A new fuzzy peer assessment methodology for cooperative learning of students

Kok Chin Chai a,*, Kai Meng Tay a, Chee Peng Lim b

a Faculty of Engineering, Universiti Malaysia Sarawak, Malaysia
b Centre for Intelligent Systems Research, Deakin University, Australia

A R T I C L E   I N F O

Article history:
Received 23 December 2014
Received in revised form 1 March 2015
Accepted 31 March 2015
Available online 14 April 2015

Keywords:
Perceptual computing
Peer assessment
Fuzzy preference relations
Linguistic grades
Individual weighting factor

A B S T R A C T

In this paper, a new fuzzy peer assessment methodology that considers vagueness and imprecision of words used throughout the evaluation process in a cooperative learning environment is proposed. Instead of numerals, words are used in the evaluation process, in order to provide greater flexibility. The proposed methodology is a synthesis of perceptual computing (Per-C) and a fuzzy ranking algorithm. Per-C is adopted because it allows uncertainties of words to be considered in the evaluation process. Meanwhile, the fuzzy ranking algorithm is deployed to obtain appropriate performance indices that reflect a student’s contribution in a group, and subsequently rank the student accordingly. A case study to demonstrate the effectiveness of the proposed methodology is described. Implications of the results are analyzed and discussed. The outcomes clearly demonstrate that the proposed fuzzy peer assessment methodology can be deployed as an effective evaluation tool for cooperative learning of students.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Cooperative learning is an educational approach or strategy in which students work in small groups to help each other to learn academic content [1]. The importance of cooperative learning in engineering disciplines has been explained and highlighted in [2–4]. Cooperative learning plays an important role to improve students’ soft skills, e.g., communication and teamwork, which is the essence in engineering studies [2]. While cooperative learning offers remarkable benefits (e.g., improving collaborative and critical thinking skills) to students at the tertiary level (i.e., the third stage of learning after graduating from the secondary school) [1], it is yet to be widely adopted owing to a number of practical challenges [5]. A search in the literature reveals that the assessment of an individual student in a group is not an easy task, since a group mark is often not a clear and fair reflection of each individual’s effort [5–7]. Besides that, it is difficult for an instructor to closely monitor each student’s efforts in a group; therefore it is not suitable for the instructor to assess each student’s contribution [8,9]. To tackle these challenges, peer assessment has been introduced to evaluate each student’s contribution in a group work [5–13]. Several successful case studies in peer assessment have been reported, e.g., in civil engineering [5], biological sciences [10], primary mathematics education [11], and computer studies [13]. In addition, substantial evidence to show that peer assessment can lead to improvements in quality and effectiveness learning is available [14].

Generally, there are two types of peer assessment [10]: (i) involving students in a class to assess other students’ work; (ii) involving students to assess the contribution/performance of other students within the same group. These two types of peer assessment can be further classified into two; i.e., formative or summative assessment [15,16]. The goals of formative assessment are to monitor students’ learning capabilities, gather their ongoing feedbacks, and improve their learning experience [15,16]. On the other hand, summative assessment evaluates students’ learning capabilities at the end of an instructional unit [15,16]. Typically, an instructor needs to decide whether to use the formative or summative form of peer assessment [14]. This paper focuses on summative assessment, which focuses on the outcome of a learning process [9]. The procedure for summative assessment is further detailed in Section 2.4.

Traditionally, the Likert scale (a numerical grading scale) is used in a way equivalent to psychological measurement [6,7,10,12]. As an example, a numerical grading scale (e.g., 1 to 5) can be used for assessment of group members [6,7,12], whereby “1” indicates “didn’t contribute”, “2” indicates “willing but not successful”, “3” indicates “average”, “4” indicates “above average”, and “5” indicates “outstanding” [6,7,12]. Even though the use of numerals in peer assessment is popular, it suffers from problems associated with
psychological measurement in terms of the meaning pertaining to the numerals used (see [17] for a study on the theoretical relationship between measurement and marking). It would be more natural to define assessment grades using subjective and vague linguistic terms [17]. Furthermore, the conventional method aggregates individual scores to produce a total score. In some situations (as illustrated in Section 4.3), it is difficult to distinguish the ranking order of students using the same numeral score.

Fuzzy set theory has been used in education assessment [9,18–23]. It is useful to deal with linguistic grades such as “didn’t contribute”, “willing but not successful”, and “average” in a grading system, which involve a substantial amount of fuzziness and vagueness [18]. It is worth mentioning that fuzzy set theory is an efficient and effective method to represent uncertainties [18,19]. Comparing with methods based on numerical grading scores [6,7,12], fuzzy set theory offers an alternative to linguistic evaluation in which “fuzzy” words, instead of numerals, are used during the assessment procedure [9,18–23]. Besides that, “computing with words”, as coined by Zadeh, is also a methodology related to fuzzy set theory, whereby the objects of computation are words and propositions drawn from a natural language [24,25].

Motivated by the success of fuzzy set theory in education assessment [9,18–23], this paper aims to propose a fuzzy peer assessment methodology that evaluates each student's contribution in a group work. The proposed methodology is a synthesis of perceptual computing (Per-C) [25–29] and a fuzzy ranking algorithm that uses fuzzy preference relations [30]. The rationale for the proposed methodology hinges on a number of imperatives. Firstly, the available information is too imprecise to be justified with numerals, which is more suitable to be represented using words [25]. In this paper, Per-C is adopted owing to its effectiveness in handling inherent uncertainties in words [25]. Specifically, Per-C is able to handle subjectivity, vagueness, imprecision, and uncertainty while achieving tractability and robustness in modeling human decision-making behaviors [25–29]. Comparing with type-1 fuzzy models [9,18–23], Per-C adopts interval type-2 fuzzy sets (IT2FSs) in tackling a decision-making problem [25–29]. IT2FSs have more flexibility in preserving and processing uncertainties than type-1 fuzzy sets [25]. Indeed, Per-C has been successfully implemented to undertake a number of fuzzy multiple criteria hierarchical decision making problems [25–29]. In [25–29], Per-C focuses on ranking the sequence of outcomes, mapping the outcomes into words and/or classifying the outcomes into different categories. Nevertheless, the use of Per-C in peer assessment is still new. In this paper, the relative importance of the outcomes, i.e., the contribution of each student with respect to those from other students, is examined in detail, which is yet to be investigated in the literature, e.g. [25–29]. In this aspect, our preliminary work [30], as discussed in Section 2.3, is further extended to serve this purpose. Then, the effectiveness and practicability of the proposed methodology are evaluated with a case study in an engineering course (i.e., Multiprocessors Architecture) at Universiti Malaysia Sarawak. The results from the conventional and proposed methodologies are analyzed and discussed. In essence, this paper contributes to a new fuzzy peer assessment methodology in which human linguistic words are adopted in the entire assessment process. Besides that, the proposed methodology provides an insight pertaining to each individual's contribution; therefore providing personalized assessment in a cooperative learning environment.

The rest of this paper is organized as follows. In Section 2, the background of fuzzy sets, perceptual computing, fuzzy ranking algorithms and peer assessment in problem-based learning is presented. In Section 3, a new technique for fuzzy peer assessment is explained in detail. In Section 4, a case study is conducted to demonstrate the usefulness of the proposed fuzzy peer assessment methodology. Concluding remarks and suggestions for future research are presented in Section 5.

