} \end{cases} \quad (5)$$

where \mathcal{A} is the evaluative expression, whose intension is the *sharpest* one in the sense of the lexicographic ordering \lll . Function (5) provides a linguistic characterization of the value $u \in \mathbb{R}$ that can be understood as a certain kind of “measurement” done by people in a concrete situation (modeled by w).

2.2. Intermediate quantifiers

Intermediate quantifiers are expressions of natural language such as *most, many, almost all, a few, a large part of*, etc. A detailed analysis of their meaning was presented in [15]. A mathematical model of their meaning was developed using means of higher-order fuzzy logic in a series of papers [10, 11, 12, 16]. Let us remark that these quantifiers are belong to type $\langle 1, 1 \rangle$ fuzzy generalized quantifiers (cf. [9, 10, 17]).

Intermediate quantifiers have the general form “ QB is A ” and are expressed by means of the following formulas:

$$\begin{aligned} (Q_{Ev}^{\forall} x)(B, A) &:= (\exists z)((\Delta(z \subseteq B) \\ &\quad \& (\forall x)(z x \Rightarrow Ax)) \wedge Ev((\mu B)z)), \quad (6) \\ (Q_{Ev}^{\exists} x)(B, A) &:= (\exists z)((\Delta(z \subseteq B) \\ &\quad \& (\exists x)(z x \wedge Ax)) \wedge Ev((\mu B)z)) \quad (7) \end{aligned}$$

where $z, B, A \in \text{Form}_{o\alpha}$ are formulas of type “fuzzy set”[†], $x \in \text{Form}_{\alpha}$ is a variable of type α . Interpretation of (6) and (7) in a model is simple (informally): we take the largest fuzzy set $z \subseteq B$ such

[†]Their interpretation is a function from a set of objects to a set of truth values.

that each element x having the property represented by z (and so, having also the property B) has also the property A and, at the same time, the size of z w.r.t. B is evaluated by the evaluative expression Ev (for example, it can be *very big*, *not small*, etc.). Few concrete intermediate quantifiers studied in the mentioned papers are the following:

- A:** All B are $A := Q_{Bi\Delta}^{\forall}(B, A)$
(the biggest possible part of B has A)
- P:** Almost all B are $A := Q_{Bi\ Ex}^{\forall}(B, A)$
(extremely big part of B has A)
- B:** Few B are $A := Q_{Bi\ Ex}^{\forall}(B, \neg A)$
(extremely big part of B has no A)
- T:** Most B are $A := Q_{Bi\ Ve}^{\forall}(B, A)$
(very big part of B has A)
- K:** Many B are $A := Q_{\neg(Sm\ \forall)}^{\forall}(B, A)$
(not small part of B has A)
- I:** Some B are $A := Q_{Bi\Delta}^{\exists}(B, A)$

where Δ is a unary connective interpreted by a function $\Delta(a) = 1$ if $a = 1$, otherwise $\Delta(a) = 0$ for all $a \in [0, 1]$.

The quantifiers are linearly ordered as presented in this table. The higher quantifiers are called *superalterns* and lower ones *subalterns*. Thus, for example, **P** is superaltern of **B** and the latter is subaltern of the former.

The size of z w.r.t. B is mathematically characterized by a measure $(\mu B)z$. Example of such measure for the fuzzy set B with finite support is

$$(\mu B)z = \begin{cases} 1 & \text{if } z = B, \\ \frac{|z|}{|B|} & \end{cases} \quad (8)$$

where $|B| = \sum_{x \in \text{Supp}(B)} Bx$ where $\text{Supp}(B)$ is the support of the fuzzy set B^* .

We can also introduce type $\langle 1 \rangle$ intermediate quantifiers:

$$(Q_{Ev}^{\forall} x)Ax := (\forall x)(Ax \ \& \ x \in \text{Supp}(A)) \wedge Ev((\mu V)A), \quad (9)$$

$$(Q_{Ev}^{\exists} x)Ax := (\exists x)(Ax \ \& \ x \in \text{Supp}(A)) \wedge Ev((\mu V)A) \quad (10)$$

where V is a universe.

