Wang, H.-Y., & Chen, S. M. (2007). Artificial Intelligence Approach to Evaluate Students' Answerscripts Based on the Similarity Measure between Vague Sets. *Educational Technology & Society*, 10 (4), 224-241.

Artificial Intelligence Approach to Evaluate Students' Answerscripts Based on the Similarity Measure between Vague Sets

Hui-Yu Wang
Department of Education, National Chengchi University, Taiwan // 94152514@@nccu.edu.tw

Shyi-Ming Chen
Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan // Tel: +886-2-27376417 // smchen@mail.ntust.edu.tw

ABSTRACT

In this paper, we present two new methods for evaluating students' answerscripts based on the similarity measure between vague sets. The vague marks awarded to the answers in the students' answerscripts are represented by vague sets, where each element u_i in the universe of discourse U belonging to a vague set is represented by a vague value. The grade of membership of u_i in the vague set $\tilde{A}$ is bounded by a subinterval $[t_{\tilde{A}}(u_i), 1 - f_{\tilde{A}}(u_i)]$ of [0, 1]. It indicates that the exact grade of membership $\mu_{\tilde{A}}(u_i)$ of u_i belonging the vague set $\tilde{A}$ is bounded by $t_{\tilde{A}}(u_i) \leq \mu_{\tilde{A}}(u_i) \leq 1 - f_{\tilde{A}}(u_i)$, where $t_{\tilde{A}}(u_i)$ is a lower bound of the grade of membership of u_i derived from the evidence for u_i, $f_{\tilde{A}}(u_i)$ is a lower bound of the negation of u_i derived from the evidence against u_i, $t_{\tilde{A}}(u_i) + f_{\tilde{A}}(u_i) \leq 1$, and $u_i \in U$. An index of optimism λ determined by the evaluator is used to indicate the degree of optimism of the evaluator, where $\lambda \in [0, 1]$. Because the proposed methods use vague sets to evaluate students' answerscripts rather than fuzzy sets, they can evaluate students' answerscripts in a more flexible and more intelligent manner. Especially, they are particularly useful when the assessment involves subjective evaluation. The proposed methods can evaluate students' answerscripts more stable than Biswas's methods (1995).

Keywords

Similarity functions, Students' answerscripts, Vague grade sheets, Vague membership values, Vague sets, Index of optimism

Introduction

In recent years, some methods have been presented for students' evaluation (Biswas, 1995; Chang & Sun, 1993; Chen & Lee, 1999; Cheng & Yang, 1998; Chiang and Lin, 1994; Frair, 1995; Echauz & Vachtsevanos, 1995; Hwang, Lin, & Lin, 2006; Kaburlasos, Marinagi, & Tsoukalas, 2004; Law, 1996; Ma & Zhou, 2000; Liu, 2005; McMartin, Mckenna, & Youssefi, 2000; Nykanen, 2006; Pears, Daniels, Berglund, & Erickson, 2001; Wang & Chen 2006a; Wang & Chen, 2006b; Wang & Chen, 2006c; Wang & Chen, 2006d; Weon & Kim, 2001; Wu, 2003). Chang and Sun (1993) presented a method for fuzzy assessment of learning performance of junior high school students. Chen and Lee (1999) presented two methods for evaluating students' answerscripts using fuzzy sets. Cheng and Yang (1998) presented a method for using fuzzy sets in education grading systems. Chiang and Lin (1994) presented a method for applying the fuzzy set theory to teaching assessment. Frair (1995) presented a method for student peer evaluations using the analytic hierarchy process method. Echauz and Vachtsevanos (1995) presented a fuzzy grading system to translate a set of scores into letter grades. Hwang, Lin and Lin, (2006) presented an approach for test-sheet composition with large-scale item banks. Kaburlasos, Marinagi, and Tsoukalas (2004) presented a software tool, called PARES, for computer-based testing and evaluation used in the Greek higher education system. Law (1996) presented a method for applying fuzzy numbers in education grading systems. Liu (2005) presented a method for using mutual information for adaptive item comparison and student assessment. Ma and Zhou (2000) presented a fuzzy set approach for the assessment of student-centered learning. McMartin, Mckenna and Youssefi (2000) used scenario assignments as assessment tools for undergraduate engineering education. Nykanen (2006) presented inducing fuzzy models for student classification. Pears, Daniels, Berglund, and Erickson (2001) presented a method for student evaluation in an international collaborative project course. Wang and Chen (2006a) presented two methods for students' answerscripts evaluations using fuzzy sets. Wang and Chen (2006b) presented two methods for evaluating students' answerscripts using fuzzy numbers associated with degrees of confidence. Wang and Chen (2006c) presented two methods for students' answerscripts evaluation using vague sets. Weon and Kim (2001) presented a leaning achievement evaluation strategy in student's learning procedure using fuzzy membership functions. Wu (2003) presented a method for applying the fuzzy set theory and the item response theory to evaluate learning performance.

