A spherical fuzzy correlation coefficient based on statistical viewpoint with its applications in classification and bidirectional approximate reasoning

Abdul Haseeb GANIE and Debashis DUTTA

Spherical fuzzy sets are more powerful in modelling the uncertain situations than picture fuzzy sets, fermatean fuzzy sets, Pythagorean fuzzy sets, intuitionistic fuzzy sets, and fuzzy sets. In this paper, we first define the variance and covariance of spherical fuzzy sets. Then, using variance and covariance, we define the unique spherical fuzzy set correlation metric in line with the statistical coefficient of correlation. Two spherical fuzzy sets are correlated in both direction and strength using the provided measure of correlation. We discussed its many characteristics. We compared the measure of correlation with the current ones through linguistic variables. We established its validity by showing its application in bidirectional approximate reasoning. We also resolve a pattern identification issue in the spherical fuzzy environment using the provided correlation function, and we compare the results with several current measurements.

Key words: correlation coefficient, fuzzy set, picture fuzzy set, spherical fuzzy set, pattern recognition, bidirectional approximate reasoning

1. Introduction

Cuong and Kreinvoch [5] first proposed the idea of a picture fuzzy set (PFS) in 2013 for handling queries that call for responses of the types no, abstain, yes, and rejection. Each component of a PFS has the levels of membership, nonmembership, and neutrality that define it. The sum of the membership grades in a PFS is at most one and due to this restriction, the scope of PFSs is limited. Therefore, the idea of spherical fuzzy sets (SFSs) was developed by Mahmood et al. [15] to handle the circumstances in which the sum of membership grades is
more than one. In an SFS, the square sum of membership grades is at most one. Since the squared sum of the spherical parameters can only be at most 1.0 and decision makers can define their hesitation information independently, the unique idea of SFS gives decision makers a wider preference domain. A decision maker, for instance, might indicate their preference for one possibility over another in relation to a criterion with (0.8, 0.2, 0.4). It is obvious that the parameter sum is greater than 1, although the squared sum is only 0.84. Some spherical fuzzy (SF) measures of similarity based on cosine function and their usability in medical diagnosis and problems of classification were introduced by Wei et al. [26]. Aydogdu and Gul [2] proposed an entropy measure for SFSs with its applicability in decision-making. Applications of some novel SF similarity measures in green supplier selection and medical diagnosis were shown by Shishavan et al. [17]. Gundogdu and Kahraman [13] extended the classical analytic hierarchy process to the SF environment. Fatma and Cengiz [9] extended the VIKOR (classical VlseKriterijumska Optimizacija I Kompromisno Resenje) method to SF setting and demonstrated its application in the selection of warehouse site. Zhang et al. [28] proposed some SF aggregation operators and their application in decision-making. The correlation coefficient of SFSs is the subject of the current paper.

A coefficient of correlation is a statistical tool that can be used in a number of analytical and experimental studies. In fuzzy and its various generalizations, the correlation coefficient has been utilized in many fields like classification, decision-making, clustering, medical diagnosis, etc. Many studies [1, 4, 8, 10–12, 16, 18–20, 24] concerning the fuzzy/non-standard fuzzy functions of correlation and their applications are available in the literature.

The primary reasons we were inspired to conduct this study are listed below.

• The correlation coefficients in the SF environment have not been properly investigated.
• The current correlation measures are unable to handle the linguistic hedges properly.
• The available correlation functions are not based on statistical viewpoint.
• The applicability of correlation functions in bidirectional approximate reasoning has not been discussed yet.

The following list outlines this paper’s contribution:

• We defined the variance and co-variance for SFSs.
• We proposed a novel correlation measure for SFSs based on variance and co-variance.
• We discussed the properties of the proposed correlation metric.
• We contrasted its performance with the available ones in terms of linguistic variables.
• We established its applicability in bidirectional approximate reasoning.
• We solved a classification problem in SF area with the aid of developed correlation function and compared the findings with the current compatibility functions.

The structure of the paper is as: The preliminary part of this paper is Section 2. In Section 3, we offer a novel function of correlation for SFSs along with its characteristics. An assessment of the offered SFS correlation metric with the existing compatibility metrics through linguistic hedges is shown in Section 4. In Section 5, the recommended correlation measure’s application to bidirectional approximate reasoning and pattern recognition is demonstrated. Section 6 presents the conclusion along with future studies.

2. Preliminaries

Let $B = \{b_1, b_2, \ldots, b_p\}$ denote the set of universe.

**Definition 1.** [27] A fuzzy set $C_1$ in $B$ is given by

$$C_1 = \{(b_t, \sigma_{C_1}(b_t)), b_t \in B\},$$

where $0 \leq \sigma_{C_1}(b_t) \leq 1$ is the grade of satisfaction of $b_t \in B$ in the set $C_1$.

**Definition 2.** [5] A picture fuzzy set $C_1$ in $B$ is given by

$$C_1 = \{(b_t, \sigma_{C_1}(b_t), \zeta_{C_1}(b_t), \eta_{C_1}(b_t))b_t \in B\},$$

where $0 \leq \sigma_{C_1}(b_t) \leq 1$, $0 \leq \zeta_{C_1}(b_t) \leq 1$, and $0 \leq \eta_{C_1}(b_t) \leq 1$ are the grades of satisfaction, non-satisfaction, and neutrality respectively of $b_t \in B$ in the set $C_1$ such that $0 \leq \sigma_{C_1}(b_t) + \zeta_{C_1}(b_t) + \eta_{C_1}(b_t) \leq 1$. Also, $\theta_{C_1}(b_t) = 1 - (\sigma_{C_1}(b_t) + \zeta_{C_1}(b_t) + \eta_{C_1}(b_t))$ is the refusal degree for the element $b_t \in B$ in the set $C_1$.