3. The principle of F-transform

The fuzzy (F-)transform is a universal technique introduced in [18, 19] that has many kinds of applications. Its fundamental idea is to map a bounded

^{*}For precise formulation of the interpretation of formulas (6) and (7) see the cited papers. The evaluation is realized w.r.t. the standard context $\bar{w} = \langle 0, 0.5, 1 \rangle$

continuous function $f : [a, b] \rightarrow \mathbb{R}$ to a finite vector of numbers and then to transform it back. The former is called a *direct F-transform* and the latter an *inverse one*. The result of inverse F-transform is a function \hat{f} that approximates the original function f . Parameters of the F-transform can be set in such a way that the approximating function \hat{f} has desired properties.

The power of F-transform stems from the proof of its approximation abilities, its ability to filter out high frequencies, to reduce noise [20] and the ability to estimate average values of first and second derivatives in a given area [21].

The first step of the F-transform procedure is to form a *fuzzy partition* of the domain $[a, b]$. It consists of a finite set of fuzzy sets

$$\mathcal{A} = \{A_0, \dots, A_n\}, \quad n \geq 2, \quad (11)$$

defined over nodes $a = c_0, \dots, c_n = b$. Properties of the fuzzy sets from \mathcal{A} are specified by five axioms, namely: *normality*, *locality*, *continuity*, *unimodality*, and *orthogonality* that is formally defined by $\sum_{i=0}^n A_i(x) = 1$, $x \in [a, b]$.

A fuzzy partition \mathcal{A} is called *h-uniform* if the nodes c_0, \dots, c_n are *h-equidistant*, i.e., for all $k = 0, \dots, n-1$, $c_{k+1} = c_k + h$, where $h = (b-a)/n$. The membership functions A_0, \dots, A_n of fuzzy sets forming the fuzzy partition \mathcal{A} are usually called *basic functions*.

A *direct F-transform* of a continuous function f is a vector $\mathbf{F}[f] = (F_0[f], \dots, F_n[f])$, where each k -th component $F_k[f]$ is equal to

$$F_k[f] = \frac{\int_a^b f(x)A_k(x) dx}{\int_a^b A_k(x) dx}, \quad k = 0, \dots, n. \quad (12)$$

The *inverse F-transform* of f with respect to $\mathbf{F}[f]$ is a continuous function $\hat{f} : [a, b] \rightarrow \mathbb{R}$ such that

$$\hat{f}(x) = \sum_{k=0}^n F_k[f] \cdot A_k(x), \quad x \in [a, b].$$

All the details and full proofs of many theorems characterizing properties of the F-transform can be found in [18, 19].

The F-transform introduced above is F^0 -transform (i.e., zero-degree F-transform). Its components are real numbers. If we replace them by polynomials of arbitrary degree $m \geq 0$, we arrive at the higher degree F^m transform.

The *direct F^1 -transform* of f with respect to A_1, \dots, A_{n-1} is a vector $F^1[f] = (F_1^1[f], \dots, F_{n-1}^1[f])$ where the components $F_k^1[f]$, $k = 1, \dots, n-1$ are linear functions

$$F_k^1[f](x) = \beta_k^0[f] + \beta_k^1[f](x - c_k) \quad (13)$$

with the coefficients β_k^0, β_k^1 given by

$$\beta_k^0[f] = \frac{\int_{c_{k-1}}^{c_{k+1}} f(x)A_k(x)dx}{\int_{c_{k-1}}^{c_{k+1}} A_k(x)dx}, \quad (14)$$

$$\beta_k^1[f] = \frac{\int_{c_{k-1}}^{c_{k+1}} f(x)(x - c_k)A_k(x)dx}{\int_{c_{k-1}}^{c_{k+1}} (x - c_k)^2 A_k(x)dx}. \quad (15)$$

Note that $\beta_k^0 = F_k[f]$, i.e. the coefficients β_k^0 are just the components of the F^0 transform given in (12).