Biswas (1995) pointed out that the chief aim of education institutions is to provide students with the evaluation reports regarding their test/examination as sufficient as possible and with the unavoidable error as small as possible. Therefore, Biswas (1995) presented a fuzzy evaluation method (*fem*) for applying fuzzy sets in students' answerscripts evaluation. He also generalized the fuzzy evaluation method to propose a generalized fuzzy evaluation method (*gfem*) for students' answerscripts evaluation. In (Biswas, 1995), the fuzzy marks awarded to answers in the students' answerscripts are represented by fuzzy sets (Zadeh, 1965). In a fuzzy set, the grade of membership of an element u_i in the universe of discourse U belonging to a fuzzy set is represented by a real value between zero and one, However, Gau and Buehrer (1993) pointed out that this single value between zero and one combines the evidence for $u_i \in U$ and the evidence against $u_i \in U$. They pointed out that it does not indicate the evidence for $u_i \in U$ and the evidence against $u_i \in U$, respectively, and it does not indicate how much there is of each. Gau and Buehrer (1993) also pointed out that the single value between zero and one tells us nothing about its accuracy. Thus, they proposed the theory of vague sets, where each element in the universe of discourse belonging to a vague set is represented by a vague value. Therefore, if we can allow the marks awarded to the questions of the students' answerscripts to be represented by vague sets, then there is room for more flexibility.

In this paper, we present two new methods for evaluating students' answerscripts based on the similarity measure between vague sets. The vague marks awarded to the answers in the students' answerscripts are represented by vague sets, where each element belonging to a vague set is represented by a vague value. An index of optimism λ(Cheng and Yang, 1998) determined by the evaluator is used to indicate the degree of optimism of the evaluator, where $\lambda \in [0, 1]$. If $0 \leq \lambda < 0.5$, then the evaluator is a pessimistic evaluator. If $\lambda = 0.5$, then the evaluator is a normal evaluator. If $0.5 < \lambda \leq 1.0$, then the evaluator is an optimistic evaluator. Because the proposed methods use vague sets to evaluate students' answerscripts rather than fuzzy sets, they can evaluate students' answerscripts in a more flexible and more intelligent manner. Especially, they are particularly useful when the assessment involves subjective evaluation. The proposed methods can evaluate students' answerscripts more stable than Biswas' methods (1995).

In this paper, we present two new methods for students' answerscripts evaluation based on the similarity measure between vague sets. The vague marks awarded to the answers in the students' answerscripts are represented by vague sets. An index of optimism λ(Cheng and Yang, 1998) determined by the evaluator is used to indicate the degree of optimism of the evaluator, where $\lambda \in [0, 1]$. If $0 \leq \lambda < 0.5$, then the evaluator is a pessimistic evaluator. If $\lambda = 0.5$, then the evaluator is a normal evaluator. If $0.5 < \lambda \leq 1.0$, then the evaluator is an optimistic evaluator. The proposed methods can evaluate students' answerscripts in a more flexible and more intelligent manner. Especially, they are particularly useful when the assessment involves subjective evaluation. The proposed methods can evaluate students' answerscripts more stable than Biswas's methods (1995).