2. Preliminaries

A number of notations and definitions related to type-1 fuzzy sets (T1FSs) and interval type-2 fuzzy sets (IT2FSs) are presented in Section 2.1. A review on perceptual computing is presented in Section 2.2. Our preliminary work related to a fuzzy ranking algorithm is reviewed in Section 2.3. Finally, an overview on peer assessment in problem-based learning is presented in Section 2.4.

2.1. Definitions

Consider a set of trapezoidal T1FSs, i.e., $\mathcal{A}_i$, where $i = 1, 2, ..., m$, in the universe of discourse, $U$. A trapezoidal T1FS $\mathcal{A}_i$ is parameterized as $\mathcal{A}_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, H_i)$, as illustrated in Fig. 1.

**Definition 1.** [30,31]: A fuzzy membership function, $\mu_{\mathcal{A}_i}$, of $\mathcal{A}_i$, as shown in Fig. 1, is defined as follows:

$$\mu_{\mathcal{A}_i}(X) = \begin{cases} \mu_{H_i}^L(X), & a_{i1} \leq X \leq a_{i2}, \\ H_i, & a_{i2} \leq X \leq a_{i3}, \\ \mu_{H_i}^R(X), & a_{i3} \leq X \leq a_{i4}, \\ 0, & \text{otherwise.} \end{cases}$$

where $\mu_{H_i}^L$ is continuous and strictly increasing in interval $[a_{i1}, a_{i2}]$, as defined in Eq. (2), $\mu_{H_i}^R$ is continuous and strictly decreasing in interval $[a_{i3}, a_{i4}]$, as defined in Eq. (3), and $H_i \in [0,1]$. Besides that, $a_{i1}, a_{i2}, a_{i3}, a_{i4}$ are real values, i.e., $\forall a_{i1}, a_{i2}, a_{i3}, a_{i4} \in \mathbb{R}$, such that $a_{i1} \leq a_{i2} \leq a_{i3} \leq a_{i4}$, and $\exists x \in U$.

$$\begin{align*} \mu_{H_i}^L & : [a_{i1}, a_{i2}] \rightarrow [(x - a_{i1})/(a_{i2} - a_{i1})]H_i \\ \mu_{H_i}^R & : [a_{i3}, a_{i4}] \rightarrow [(a_{i4} - x)/(a_{i4} - a_{i3})]H_i. \end{align*}$$

**Definition 2.** [30,32,33]: An IT2FS $\tilde{\mathcal{A}}_i$ is denoted as $\tilde{\mathcal{A}}_i = (\mathcal{A}_i, \bar{\mathcal{A}}_i)$, and $\bar{\mathcal{A}}_i$ is parameterized in Eq. (4). The upper and lower membership functions of $\mathcal{A}_i$ (i.e., $\mu_{\mathcal{A}_i}$ and $\mu_{\bar{\mathcal{A}}_i}$, respectively) are represented by type-1 membership functions.

$$\bar{\mathcal{A}}_i = (\tilde{\mathcal{A}}_i, \mathcal{A}_i) = (\tilde{\mathcal{A}}_i, \mathcal{A}_i(H_i)), (\tilde{\mathcal{A}}_i, \tilde{\mathcal{A}}_i, \mathcal{A}_i, \mathcal{A}_i(H_i)) \quad (4)$$

**Definition 3.** [32]: The fuzzy addition operation between two IT2FSs is defined as follows:

![Fig. 1. The membership function of a trapezoidal fuzzy set.](image)
\[ A_{i} \otimes \bar{A}_{i} = \langle (\bar{a}_{11}, \bar{a}_{12}, \bar{a}_{13}, \bar{a}_{14}; H_{1}), (\bar{q}_{11}, \bar{q}_{12}, \bar{q}_{13}, \bar{q}_{14}; H_{2}) \rangle, \]
\[ \ominus \left( \langle \bar{a}_{21}, \bar{a}_{22}, \bar{a}_{23}, \bar{a}_{24}; H_{1} \rangle, \langle \bar{q}_{21}, \bar{q}_{22}, \bar{q}_{23}, \bar{q}_{24}; H_{2} \rangle \right) \]
\[ = \left( \left( \bar{a}_{11} + \bar{a}_{21}, \bar{a}_{12} + \bar{a}_{22}, \bar{a}_{13} + \bar{a}_{23}, \bar{a}_{14} + \bar{a}_{24}; \min \left( H_{1}, H_{2} \right) \right), (\bar{q}_{11} + \bar{q}_{21}, \bar{q}_{12} + \bar{q}_{22}, \bar{q}_{13} + \bar{q}_{23}, \bar{q}_{14} + \bar{q}_{24}; \min \left( H_{1}, H_{2} \right) \right) \right). \]

**Definition 4.** [32]: The fuzzy multiplication operation between an IT2FS and crisp value \( k \) is defined as follows:

\[ k \bar{A}_{i} = \langle (k \times \bar{a}_{11}, k \times \bar{a}_{12}, k \times \bar{a}_{13}, k \times \bar{a}_{14}; H_{1}), (k \times \bar{q}_{11}, k \times \bar{q}_{12}, k \times \bar{q}_{13}, k \times \bar{q}_{14}; H_{2}) \rangle. \]

### 2.2. Review on perceptual computing

The general structure of Per-C [25–29] is depicted in Fig. 2. It consists of three components [25–29], i.e., an encoder, a computing-with-words (CWW) engine, and a decoder. Linguistic grades or words from humans are converted into IT2FSs through the encoder. The CWW engine aggregates the outputs from the encoder. The decoder maps the outputs of the CWW engine into a recommendation, which can be in the form of a word, rank, or class. However, the decoder ranks the outputs independently without considering their relative importance with respect to the recommendation. Nevertheless, it is imperative to rank each student’s contribution and derive a set of performance indices that reflect the student’s relative contribution in fuzzy peer assessment (i.e., peer assessment scores (PA)). While many fuzzy ranking algorithms are available in the literature, to the best of our knowledge, only a few solutions (e.g., [30] and [32]) are focused on the relative importance of two or more IT2FSs. The details are presented in Section 2.3.

### 2.3. Review on fuzzy ranking algorithms

As discussed earlier, our focus is on investigating the relative importance of two or more IT2FSs. As such, the methods in [30] and [32] are considered. Specifically, our preliminary work [30], which is an extension of that in [32], forms the foundation of the proposed method. In [30], we studied the rationality, i.e., the fulfillment of the six reasonable ordering properties as stated in [25, 34], and presented a number of improvements as compared with those in [32]. Assume that a set of IT2FSs, \( \bar{A}_{i} \), where \( i = 1, 2, \ldots, m \). To better clarify the explanation, a simulated example with two IT2FSs (i.e., \( \bar{A}_{1} = ((5.03,7.03,8.25,9.38;1),(6.03,7.75,7.93,8.82;0.75)) \) and \( \bar{A}_{2} = ((4.03,6.03,7.07,8.93;1),(5.03,6.57,6.60,8.00;0.75)) \), as illustrated in Fig. 3, is considered. The fuzzy ranking algorithm in [30] is summarized in five steps, as follows.

**Step 1.** Discretize the support of \( \bar{A}_{1} \) and \( \bar{A}_{2} \) of \( \bar{A}_{i} \) into \( N \) points, i.e., \( x_{\bar{A}_{i},k} \) and \( x_{\bar{A}_{j},k} \), \( k = 1, 2, 3, \ldots, N \), and obtain \( \mu_{\bar{A}_{i}}(x_{\bar{A}_{i},k}) \) and \( \mu_{\bar{A}_{j}}(x_{\bar{A}_{j},k}) \), respectively.