Theorem 1

If f is four-times continuously differentiable on $[a, b]$ then for each $k = 1, \dots, n - 1$,

$$\beta_k^0[f] = f(c_k) + O(h^2), \quad (16)$$

$$\beta_k^1[f] = f'(c_k) + O(h^2). \quad (17)$$

Thus, the F-transform components provide a weighted average of values of the function f in the area around the node c_k (16), and also a weighted average of slopes (17) of f in the same area.

4. Mining information from time series

4.1. Time series

A time series is a stochastic process (see [22, 23]) $X : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$ where Ω is a set of elementary random events and $\mathbb{T} = \{0, \dots, p\} \subset \mathbb{N}$ is a finite set whose elements are interpreted as time moments. Our basic assumption is that the time series can be decomposed as follows:

$$X(t, \omega) = TC(t) + S(t) + R(t, \omega), \quad t \in \mathbb{T}, \omega \in \Omega, \quad (18)$$

where Ω is a set of elementary random events, $TC(t)$ is a *trend-cycle*, $S(t)$ is a seasonal component that is a mixture of periodic functions and $R(t, \omega)$ is a random noise[†]. It was proved in [20] that the F-transform can be used for extraction of the trend-cycle TC with error close to zero.

Let $\bar{\mathbb{T}} \subseteq \mathbb{T}$ be a subinterval of \mathbb{T} . Then $X|\bar{\mathbb{T}}$ denotes the time series X in (18) restricted to $\bar{\mathbb{T}}$.

4.2. Linguistic evaluation of local trend

The trend-cycle TC of a time series X is the component that represents variations of low frequency in a time series, the high frequency fluctuations having been filtered out. Thus, TC has two subcomponents: *trend (tendency)* and *cycle*. Trend is a slope (a general direction) of the time series, namely whether it is increasing, stagnating, or decreasing. We can distinguish trend of the whole time series but also a *local trend* of $X|\bar{\mathbb{T}}$ in a time interval $\bar{\mathbb{T}} \subseteq \mathbb{T}$.

[†]Each $R(t)$ for $t \in \mathbb{T}$ is a random variable with zero mean and finite variance.

Surprisingly, recognition of trend may not be a trivial task even when watching the graph (cf., e.g., the course of the time series from Figure 1 — what is its trend?). A convenient solution is provided by the F^1 -transform because it provides estimation of the average slope (tangent) of a given function in a specified subset of its domain. Moreover, in [2], we described a method using which we automatically generate evaluation of trend in the form of natural language expressions.

Definition 1

Let $\bar{\mathbb{T}} \subset \mathbb{T}$ be a time interval and A a basic function defined over $\bar{\mathbb{T}}$. Let $X|\bar{\mathbb{T}}$ be a time series restricted to $\bar{\mathbb{T}}$ and $\beta^1[X|\bar{\mathbb{T}}]$ be the coefficient (15) computed using A . Finally, let w_{tg} be context for slope of the time series X . Then intension of the evaluative expression that evaluates trend of X in the time interval $\bar{\mathbb{T}}$ is

$$\text{Tr}(X|\bar{\mathbb{T}}) = LPerc(\beta^1[X|\bar{\mathbb{T}}], w_{tg}). \quad (19)$$

Since (19) is intension, its extension in the context w_{tg} is $\text{Tr}(X|\bar{\mathbb{T}})_{w_{tg}} \subseteq \mathbb{R}$ that is a fuzzy set of real numbers being values of the tangent (15). The way how $\text{Tr}(X|\bar{\mathbb{T}})$ is transformed into the surface form of genuine natural language expressions is described in [2].