Basic Concepts of the Vague Set Theory

Gau and Buehrer (1993) presented the theory of vague sets. Chen (1995a) presented the arithmetic operations between vague sets. In (Chen, 1995b) and (Chen, 1997), Chen presented similarity measures between vague sets. A vague set $\tilde{A}$ in the universe of discourse U is characterized by a truth-membership function $t_{\tilde{A}}$ and a false-membership function $f_{\tilde{A}}$, where $t_{\tilde{A}}: U \rightarrow [0, 1]$, $f_{\tilde{A}}: U \rightarrow [0, 1]$, $t_{\tilde{A}}(u_i)$ is a lower bound of the grade of membership of u_i derived from the evidence for u_i, $f_{\tilde{A}}(u_i)$ is a lower bound of the negation of u_i derived from the evidence against u_i, $t_{\tilde{A}}(u_i) + f_{\tilde{A}}(u_i) \leq 1$, and $u_i \in U$. The grade of membership of u_i in the vague set $\tilde{A}$ is bounded by a subinterval $[t_{\tilde{A}}(u_i), 1 - f_{\tilde{A}}(u_i)]$ of [0, 1]. The vague value $[t_{\tilde{A}}(u_i), 1 - f_{\tilde{A}}(u_i)]$ indicates that the exact grade of membership $\mu_{\tilde{A}}(u_i)$ of u_i is bounded by $t_{\tilde{A}}(u_i) \leq \mu_{\tilde{A}}(u_i) \leq 1 - f_{\tilde{A}}(u_i)$, where $t_{\tilde{A}}(u_i) + f_{\tilde{A}}(u_i) \leq 1$. An example of a vague set $\tilde{A}$ in the universe of discourse U is shown in Fig. 1.

If the universe of discourse U is a finite set, then a vague set $\tilde{A}$ of the universe of discourse U can be represented as follows:

$$\tilde{A} = \sum_{i=1}^{n} \left[t_{\tilde{A}}(u_i), 1 - f_{\tilde{A}}(u_i) \right] / u_i . \quad (1)$$

If the universe of discourse U is an infinite set, then a vague set $\tilde{A}$ of the universe of discourse can be represented as

$$\tilde{A} = \int_U \left[t_{\tilde{A}}(u_i),\ 1 - f_{\tilde{A}}(u_i)\right]/u_i, \quad u_i \in U, \qquad (2)$$

where the symbol $\int$ denotes the union operator.

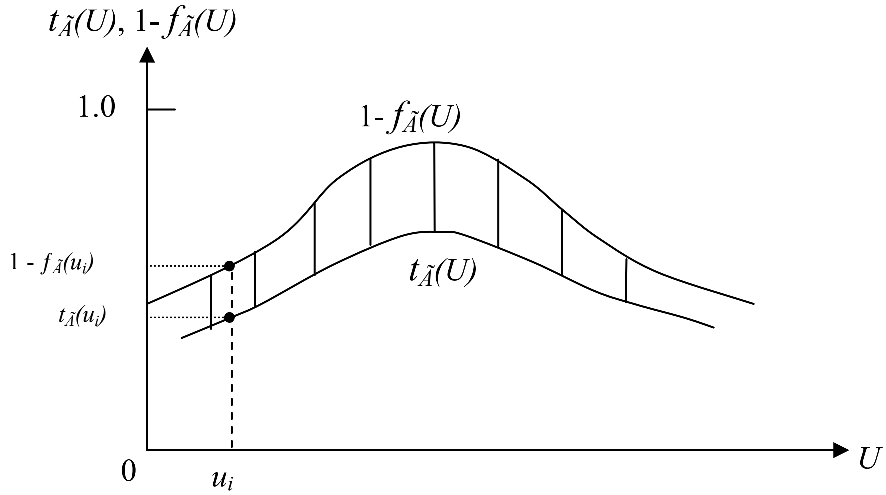

Figure 1. A vague set

Definition 1: Let $\tilde{A}$ be a vague set of the universe of discourse U with the truth-membership function $t_{\tilde{A}}$ and the false-membership function $f_{\tilde{A}}$, respectively. The vague set $\tilde{A}$ is convex if and only if for all u_1, u_2 in U,

