On Type-2 Fuzzy Logic and Linguistic Summarization of Databases

Abstract

Type-2 fuzzy logic is a generalization of fuzzy logic by Zadeh [16], [17]. The keypoint of this generalization is that membership degrees in type-2 fuzzy sets (T2FSs) are traditional fuzzy sets in [0, 1], while in traditional fuzzy sets (type-1 fuzzy sets, T1FSs), membership degrees are real numbers in [0, 1]. This provides more versatile means for modelling imprecise information, e.g. natural language expressions.

Computational methods based on T2FSs are broadly applied in many domains related to computer science. In this approach, we present the calculus of linguistically quantified statements in which linguistic expressions are represented by T2FSs. This provides techniques to define and generate linguistic summaries of databases in the sense of Yager [14] but based on T2FSs. Linguistic summaries are natural language sentences generated by dedicated software from large databases (e.g. a few million of records). Original elements are introduced by the author in a generalization of the calculus of linguistically quantified statements from type-1 to type-2, and definitions of cardinalities of T2FSs which cover real cardinalities of T1FSs. Hence, a point of departure to formalize and to apply type-2 linguistic summaries of databases, their generation, their degrees of truth as measures of their informativeness, is presented.

Keywords: type-2 fuzzy sets, type-2 fuzzy logic, type-2 membership functions, linguistically quantified statements, type-2 fuzzy quantifiers, fuzzy support of a type-2 fuzzy set, type-2 linguistic summaries of databases
1. Type-2 fuzzy sets and type-2 fuzzy logic

The concept of a Type-2 Fuzzy Set (T2FS) was introduced by Zadeh in 1975 [16]. Karnik and Mendel developed this idea and presented the foundations of type-2 fuzzy logic systems [6], [9]. Various applications and operations on T2FSs are discussed [2], [5], [8], [11]. Besides, classes of T2FSs, like Interval or Triangular T2FSs, are distinguished [7], [12].

A T2FS \( \tilde{A} \) in a finite universe of discourse \( \mathcal{X} \) is defined as:

\[
\tilde{A} = \sum_{x \in \mathcal{X}} \mu_{\tilde{A}}(x)/x
\]  \hspace{1cm} (1.1)

where the membership degree \( \mu_{\tilde{A}}(x) \) is a T1FS in \([0, 1]\). Eq. (1.1) may apply to an infinite \( \mathcal{X} \), if \( \sum \) is replaced with \( \int \). Another useful notation of a T2FS \( \tilde{A} \) in \( \mathcal{X} \) is:

\[
\tilde{A} = \{(x, u_{\tilde{A}}, \mu_x(u_{\tilde{A}})) : x \in \mathcal{X}, u_{\tilde{A}} \in J_x \} \] \hspace{1cm} (1.2)

in which \( J_x \subseteq [0, 1] \) is the set of all primary membership degrees of \( x \) to \( \tilde{A} \), and \( \mu_x(u_{\tilde{A}}) \in [0, 1] \) denotes the secondary membership degree of \( x \) to \( \tilde{A} \).

The point of using fuzzy membership degrees, instead of real, is that type-1 membership functions may not support many kinds of uncertainty that appear in natural language and in other sources of imperfect information, like measurements or human perceptions. Fuzzy membership degrees express a measure of dispersion for real-valued membership functions, similarly to the sense of standard deviation (second-order-moment) which describes, additionally, the mean (first-order-moment) in statistic. Primary and secondary membership functions and degrees are presented in Fig. 1.

The complement of a T2FS \( \tilde{A} \) in \( \mathcal{X} \) is denoted by \( \tilde{A}^c \) and is a T2FS in \( \mathcal{X} \) with the membership function defined as

\[
\mu_{\tilde{A}^c}(x) = \sum_{u_{\tilde{A}} \in J_x} \mu_x(u_{\tilde{A}})/(1 - u_{\tilde{A}})
\] \hspace{1cm} (1.3)

where \( u_{\tilde{A}} \) is a primary, and \( \mu_x(u_{\tilde{A}}) \) is a secondary membership degree of \( x \) to \( \tilde{A} \). This definition is consistent with the complement of a T1FS.