General characteristics of the trend is its increase, decrease or stagnation. We will omit precise definition and say informally that $X|\bar{\mathbb{T}}$ is *stagnating* if in each context w_{tg} , $\text{Tr}(X|\bar{\mathbb{T}})$ is *Ze* or *Sm Ex* (extremely small). We will write *Stag*($X|\bar{\mathbb{T}}$) $_{w_{tg}}$. The trend is *increasing (decreasing)* if it is not stagnating and $\beta^1[X|\bar{\mathbb{T}}] > 0$ ($\beta^1[X|\bar{\mathbb{T}}] < 0$).

Finally, we will put

$$\text{Sign}(\text{Tr}(X|\bar{\mathbb{T}})) = \begin{cases} 1, & \text{if } \text{Tr}(X|\bar{\mathbb{T}}) \text{ is increasing,} \\ 0, & \text{if } \text{Tr}(X|\bar{\mathbb{T}}) \text{ is stagnating,} \\ -1, & \text{if } \text{Tr}(X|\bar{\mathbb{T}}) \text{ is decreasing.} \end{cases} \quad (20)$$

Our goal now is to find intervals $\bar{\mathbb{T}}_i$, $i = 1, \dots, s$ in which the time series X has monotonous behavior and to provide linguistic evaluation of their trend. Let us remark that this task belongs to the general problem of time series segmentation (cf. [1]).

First, we will set nodes $c_0, c_1, \dots, c_m \in \mathbb{T}$ where $c_0 = 0$ is the first and $c_m = p$ the last time moment of \mathbb{T} . Then we introduce auxiliary intervals $\bar{t}_j = c_j - c_{j-1}$, $j = 1, \dots, m$. Each interval $\bar{\mathbb{T}}_i \subseteq \mathbb{T}$ will consist of one or more adjacent intervals \bar{t}_j .

The possible algorithm for finding intervals $\bar{\mathbb{T}}$ with monotonous behavior of the time series X is described as Algorithm 1 below. Its idea is simple: we start with the last interval \bar{t}_m . The processed $\bar{\mathbb{T}}_i$ is extended to the left by \bar{t}_j and tested whether its trend has not changed or is different from the trend of the following \bar{t}_{j-1} (to the left). If yes, we take $\bar{\mathbb{T}}_i$ as completed, generate linguistic characterization of the trend of $X|\bar{\mathbb{T}}_i$ and continue with forming next interval $\bar{\mathbb{T}}_{i+1}$.

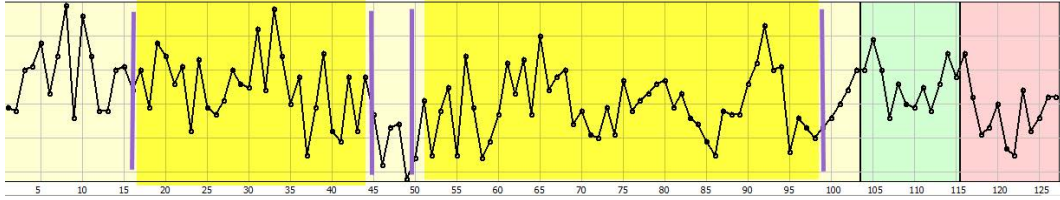


Figure 1: Linguistic evaluation of local trend of time series in the time intervals $\bar{\mathbb{T}}_1 = [17 - 44]$ (*negligible decrease*), $\bar{\mathbb{T}}_2 = [50 - 97]$ (*stagnating*), $\bar{\mathbb{T}}_3 = [104 - 114]$ (*clear decrease*), $\bar{\mathbb{T}}_4 = [116 - 126]$ (*somewhat decrease*). Trend of the whole time series is *stagnating*. The context for the slope is $w_{tg} = \langle 0, 1200/12, 3000/12 \rangle$.

Recall the definition of the property (intension) $Stag(X|\bar{\mathbb{T}})$ of the time series $X|\bar{\mathbb{T}}$ being stagnating in a given context w_{tg} and time interval $\bar{\mathbb{T}}$. We will write its truth value as $\|Stag(X|\bar{\mathbb{T}})_{w_{tg}}\|$.