As an example, the discretized points of \( \bar{A}_{1} \) are expressed in a sequence of \( x_{\bar{A}_{1},1}, x_{\bar{A}_{1},2}, \ldots, x_{\bar{A}_{1},N} \), where \( X_{\bar{A}_{1}}, 1 \) and \( X_{\bar{A}_{1}}, N \) are the left and right-end points of \( \bar{A}_{1} \), respectively. The discretized points in the horizontal component of \( \bar{A}_{1} \) and \( \bar{A}_{2} \), i.e., \( x_{\bar{A}_{1},k} \) and \( x_{\bar{A}_{2},k} \) are computed using Eqs. (7) and (8), respectively. On the other hand, the discretized points in the vertical component of \( \bar{A}_{1} \) and \( \bar{A}_{2} \), i.e., \( \mu_{\bar{A}_{1}}(x_{\bar{A}_{1},k}) \) and \( \mu_{\bar{A}_{2}}(x_{\bar{A}_{2},k}) \) are computed using Eq. (1). The discretized points of \( \bar{A}_{2} \) are expressed in Eqs. (9) and (11) as well as \( \bar{A}_{j} \) are expressed in Eqs. (10) and (12), as follows.

\[ x_{\bar{A}_{1},k} = x_{\bar{A}_{1},1} + (k - 1) \left( \frac{x_{\bar{A}_{1},N} - x_{\bar{A}_{1},1}}{N - 1} \right) \]  
\[ x_{\bar{A}_{2},k} = x_{\bar{A}_{2},1} + (k - 1) \left( \frac{x_{\bar{A}_{2},N} - x_{\bar{A}_{2},1}}{N - 1} \right) \]  
\[ x_{\bar{A}_{1}} = \begin{bmatrix} x_{\bar{A}_{1},1} \\ \vdots \\ x_{\bar{A}_{1},N} \end{bmatrix} \]  
\[ x_{\bar{A}_{2}} = \begin{bmatrix} x_{\bar{A}_{2},1} \\ \vdots \\ x_{\bar{A}_{2},N} \end{bmatrix} \]  
\[ \mu_{\bar{A}_{1}}(x_{\bar{A}_{1},k}) \]  
\[ \mu_{\bar{A}_{2}}(x_{\bar{A}_{2},k}) \]  

With the example in Fig. 3, the left- and right-end points of \( \bar{A}_{1} \) (i.e., 5.03 and 9.38) and \( \bar{A}_{2} \) (i.e., 4.03 and 8.93) are discretized into \( N = 1000 \) using Eq. (7). The left- and right-end points of \( \bar{A}_{1} \) and \( \bar{A}_{2} \) are computed using Eq. (8). The discretized points for \( \bar{A}_{1} \) and \( \bar{A}_{2} \) as well as \( \bar{A}_{1} \) and \( \bar{A}_{2} \) are represented as matrices, i.e., Eqs. (9) and (11) as well as Eqs. (10) and (12), respectively.

**Step 2.** Compute \( P(\bar{A}_{1} \geq \bar{A}_{2}) \) and \( P(\bar{A}_{2} \geq \bar{A}_{1}) \) using Eqs. (13) and (14), respectively. Note that \( P(\bar{A}_{1} \geq \bar{A}_{2}) \) and \( P(\bar{A}_{2} \geq \bar{A}_{1}) \) are the ratios of the distance between two FSs of favourable outcomes to the total distance pertaining to two FSs of the entire possible outcomes.

\[ P(\bar{A}_{1} \geq \bar{A}_{2}) = \max \left( 1 - \frac{\bar{F}_{1}}{\bar{F}_{2}}, 0 \right) \]  
\[ P(\bar{A}_{2} \geq \bar{A}_{1}) = \max \left( 1 - \frac{\bar{F}_{2}}{\bar{F}_{1}}, 0 \right) \]
where,
\[
E^u_{ji} = \max_{k=1,2,...,N} (x_{A_i,k}) - \min_{k=1,2,...,N} (x_{A_j,k}) + \frac{1}{N} \sum_{k=1}^{N} \left[ \max(x_{A_i,k} - x_{A_j,k}, 0) + \max(\mu_{A_i}(x_{A_i,k}) - \mu_{A_j}(x_{A_j,k}), 0) \right]
\]
\[
E^l_{ji} = (x_{A_i,N} - x_{A_j,1}) + \frac{1}{N} \sum_{k=1}^{N} \left[ x_{A_i,k} - x_{A_j,k} + \left| \mu_{A_i}(x_{A_i,k}) - \mu_{A_j}(x_{A_j,k}) \right| \right]
\]
i, j = 1, 2, 3, ..., m

\[
P(A_j \geq A_i) = \max \left( 1 - \max \left( \frac{E^u_{ji}}{E^l_{ji}}, 0 \right), 0 \right).
\]

where,
\[
E^u_{ji} = \max_{k=1,2,...,N} (x_{A_j,k}) - \min_{k=1,2,...,N} (x_{A_i,k}) + \frac{1}{N} \sum_{k=1}^{N} \left[ \max(x_{A_j,k} - x_{A_i,k}, 0) + \max(\mu_{A_j}(x_{A_j,k}) - \mu_{A_i}(x_{A_i,k}), 0) \right]
\]
\[
E^l_{ji} = (x_{A_j,N} - x_{A_i,1}) + \frac{1}{N} \sum_{k=1}^{N} \left[ x_{A_j,k} - x_{A_i,k} + \left| \mu_{A_j}(x_{A_j,k}) - \mu_{A_i}(x_{A_i,k}) \right| \right]
\]
i, j = 1, 2, 3, ..., m

Step 3. Generate the fuzzy preference matrices as in Eqs. (15) and (16).

\[
P = \left[ \begin{array}{cccc}
p(A_1 \geq A_1) & \ldots & p(A_1 \geq A_j) & \ldots & p(A_1 \geq A_m) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
p(A_i \geq A_1) & \ldots & p(A_i \geq A_j) & \ldots & p(A_i \geq A_m) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
p(A_m \geq A_1) & \ldots & p(A_m \geq A_j) & \ldots & p(A_m \geq A_m)
\end{array} \right].
\]

\[
P^* = \left[ \begin{array}{cccc}
p(A_1 \geq \bar{A}_1) & \ldots & p(A_1 \geq \bar{A}_j) & \ldots & p(A_1 \geq \bar{A}_m) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
p(A_i \geq \bar{A}_1) & \ldots & p(A_i \geq \bar{A}_j) & \ldots & p(A_i \geq \bar{A}_m) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
p(A_m \geq \bar{A}_1) & \ldots & p(A_m \geq \bar{A}_j) & \ldots & p(A_m \geq \bar{A}_m)
\end{array} \right].
\]

Based on the aforementioned example, the fuzzy preference matrices for \(A_1\) and \(A_2\) (i.e., \(P\)) and for \(\bar{A}_1\) and \(\bar{A}_2\) (i.e., \(P^*\)) are computed using Eq. (13) and Eq. (14) in Step 2, respectively. The simulation outcomes from Step 2 are compiled and written in Eqs. (15) and (16), as follows.

\[
P = \begin{bmatrix} 0.5000 & 0.6098 \\ 0.3902 & 0.5000 \end{bmatrix}, \quad P^* = \begin{bmatrix} 0.5000 & 0.7036 \\ 0.2964 & 0.5000 \end{bmatrix}
\]

1. Divide all students into several groups
   - Group #1
   - Group #2
   - ... Groups
2. Generate peer assessment questionnaire by an instructor of the course and students
3. Proceed with cooperative learning in each group
4. Assess each group’s work by the instructor
5. Assess each student’s work by other students in the same group
6. Calculate the Individual Weighting Factor (IWF)
7. Calculate the score that reflects each student’s efforts

Fig. 4. The procedure to assess an individual student’s contribution in a group project [9].