Let \( \tilde{A}, \tilde{B} \) be T2FSs in \( \mathcal{X} \). Let \( t_1, t_2 \) be \( t \)-norms, and \( s \) be a \( t \)-conorm. Union of \( \tilde{A} \) and \( \tilde{B} \) is the T2FS in \( \mathcal{X} \), denoted as \( \tilde{A} \cup \tilde{B} \):

\[
\mu_{\tilde{A} \cup \tilde{B}}(x) = \sum_{u_{\tilde{A}}} \sum_{u_{\tilde{B}}} (\mu_x(u_{\tilde{A}}) t_1 \mu_x(u_{\tilde{B}}))/(u_{\tilde{A}} s u_{\tilde{B}})
\] \hspace{1cm} (1.4)
\[ J_{x} \]

\[ \begin{align*}
\mu_{\tilde{A} \cap B}(x) &= \sum_{u_{\tilde{A}}} \sum_{u_{B}} \left( \mu_{x}(u_{\tilde{A}}) \mu_{x}(u_{B}) \right) / \left( u_{\tilde{A}} t_{1} u_{B} \right) \\
\end{align*} \]  

(1.5)

The intersection and union operations for T2FSs are commutative, idempotent, and associative. They cover analogous operations for T1FSs, hence, they are generalizations of traditional \( \cup, \cap, \) and \( \cdot \), cf. [15].

**Cylindric extensions of type-2 fuzzy sets** Let \( \mathcal{X}_{1}, \ldots, \mathcal{X}_{N} \) are universes of discourse, and \( \tilde{A} \) is a type-2 fuzzy set in \( \mathcal{X}_{j} \), \( j \in \{1, \ldots, N\} \). The **cylindric extension** of \( \tilde{A} \) to \( \mathcal{X}_{1} \times \ldots \times \mathcal{X}_{N} \), is the type-2 fuzzy set \( ce(\tilde{A}) \) in \( \mathcal{X}_{1} \times \ldots \times \mathcal{X}_{N} \), defined as:

\[ ce(\tilde{A}) = \sum_{x_{1}, \ldots, x_{N} \in \mathcal{X}_{1} \times \ldots \times \mathcal{X}_{N}} \mu_{ce(\tilde{A})}(x_{1}, \ldots, x_{N}) / (x_{1}, \ldots, x_{N}), x_{1} \in \mathcal{X}_{1}, \ldots, x_{N} \in \mathcal{X}_{N} \]  

(1.6)

where \( \mu_{ce(\tilde{A})}(x_{1}, \ldots, x_{N}) = \mu_{\tilde{A}}(x_{j}), x_{j} \in \mathcal{X}_{j} \).

**Embedded fuzzy sets** For each T2FS \( \tilde{A} \) in \( \mathcal{X} \), one can define **embedded type-2 fuzzy sets** and **embedded T1FSs** [9]. Let \( \forall x \in \mathcal{X} \delta_{x} \in J_{x} \subseteq [0,1] \). An
Fig. 2. A footprint of uncertainty (FOU) and lower and upper membership functions of a T2FS in $\mathbb{R}$.

The embedded T2FS $\tilde{A}_\delta$ in $\tilde{A}$ is defined: $\mu_{\tilde{A}_\delta}(x) = \mu_x(\delta_x) / \delta_x$, $\delta_x \in J_x$, where $\mu_x$ is the secondary membership function for a fixed $x \in X$.

An embedded T1FS $A_\delta$ for a T2FS $\tilde{A}$ in $X$ is defined: let $\forall x \in X \delta_x \in J_x \subseteq [0, 1]$. The membership function for $A_\delta$ is given as $\mu_{A_\delta}(x) = \delta_x$.