**Definition 3.** [15] A SFS $C_1$ in $B$ is given by

$$C_1 = \{(b_t, \sigma_{C_1}(b_t), \zeta_{C_1}(b_t), \eta_{C_1}(b_t))b_t \in B\},$$

where $0 \leq \sigma_{C_1}(b_t) \leq 1$, $0 \leq \zeta_{C_1}(b_t) \leq 1$, and $0 \leq \eta_{C_1}(b_t) \leq 1$ are the grades of satisfaction, non-satisfaction, and neutrality respectively of $b_t \in B$ in the set $C_1$ such that $0 \leq \sigma_{C_1}^2(b_t) + \zeta_{C_1}^2(b_t) + \eta_{C_1}^2(b_t) \leq 1$. Also, $\theta_{C_1}(b_t) = \sqrt{1 - (\sigma_{C_1}^2(b_t) + \zeta_{C_1}^2(b_t) + \eta_{C_1}^2(b_t))}$ is the refusal degree for the element $b_t \in B$ in the set $C_1$. 


3. A novel SF correlation coefficient

We first give the definition of the variance and covariance of SFSs. Let $SFS(B)$ denote the set of all SFSs on $B$.

**Definition 4.** For any $C_1 \in SFS(B)$, the variance is given by

$$ VAR(C_1) = \frac{1}{p} \sum_{t=1}^{p} \left[ \left( \sigma_{C_1}(b_t) \right)^2 - \left( \overline{\sigma_{C_1}} \right)^2 \right] + \left[ \left( \zeta_{C_1}(b_t) \right)^2 - \left( \overline{\zeta_{C_1}} \right)^2 \right] + \left[ \left( \eta_{C_1}(b_t) \right)^2 - \left( \overline{\eta_{C_1}} \right)^2 \right], $$

where $\overline{\sigma_{C_1}} = \frac{1}{p} \sum_{t=1}^{p} \sigma_{C_1}(b_t)$, $\overline{\zeta_{C_1}} = \frac{1}{p} \sum_{t=1}^{p} \zeta_{C_1}(b_t)$ and $\overline{\eta_{C_1}} = \frac{1}{p} \sum_{t=1}^{p} \eta_{C_1}(b_t)$.

**Definition 5.** For any $C_1, C_2 \in SFS(B)$, the covariance is given by

$$ COV(C_1, C_2) = \frac{1}{p} \sum_{t=1}^{p} \left[ \left( \sigma_{C_1}(b_t) \right)^2 - \left( \overline{\sigma_{C_1}} \right)^2 \right] \left( \sigma_{C_2}(b_t) \right)^2 - \left( \overline{\sigma_{C_2}} \right)^2 \right] + \left[ \left( \zeta_{C_1}(b_t) \right)^2 - \left( \overline{\zeta_{C_1}} \right)^2 \right] \left( \zeta_{C_2}(b_t) \right)^2 - \left( \overline{\zeta_{C_2}} \right)^2 \right] + \left[ \left( \eta_{C_1}(b_t) \right)^2 - \left( \overline{\eta_{C_1}} \right)^2 \right] \left( \eta_{C_2}(b_t) \right)^2 - \left( \overline{\eta_{C_2}} \right)^2 \right]. $$

Now, we define the correlation coefficient for SFSs.

**Definition 6.** For any $C_1, C_2 \in SFS(B)$, a correlation coefficient is given by

$$ C_{GD}(C_1, C_2) = \frac{COV(C_1, C_2)}{\sqrt{VAR(C_1) \times VAR(C_2)}}. \quad (1) $$

Now, we discuss some essential properties of the suggested correlation coefficient $C_{GD}$ given in Eq. (1).

**Theorem 1.** For any $C_1, C_2 \in SFS(B)$, we have

1. $C_{GD}(C_1, C_2) = C_{GD}(C_2, C_1),$
2. $-1 \leq C_{GD}(C_1, C_2) \leq 1,$
3. if $C_1 = cC_2$ for some $c$, then $C_{GD}(C_1, C_2) = 1.$
Proof. (1)

\[
COV(C_1, C_2) = \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_1}(b_t))^2 - (\overline{\sigma_{C_1}})^2 \right) \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) \right] \\
+ \left[ \left( (\zeta_{C_1}(b_t))^2 - (\overline{\zeta_{C_1}})^2 \right) \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \right] \\
+ \left[ \left( (\eta_{C_1}(b_t))^2 - (\overline{\eta_{C_1}})^2 \right) \left( (\eta_{C_2}(b_t))^2 - (\overline{\eta_{C_2}})^2 \right) \right]
\]

\[
= \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) \left( (\sigma_{C_1}(b_t))^2 - (\overline{\sigma_{C_1}})^2 \right) \right] \\
+ \left[ \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \left( (\zeta_{C_1}(b_t))^2 - (\overline{\zeta_{C_1}})^2 \right) \right] \\
+ \left[ \left( (\eta_{C_2}(b_t))^2 - (\overline{\eta_{C_2}})^2 \right) \left( (\eta_{C_1}(b_t))^2 - (\overline{\eta_{C_1}})^2 \right) \right]
\]

\[
= COV(C_2, C_1).
\]

So,

\[
C_{GD}(C_1, C_2) = \frac{COV(C_1, C_2)}{\sqrt{VAR(C_1) \times VAR(C_2)}}
\]

\[
= \frac{COV(C_2, C_1)}{\sqrt{VAR(C_2) \times VAR(C_1)}} = C_{GD}(C_2, C_1).
\]