Step 4. Compute the ranking indices of \(\bar{A}_i\) and \(\bar{A}_j\) using Eqs. (17) and (18), respectively.

\[
RI(\bar{A}_i) = \frac{1}{m(m-1)} \left( \sum_{j=1}^{m} p(\bar{A}_i \geq \bar{A}_j) + \frac{m}{2} - 1 \right)
\]

\[
RI(\bar{A}_j) = \frac{1}{m(m-1)} \left( \sum_{j=1}^{m} p(\bar{A}_j \geq \bar{A}_i) + \frac{m}{2} - 1 \right)
\]

Again, based on the same example, using Eq. (17), the ranking indices for \(\bar{A}_1\) and \(\bar{A}_2\) are 0.5549 and 0.4451, respectively. Using Eq. (18), the ranking indices for \(\bar{A}_1\) and \(\bar{A}_2\) are 0.6018 and 0.3982, respectively.

Step 5. The ranking index of \(\bar{A}_i\) is computed using Eq. (19).

\[
RI(\bar{A}_i) = \frac{RI(\bar{A}_i) + RI(\bar{A}_j)}{2}
\]

A higher ranking index of \(\bar{A}_i\) indicates a higher ranking of \(\bar{A}_i\). With the same example, using Eq. (19), the ranking indices of \(\bar{A}_1\) and \(\bar{A}_2\) are 0.5784 and 0.4216, respectively. Therefore, \(\bar{A}_1\) has a higher ranking than \(\bar{A}_2\).

2.4. Overview on peer assessment in problem-based learning

2.4.1. Background

A summative assessment procedure, as discussed in [9], is presented in Fig. 4. Firstly, students are divided into several groups, and a set of peer assessment questionnaires that has been agreed upon by the instructor and students is distributed. Then, each group proceeds with the process of cooperative learning. At the end of the cooperative learning process, the students are required to submit and present their findings for assessment. Their achievements are evaluated by the instructor as well as by their peers, as shown in the 5th stage in Fig. 4. The results from peer assessment (denoted as the PA scores) are aggregated and the Individual Weighting Factor (IWF) is derived, as in the 6th stage. IWF is a measure of each individual’s contribution in his/her group. There are several methods to compute IWF [5,6]. In this paper, the method in [5] is adopted owing to a number of advantages, i.e., it discourages free-riders, encourages above-average contributions, and discourages individualistic behaviours in a team. Finally, the result that reflects each student’s effort is computed.
In this paper, a cooperative learning activity pertaining to an engineering course (i.e., Multiprocessors Architecture) at Universiti Malaysia Sarawak, was studied. The goals of cooperative learning were to improve students' problem-solving skills in groups. Students were required to focus on a task related to computing problems, i.e., designing a multiprocessor-based system for data analysis.

A total of 28 students participated in the cooperative learning activity. In accordance with Fig. 4, they were divided into 7 groups by the course instructor. Each group consists of four members with different proficiency levels and background. The students were required to complete their project within a given deadline. Then, they were required to submit their project report and present the software that they developed. Peer assessment was conducted using an evaluation form provided by the instructor. Each student was assessed based on seven criteria, as shown in Table 1. Each criterion was tagged with a weight predefined by the instructor, i.e., W1, W2, W3, W4, W5, W6, and W7, corresponding to 20%, 20%, 15%, 15%, 10%, 10%, and 10%, respectively. An example of the students' peer assessment results is shown in Table 1.

Each group (i.e., #g, where g = 1, 2, 3, ..., 7) consisted of four students (i.e., SgA, SgB, SgC, SgD). SgR and SgS represented the student being assessed and the (three) evaluators, where SgR ∈ SgA, SgB, SgC, SgD), respectively. Self-evaluation was not practiced, i.e., SgR & SgS indicated the evaluator's weight (i.e., a measure of the evaluator's expertise level). In this study, each evaluator was given an equal weight. As an example of the peer assessment structure of group #1 is depicted in Fig. 5.

Besides that, each group was assessed by the instructor based on two general criteria, i.e., project report and project presentation, as shown in Table 2. These two general criteria were further divided into a number of specific criteria. An example of the group assessment structure (i.e., for group #1) is shown in Table 2, whereby group #1 scored 83.2% out of 100.

### Table 1

<table>
<thead>
<tr>
<th>Student Being Assessed Specific Criteria, i</th>
<th>Evaluators Weighing</th>
<th>#1A</th>
<th>#1B</th>
<th>#1C</th>
<th>#1D</th>
<th>#1A</th>
<th>#1B</th>
<th>#1C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Participated in group meetings</td>
<td>(20%)</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>0.60</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>2. Communicated constructively to discussion</td>
<td>(20%)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3. Generally was cooperative in group activities</td>
<td>(15%)</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.75</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>4. Contributed to good problem-solving skills</td>
<td>(15%)</td>
<td>0.60</td>
<td>0.45</td>
<td>0.45</td>
<td>0.60</td>
<td>0.60</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>5. Contributed useful ideas</td>
<td>(10%)</td>
<td>0.50</td>
<td>0.30</td>
<td>0.30</td>
<td>0.50</td>
<td>0.40</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>6. Demonstrated good interest to task given</td>
<td>(10%)</td>
<td>0.40</td>
<td>0.40</td>
<td>0.50</td>
<td>0.40</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>7. Prepared drafts of report in good quality</td>
<td>(10%)</td>
<td>0.50</td>
<td>0.40</td>
<td>0.40</td>
<td>0.50</td>
<td>0.50</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>

**Aggregated Results from each evaluator**

- #1A: 4.20
- #1B: 4.35
- #1C: 4.25

### Table 2

<table>
<thead>
<tr>
<th>Project Marking Form</th>
<th>Specific Criteria</th>
<th>Group Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Aspects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Report (60%)</td>
<td>1. Correct use of methods (25%)</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>2. Design, simulate and analyze the system to meet the requirements of the project (25%)</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>3. Delivery the outcomes of the project in written forms (25%)</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>4. Structure and presentation of the project report (25%)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td><strong>Total mark (out of 60%)</strong></td>
<td><strong>50.4</strong></td>
</tr>
<tr>
<td>Project Presentation (40%)</td>
<td>5. Deliver the outcomes of the project in oral forms (25%)</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>6. Reliability of the proposed method (25%)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>7. Efficiency and flexibility of the proposed method (25%)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>8. Responses to questions and defend their projects with valid reasoning and arguments (25%)</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td><strong>Total mark (out of 40%)</strong></td>
<td><strong>32.8</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Total group score (out of 100%)</strong></td>
<td><strong>83.2</strong></td>
</tr>
</tbody>
</table>

2.4.2. The conventional peer assessment methodology

The conventional peer assessment methodology in [5,6] is reviewed. The methodologies, as presented in [5,6], can be represented by Eqs. (20) and (21). Firstly, a peer assessment score (i.e., PA), which is the ratio of the actual sum of scores to the highest possible score, is calculated. Then, the IWF that reflects each student's contribution is computed using Eq. (21) [5]. Finally, each individual's mark is computed using the parabolic equation of Eq. (22), where \( \alpha \) is the scale factor that has an impact on the spread of the individual marks. In this paper, \( \alpha = 1.2 \) is adopted (as indicated in [5]).

\[
PA = \frac{\text{Actual sum scored}}{\text{Highest possible score}} \tag{20}
\]

\[
IWF = \frac{PA}{\text{Average PA score}} \tag{21}
\]

\[
IM = GM \times \begin{cases} 
IWF - \frac{(IWF - 1)^2}{2\beta}, & 1 < IWF < 1 + \beta \\
1 + 0.5\beta, & IWF \geq 1 + \beta 
\end{cases} \tag{22}
\]

where \( \beta = \alpha(1 - 0.01 GM) \). GM denotes the group mark from the instructor.