**Footprint of uncertainty, lower and upper membership functions**

We denote $J_X = \bigcup_{x \in X} J_x$. The footprint of uncertainty (FOU) of $\tilde{A}$ is a crisp set in $X \times J_X$ consisting of all pairs $\langle x, u_{\tilde{A}} \rangle$, such that the secondary membership degrees for $u_{\tilde{A}}$ are greater than zero:

$$FOU(\tilde{A}) = \{ \langle x, u_{\tilde{A}} \rangle : x \in X, u_{\tilde{A}} \in J_x, \mu_x(u_{\tilde{A}}) > 0 \}$$  \hspace{1cm} (1.7)

The lower membership function (LMF) of $\tilde{A}$ and upper membership function (UMF) of $\tilde{A}$ are just lower and upper bounds of $FOU(\tilde{A})$, respectively:

$$LMF(\tilde{A}) = \{ \langle x, u_{\tilde{A}} \rangle : x \in X, u_{\tilde{A}} = \inf J_x \}$$  \hspace{1cm} (1.8)

$$UMF(\tilde{A}) = \{ \langle x, u_{\tilde{A}} \rangle : x \in X, u_{\tilde{A}} = \sup J_x \}$$  \hspace{1cm} (1.9)

See Fig. 2, in which $FOU$, $LMF$, and $UMF$ of a sample T2FS are shown. Notice, that both $LMF(\tilde{A})$ and $UMF(\tilde{A})$ are embedded T1FSs in $\tilde{A}$. 
Cardinalities of type-2 fuzzy sets

In general, cardinalities of T2FSs should be consistent with cardinalities of T1FSs. The most popular real (scalar) cardinality of a type-1 fuzzy sets $A$ in a finite $X$ is \textit{sigma-count}: $\Sigma \text{Count}(A) = \sum_{x \in X} \mu_A(x)$ [3]. Hence, the definition of the \textit{non-fuzzy sigma-count of $A$} in $X$, $n\sigma$-count($\tilde{A}$), is now proposed, based on [4]:

$$|\tilde{A}| = n\sigma\text{-count}(\tilde{A}) = \sum_{x \in X} \sup \{u_x \in J_x : \mu_x(u_x) = 1\} \quad (1.10)$$

When all secondary membership functions in $\tilde{A}$ have unique maxima, the supremum in (1.10) can be omitted. What is the most important, for a T1FS $A$, $n\sigma$-count($A$) reduces to $\Sigma \text{Count}(A)$.

Another scalar cardinality is proposed for Interval T2FSs [13]:

$$|\tilde{A}| = \frac{1}{2} \sum_{x \in X} (LMF_{\tilde{A}}(x) + UMF_{\tilde{A}}(x)) \quad (1.11)$$

Eq. (1.11) takes into account the mean of lower and upper membership degrees of $x$, which helps to avoid misleading representation of an Interval T2FS by its $UMF$ only (as (1.10) could suggest). This cardinality also reduces to $\Sigma \text{Count}(A)$, if $A$ is a T1FS.

A different concept of a scalar cardinality for a T2FS is based on $\alpha$-cuts of T1FSs determined by secondary membership functions in $\tilde{A}$:

$$|\tilde{A}| = \frac{1}{2} \sum_{x \in X} \left( \inf \{u_x \in J_x : \mu_x(u_x) \geq \alpha\} + \sup \{u_x \in J_x : \mu_x(u_x) \geq \alpha\} \right) \quad (1.12)$$

The version using strong $\alpha$-cuts, i.e. $A_{\alpha} = \{x \in X : \mu_A(x) > \alpha\}$, can also be taken into account. We notice that (1.12) for $\alpha = 1$ extends (1.10) if only all secondary membership functions in $\tilde{A}$ are normal and have regular maxima in $[0, 1]$. Besides, one may also use the weighted average in (1.12):

$$|\tilde{A}| = \sum_{x \in X} \left( w_1 \inf \{u_x \in J_x : \mu_x(u_x) > \alpha\} + w_2 \sup \{u_x \in J_x : \mu_x(u_x) > \alpha\} \right) \quad (1.13)$$

where $w_1 + w_2 = 1$, $w_1, w_2 \in [0, 1]$. As far as (1.11), also (1.12) for $\alpha = 1$ reduces to $\Sigma \text{Count}$.