(2) Using Cauchy-Schwartz inequality, we have

\[
(COV(C_1, C_2))^2 \leq \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_1}(b_t))^2 - (\overline{\sigma_{C_1}})^2 \right)^2 + \left( (\zeta_{C_1}(b_t))^2 - (\overline{\zeta_{C_1}})^2 \right)^2 \right] \\
+ \left[ \left( (\eta_{C_1}(b_t))^2 - (\overline{\eta_{C_1}})^2 \right)^2 \right]\]

\[
\times \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right)^2 + \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right)^2 \right] \\
+ \left[ \left( (\eta_{C_2}(b_t))^2 - (\overline{\eta_{C_2}})^2 \right)^2 \right]
\]

\[
= VAR(C_1) \times VAR(C_2).
\]
\[ |\text{COV}(C_1, C_2)| \leq \sqrt{\text{VAR}(C_1) \times \text{VAR}(C_2)} \]

or

\[ -\sqrt{\text{VAR}(C_1) \times \text{VAR}(C_2)} \leq \text{COV}(C_1, C_2) \leq \sqrt{\text{VAR}(C_1) \times \text{VAR}(C_2)} \]

or

\[ -1 \leq \frac{\text{COV}(C_1, C_2)}{\sqrt{\text{VAR}(C_1) \times \text{VAR}(C_2)}} \leq 1 \]

or

\[ -1 \leq C_{GD}(C_1, C_2) \leq 1. \]

(3) Let \( C_1 = cC_2 \), then

\[ \sigma_{C_1}(b_t) = c\sigma_{C_2}(b_t), \ \zeta_{C_1}(b_t) = c\zeta_{C_2}(b_t), \ \eta_{C_1}(b_t) = c\eta_{C_2}(b_t), \ \forall t = 1, 2, ..., p. \]

\[
\text{COV}(C_1, C_2) = \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_1}(b_t))^2 - (\overline{\sigma_{C_1}})^2 \right) \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) \right. \\
+ \left. \left( (\zeta_{C_1}(b_t))^2 - (\overline{\zeta_{C_1}})^2 \right) \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \right] \\
= \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (c\sigma_{C_2}(b_t))^2 - (c\overline{\sigma_{C_2}})^2 \right) \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) \right. \\
+ \left. \left( (c\zeta_{C_2}(b_t))^2 - (c\overline{\zeta_{C_2}})^2 \right) \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \right] \\
= \frac{c^2}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \right] \\
= c^2 \text{VAR}(C_2). \]

Also,

\[
\text{VAR}(C_1) = \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_1}(b_t))^2 - (\overline{\sigma_{C_1}})^2 \right) + \left( (\zeta_{C_1}(b_t))^2 - (\overline{\zeta_{C_1}})^2 \right) \right] \\
= \frac{1}{p} \sum_{t=1}^{p} \left[ \left( (c\sigma_{C_2}(b_t))^2 - (c\overline{\sigma_{C_2}})^2 \right) + \left( (c\zeta_{C_2}(b_t))^2 - (c\overline{\zeta_{C_2}})^2 \right) \right] \\
= \frac{c^4}{p} \sum_{t=1}^{p} \left[ \left( (\sigma_{C_2}(b_t))^2 - (\overline{\sigma_{C_2}})^2 \right) + \left( (\zeta_{C_2}(b_t))^2 - (\overline{\zeta_{C_2}})^2 \right) \right] \\
= c^4 \text{VAR}(C_2). \]
So,

\[
C_{GD}(C_1, C_2) = \frac{COV(C_1, C_2)}{\sqrt{VAR(C_1) \times VAR(C_2)}}
\]

\[
= \frac{c^2 VAR(C_2)}{\sqrt{c^4 VAR(C_2) \times VAR(C_2)}} = 1.
\]

4. A comparative study based on Linguistic hedges

**Definition 7.** For any \( C_1 \in SFS(B) \), \( C_1^\mu \) is defined as

\[
C_1^\mu = \left\{ \left( f_t, (\sigma C_1(f_t) + \eta C_1(f_t))^\mu - (\eta C_1(f_t))^\mu, \sqrt{1 - \left(1 - (\xi C_1(f_t))^2\right)^\mu}, (\eta C_1(f_t))^\mu \right), f_t \in B \right\}, \mu > 0.
\]

**Example 1.** Consider SFS \( C_1 \in B = \{f_1, f_2, f_3, f_4, f_5\} \) given as

\[
C_1 = \left\{ (f_1, 0.7, 0.1, 0.1), (f_2, 0.2, 0.3, 0.4), (f_3, 0.2, 0.1, 0.5), (f_4, 0.1, 0.5, 0.2), (f_5, 0.2, 0.2, 0.2) \right\}.
\]

With the help of above Definition, we define the SFSs More or less LARGE = \( C_1^\frac{1}{3} \), Not very LARGE = \( (C_1^2)^c \), very very LARGE = \( C_1^4 \), very LARGE = \( C_1^2 \), and LARGE = \( C_1 \), as follows.