$$t_{\tilde{A}}(\lambda\ u_1 + (1 - \lambda\ )u_2) \geq \text{Min}(t_{\tilde{A}}(u_1), t_{\tilde{A}}(u_2)), \qquad (3)$$

$$1 - f_{\tilde{A}}\ (\lambda\ u_1 + (1 - \lambda\ )\ u_2) \geq \text{Min}(1 - f_{\tilde{A}}(u_1), 1 - f_{\tilde{A}}(u_2)), \qquad (4)$$

where $\lambda \in [0, 1]$.

Definition 2: A vague set $\tilde{A}$ of the universe of discourse U is called a normal vague set if $\exists\ u_i \in U$, such that $1 - f_{\tilde{A}}(u_i) = 1$. That is, $f_{\tilde{A}}(u_i) = 0$.

Definition 3: A vague number is a vague subset in the universe of discourse U that is both convex and normal.

Chen (1995b) presented a similarity measure between vague values. Let $X = [t_x, 1 - f_x]$ be a vague value, where $t_x \in [0, 1]$, $f_x \in [0, 1]$ and $t_x + f_x \leq 1$. The score of the vague value X can be evaluated by the score function S shown as follows:

$$S(X) = t_x - f_x, \qquad (5)$$

where $S(X) \in [-1, 1]$. Let X and Y be two vague values, where $X = [t_x, 1 - f_x]$, $Y = [t_y, 1 - f_y]$, $t_x \in [0, 1]$, $f_x \in [0, 1]$, $t_x + f_x \leq 1$, $t_y \in [0, 1]$, $f_y \in [0, 1]$, and $t_y + f_y \leq 1$. The degree of similarity $M(X, Y)$ between the vague values X and Y can be evaluated by the function M,

$$M(X,Y) = 1 - \left|\frac{S(X) - S(Y)}{2}\right|, \qquad (6)$$

where $S(X) = t_x - f_x$ and $S(Y) = t_y - f_y$. The larger the value of $M(X, Y)$, the higher the degree of similarity between the

vague values X and Y. It is obvious that if X and Y are identical vague values (i.e., $X = Y$), then $S(X) = S(Y)$. By applying Eq. (6), we can see that $M(X, Y) = 1$, i.e., the degree of similarity between the vague values X and Y is equal to 1.
Table 1 shows some examples of the degree of similarity $M(X, Y)$ between X and Y.

Table 1. Some examples of the degree of similarity $M(X, Y)$ between the vague values X and Y

X	Y	$M(X, Y)$
[1, 1]	[0, 0]	0
[1, 1]	[1, 0]	$\frac{1}{2}$
[1, 0]	[1, 1]	$\frac{1}{2}$
[0, 1]	[0, 1]	1

Let X and Y be two vague values, where $X = [t_x, 1 - f_x]$, $Y = [t_y, 1 - f_y]$, $t_x \in [0, 1]$, $f_x \in [0, 1]$, $t_x + f_x \leq 1$, $t_y \in [0, 1]$, $f_y \in [0, 1]$, and $t_y + f_y \leq 1$. The proposed similarity measure between vague values has the following properties:

Property 1: Two vague values X and Y are identical if and only if $M(X, Y) = 1$.
Proof:
(i) If X and Y are identical, then $t_x = t_y$ and $1 - f_x = 1 - f_y$ (i.e., $f_x = f_y$). Because $S(X) = t_x - f_x$ and $S(Y) = t_y - f_y = t_x - f_x$, the degree of similarity between the vague values X and Y is calculated as follows:

$$M(X,Y) = 1 - \left|\frac{S(X) - S(Y)}{2}\right|$$

$$= 1 - \left|\frac{(t_x - f_x) - (t_y - f_y)}{2}\right|$$

$$= 1 - \left|\frac{(t_x - f_x) - (t_x - f_x)}{2}\right|$$

$$= 1.$$

(ii) If $M(X, Y) = 1$, then

$$M(X,Y) = 1 - \left|\frac{S(X) - S(Y)}{2}\right|$$

$$= 1 - \left|\frac{(t_x - f_x) - (t_y - f_y)}{2}\right|$$

$$= 1.$$

It implies that $t_x = f_x$ and $t_y = f_y$ (i.e., $1 - t_y = 1 - f_y$). Therefore, the vague values X and Y are identical. **Q. E. D.**