Conventionally, the five-point Likert scale is commonly used, whereby "1", "2", "3", "4", and "5" indicate "poor", "below average", "average", "above average", and "excellent", respectively. Consider the example shown in Table 1. Students #1A and #1D are evaluated. Student #1A is rated 4, 5, and 5 by Evaluators #1B, #1C, and #1D, respectively, for the first specific criterion i.e., "participated in group meetings". Since the criterion is weighted at 20%, the ratings (i.e., 4, 5, and 5) are normalized to 0.80, 1.00, and 1.00, respectively, as presented in Table 1. By applying Eqs. (20) and (21), PA and IWF are computed. The results are tabulated in Table 3.
Table 3
The outcomes of Group 1 using the conventional methodology.

<table>
<thead>
<tr>
<th>Students</th>
<th>#1A</th>
<th>#1D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer assessment score, PA</td>
<td>0.8533</td>
<td>0.8533</td>
</tr>
<tr>
<td>Individual Weighting Factor, IWF</td>
<td>1.0503</td>
<td>1.503</td>
</tr>
</tbody>
</table>

3. A new methodology for fuzzy peer assessment

In this section, the proposed new methodology for fuzzy peer assessment is described. Fig. 6 shows the overall fuzzy peer assessment framework. In this paper, the focus is on enhancing the 5th and 6th stages of the conventional methodology (as illustrated in Fig. 4) by using the Per-C paradigm [25–28]. The details of the proposed methodology are summarized in stages 5.1–5.4 and stage 6.0, as follows.

3.1. Encoder

The encoder converts human linguistic words into IT2FSs. This operation is performed based on the Interval Approach (IA), whereby the IA codebook can be found in [25–27]. In this paper, five words (i.e., Poor, Below Average, Average, Above Average, and Excellent, abbreviated as P, BA, A, AA, and EX) are retrieved. The details of these five words (namely linguistic grades) are summarized in Table 4 and illustrated in Fig. 7. Each linguistic grade pertaining to the specific criterion is further described by the instructor. The interpretation of these linguistic grades is explained in Table A1 (Appendix A), which provides the details to each evaluator to aid his/her assessments.

3.2. Computing-with-words

Table 5 depicts the assessment of students #1A and #1D with linguistic grades. Using the CWW engine, the assessments from all evaluators are firstly aggregated with Eq. (23).

$$\bar{y}_{g,r} = \frac{1}{7} \sum_{i=1}^{7} \tilde{x}_{g,r,i} w_i.$$  \hfill (23)

where $\tilde{x}_{g,r,i}$ denotes the linguistic grade given by an evaluator (i.e., $r$), $g$ indicates the group that $r$ belongs to (i.e., $g = 1, 2, 3, \ldots, 7$), and
Table 4
IT2FS for linguistic terms [25,26].

<table>
<thead>
<tr>
<th>Linguistic Grades</th>
<th>Abbreviations</th>
<th>UMF ($\delta_1$, $\delta_2$, $\delta_3$, $\delta_4$; $H$)</th>
<th>LMF ($\delta_1$, $\delta_2$, $\delta_3$, $\delta_4$; $H$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Poor’</td>
<td>$P$</td>
<td>$[0.00, 0.00, 1.50, 3.50; 1]$</td>
<td>$[0.00, 0.00, 2.50, 1.00]$</td>
</tr>
<tr>
<td>‘Below Average’</td>
<td>$BA$</td>
<td>$[0.50, 2.50, 3.50, 5.50; 1]$</td>
<td>$[1.50, 0.00, 3.00, 4.50; 0.75]$</td>
</tr>
<tr>
<td>‘Average’</td>
<td>$A$</td>
<td>$[2.50, 4.50, 5.50, 7.50; 1]$</td>
<td>$[3.50, 5.00, 6.50, 5.00; 0.75]$</td>
</tr>
<tr>
<td>‘Above Average’</td>
<td>$AA$</td>
<td>$[4.50, 6.50, 7.50, 9.50; 1]$</td>
<td>$[5.50, 7.00, 8.50, 8.00; 0.75]$</td>
</tr>
<tr>
<td>‘Excellent’</td>
<td>$E$</td>
<td>$[6.50, 8.50, 10.00, 10; 1]$</td>
<td>$[7.50, 9.50, 10.00, 10; 1.00]$</td>
</tr>
</tbody>
</table>

![Fig. 7. Linguistic grades (a) Poor; (b) Below Average; (c) Average; (d) Above Average and (e) Excellent.](image)

Table 5
An example of peer assessment using the proposed methodology.

<table>
<thead>
<tr>
<th>Student Criteria Assessed</th>
<th>Evaluators Weighing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#1A</td>
</tr>
<tr>
<td></td>
<td>#1B</td>
</tr>
<tr>
<td></td>
<td>#1C</td>
</tr>
<tr>
<td></td>
<td>#1D</td>
</tr>
<tr>
<td>1. Participated in group meetings</td>
<td>(20%) AA</td>
</tr>
<tr>
<td>2. Communicated constructively to discussion</td>
<td>(20%) EX</td>
</tr>
<tr>
<td>3. Generally was cooperative in group activities</td>
<td>(15%) AA</td>
</tr>
<tr>
<td>4. Contributed to good problem-solving skills</td>
<td>(15%) AA</td>
</tr>
<tr>
<td>5. Contributed useful ideas</td>
<td>(10%) EX</td>
</tr>
<tr>
<td>6. Demonstrated good interest to task given</td>
<td>(10%) AA</td>
</tr>
<tr>
<td>7. Prepared drafts of report in good quality</td>
<td>(10%) EX</td>
</tr>
</tbody>
</table>


i denotes the specific criterion (i.e., $i = 1, 2, 3, \ldots, 7$). As an example, $X_{1, B, 1}$ is the linguistic grade given to the first specific criterion (i.e., “participated in group meetings”) by Evaluator B in group #1 (i.e., #1B) and $X_{1, B, 1}$ is $A$ in Table 5. Note that $w_i$ denotes the normalized weight of $W_i$ (i.e., $w_i \in [0, 1]$). As an example, $W_1 = 20\%$ and $W_1 = 0.20$. On the other hand, $y_{E, r}$ denotes the aggregated outcomes of Evaluator $r$ in group #g.

The results from all evaluators are aggregated using Eq. (24).

$$\hat{Y}_{E, g} = \frac{\sum_{r \in g, R \in R_{E, g}} W_{g, r}}{\sum_{r \in g, R \in R_{E, g}} W_{g, r}}$$  \hspace{1cm} (24)

where $\hat{W}_{g, r}$ indicates the expertise level (or weight) of Evaluator $r$ in group #g. In this paper, $\hat{W}_{g, A} = \hat{W}_{g, B} = \hat{W}_{g, C} = \hat{W}_{g, D}$ i.e., all students are equally weighted. Therefore, Eq. (24) can be simplified to

$$\hat{Y}_{E, g} = \frac{1}{|R|} \sum_{r \in g, R \in g} y_{E, r}$$  \hspace{1cm} (25)

where $|R|$ refers to the number of evaluators involved. Notice that fuzzy addition and multiplication operations (as in Definitions 3 and 4) are used in Eqs. (23), (24) and (25).