\[
C_1 = \left\{ (f_1, 0.7, 0.1, 0.1), (f_2, 0.2, 0.3, 0.4), (f_3, 0.2, 0.1, 0.5), (f_4, 0.1, 0.5, 0.2), (f_5, 0.2, 0.2, 0.2) \right\},
\]

\[
C_1^2 = \left\{ (f_1, 0.6300, 0.1411, 0.0100), (f_2, 0.2000, 0.4146, 0.1600), (f_3, 0.2400, 0.1411, 0.2500), (f_4, 0.0500, 0.6614, 0.0400), (f_5, 0.1200, 0.2800, 0.0400) \right\},
\]

\[
C_1^4 = \left\{ (f_1, 0.4095, 0.1985, 0.0), (f_2, 0.1040, 0.5606, 0.0256), (f_3, 0.1776, 0.1985, 0.0625), (f_4, 0.0065, 0.8268, 0.0016), (f_5, 0.0240, 0.3881, 0.0016) \right\},
\]

\[
C_1^{\frac{1}{3}} = \left\{ (f_1, 0.5782, 0.0708, 0.3162), (f_2, 0.1421, 0.2146, 0.6325), (f_3, 0.1296, 0.0708, 0.7071), (f_4, 0.1005, 0.3660, 0.4472), (f_5, 0.1852, 0.1421, 0.4472) \right\},
\]
\[
(C_1^2)^C = \left\{ (f_1, 0.1411, 0.6300, 0.0100), (f_2, 0.4146, 0.2000, 0.1600), \\
(f_3, 0.1411, 0.2400, 0.2500), (f_4, 0.6614, 0.0500, 0.0400), \\
(f_5, 0.2800, 0.1200, 0.0400) \right\}.
\]

We compare our suggested method with a few other ones for estimating the correlation coefficients using these SFSs. Tables 1–3 show the results and utilize the following notations.

Not very LARGE: N.V.L., more or less LARGE: M.L.L., very very LARGE: V.V.L., very LARGE: V.L., LARGE: L.

### Table 1: Correlation coefficients using $C_{UGMJA1}$ [25]

<table>
<thead>
<tr>
<th></th>
<th>M.L.L.</th>
<th>L.</th>
<th>V.L.</th>
<th>V.V.L.</th>
<th>N.V.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.L.L.</td>
<td>1.0000</td>
<td>0.8991</td>
<td>0.6519</td>
<td>0.4339</td>
<td>0.4768</td>
</tr>
<tr>
<td>L.</td>
<td>0.8991</td>
<td>1.0000</td>
<td>0.9045</td>
<td>0.7167</td>
<td>0.5995</td>
</tr>
<tr>
<td>V.L.</td>
<td>0.6519</td>
<td>0.9045</td>
<td>1.0000</td>
<td>0.9307</td>
<td>0.5866</td>
</tr>
<tr>
<td>V.V.L.</td>
<td>0.4339</td>
<td>0.7165</td>
<td>0.9307</td>
<td>1.0000</td>
<td>0.4950</td>
</tr>
<tr>
<td>N.V.L.</td>
<td>0.4768</td>
<td>0.5995</td>
<td>0.5866</td>
<td>0.4950</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

### Table 2: Correlation coefficients using $C_{UGMJA2}$ [25]

<table>
<thead>
<tr>
<th></th>
<th>M.L.L.</th>
<th>L.</th>
<th>V.L.</th>
<th>V.V.L.</th>
<th>N.V.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.L.L.</td>
<td>1.0000</td>
<td>0.8361</td>
<td>0.6476</td>
<td>0.3822</td>
<td>0.4737</td>
</tr>
<tr>
<td>L.</td>
<td>0.8361</td>
<td>1.0000</td>
<td>0.8467</td>
<td>0.5870</td>
<td>0.5612</td>
</tr>
<tr>
<td>V.L.</td>
<td>0.6476</td>
<td>0.8467</td>
<td>1.0000</td>
<td>0.8144</td>
<td>0.5866</td>
</tr>
<tr>
<td>V.V.L.</td>
<td>0.3822</td>
<td>0.5870</td>
<td>0.8144</td>
<td>1.0000</td>
<td>0.4332</td>
</tr>
<tr>
<td>N.V.L.</td>
<td>0.4737</td>
<td>0.5612</td>
<td>0.5866</td>
<td>0.4332</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

### Table 3: Correlation coefficients using $C_{GD}$

<table>
<thead>
<tr>
<th></th>
<th>M.L.L.</th>
<th>L.</th>
<th>V.L.</th>
<th>V.V.L.</th>
<th>N.V.L.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.L.L.</td>
<td>1.0000</td>
<td>0.9129</td>
<td>0.6862</td>
<td>0.4026</td>
<td>−0.1484</td>
</tr>
<tr>
<td>L.</td>
<td>0.9129</td>
<td>1.0000</td>
<td>0.8947</td>
<td>0.6006</td>
<td>−0.2789</td>
</tr>
<tr>
<td>V.L.</td>
<td>0.6862</td>
<td>0.8947</td>
<td>1.0000</td>
<td>0.8710</td>
<td>−0.3732</td>
</tr>
<tr>
<td>V.V.L.</td>
<td>0.4026</td>
<td>0.6006</td>
<td>0.8710</td>
<td>1.0000</td>
<td>−0.3481</td>
</tr>
<tr>
<td>N.V.L.</td>
<td>−0.1484</td>
<td>−0.2789</td>
<td>−0.3732</td>
<td>−0.3481</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
A coefficient of correlation $C$ should satisfy the following conditions due to the characterization of linguistic variables.

$$
C(M.L.L., V.V.L.) < C(M.L.L., V.L.) < C(M.L.L., L.)
$$
$$
C(V.V.L., M.L.L.) < C(V.V.L., L.) < C(V.V.L., V.L.)
$$
$$
C(L., V.V.L.) < C(L., V.L.) < C(L., M.L.L.)
$$
$$
$$

From Tables 1–3, we find out that the correlation coefficients $C_{UGMJA_1}$ [25] and $C_{UGMJA_2}$ [25] do not satisfy all the conditions of Eq. (2) whereas our suggested correlation fulfills all the required conditions as shown in Figs. 1–4. Also because

---

Figure 1: Condition first of Eq. (2)

Figure 2: Condition second of Eq. (2)
of the characterization of linguistic variables, the correlation of N.V.L. with all other linguistic hedges should be negative and this is shown by only the offered correlation measure.