The calculus of linguistically quantified statements uses \textit{relative cardinality} of a finite T1FS $A$ in $X$ with respect to a finite T1FS $B$ in $X$:

$^1$A T1FS $A$ in $X$ is normal iff $\sup_{x \in X} \mu_A(x) = 1$. 

---

\textbf{On Type-2 Fuzzy Logic and Linguistic Summarization of Databases} 219
\[ \sum \text{Count}(A \mid B) = \frac{\sum \text{Count}(A \cap B)}{\text{Count}(B)} \]

Analogously, we propose the definition of relative cardinality of \( \tilde{A} \) w.r.t. \( \tilde{B} \):

\[ |\tilde{A} \cap \tilde{B}| = \frac{|\tilde{A} \cap \tilde{B}|}{|\tilde{B}|}, \quad |\tilde{B}| \neq 0 \quad (1.14) \]

in which \( \tilde{A} \cap \tilde{B} \) is obtained via (1.5). The symbol \( |\cdot| \) can be interpreted in terms of any of (1.10), (1.11), (1.12), or (1.13). We apply \( |\tilde{A} \cap \tilde{B}| \) in evaluating degrees of truth of linguistically quantified statements, see (1.19).

For a T2FS \( \tilde{A} \) in a finite \( \mathcal{X} \), the inequality

\[ 0 \leq |\tilde{A}| \leq |\mathcal{X}| \quad (1.15) \]

holds. Again, \( |\cdot| \) can be interpreted in terms of one of (1.10) \( \div \) (1.13).

**Normal and convex type-2 fuzzy sets** Let \( \tilde{A} \) be a T2FS in \( \mathcal{X} \). \( \tilde{A} \) is normal iff

\[ \exists x' \in \mathcal{X} \quad u_{\tilde{A}} = 1 \land \mu_{x'}(u_{\tilde{A}} = 1) = 1, \quad u_{\tilde{A}} \in J_{x'} \quad (1.16) \]

This definition covers also the definition of normal T1FSs; the proof for this was presented in [10], [11].

**Convexity** of a T2FS in \( \mathbb{R} \), is based on embedded T1FSs. We denote the set of all non-zero secondary membership degrees of a T2FS \( \tilde{A} \) in \( \mathcal{X} \) as

\[ \mu_{\mathcal{X}}(\tilde{A}) = \{ r \in (0, 1] \mid \exists x \in \mathcal{X} \land \exists u_{\tilde{A}} \in J_x, \mu_x(u_{\tilde{A}}) = r \} \]

Let \( \tilde{A} \) be a T2FS in \( \mathbb{R} \). \( \tilde{A} \) is convex iff for each \( c \in \mu_{\mathcal{X}}(\tilde{A}) \) the embedded T1FSs in \( \tilde{A} \) such that

\[ \mu_{\text{min}}(x) = \min\{ u_{\tilde{A}} \in J_x : \mu_x(u_{\tilde{A}}) = c \} \]
\[ \mu_{\text{max}}(x) = \max\{ u_{\tilde{A}} \in J_x : \mu_x(u_{\tilde{A}}) = c \} \quad (1.17) \]

are convex\(^2\). (1.17) extends convexity for T1FSs, for the proof – see [10], [11].

**Type-2 linguistically quantified statements** In the Zadeh calculus, two forms of linguistically quantified statements are considered. We enhance this calculus using the same forms, but we represent linguistic terms by T2FSs, not by T1FS. Let \( \tilde{S}_1, \tilde{S}_2 \) be linguistic expressions represented