Property 2: $M(X, Y) = M(Y, X)$.
Proof:
Because

$$M(X,Y) = 1 - \left|\frac{S(X) - S(Y)}{2}\right|,$$

$$M(X,Y) = 1 - \left|\frac{S(Y) - S(X)}{2}\right|,$$

and

$$\left|\frac{S(X)-S(Y)}{2}\right| = \left|\frac{S(Y)-S(X)}{2}\right|,$$

we can see that

$$1-\left|\frac{S(X)-S(Y)}{2}\right| = 1-\left|\frac{S(Y)-S(Y)}{2}\right|$$

and $M(X, Y) = M(Y, X)$. **Q. E. D.**

Let $\tilde{A}$ and $\tilde{B}$ be two vague sets in the universe of discourse U, $U = \{u_1, u_2, \ldots, u_n\}$, where

$$\tilde{A} = [t_{\tilde{A}}(u_1), 1-f_{\tilde{A}}(u_1)]/u_1 + [t_{\tilde{A}}(u_2), 1-f_{\tilde{A}}(u_2)]/u_2 + \ldots + [t_{\tilde{A}}(u_n), 1-f_{\tilde{A}}(u_n)]/u_n,$$

and

$$\tilde{B} = [t_{\tilde{B}}(u_1), 1-f_{\tilde{B}}(u_1)]/u_1 + [t_{\tilde{B}}(u_2), 1-f_{\tilde{B}}(u_2)]/u_2 + \ldots + [t_{\tilde{B}}(u_n), 1-f_{\tilde{B}}(u_n)]/u_n.$$

Let $V_{\tilde{A}}(u_i) = [t_{\tilde{A}}(u_i), 1-f_{\tilde{A}}(u_i)]$ be the vague membership value of u_i in the vague set $\tilde{A}$, and let $V_{\tilde{B}}(u_i) = [t_{\tilde{B}}(u_i), 1-f_{\tilde{B}}(u_i)]$ be the vague membership value of u_i in the vague set $\tilde{B}$. By applying Eq. (5), we can see that the score $S(V_{\tilde{A}}(u_i))$ of the vague membership value $V_{\tilde{A}}(u_i)$ can be evaluated as follows:

$$S(V_{\tilde{A}}(u_i)) = t_{\tilde{A}}(u_i) - f_{\tilde{A}}(u_i),$$

and the score $S(V_{\tilde{B}}(u_i))$ of the vague membership value $V_{\tilde{B}}(u_i)$ can be evaluated as follows:

$$S(V_{\tilde{B}}(u_i)) = t_{\tilde{B}}(u_i) - f_{\tilde{B}}(u_i),$$

where $1 \le i \le n$. Then, the degree of similarity $H(\tilde{A}, \tilde{B})$ between the vague sets $\tilde{A}$ and $\tilde{B}$ can be evaluated by the function H,

$$H(\tilde{A}, \tilde{B}) = \frac{1}{n}\sum_{i=1}^{n} M(V_{\tilde{A}}(u_i), V_{\tilde{B}}(u_i))$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(1-\left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{B}}(u_i))}{2}\right|\right), \quad (7)$$

where $H(\tilde{A}, B) \in [0, 1]$. The larger the value of $H(\tilde{A}, \tilde{B})$, the higher the similarity between the vague sets $\tilde{A}$ and $\tilde{B}$.