5. Applications

Here we establish the utility of the developed metric in bidirectional approximate reasoning and pattern investigation.
In this model, we start by taking a look at the single-input, single-output, forward approximate reasoning system for SFSs to check the accuracy of the correlation function.

Let us apply the proposed correlation function to a bidirectional approximate reasoning system for SFSs to check the accuracy of the correlation function.

Here, we apply the proposed correlation function to a bidirectional approximate reasoning system for SFSs to check the accuracy of the correlation function.

\[
P_1 : \text{If } A \text{ is } C_1, \text{ then } E \text{ is } D_1.
\]

\[
P_2 : \text{If } A \text{ is } C_2, \text{ then } E \text{ is } D_2.
\]

\[
P_3 : \text{If } A \text{ is } C_3, \text{ then } E \text{ is } D_3.
\]

... 

\[
P_p : \text{If } A \text{ is } C_p, \text{ then } E \text{ is } D_p.
\]

**Fact:** \( A \) is \( C^* \).

**Consequence:** \( E \) is \( D^* \).

In this model, \( C_t \) and \( C^* \) are SFSs of the Universe \( B = \{b_1, b_2, \ldots, b_p\} \) and \( P_t(1 \leq t \leq p) \) is the \( t \)-th output rule and \( D_t \) and \( D^* \) are SFSs of the Universe \( F = \{f_1, f_2, \ldots, f_q\} \). With the help of the Chen et al. [3], we have the following:

\[
C_{GD}(C_1, C^*) = l_1 \implies \text{“} E \text{ is } D_1^* \text{“} \text{ is the derived result of the Rule } P_1, \text{ where } C_{GD}(C_1, C^*) \text{ indicates the correlation between } C_1 \text{ and } C^*, \text{ and }
\]

\[
D_1^* = |l_1| \times D_1 = \left\{ \begin{array}{l}
(f_1, |l_1| \sigma_{D_1}(f_1), 1 - |l_1| + |l_1| \zeta_{D_1}(f_1), |l_1| \eta_{D_1}(f_1)), \\
(f_2, |l_1| \sigma_{D_1}(f_2), 1 - |l_1| + |l_1| \zeta_{D_1}(f_2), |l_1| \eta_{D_1}(f_2)), \\
\ldots \\
(f_q, |l_1| \sigma_{D_1}(f_q), 1 - |l_1| + |l_1| \zeta_{D_1}(f_q), |l_1| \eta_{D_1}(f_q)).
\end{array} \right.
\]

\[
C_{GD}(C_2, C^*) = l_2 \implies \text{“} E \text{ is } D_2^* \text{“} \text{ is the desired result of the Rule } P_2, \text{ where }
\]

\[
D_2^* = |l_2| \times D_2 = \left\{ \begin{array}{l}
(f_1, |l_2| \sigma_{D_2}(f_1), 1 - |l_2| + |l_2| \zeta_{D_2}(f_1), |l_2| \eta_{D_2}(f_1)), \\
(f_2, |l_2| \sigma_{D_2}(f_2), 1 - |l_2| + |l_2| \zeta_{D_2}(f_2), |l_2| \eta_{D_2}(f_2)), \\
\ldots \\
(f_q, |l_2| \sigma_{D_2}(f_q), 1 - |l_2| + |l_2| \zeta_{D_2}(f_q), |l_2| \eta_{D_2}(f_q)).
\end{array} \right.
\]

\[
C_{GD}(C_p, C^*) = l_p \implies \text{“} E \text{ is } D_p^* \text{“} \text{ is the desired result of the Rule } P_p, \text{ where }
\]

\[
D_p^* = |l_p| \times D_p = \left\{ \begin{array}{l}
(f_1, |l_p| \sigma_{D_p}(f_1), 1 - |l_p| + |l_p| \zeta_{D_p}(f_1), |l_p| \eta_{D_p}(f_1)), \\
(f_2, |l_p| \sigma_{D_p}(f_2), 1 - |l_p| + |l_p| \zeta_{D_p}(f_2), |l_p| \eta_{D_p}(f_2)), \\
\ldots \\
(f_q, |l_p| \sigma_{D_p}(f_q), 1 - |l_p| + |l_p| \zeta_{D_p}(f_q), |l_p| \eta_{D_p}(f_q)).
\end{array} \right.
\]
Example 2. Let us have a look at the following SFS-based forward approximation reasoning system.

\[ P_1 : \text{If } A \text{ is } C_1, \text{ then } E \text{ is } D_1. \]
\[ P_2 : \text{If } A \text{ is } C_2, \text{ then } E \text{ is } D_2. \]
\[ P_3 : \text{If } A \text{ is } C_3, \text{ then } E \text{ is } D_3. \]
\[ \text{Fact : } A \text{ is } C^*. \]
\[ \text{Consequence : } E \text{ is } D^*. \]

In this system

\[ C_1 = \{(b_1, 0.2, 0.1, 0.4), (b_2, 0.3, 0.5, 0.1), (b_3, 0.7, 0, 0.2)\}, \]
\[ C_2 = \{(b_1, 0.5, 0.1, 0.3), (b_2, 0.8, 0, 0.1), (b_3, 0.4, 0.5, 0)\}, \]
\[ C_3 = \{(b_1, 0.7, 0.2, 0.1), (b_2, 1, 0, 0), (b_3, 0.1, 0.3, 0.5)\}, \]
\[ C^* = \{(b_1, 0.8, 0, 0.1), (b_2, 0.4, 0.5, 0), (b_3, 0.2, 0.3, 0.3)\}, \]
\[ D_1 = \{(f_1, 0.1, 0.3, 0.4), (f_2, 0.1, 0.4, 0.2), (f_3, 0.3, 0.4, 0.2)\} \]
\[ D_2 = \{(f_1, 0.1, 0.9, 0), (f_2, 0, 0.5, 0), (f_3, 0.1, 0, 0.1)\}, \]
\[ D_3 = \{(f_1, 0.3, 0.1, 0.4), (f_2, 0.5, 0.3, 0.1), (f_3, 0.4, 0, 0.3)\}. \]