Furthermore, we define also other properties of the local trend, namely $SIInc, SDec$ (slightly increasing, slightly decreasing)[†] and the truth values $\|SIInc(X|\bar{\mathbb{T}})_{w_{tg}}\|, \|SDec(X|\bar{\mathbb{T}})_{w_{tg}}\|$, and similarly also $NeInc, NeDec$ (negligibly), $SoInc, SoDec$ (somewhat), $ClInc, ClDec$ (clearly), $RoInc, RoDec$ (roughly), $ShInc, ShDec$ (sharply), $SiInc, SiDec$ (significantly). Using techniques described in this

Algorithm 1 Finding intervals of monotonous behavior of time series

```

j := m; i := 1; k := 0;  $\bar{\mathbb{T}}_i^{(k)} := \bar{t}_m$ 
repeat
   $\bar{\mathbb{T}}_i^{(k+1)} := \bar{t}_{j-1} \cup \bar{\mathbb{T}}_i^{(k)}$ 
  if  $\text{Sign}(\text{Tr}(X|\bar{\mathbb{T}}_i^{(k+1)})) \neq \text{Sign}(\text{Tr}(X|\bar{t}_{j-1}))$ 
  or  $\text{Sign}(\text{Tr}([X|\bar{\mathbb{T}}_i^{(k+1)}])) \neq \text{Sign}(\text{Tr}([X|\bar{\mathbb{T}}_i^{(k)}]))$ 
  then
     $\bar{\mathbb{T}}_i := \bar{\mathbb{T}}_i^{(k)}$ 
    print  $\text{Tr}(X|\bar{\mathbb{T}}_i)$ 
    k := 0
    i := i + 1
    j := j - 1
     $\bar{\mathbb{T}}_i^{(k)} := \bar{t}_j$ 
    if j = 1 then
       $\bar{\mathbb{T}}_i := \bar{\mathbb{T}}_i^{(k)}$ 
      print  $\text{Tr}(X|\bar{\mathbb{T}}_i)$ 
    end if
  end if
else
  k := k + 1
  j := j - 1
end if
until j = 1
s := i

```

section, we can get the following linguistic evaluation of the behavior of time series X :

(i) Linguistic evaluation $\text{Tr}(X|\bar{\mathbb{T}})$ of the global

[†]To see how these formulas are obtained, we give precise formal definition of the property “slightly increasing” in the language of FTT: $SIInc := \lambda w_\omega \lambda x_\sigma \lambda x_{\bar{\tau}} \cdot (\Delta(LPerc(x_\sigma x_{\bar{\tau}})w_\omega) \equiv Sm Ve) \& ((Sm Ve)w_\omega)(x_\sigma x_{\bar{\tau}}) \& (x_\sigma x_{\bar{\tau}}) > 0$ where $\sigma, \bar{\tau}, \omega$ are types of the trend of time series, time interval and context, respectively.

trend of the whole time series X .

This is obtained as follows:

- (a) Compute the coefficient $\beta^1[X|\bar{\mathbb{T}}]$, set the context w_{tg} for the slope, and generate the local perception $LPerc(\beta^1[X|\bar{\mathbb{T}}], w_{tg})$. Using translation table defined in [2] find on the surface level an evaluative linguistic expression \mathcal{A} , for example *sharply increasing, slightly decreasing, stagnating*, etc. If $\beta^1[X|\bar{\mathbb{T}}] > 0$ then we the statement concerns increase, in the opposite case it concerns decrease.
- (b) Generate the linguistic statement

Global trend of time series X is \mathcal{A} .

- (ii) A list of linguistic evaluations $\text{Tr}(X|\bar{\mathbb{T}}_i)$, $i = 1, \dots, s$ of the local trend of X in the time intervals $\bar{\mathbb{T}}_i$ generated using Algorithm 1. The linguistic statement is

Local trend of time series X

in the interval $\bar{\mathbb{T}}_i$ is \mathcal{A} .

where \mathcal{A} is obtained in the same way as in item (i) on the basis of the value $\beta^1[X|\bar{\mathbb{T}}_i]$.