Let $\tilde{A}$ and $\tilde{B}$ be two vague sets in the universe of discourse U, $U = \{u_1, u_2, \ldots, u_n\}$, where

$$\tilde{A} = [t_{\tilde{A}}(u_1), 1-f_{\tilde{A}}(u_1)]/u_1 + [t_{\tilde{A}}(u_2), 1-f_{\tilde{A}}(u_2)]/u_2 + \ldots + [t_{\tilde{A}}(u_n), 1-f_{\tilde{A}}(u_n)]/u_n,$$

and

$$\tilde{B} = [t_{\tilde{B}}(u_1), 1-f_{\tilde{B}}(u_1)]/u_1 + [t_{\tilde{B}}(u_2), 1-f_{\tilde{B}}(u_2)]/u_2 + \ldots + [t_{\tilde{B}}(u_n), 1-f_{\tilde{B}}(u_n)]/u_n.$$

The proposed similarity measure between vague sets has the following properties:

Property 3: Two vague sets $\tilde{A}$ and $\tilde{B}$ are identical if and only if $H(\tilde{A}, \tilde{B}) = 1$.

Proof:

(i) If $\tilde{A}$ and $\tilde{B}$ are identical, then

$$[t_{\tilde{A}}(u_i), 1 - f_{\tilde{A}}(u_i)] = [t_{\tilde{B}}(u_i), 1 - f_{\tilde{B}}(u_i)],$$ where $1 \le i \le n$.

That is, $t_{\tilde{A}}(u_i) = t_{\tilde{B}}(u_i)$, $f_{\tilde{A}}(u_i) = f_{\tilde{B}}(u_i)$, and $1 \le i \le n$. Because

$$S(V_{\tilde{A}}(u_i)) = t_{\tilde{A}}(u_i) - f_{\tilde{A}}(u_i)$$

and

$$S(V_{\tilde{B}}(u_i)) = t_{\tilde{B}}(u_i) - f_{\tilde{B}}(u_i) = t_{\tilde{A}}(u_i) - f_{\tilde{A}}(u_i) = S(V_{\tilde{A}}(u_i)).$$

Therefore, we can see that

$$H(\tilde{A}, \tilde{B}) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{B}}(u_i))}{2}\right|\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{A}}(u_i))}{2}\right|\right)$$

$$= 1.$$

(ii) If $H(\tilde{A}, \tilde{B}) = 1$, then

$$H(\tilde{A}, \tilde{B}) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{B}}(u_i))}{2}\right|\right) = 1.$$

It implies that $S(V_{\tilde{A}}(u_i)) = S(V_{\tilde{B}}(u_i))$, where $1 \le i \le n$. Because $S(V_{\tilde{A}}(u_i)) = S(V_{\tilde{B}}(u_i))$ and $S(V_{\tilde{B}}(u_i)) = t_{\tilde{B}}(u_i) - f_{\tilde{B}}(u_i)$, where $1 \le i \le n$, we can see that

$$t_{\tilde{A}}(u_i) = t_{\tilde{B}}(u_i) \text{ and } f_{\tilde{A}}(u_i) = f_{\tilde{B}}(u_i) \text{ (i.e., } 1 - f_{\tilde{A}}(u_i) = 1 - f_{\tilde{B}}(u_i)\text{)},$$

where $1 \le i \le n$. Therefore, the vague sets $\tilde{A}$ and $\tilde{B}$ are identical. **Q. E. D.**

Property 4: $H(\tilde{A}, \tilde{B}) = H(\tilde{B}, \tilde{A})$.
Proof:
Because

$$H(\tilde{A}, \tilde{B}) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{B}}(u_i))}{2}\right|\right)$$

and

$$H(\tilde{B}, \tilde{A}) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{B}}(u_i) - S(V_{\tilde{A}}(u_i))}{2}\right|\right),$$

and because

$$\frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\tilde{B}}(u_i))}{2}\right|\right) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - \left|\frac{S(V_{\tilde{B}}(u_i) - S(V_{\tilde{A}}(u_i))}{2}\right|\right),$$

we can see that $H(\tilde{A}, \tilde{B}) = H(\tilde{B}, \tilde{A})$. **Q. E. D.**