Using the suggested correlation metric, we have \( C_{GD}(C_1, C^*) = -0.3071, \)
\( C_{GD}(C_2, C^*) = -0.1073, \) and \( C_{GD}(C_3, C^*) = 0.2335. \) So, we obtain

\[ D_1^* = \{(f_1, 0.0307, 0.7850, 0.1228), (f_2, 0.0307, 0.8158, 0.0614), (f_3, 0.0921, 0.8158, 0.0614)\}, \]
\[ D_2^* = \{(f_1, 0.0104, 0.9896, 0), (f_2, 0, 0.9482, 0), (f_3, 0.0104, 0.8963, 0.0104)\} \]
\[ D_3^* = \{(f_1, 0.0700, 0.7899, 0.0934), (f_2, 0.1167, 0.8366, 0.0233), (f_3, 0.0934, 0.7665, 0.0700)\}. \]
So,

\[ D^* = D_1^* \cup D_2^* \cup \ldots \cup D_p^* = \left\{ (f_1, 0.0700, 0.7850, 0), (f_2, 0.1167, 0.8158, 0), (f_3, 0.0934, 0.7665, 0.0104) \right\}. \]

Now, using the suggested correlation metric \( C_{GD} \), we have

\[ C_{GD}(D_1, D^*) = 0.0612, \quad C_{GD}(D_2, D^*) = 0.1471, \quad \text{and} \quad C_{GD}(D_3, D^*) = 0.4562. \]

The results indicate that \( C_{GD}(C_1, C^*) < C_{GD}(C_2, C^*) < C_{GD}(C_3, C^*) \) and \( C_{GD}(D_1, D^*) < C_{GD}(D_2, D^*) < C_{GD}(D_3, D^*) \), i.e., \( C_{GD}(C_1, C^*) \) and \( C_{GD}(C_3, C^*) \) have the smallest and maximum values respectively among the values of \( C_{GD}(C_1, C^*), C_{GD}(C_2, C^*), \) and \( C_{GD}(C_3, C^*) \), whereas \( D_1 \) and \( D_3 \) share the least and greatest correlation with \( D^* \) respectively (see Fig. 5).

![CGD](image)

**Figure 5:** Forward approximate reasoning

On the other hand, consider the following backward approximation-based reasoning model:

- \( P_1 \): If \( A \) is \( C_1 \), then \( E \) is \( D_1 \).
- \( P_2 \): If \( A \) is \( C_2 \), then \( E \) is \( D_2 \).
- \( P_3 \): If \( A \) is \( C_3 \), then \( E \) is \( D_3 \).

\[ \ldots \]

- \( P_p \): If \( A \) is \( C_p \), then \( E \) is \( D_p \).

Fact: \( E \) is \( D^* \).

Consequence: \( A \) is \( C^* \).
In this model, \( C_t \) and \( C^* \) are SFSs of the Universe \( B = \{b_1, b_2, \ldots, b_p\} \) and \( P_t (1 \leq t \leq p) \) is the \( t \)-th output rule and \( D_t \) and \( D^* \) are SFSs of the Universe \( F = \{f_1, f_2, \ldots, f_q\} \). With the help of the Chen et al. [3], we have the following:

\[
C_{GD}(D_1, D^*) = k_1 \implies \text{“A is } C_1^* \text{” is the derived result of the Rule } P_1, \text{ where } C_{GD}(D_1, D^*) \text{ indicates the correlation between } D_1 \text{ and } D^*, \text{ where }
\]

\[
C_1^* = |k_1| \times C_1 = \begin{cases} (b_1, |k_1|\sigma_{C_1}(b_1), 1 - |k_1| + |k_1|\zeta_{C_1}(b_1), |k_1|\eta_{C_1}(b_1)), \\ (b_2, |k_1|\sigma_{C_1}(b_2), 1 - |k_1| + |k_1|\zeta_{C_1}(b_2), |k_1|\eta_{C_1}(b_2)), \\ \vdots \\ (b_q, |k_1|\sigma_{C_1}(f_q), 1 - |k_1| + |k_1|\zeta_{D_1}(f_q), |k_1|\eta_{D_1}(f_q)) \end{cases}
\]

\[
C_{GD}(D_2, D^*) = k_2 \implies \text{“A is } C_2^* \text{” is the derived result of the Rule } P_2, \text{ where }
\]

\[
C_2^* = |k_2| \times C_2 = \begin{cases} (b_1, |k_2|\sigma_{C_2}(f_1), 1 - |k_2| + |k_2|\zeta_{C_2}(b_1), |k_2|\eta_{C_2}(b_1)), \\ (b_2, |k_2|\sigma_{C_2}(b_2), 1 - |k_2| + |k_2|\zeta_{C_2}(b_2), |k_2|\eta_{C_2}(b_2)), \\ \vdots \\ (b_q, |k_2|\sigma_{C_2}(f_q), 1 - |k_2| + |k_2|\zeta_{D_2}(f_q), |k_2|\eta_{D_2}(f_q)) \end{cases}
\]

\[
C_{GD}(D_p, D^*) = k_p \implies \text{“A is } C_p^* \text{” is the derived result of the Rule } P_p, \text{ where }
\]

\[
C_p^* = |k_p| \times C_p = \begin{cases} (b_1, |k_p|\sigma_{C_p}(b_1), 1 - |k_p| + |k_p|\zeta_{C_p}(b_1), |k_p|\eta_{C_p}(b_1)), \\ (b_2, |k_p|\sigma_{C_p}(b_2), 1 - |k_p| + |k_p|\zeta_{C_p}(b_2), |k_p|\eta_{C_p}(b_2)), \\ \vdots \\ (b_q, |k_p|\sigma_{C_p}(f_q), 1 - |k_p| + |k_p|\zeta_{D_p}(f_q), |k_p|\eta_{D_p}(f_q)) \end{cases}
\]