- (iii) Linguistic statements such as: “In the first (second, third, fourth) quarter of the year, the trend was slightly decreasing (increasing, stagnating)”.
- (iv) Linguistic statements such as: “The longest period of *stagnating* (increasing, slightly decreasing, etc.) lasted $\bar{\mathbb{T}}$ time moments (days, weeks, months, etc.)” where $\bar{\mathbb{T}}$ is the longest of the intervals $\bar{\mathbb{T}}_1, \dots, \bar{\mathbb{T}}_s$ for which $\text{Tr}(X|\bar{\mathbb{T}}_i)$ is stagnating (sharply increasing, slightly decreasing, etc.).

Linguistic statements both in (iii) as well as in (iv) are obtained analogously as in (i). If we are interested in truth values of the above statements, we can use the formulas $Stag, SIInc, SDec, ShInc$, etc. defined above.

In Figure 1, an example of the generated linguistic evaluations[‡] in 4 different parts of the given time

[‡]The results were obtained using experimental software LFL Forecaster (see http://irafm.osu.cz/en/c110_lfl-forecaster/) which implements the described method.

series is provided. We computed the corresponding coefficients $\beta^1[X|\mathbb{T}_i]$, $i = 1, \dots, 4$ by (15) and generated their linguistic evaluation w.r.t. the context $w_{tg} = \langle 0, 1200/12, 3000/12 \rangle$. One may agree that unambiguous evaluation of the trend when viewing the graph is quite difficult. We argue that our method conforms well with the intuition and makes the evaluation objective.

4.3. Summarization

Summarization of knowledge about time series belongs among interesting tasks that was addressed by several authors (see, e.g., [24, 25]). In this paper, we suggest to apply the sophisticated formal theory of *intermediate quantifiers*, which is a constituent of FNL. We can summarize information within one time series or over a set of time series.

4.3.1. One time series

First we explain how are summarizing linguistic statements interpreted. Let us consider the following one:

(*) **In *most* cases, trend of the time series was *slightly increasing*.**

This proposition can be formalized by the following formula:

$$(Q_{Bi\ Ve}^{\forall \mathbb{T}})(SIInc(\text{Tr}(X|\mathbb{T}_i))) \quad (21)$$

where in (21), the quantifier \mathbf{T} from Section 2.2 restricted to type $\langle 1 \rangle$ (cf. definition (9)) is used.

Interpretation of (21) is the following: The universe is given by the set of time intervals $\mathcal{T} = \{\mathbb{T}_i \mid i = 1, \dots, s\}$. For all \mathbb{T}_i we find truth value $\|SIInc(\text{Tr}(X|\mathbb{T}_i))\|$, determine the support $\text{Supp}(SIInc(\text{Tr}(X|\mathbb{T}_i)))$, compute the relative cardinality (8)

$$(\mu V)(SIInc(\text{Tr}(X|\mathbb{T}_i))) = \frac{\sum_{i=1}^s \|SIInc(\text{Tr}(X|\mathbb{T}_i))\|}{s}$$

and determine the truth value $\|Bi\ Ve((\mu V)SIInc(\text{Tr}(X|\mathbb{T}_i)))\|$ (i.e., the truth value of the statement “the relative cardinality $(\mu V)(SIInc(\text{Tr}(X|\mathbb{T}_i)))$ is *very big*” in the standard context $w = \langle 0, 0.5, 1 \rangle$). The truth value of (21) is then obtained as

$$\min\{\|SIInc(\text{Tr}(X|\mathbb{T}_i))\|, \|Bi\ Ve((\mu V)(SIInc(\text{Tr}(X|\mathbb{T}_i))))\|\}. \quad (22)$$