3.3. Decoder

The roles of the decoder are two-fold, i.e., to map the aggregated outcomes (represented in IT2FS, i.e., $\hat{Y}_{E, g}$) from CWB to recommendation in term of words (Stage 3) and to rank the aggregated outcomes of students in the same group (Stage 4). The ranking indices are used to derive JIFW in the latter stage. The former attempts to map the aggregated outcomes, i.e., $\hat{Y}_{E, g}$, to the words listed in Table 4. The associated IT2FSs of these linguistic grades, i.e., poor, below average, average, above average and excellent, are denoted as $\hat{Y}_{P}$, $\hat{Y}_{BA}$, $\hat{Y}_{A}$, $\hat{Y}_{AA}$, and $\hat{Y}_{EX}$, respectively, and their corresponding IT2FSs are shown in Fig. 7. Using the Jaccard similarity indicator [25], the similarity indicators between $\hat{Y}_{E, g}$ and $\hat{Y}_{C}$ are denoted as $SJ(\hat{Y}_{E, g}, \hat{Y}_{C})$, where $\hat{Y}_{C} \in \{\hat{Y}_{P}, \hat{Y}_{BA}, \hat{Y}_{A}, \hat{Y}_{AA}, \hat{Y}_{EX}\}$. Eq. (26) shows the similarity indicator. $SJ(\hat{Y}_{E, g}, \hat{Y}_{C}) \approx 1$ indicates $\hat{Y}_{E, g}$ is similar to $\hat{Y}_{C}$.

$$SJ(\hat{Y}_{E, g}, \hat{Y}_{C}) \approx \frac{\sum_{t=1}^{T} \min(\mu_{\hat{Y}_{E, g}}(y_{t}), \mu_{\hat{Y}_{C}}(y_{t})) + \sum_{t=1}^{T} \min(\mu_{\hat{Y}_{E, g}}(y_{t}), 1 - \mu_{\hat{Y}_{C}}(y_{t}))}{\sum_{t=1}^{T} \max(\mu_{\hat{Y}_{E, g}}(y_{t}), \mu_{\hat{Y}_{C}}(y_{t})) + \sum_{t=1}^{T} \max(\mu_{\hat{Y}_{E, g}}(y_{t}), 1 - \mu_{\hat{Y}_{C}}(y_{t}))}$$  \hspace{1cm} (26)

where $\mu_{\hat{Y}_{E, g}}(y_{t})$ and $\mu_{\hat{Y}_{C}}(y_{t})$ are the upper and lower membership functions of $\hat{Y}_{E, g}$ and $\hat{Y}_{C}$, respectively. $y_{t} (t = 1, 2, \ldots, T)$ are equally spaced in the support region of $\hat{Y}_{E, g} \cup \hat{Y}_{C}$. Consider an output space $Y$, $\hat{Y}_{E, g}$, $\hat{Y}_{C}$, and $y_{t}$ are the elements in $Y$ (i.e., $\hat{Y}_{E, g}, \hat{Y}_{C}, y_{t} \in Y$). In this paper, each aggregated outcome, i.e., $\hat{Y}_{E, g}$, is mapped to two linguistic grades with the two largest similarity measures. $SJ(\hat{Y}_{E, g}, \hat{Y}_{C})$ is represented
in percentage (%) by multiplying $S(\tilde{Y}_{g,R}, \tilde{Y}_{g,C})$ with 100. These similarity indicators map the outcomes, $\tilde{Y}_{g,R}$, into recommendations as described in Table 6.

To rank the aggregated outcomes of students in the same group, consider group #g consists of four students and the assessment outcomes are denoted as $\tilde{Y}_{g,R} \in \{\tilde{Y}_{g,A}, \tilde{Y}_{g,B}, \tilde{Y}_{g,C}, \tilde{Y}_{g,D}\}$. $\tilde{Y}_{g,R}$ are the fuzzy outcomes represented by IT2FSs. These fuzzy outcomes are ranked based on the algorithm proposed in [30] and reviewed in Section 2.3. The results are denoted as $R(\tilde{Y}_{g,R}) \in [R(\tilde{Y}_{g,A}), R(\tilde{Y}_{g,B}), R(\tilde{Y}_{g,C}), R(\tilde{Y}_{g,D})]$ in which $R(\tilde{Y}_{g,A}) + R(\tilde{Y}_{g,B}) + R(\tilde{Y}_{g,C}) + R(\tilde{Y}_{g,D}) = 1$. Based on the ranking index, $R(\tilde{Y}_{g,R})$, the IWF of each student (i.e., $IWF(\tilde{Y}_{g,R})$) is obtained using Eq. (27).

$$IWF(\tilde{Y}_{g,R}) = |S| \times R(\tilde{Y}_{g,R}).$$  

(27)

where $|S|$ denotes the number of students in the group.

Subsequently, each student’s mark, IM, can be derived by substituting the results from Eq. (27) into Eq. (22).

### 4. A case study

In this section, a real case study is conducted. A variety of aspects from the initial survey to the end results are discussed and analyzed.

#### 4.1. Survey

A survey on the selection of the type of assessment grades (i.e., linguistic grades or numerals) for peer assessment was firstly conducted. The responses indicated that the use of both linguistic grades and numeral was acceptable. While linguistic grades were perceived as more natural than numerals for peer assessment, a clear description for each linguistic grade should be provided. Therefore, a guideline, as presented in Table A1 (Appendix A), was prepared.

#### 4.2. Results

The simulation results from the proposed fuzzy peer assessment methodology are represented as IT2FSs, as shown in Fig. 8. As an example, Fig. 8(a) presents the result of students in group #1. The legends (i.e., #1A, #1B, #1C, and #1D) indicate the four students in group #1. The mark of student #1A is as represented as an IT2FS. The detailed results from both conventional and proposed methodologies are presented in Table 7. Column “Students” indicates the student identification, columns “Conventional IWF” and “Proposed IWF” show the results from Eqs. (21) and (27), respectively. Column “CM” indicates the group marks from the instructor. Columns “IM with Conventional IWF (%)” and “IM with Proposed IWF (%)” indicate each student’s marks, as the outcome of Eq. (22). Column “Peer Rating” indicates the results from Eq. (26). As an example, row “#1A” represents student A in group #1. He/she attains IMF = 1.0503 and IMF = 1.0869 using the conventional and proposed methodologies. His/her group is awarded an overall assessment of 83.20%. The conventional methodology indicates that student #1A’s individual mark is 86.90%. Meanwhile, the proposed methodology awards student #1A 88.87%, and he/she is categorized as an “above average” (65.1%) and “excellent” (25.5%) student.

#### 4.3. Analysis

The results in Table 7 are further illustrated in Fig. 9 for the purpose of visualization. Fig. 9 shows the group marks as well as the scores of each student based on the conventional and proposed methodologies. The green, blue, and red rectangular bars, indicate the group marks, IM (individual mark) using the conventional methodology, and IM using the proposed methodologies, respectively. The group marks do not reflect individual contributions as all students in the same group receive the same score. As an example, students #1A, #1B, #1C, and #1D attain the same score of 83.20%. Such assessment approach can be supplemented with the conventional and proposed methodologies. As can be seen in Fig. 9, the outcomes from the conventional and proposed methodologies are well-correlated; but the proposed methodology is able to distinguish students with similar scores. As an example, students #1A, #1B, #1C, and #1D are given 86.90%, 77.14%, 80.90%, and 86.90% respectively, using the conventional methodology; therefore they are ranked 1, 4, 3, and 1, respectively. Meanwhile, these students are given 88.87%, 72.79%, 79.01%, and 89.95%, respectively, using the proposed methodology; therefore they are ranked 2, 4, 3, and 1, respectively. The conventional methodology provides the same IM for students #1A and #1D (i.e., 86.90%). However, their performance can be distinguished with the proposed methodology.