\(^2\)A T1FS \( A \) in \( \mathcal{X} \) is convex iff each its \( \alpha \)-cut is convex in the classical sense.
by type-2 fuzzy sets in finite $\mathcal{X}$, and $\tilde{Q}$ — a linguistic quantifier represented by a normal and convex T2FS in $\mathcal{X}_{\tilde{Q}} \subseteq \mathbb{R}^+ \cup \{0\}$. The formula

$$\tilde{Q} \, x's \, are \, \tilde{S}_1$$

is the first form ($\tilde{Q}^I$) of the type-2 quantified statement, and

$$\tilde{Q} \, x's \, being \, \tilde{S}_2 \, are \, \tilde{S}_1$$

is the second form ($\tilde{Q}^{II}$). Degrees of truth of (1.18), (1.19) are T1FSs in $[0, 1]$ computed as values of the type-2 membership functions $\mu_{\tilde{Q}}$:

$$T\left(\tilde{Q} \, x's \, are \, \tilde{S}_1\right) = \mu_{\tilde{Q}} \left(\frac{|\tilde{S}_1|}{M}\right)$$

where the cardinality of $\tilde{S}_1$ in $\mathcal{X}$ is evaluated by one of (1.10)÷(1.13), $M = |\mathcal{X}|$ if $\tilde{Q}$ is relative\(^3\), or $M = 1$ if $\tilde{Q}$ is absolute, and

$$T\left(\tilde{Q} \, x's \, being \, \tilde{S}_2 \, are \, \tilde{S}_1\right) = \mu_{\tilde{Q}} \left(\frac{|\tilde{S}_1 \cap \tilde{S}_2|}{|\tilde{S}_2|}\right)$$

We notice that (1.20), (1.21) reduce to their counterparts for T1FSs, cf. [17].

2. Type-2 linguistic summaries of databases

A type-2 linguistic summary of a database is a quasi-natural sentence [10]

$$\tilde{Q} \, P \, are/have \, \tilde{S} \, [T]$$

where the symbols are interpreted as follows:

- $\tilde{Q}$ is a linguistic quantifier represented by a T2FS.
- $P$ is the subject of the summary.
- $\tilde{S}$ is a summarizer represented by a T2FS.
- $T$ is a fuzzy set in $[0, 1]$ – the degree of truth of the summary.

We consider a classic database model proposed by Codd [1]. Let $\mathcal{Y} = \{y_1, \ldots, y_m\}$ be a set of objects, e.g. cars. Let $V_1, \ldots, V_n$ be attributes

\(^3\)Properties "absolute" and "relative" for type-2 fuzzy quantifiers can be easily defined on the base of the definitions of these properties for type-1 fuzzy quantifiers, cf. [11], [17].
manifested by the objects from $\mathcal{Y}$, e.g. price, age. Let $\mathcal{X}_1, \ldots, \mathcal{X}_n$ be the domains of $V_1, \ldots, V_n$, respectively. We denote a value of the attribute $V_j$ for object $y_i$ as $V_j(y_i)$, $i \leq m$, $j \leq n$, e.g. let $V_j = \text{Age}$, $y_i = \text{'Opel'}$, $V_j(y_i) = 15 \in \mathcal{X}_j$. Hence, the database $\mathcal{D}$ collecting information about elements from $\mathcal{Y}$, is in the form of: $\mathcal{D}^T = \{d_1, \ldots, d_m\}$, where $d_i = (V_1(y_i), \ldots, V_n(y_i)) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$, $i = 1, \ldots, m$. $d_i$ is a record describing object $y_i$.

Type-2 summarizers The goal is to find the degree of truth of a summary (2.1), based on (1.18). The summarizer $\tilde{S}$ is represented by a T2FS in $\mathcal{X}_j$, but, to be more specific, we use the cylindric extension of $\tilde{S}_j$ to $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$. The degree of truth is a T1FS $T$ in $[0, 1]$:

$$T\left(\tilde{Q} \ P \ \text{are/have} \ \tilde{S}\right) = \mu_{\tilde{Q}}\left(\frac{|\text{oc}(\tilde{S}) \cap \mathcal{D}|}{M}\right) \tag{2.2}\label{2.2}$$

where $M = 1$ if $\tilde{Q}$ is absolute, or $M = m = |\mathcal{D}|$ – the number of records in the summarized database, if $\tilde{Q}$ is relative.