Mining summarizing information in natural language of the form (*) proceeds as follows. We compute truth value of the formula (21) for each subformula *Stag*, *SIInc*, *SIDec*, *NeInc*, *NeDec*, *ShInc*, *ShDec*, etc. and for each intermediate quantifier \mathbf{A} , \mathbf{P} , \mathbf{B} , \mathbf{T} , \mathbf{K} , \mathbf{I} considered in Section 2.2. The resulting mined information is a set of formulas

(21) determined by subformulas from the former list and the highest superaltern from the latter list, provided that the corresponding truth value (22) is sufficiently high (higher than a given threshold). Finally, each mined formula is transformed into the corresponding linguistic statement of the form (*), for example “In *most* (*many*, *few*) cases, the time series was *stagnating* (*sharply decreasing*, *roughly increasing*)”, etc..

We can also consider a special quantifier *often* that can be modeled by $(Q_{Bi}^{\forall \mathbb{T}})$. This enables us to mine information such as “In the past few years, the trend was *often stagnating* (*increasing*, *negligibly decreasing*), etc.

4.3.2. Set of time series and syllogistic reasoning

Let a set $\{X_i \mid i = 1, \dots, r\}$ of time series (18) be given. Analogously as in the previous section, we can model the meaning of sentences, such as

- “*Most* (*many*, *few*) analyzed time series stagnated recently but their future trend is *slightly increasing*.”
- “There is an evidence of *huge* (*slight*, *clear*) decrease of trend of almost all time series in the recent quarter of the year.”

The corresponding formulas are constructed similarly as above. The difference is that the quantified variable runs over the set of time series. It is also possible to combine more quantifiers, i.e., to mine linguistic information, for example “*Few* analyzed time series *often stagnated* during the past n years.”

There are unlimited possibilities for mine interesting information from the given set of time series especially if we include also forecast of their future development. Then we can generate comments to interesting time intervals, or we can also determine time intervals in which behavior of the time series is interesting for us, for example, “in which period was the time series sharply increasing”, “how long was the time series stagnating or decreasing before sharp increase”, etc.

Exciting possibility is to derive new information on the basis of the formal theory of generalized syllogistic reasoning developed, e.g., in [11, 12] where validity of over 120 generalized Aristotle’s syllogisms with intermediate quantifiers was formally proved.

Example of a valid syllogism:

No sharply increasing trend is long

Most interesting trends are long

Most interesting trends are not sharply increasing

It is important to note that if a syllogism is valid then it is true in all models. In our case this means that it is true for arbitrary time series and in arbitrary context.

5. Conclusion

In this paper, we focused on the problem of mining information in the form of sentences of natural language from time series. Not that there are many methods for mining information from time series but very few providing it in natural language (see [1] and the citations therein). Our main contribution consists in utilization of the coherent techniques of F-transform and fuzzy natural logic. With the help of them we are able to disclose various kinds of information and put them together using natural language. Strength and reliability of these techniques are based on their properties proved mathematically.

Our future work will be focused on refinement and extension of the above presented techniques. Notice that precise and detailed formal analysis of otherwise relatively simple statements is quite complicated.