It is also worth mentioning that the proposed methodology provides both crisp scores and recommendations. As an example, IT2FS for student #1A, as in Fig. 8(a), can be defuzzified to provide two meaningful interpretations i.e., a crisp score of 88.87% and a recommendation of the achievement being “above average” (65.1%) and “excellent” (25.5%). From Table 6, student #1A is most likely a strong group member who tries hard to complete the project (“above average”). Student #1A is also potentially a group leader who works hard to complete the project (“excellent”). Compared with the conventional methodology which only provides an individual mark, the proposed methodology provides more meaningful evaluation outcomes.

#### 4.4. Similarity measure

The results from the fuzzy ranking algorithm [30] and Jaccard similarity measure algorithm [25] are in good agreement. As an example, $IWF(\tilde{Y}_{1,A})$ and $IWF(\tilde{Y}_{1,D})$ are 1.0869 and 1.0886, respectively. This shows that $IWF(\tilde{Y}_{1,D})$ is slightly larger than $IWF(\tilde{Y}_{1,A})$, which is in agreement with both students’ scores, i.e., student #1A with AA (65.1%), EX (25.5%) and student #1D with AA (65.1%), EX (25.6%). The results of students #1A and #1D are indeed very close, but their contributions are slightly different.

#### 4.5. Detection of possible free riders

Another key benefit of the proposed methodology is its ability to detect free rider(s) in a group. An illustrative example is as follows. Assume that student #1A is a free rider, and his/her grade is “poor”, as shown in Table 8.

The aggregated results using the proposed methodology are illustrated in Fig. 10 and Table 9. The group marks do not reflect individual contributions as students from the same group are given...
Fig. 8. Aggregated results of the students in each group: (a) Group #1; (b) Group #2; (c) Group #3; (d) Group #4; (e) Group #5; (f) Group #6; (g) Group #7.
Table 7
The outcomes using the conventional and proposed methodologies.

<table>
<thead>
<tr>
<th>Students</th>
<th>Conventional IWF</th>
<th>Proposed IWF</th>
<th>GM (%)</th>
<th>IM with Conventional IWF (%)</th>
<th>IM with proposed IWF (%)</th>
<th>Peer Rating (Recommendations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1A</td>
<td>1.0503</td>
<td>1.0869</td>
<td>83.20</td>
<td>86.900</td>
<td>88.870</td>
<td>AA(65.1%), EX(25.5%)</td>
</tr>
<tr>
<td>#1B</td>
<td>0.9272</td>
<td>0.8749</td>
<td>83.20</td>
<td>77.141</td>
<td>72.790</td>
<td>A(27.4%), AA(67.4%)</td>
</tr>
<tr>
<td>#1C</td>
<td>0.9723</td>
<td>0.9497</td>
<td>83.20</td>
<td>80.896</td>
<td>79.013</td>
<td>A(19.8%), AA(90.7%)</td>
</tr>
<tr>
<td>#1D</td>
<td>1.0503</td>
<td>1.0866</td>
<td>83.20</td>
<td>86.900</td>
<td>88.951</td>
<td>AA(65.1%), EX(25.5%)</td>
</tr>
<tr>
<td>#2A</td>
<td>0.9857</td>
<td>0.9686</td>
<td>87.60</td>
<td>86.349</td>
<td>84.849</td>
<td>AA(37.6%), EX(48.3%)</td>
</tr>
<tr>
<td>#2B</td>
<td>0.9785</td>
<td>0.9533</td>
<td>87.60</td>
<td>85.723</td>
<td>83.513</td>
<td>A(39.8%), EX(45.5%)</td>
</tr>
<tr>
<td>#2C</td>
<td>1.0250</td>
<td>1.0549</td>
<td>87.60</td>
<td>85.620</td>
<td>91.520</td>
<td>A(27.6%), EX(68.0%)</td>
</tr>
<tr>
<td>#2D</td>
<td>1.0107</td>
<td>1.0232</td>
<td>87.60</td>
<td>88.507</td>
<td>89.474</td>
<td>A(30.9%), EX(60.1%)</td>
</tr>
<tr>
<td>#3A</td>
<td>1.0052</td>
<td>1.0078</td>
<td>79.00</td>
<td>79.405</td>
<td>79.609</td>
<td>A(16.5%), EX(89.3%)</td>
</tr>
<tr>
<td>#3B</td>
<td>0.8769</td>
<td>0.7922</td>
<td>79.00</td>
<td>69.278</td>
<td>62.587</td>
<td>A(41.1%), AA(45.7%)</td>
</tr>
<tr>
<td>#3C</td>
<td>1.0590</td>
<td>1.1008</td>
<td>79.00</td>
<td>83.154</td>
<td>85.372</td>
<td>AA(65.1%), EX(25.7%)</td>
</tr>
<tr>
<td>#3D</td>
<td>1.0590</td>
<td>1.0991</td>
<td>79.00</td>
<td>83.154</td>
<td>85.290</td>
<td>AA(65.1%), EX(25.5%)</td>
</tr>
<tr>
<td>#4A</td>
<td>1.0741</td>
<td>1.1215</td>
<td>80.00</td>
<td>83.084</td>
<td>87.259</td>
<td>A(76.0%), EX(21.8%)</td>
</tr>
<tr>
<td>#4B</td>
<td>1.0569</td>
<td>1.0934</td>
<td>80.00</td>
<td>84.056</td>
<td>86.017</td>
<td>A(83.8%), EX(19.4%)</td>
</tr>
<tr>
<td>#4C</td>
<td>0.9753</td>
<td>0.9636</td>
<td>80.00</td>
<td>78.024</td>
<td>77.086</td>
<td>A(26.6%), AA(69.5%)</td>
</tr>
<tr>
<td>#4D</td>
<td>0.8937</td>
<td>0.8216</td>
<td>80.00</td>
<td>71.493</td>
<td>65.725</td>
<td>A(46.0%), AA(41.1%)</td>
</tr>
<tr>
<td>#5A</td>
<td>0.9002</td>
<td>0.8338</td>
<td>68.60</td>
<td>61.755</td>
<td>57.200</td>
<td>A(41.1%), AA(46.0%)</td>
</tr>
<tr>
<td>#5B</td>
<td>1.0531</td>
<td>1.0889</td>
<td>68.60</td>
<td>72.004</td>
<td>73.980</td>
<td>A(79.8%), EX(20.5%)</td>
</tr>
<tr>
<td>#5C</td>
<td>1.0786</td>
<td>1.1304</td>
<td>68.60</td>
<td>73.470</td>
<td>75.997</td>
<td>A(68.6%), EX(24.3%)</td>
</tr>
<tr>
<td>#5D</td>
<td>0.9682</td>
<td>0.9469</td>
<td>68.60</td>
<td>66.415</td>
<td>64.956</td>
<td>A(25.9%), AA(71.0%)</td>
</tr>
<tr>
<td>#6A</td>
<td>0.9260</td>
<td>0.8701</td>
<td>72.00</td>
<td>66.669</td>
<td>62.644</td>
<td>A(30.8%), AA(60.3%)</td>
</tr>
<tr>
<td>#6B</td>
<td>1.0344</td>
<td>1.0614</td>
<td>72.00</td>
<td>74.360</td>
<td>76.017</td>
<td>A(79.8%), EX(20.5%)</td>
</tr>
<tr>
<td>#6C</td>
<td>0.9885</td>
<td>0.9778</td>
<td>72.00</td>
<td>71.174</td>
<td>70.404</td>
<td>A(19.8%), AA(90.7%)</td>
</tr>
<tr>
<td>#6D</td>
<td>1.0511</td>
<td>1.0907</td>
<td>72.00</td>
<td>75.421</td>
<td>77.650</td>
<td>A(72.3%), EX(23.0%)</td>
</tr>
<tr>
<td>#7A</td>
<td>0.8665</td>
<td>0.7805</td>
<td>63.00</td>
<td>54.587</td>
<td>49.170</td>
<td>A(48.7%), AA(38.6%)</td>
</tr>
<tr>
<td>#7B</td>
<td>1.0599</td>
<td>1.1009</td>
<td>63.00</td>
<td>66.541</td>
<td>68.632</td>
<td>A(72.3%), EX(23.0%)</td>
</tr>
<tr>
<td>#7C</td>
<td>1.0768</td>
<td>1.1262</td>
<td>63.00</td>
<td>67.450</td>
<td>69.820</td>
<td>AA(65.1%), EX(25.5%)</td>
</tr>
<tr>
<td>#7D</td>
<td>0.9969</td>
<td>0.9925</td>
<td>63.00</td>
<td>62.801</td>
<td>62.526</td>
<td>A(19.8%), AA(90.7%)</td>
</tr>
</tbody>
</table>