Eq. (2.2) presents $T$ for summaries in which $\tilde{S}$ is represented by single T2FS, related to one of $V_1, \ldots, V_n$. But summarizers of type-2 can also be based on more than one attribute, e.g. young and well educated. Hence, if $n$ T2FSs $\tilde{S}_1, \ldots, \tilde{S}_n$ are chosen to represent the type-2 summarizer, its membership function is: $\tilde{S} = \tilde{S}_1 \text{ AND/} \text{OR} \tilde{S}_2 \text{ AND/} \text{OR} \cdots \text{ AND/} \text{OR} \tilde{S}_n$ in which OR, AND are given by (1.4) and (1.5). Thus, the degree of truth is similar to (2.2): $T = \mu_{\tilde{Q}}\left(\frac{|\text{oc}(\tilde{S}) \cap \mathcal{D}|}{M}\right)$.

Type-2 qualifiers Linguistic summaries based on the second form of a linguistically quantified statement (1.19), are now described. In particular, we are interested in summaries

$$\tilde{Q} \ P \ \text{being/having} \ \tilde{W} \ \text{are/have} \ \tilde{S} \ [T] \tag{2.3}\label{2.3}$$

in which $\tilde{W}$ – a type-2 qualifier – is represented by a T2FS in $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$;

$$\tilde{W} = \tilde{W}_{g_1} \text{ AND/} \text{OR} \cdots \text{ AND/} \text{OR} \tilde{W}_{g_s} \tag{2.4}\label{2.4}$$

where $g_1, \ldots, g_s \in \{1, \ldots, n\}$, and usually $g_1 \neq g_2 \neq \cdots \neq g_s$. A sample summary in the form of (2.3) is: Many fast and well-equipped cars are
expensive, in which: $\mathcal{Q}=$ Many, $\mathcal{W}=$ fast and well-equipped, $\mathcal{S}=$ expensive. Hence, the degree of truth is:

$$T\left(\mathcal{Q} P \text{ being/having } \mathcal{W} \text{ are/have } \mathcal{S}\right) = \mu_{\mathcal{Q}}\left(\frac{|\mathcal{S} \cap \mathcal{W} \cap \mathcal{D}|}{|\mathcal{W} \cap \mathcal{D}|}\right)$$

(2.5)

where $\mu_{\mathcal{S} \cap \mathcal{W} \cap \mathcal{D}}$ is given as $\mu_{\mathcal{S} \cap \mathcal{W} \mid \mathcal{D}}: \mathcal{D} \rightarrow \mathcal{F}([0, 1])$ and $(\mu_{\mathcal{S} \cap \mathcal{W} \mid \mathcal{D}})(d_i) = \mu_{\mathcal{S} \cap \mathcal{W} \mid \mathcal{D}}(d_i)$ which means that $\mathcal{S} \cap \mathcal{W}$ is considered in the set of records $\mathcal{D}$, $\mu_{\mathcal{W} \cap \mathcal{D}}$ is determined analogously, and $|\cdot|$ is evaluated via one of (1.10)$^\div$(1.13).

3. Sample application: generation of press news

The task of the experiment is to generate short textual messages in natural language from a large set of numerical data. In other words, one intends to obtain a description in natural language of the content of a database, rather than its statistical analysis, for instance. The general schema of the system is depicted in Fig. 3. The input is divided into three parts:

- database, $\mathcal{D}$ – the set of numerical data which is being analyzed by the summaries generator; the records in $\mathcal{D}$ represent objects, e.g. cars, books, etc.;
• values/attributes of interest provided by users – linguistically expressed properties possessed by objects/records, e.g. fast cars, expensive books, etc., (without fuzzy sets related to them);
• expert knowledge – type-2 fuzzy sets provided by experts as models of linguistic expressions, like many, few, fast (cars), expensive (books).

The output News ready for publishing can be modified and/or adjusted by human experts and, then, published as memos, news, RSS messages, etc. Similarly to systems that support, e.g. medical diagnosis, each automatically suggested decision or diagnosis must be corrected or verified by a human.