So, the desired result of the approximate reasoning method is “A is \( C^* \), where

\[
C^* = C_1^* \cup C_2^* \cup \ldots \cup C_p^*
\]

\[
= \begin{cases} (b_1, \max_i(|k_i|\sigma_{C_i}(b_1)), \min_i (1-|k_i|+|k_i|\zeta_{C_i}(b_1)), \min_i (|k_i|\eta_{C_i}(b_1))) \\ (b_2, \max_i(|k_i|\sigma_{C_i}(b_2)), \min_i (1-|k_i|+|k_i|\zeta_{C_i}(b_2)), \min_i (|k_i|\eta_{C_i}(b_2))) \\ \vdots \\ (b_q, \max_i(|k_i|\sigma_{C_i}(f_q)), \min_i (1-|k_i|+|k_i|\zeta_{D_i}(f_q)), \min_i (|k_i|\eta_{D_i}(f_q))) \end{cases},
\]

and \( \cup \) denotes the union operator between SFSs, \( 1 \leq t \leq p \).

**Example 3.** Let us have a look at the following SFS-based backward approximation reasoning system.

\[
P_1: \text{ If } A \text{ is } C_1, \text{ then } E \text{ is } D_1. \\
P_2: \text{ If } A \text{ is } C_2, \text{ then } E \text{ is } D_2. \\
P_3: \text{ If } A \text{ is } C_3, \text{ then } E \text{ is } D_3. \\
\text{Fact: } E \text{ is } D^*. \\
\text{Consequence: } A \text{ is } C^*.
\]
In this system
\[ C_1 = \{(b_1, 0.2, 0.7, 0.1), (b_2, 0, 0.1, 0), (b_3, 0.5, 0.3, 0.1)\}, \]
\[ C_2 = \{(b_1, 0.4, 0.1, 0.2), (b_2, 0.1, 0.5, 0.3), (b_3, 0.7, 0.1, 0.2)\}, \]
\[ C_3 = \{(b_1, 0.1, 0.5, 0.3), (b_2, 0.1, 0.6, 0.1), (b_3, 0, 0.5, 0.4)\}, \]
\[ D_1 = \{(f_1, 0.4, 0.1, 0.3), (f_2, 0.3, 0.5, 0.1), (f_3, 0.4, 0.1, 0.3)\}, \]
\[ D_2 = \{(f_1, 0, 0.9, 0.1), (f_2, 0, 0.5, 0.1), (f_3, 0.1, 0.3, 0.5)\}, \]
\[ D_3 = \{(f_1, 0.3, 0.4, 0.1), (f_2, 0.2, 0.4, 0.3), (f_3, 0.5, 0.1, 0.2)\}, \]
\[ D^* = \{(f_1, 0.6, 0.1, 0.1), (f_2, 0.5, 0.4, 0), (f_3, 0.3, 0.3, 0.2)\}. \]

Using the suggested correlation metric, we have \( C_{GD}(D_1, D^*) = 0.3919, \) \( C_{GD}(D_2, D^*) = -0.2714, \) and \( C_{GD}(D_3, D^*) = -0.4876. \) So, we obtain
\[
C^*_1 = \{ (b_1, 0.0784, 0.8824, 0.0392), (b_2, 0, 1, 0), (b_3, 0.1960, 0.7256, 0.0392) \},
\]
\[
C^*_2 = \{ (b_1, 0.1086, 0.7557, 0.0543), (b_2, 0.0271, 0.8643, 0.0814), (b_3, 0.1900, 0.7557, 0.0543) \},
\]
\[
C^*_3 = \{ (b_1, 0.0488, 0.7562, 0.1463), (b_2, 0.0488, 0.8050, 0.0488), (b_3, 0, 0.7562, 0.1950) \}.
\]

So,
\[
C^* = C^*_1 \cup C^*_2 \cup \ldots \cup C^*_p = \{ (b_1, 0.1086, 0.7557, 0.0392), (b_2, 0.0488, 0.8050, 0), (b_3, 0.1960, 0.7256, 0.0392) \}.
\]

Now, using the suggested correlation metric \( C_{GD} \), we have \( C_{GD}(C_1, C^*) = 0.9783, C_{GD}(C_2, C^*) = 0.6755, \) and \( C_{GD}(C_3, C^*) = 0.5590. \)

The results indicate that \( C_{GD}(C_1, C^*) > C_{GD}(C_2, C^*) > C_{GD}(C_3, C^*) \) and \( C_{GD}(D_1, D^*) > C_{GD}(D_2, D^*) > C_{GD}(D_3, D^*) \), i.e., \( C_{GD}(C_1, C^*) \) and \( C_{GD}(C_3, C^*) \) have the greatest and smallest values respectively among the values of \( C_{GD}(C_1, C^*), C_{GD}(C_2, C^*), \) and \( C_{GD}(C_3, C^*), \) whereas \( D_1 \) and \( D_3 \) shares the greatest and least correlation with \( D^* \) respectively (see Fig. 6).