Table 8
An illustrative example i.e., student #1A is assessed as “poor” for the entire assessment.

<table>
<thead>
<tr>
<th>Student Being Assessed</th>
<th>Evaluators</th>
<th>#1A</th>
<th>#1B</th>
<th>#1C</th>
<th>#1D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Criteria, i</td>
<td>Weighing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Participated in group meetings</td>
<td>(20%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>2. Communicated constructively to discussion</td>
<td>(20%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>3. Generally was cooperative in group activities</td>
<td>(15%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>4. Contributed to good problem-solving skills</td>
<td>(15%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>5. Contributed useful ideas</td>
<td>(10%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>6. Demonstrated good interest to task given</td>
<td>(10%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>7. Prepared drafts of report in good quality</td>
<td>(10%)</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 9. Peer assessment results for 28 students using three different methods.
the same score, i.e., 83.20%. The IM scores using the conventional methodology are 25.63%, 91.64%, 92.29% and 92.29% for students #1A, #1B, #1C and #1D, respectively. Even though student #1A obtains a low score, it is hard to ascertain the degree of the student being a free-rider. With the proposed methodology, the IM scores for students #1A, #1B, #1C and #1D are 41.60%, 91.06%, 91.06% and 91.59%, respectively. Student #1A is also assessed as “P(100%)”. According to Table 6, student #1A is identified as a ‘sleeping’ team member who relies on others to complete the overall project, and the assessment outcome of “P(100%)” clearly indicates the degree of the student being a free rider in a group activity.

5. Conclusions

In this paper, a new fuzzy peer assessment methodology is proposed. The proposed methodology is the synthesis of Per-C and a fuzzy ranking algorithm that uses fuzzy preference relations. Per-C is able to provide assessment pertaining to students’ contributions in term of recommendations in words. Meanwhile, the fuzzy ranking algorithm is able to provide indices that reflect students’ contributions in a group, and subsequently rank them accordingly. These indices are employed to derive students’ IWF scores. The group marks given by the instructor are then assigned to the students in the respective group based on their associated IWF scores. A real case study on an engineering course (i.e., Multiprocessors Architecture) has been conducted. The results from the conventional and proposed methodologies are compared, analyzed, and discussed in detail. The outcomes positively demonstrate that the proposed methodology is a potential solution to undertake fuzzy peer assessment tasks, which involve vagueness and uncertainty, in cooperative learning environments. In short, the proposed approach has shown to be a useful potential solution to several shortcomings related to the use of numerals in peer assessment, e.g., the meaning of numerals, the difficulty to distinguish the ranking order of students, and the difficulty in interpreting peer assessment outcomes.

In this paper, a small population of students is used as a case study to demonstrate the viability of the proposed approach in peer assessment. As future works, the reliability of the proposed approach will be evaluated via a comprehensive study with larger size of population. Besides, consensus in group decision making [35,36] will be incorporated into the proposed methodology. The benefits include extending the proposed methodology from summative assessment to a hybrid framework of summative and formative assessments. Besides that, the proposed methodology will be deployed in other application domains such as risk assessment [37] and energy research [38].

Acknowledgements

The financial support of the FRGS grant (i.e., FRGS/1/2013/ICT02/UNIMAS/02/1) and RACE grant (i.e., RACE/F2/TK/UNIMAS/5) is gratefully acknowledged.

Appendix A.

See Table A1.
Table A1

Peer Assessment Form (Part 1: Details of Linguistic grades to each specific criterion).

<table>
<thead>
<tr>
<th>Specific Criteria</th>
<th>Linguistic grades and their respective interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td>Below Average</td>
</tr>
<tr>
<td>1. Participated in group meetings</td>
<td>Does not provide any ideas when participating in the group and in classroom discussion. Refuses to participate.</td>
</tr>
<tr>
<td>2. Communicated constructively to discussion</td>
<td>Does not listen to other team members.</td>
</tr>
<tr>
<td>3. Generally was cooperative in group activities</td>
<td>Is always publicly critical of the project or the work of other members of the group. Has a negative attitude towards every aspect.</td>
</tr>
<tr>
<td>4. Contributed to good problem-solving skills</td>
<td>Pretends to solve problems; Causes disruption to others work.</td>
</tr>
</tbody>
</table>

Scoring Rubric: The Task

| 5. Contributed useful ideas | Never provide idea. Free rider. | Rarely provide useful idea. Frequently provides irrelevant ideas, comments and questions. | Sometimes took initiative in generating ideas. Sometimes suggest irrelevant ideas, comments and questions. | Usually took initiative in generating ideas. Suggest ideas, comments and questions. | Always took initiative in generating creative ideas. Always suggest concrete ideas, comments and questions. Concise and clear expression of ideas. |
| 6. Demonstrated good interest to task given | Does not complete any assigned tasks and uses others to complete his/her work. | Completed most of the individual tasks but did not assist other group members during the project. | Completed individual task and assisted other group members some times during the project. Maintains focus on the topic and provides adequate if minimal details. Logical organization with few lapses and acceptable transitions. | Completed most of the assigned tasks. Volunteered to assist group members in finishing the tasks. Maintains good focus on the topic and provides sufficient details. Logical organization that displays completeness with few lapses in transitions. | Completed all assigned tasks. Always assisted other group members in finishing off the tasks. Maintains exceptional focus on the topic and provides ample supporting details. Exceptional organization and provides effective transitions. |
| 7. Prepared drafts of report in good quality | Demonstrates no focus on the topic. No organization | Demonstrates insufficient focus on the topic and provides few details. Little evidence of organization with poor transitions. | Maintains focus on the topic and provides adequate if minimal details. Logical organization with few lapses and acceptable transitions. | Maintains good focus on the topic and provides sufficient details. Logical organization that displays completeness with few lapses in transitions. | 

References