Assume that a user provided \( z \in \mathbb{N} \) linguistic values which are interesting for him/her because of some particular points of view or reasons. Assume, that the expert knowledge block suggests using \( k \in \mathbb{N} \) linguistic quantifiers \( \tilde{Q}_1, \ldots, \tilde{Q}_k \). Hence, the general number of summaries to be generated is:

\[
k \sum_{i=0}^{z-1} \binom{z}{i} (2^{z-i} - 1)
\]  (3.1)

We denote the values of interest chosen by a user \( \tilde{S}_1, \ldots, \tilde{S}_z \); they are represented by type-2 fuzzy sets in \( \mathcal{X}_1, \ldots, \mathcal{X}_n \). Therefore, the algorithm generating linguistic summaries on the base of these values is (a fragment):

```csharp
// generation of summaries via \( \tilde{Q}_i \)
1. for each non-empty \( \tilde{S} \subseteq \{ \tilde{S}_1, \ldots, \tilde{S}_z \} \)
   1.1. determine \( \mu_{\tilde{S}}(d_i), \ i = 1, \ldots, m \)
   1.2. for each quantifier \( \tilde{Q}_h, \ h = 1, \ldots, k \)
       if \( \tilde{Q}_h \) is absolute
           compute: \( T_{1,h} \) via (2.2) for \( M = 1 \)
       else // i.e. if \( \tilde{Q}_h \) is relative
           compute: \( T_{1,h} \) via (2.2) for \( M = m = |\mathcal{D}| \)
   1.3. generate the summary \( \tilde{Q}_{h_{\text{max}}} P \) are/have \( \tilde{S} \) \( [T] \)

...```

The algorithm and the prototype of the news generator has been implemented on .NET platform in the language C#. The database, implemented in MS Access, consisted of about 10 000 records collecting data on workers of a company. In particular, we used database view (a read-only table combined of several working tables) which included records
(Age, Education, Salary, Gender). Experts proposed the linguistic values of the attributes $V_1$ = Age, ..., $V_4$ = Gender. Several linguistic values were assigned to each attribute: e.g., $V_1$ = Age could be described by {young, middle-aged, experienced, about 40, about 30}, or $V_4$ = Gender - by {male, female}; these linguistic values were represented by Interval T2FSs. Sample linguistic quantifiers were: "About half", "Much more than 2000", and "Many". A sample news generated via the described algorithm for $z = 3$, $m = 10000$, and $S_1$ = about 30, $S_2$ = high school, and $S_3$ = about $40000$, is presented:

About half of workers are about 30 [0.58, 0.61]. Much more than 2000 workers graduated from high school [0.74, 0.74]. About half of workers earn about $40000 [0.53, 0.53]. Many workers graduated from high schools and earn about $40000 [0.31, 0.36].

The experiment and results are broadly described in [10].

4. Conclusions

The paper presents type-2 fuzzy logic which is an extension of traditional (type-1) fuzzy logic, i.e. sentence and quantifier calculi. Type-2 fuzzy logic applies type-2 fuzzy sets to represent linguistic expressions and linguistic quantifiers appearing in linguistically quantified statements, see (1.18) and (1.19). In these terms, type-2 fuzzy logic generalizes type-1 fuzzy logic, because each type-1 fuzzy set is a special case of a T2FS.

Type-2 linguistic summaries (directly related to type-2 fuzzy logic) and their sample application, are shown. Short textual messages from large databases are generated by dedicated software, and presented to an end-user in a human-consistent and easy-to-operate form. In general, theory and applications presented in this paper can be placed in the domain of soft-computing and artificial intelligence.

References


Institute of Information Technology
Technical University of Łódź
Wólczańska 215, 90-924 Łódź, Poland
e-mail: aniewiadomski@ics.p.lodz.pl

Department of Knowledge Engineering and Computer Intelligence
Academy of Humanities and Economics
Rewolucji 52, 90–222 Łódź, Poland
e-mail: aniewiadomski@ahe.lodz.pl