Together, the outcomes of Examples 2 and 3 allow us to draw the conclusion that the suggested strategy works well for approximate reasoning.
5.2. Pattern investigation

Here, we establish how to use the offered SFS correlation metric to solve classification-related problems. Utilizing numerous criteria of compatibility, like “correlation”, “distance”, “similarity”, etc., enables pattern analysis to categorize an unexplained pattern into one of the recognized patterns. We contrast our results with several compatibility measurements as well.

In the below example, we will answer a classification-related problem.

Example 4. [11] Let $C_k, k = 1, 2, 3$ and $C$ be some patterns expressed in terms of SFSs as:

\[
C_1 = \left\{ (f_1, 0.4, 0.3, 0.1), (f_2, 0.5, 0.3, 0.2), (f_3, 0.4, 0.3, 0), (f_4, 0.7, 0, 0.2), (f_5, 0.6, 0.1, 0.1) \right\},
\]

\[
C_2 = \left\{ (f_1, 0.7, 0.1, 0.1), (f_2, 0.2, 0.3, 0.4), (f_3, 0.2, 0.1, 0.5), (f_4, 0.1, 0.5, 0.2), (f_5, 0.3, 0.3, 0.3) \right\},
\]

\[
C_3 = \left\{ (f_1, 0.1, 0.3, 0.4), (f_2, 0.4, 0.3, 0.1), (f_3, 0.3, 0.4, 0.2), (f_4, 0.2, 0.5, 0.3), (f_5, 0.5, 0.3, 0.1) \right\},
\]

and

\[
C = \left\{ (f_1, 0.6, 0.2, 0.1), (f_2, 0.3, 0.4, 0.2), (f_3, 0.4, 0.3, 0.2), (b_4, 0.7, 0.1, 0), (f_5, 0.4, 0.2, 0.2) \right\}.
\]

The task is to verify which pattern $C_t, t = 1, 2, 3$ shares the maximum similarity with $C$. For this purpose, we combine the offered SFS correlation measurement with the currently known compatibility functions. The computed results are displayed in Table 4. The most of functions of compatibility along with the developed SFS metric, make it clear that $C$ should be assigned to $C_1$ (see Table 4). After finding out the pattern to which $C$ belongs, we compute the Confidence degree (CD) [14] of each compatibility function as
A SPHERICAL FUZZY CORRELATION COEFFICIENT

\[ DoC = \sum_{k=1,k\neq j}^{m} |COR(C_k, C) - COR(C_j, C)|, \]
where \( COR \) is any measure of comparison like correlation measure, distance measure, similarity measure, etc. and \( C_j \) is the pattern to which \( C \) belongs. In comparison to the current picture fuzzy and SF compatibility measures, as shown in Fig. 7, we find that the CD of the proposed SF correlation coefficient is very high.

Table 4: Values of various PF/SF compatibility tests calculated with reference to Example 4

<table>
<thead>
<tr>
<th>Compatibility measure</th>
<th>((C_1, C))</th>
<th>((C_2, C))</th>
<th>((C_3, C))</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D_1) [6]</td>
<td>0.1000</td>
<td>0.1867</td>
<td>0.1933</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_2) [6]</td>
<td>0.1000</td>
<td>0.1833</td>
<td>0.1929</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_3) [6]</td>
<td>0.2000</td>
<td>0.2600</td>
<td>0.2600</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_4) [6]</td>
<td>0.0894</td>
<td>0.1428</td>
<td>0.1456</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_5) [7]</td>
<td>0.2000</td>
<td>0.3000</td>
<td>0.3400</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_6) [7]</td>
<td>0.4000</td>
<td>0.6557</td>
<td>0.7071</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_7) [7]</td>
<td>0.1789</td>
<td>0.2933</td>
<td>0.3162</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_8) [7]</td>
<td>0.2000</td>
<td>0.2800</td>
<td>0.3000</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_9) [21]</td>
<td>0.4000</td>
<td>0.6000</td>
<td>0.6800</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{10}) [21]</td>
<td>0.1265</td>
<td>0.2074</td>
<td>0.2236</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{11}) [21]</td>
<td>0.0500</td>
<td>0.0650</td>
<td>0.0650</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{12}) [23]</td>
<td>0.3750</td>
<td>0.5154</td>
<td>0.4755</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{13}) [23]</td>
<td>0.3491</td>
<td>0.3951</td>
<td>0.3880</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{14}) [23]</td>
<td>0.1250</td>
<td>0.1872</td>
<td>0.1775</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(D_{15}) [23]</td>
<td>0.1955</td>
<td>0.2268</td>
<td>0.2232</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(K_1) [22]</td>
<td>0.9168</td>
<td>0.7625</td>
<td>0.7138</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(K_2) [22]</td>
<td>0.8838</td>
<td>0.7500</td>
<td>0.6739</td>
<td>(C_1)</td>
</tr>
<tr>
<td>(C_{GD}) (Proposed)</td>
<td>0.5014</td>
<td>0.2235</td>
<td>-0.5555</td>
<td>(C_1)</td>
</tr>
</tbody>
</table>

Figure 7: Confidence Degree of various picture fuzzy/SF compatibility measures
6. Conclusion

The SFS correlation coefficient proposed in this paper has shown both the level of association and the correlation degree between SFSs. The offered correlation metric has handled the linguistic hedges properly and the existing ones have led to unsatisfactory results. In bidirectional approximate reasoning, the suggested metric has given accurate results. Also, the suggested SF correlation coefficient has given satisfactory results in pattern investigation and has a very high CD than some available comparison measures. In the future, we will discuss its application in decision-making, clustering, medical diagnosis, etc.

References