Acknowledgment

This paper was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

References

- [1] T-C. Fu. A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24:164–181, 2011.
- [2] V. Novák, V. Pavliska, I. Perfilieva, and M. Štěpnička. F-transform and fuzzy natural logic in time series analysis. In *Proc. Int. Conference EUSFLAT-LFA '2013*, Milano, Italy, 2013.
- [3] V. Novák, V. Pavliska, M. Štěpnička, and L. Štěpničková. Time series trend extraction and its linguistic evaluation using F-transform and fuzzy natural logic. In L.A. Zadeh, A.M. Abbasov, R.R. Yager, and S.N. Shahbazova, editors, *Recent Developments and New Directions in Soft Computing*, pages 429–442. Springer, Berlin, 2014.
- [4] V. Novák and I. Perfilieva. Time Series Mining by Fuzzy Natural Logic and F-transform. In Tung X. Bui and Jr. Ralph H. Sprague, editors, *Proc. 48th Hawaii International Conference on System Sciences*, pages 1493–1502, Kauai, U.S.A., 2015. IEEE Computer Society, CPS.
- [5] V. Novák. A comprehensive theory of trichotomous evaluative linguistic expressions. *Fuzzy Sets and Systems*, 159(22):2939–2969, 2008.
- [6] V. Novák. Mathematical fuzzy logic in modeling of natural language semantics. In P. Wang, D. Ruan, and E.E. Kerre, editors, *Fuzzy Logic – A Spectrum of Theoretical & Practical Issues*, pages 145–182. Elsevier, Berlin, 2007.
- [7] V. Novák. Perception-based logical deduction. In B. Reusch, editor, *Computational Intelligence, Theory and Applications*, pages 237–250. Springer, Berlin, 2005.
- [8] V. Novák and S. Lehmke. Logical structure of fuzzy IF-THEN rules. *Fuzzy Sets and Systems*, 157:2003–2029, 2006.
- [9] A. Dvořák and M. Holčapek. L-fuzzy quantifiers of the type $\langle 1 \rangle$ determined by measures. *Fuzzy Sets and Systems*, 160:3425–3452, 2009.
- [10] V. Novák. A formal theory of intermediate quantifiers. *Fuzzy Sets and Systems*, 159(10):1229–1246, 2008.
- [11] P. Murinová and V. Novák. A formal theory of generalized intermediate syllogisms. *Fuzzy Sets and Systems*, 186:47–80, 2012.
- [12] P. Murinová and V. Novák. The structure of generalized intermediate syllogisms. *Fuzzy Sets and Systems*, 247:18–37, 2014.
- [13] V. Novák. On fuzzy type theory. *Fuzzy Sets and Systems*, 149:235–273, 2005.
- [14] V. Novák. EQ-algebra-based fuzzy type theory and its extensions. *Logic Journal of the IGPL*, 19:512–542, 2011.
- [15] P.L. Peterson. *Intermediate Quantifiers. Logic, linguistics, and Aristotelian semantics*. Ashgate, Aldershot, 2000.
- [16] P. Murinová and V. Novák. Analysis of generalized square of opposition with intermediate quantifiers. *Fuzzy Sets and Systems*, 242:89–113, 2014.
- [17] M. Delgado, M.D. Ruiza, D. Sanchez, and M.A. Vila. Fuzzy quantification: a state of the art. *Fuzzy Sets and Systems*, 242:1–30, 2014.
- [18] I. Perfilieva. Fuzzy transforms: theory and applications. *Fuzzy Sets and Systems*, 157:993–1023, 2006.
- [19] I. Perfilieva, M. Daňková, and B. Bede. Towards a higher degree F-transform. *Fuzzy Sets and Systems*, 180:3–19, 2011.
- [20] V. Novák, I. Perfilieva, M. Holčapek, and V. Kreinovich. Filtering out high frequencies in time series using F-transform. *Information Sciences*, 274:192–209, 2014.
- [21] V. Kreinovich and I. Perfilieva. Fuzzy transforms of higher order approximate derivatives: A theorem. *Fuzzy Sets and Systems*, 180:55–68, 2011.
- [22] J. Anděl. *Statistical Analysis of Time Series*. SNTL, Praha, 1976 (in Czech).
- [23] J.D. Hamilton. *Time Series Analysis*. Princeton, Princeton University Press, 1994.
- [24] R. Castillo-Ortega, N. Marín, and D. Sánchez. A fuzzy approach to the linguistic summarization of time series. *Multiple-Valued Logic and Soft Computing*, 17(2-3):157–182, 2011.
- [25] J. Kacprzyk, A. Wilbik, and Zadrożny. Linguistic summarization of time series using a fuzzy quantifier driven aggregation. *Fuzzy Sets and Systems*, 159:1485–1499, 2008.