Example 1: Let $\tilde{A}$ and $\widetilde{B}$ be two vague sets of the universe of discourse U,

$U = \{u_1, u_2, u_3, u_4, u_5\}$,
$\tilde{A} = [0.2, 0.4]/u_1 + [0.3, 0.5]/u_2 + [0.5, 0.7]/u_3 + [0.7, 0.9]/u_4 + [0.8, 1]/u_5$
$\widetilde{B} = [0.3, 0.5]/u_1 + [0.4, 0.6]/u_2 + [0.6, 0.8]/u_3 + [0.7, 0.9]/u_4 + [0.8, 1]/u_5$,

where

$V_{\tilde{A}}(u_1) = [0.2, 0.4]$, $V_{\widetilde{B}}(u_1) = [0.3, 0.5]$,

$V_{\tilde{A}}(u_2) = [0.3, 0.5]$, $V_{\widetilde{B}}(u_2) = [0.4, 0.6]$,

$V_{\tilde{A}}(u_3) = [0.5, 0.7]$, $V_{\widetilde{B}}(u_3) = [0.6, 0.8]$,

$V_{\tilde{A}}(u_4) = [0.7, 0.9]$, $V_{\widetilde{B}}(u_4) = [0.7, 0.9]$,

$V_{\tilde{A}}(u_5) = [0.8, 1]$, $V_{\widetilde{B}}(u_5) = [0.8, 1]$.

By applying Eq. (5), we can get

$S(V_{\tilde{A}}(u_1)) = 0.2 - 0.6 = -0.4$,

$S(V_{\tilde{A}}(u_2)) = 0.3 - 0.5 = -0.2$,

$S(V_{\tilde{A}}(u_3)) = 0.5 - 0.3 = 0.2$,

$S(V_{\tilde{A}}(u_4)) = 0.7 - 0.1 = 0.6$,

$S(V_{\tilde{A}}(u_5)) = 0.8 - 0 = 0.8$,

$S(V_{\widetilde{B}}(u_1)) = 0.3 - 0.5 = -0.2$,

$S(V_{\widetilde{B}}(u_2)) = 0.4 - 0.4 = 0$,

$S(V_{\widetilde{B}}(u_3)) = 0.6 - 0.2 = 0.4$,

$S(V_{\widetilde{B}}(u_4)) = 0.7 - 0.1 = 0.6$,

$S(V_{\widetilde{B}}(u_5)) = 0.8 - 0 = 0.8$.

By applying Eq. (7), the degree of similarity $H(\tilde{A}, \widetilde{B})$ between the vague sets $\tilde{A}$ and $\widetilde{B}$ can be evaluated, shown as follows:

$$
\begin{aligned}
H(\tilde{A}, \widetilde{B}) &= \frac{1}{5}\sum_{i=1}^{5}\left(1-\left|\frac{S(V_{\tilde{A}}(u_i)) - S(V_{\widetilde{B}}(u_i))}{2}\right|\right) \\
&= \frac{1}{5}\left[\left((1-\left|\frac{-0.4-(-0.2)}{2}\right|\right)+\left(1-\left|\frac{-0.2-0}{2}\right|\right)+\left(1-\left|\frac{0.2-0.4}{2}\right|\right)+\right. \\
&\quad \left.\left(1-\left|\frac{0.6-0.6}{2}\right|\right)+\left(1-\left|\frac{0.8-0.8}{2}\right|\right)\right] \\
&= \frac{1}{5}(0.9+0.9+0.9+1+1) \\
&= 0.94.
\end{aligned}
$$

It indicates that the degree of similarity between the vague sets $\tilde{A}$ and $\widetilde{B}$ is equal to 0.94.

A Review of Biswas' Methods for Students' Answerscripts Evaluation

Biswas (1995) used the matching function S to measure the degree of similarity between two fuzzy sets (Zadeh, 1965). Let $\overline{A}$ and $\overline{B}$ be the vector representation of the fuzzy sets A and B, respectively. Then, the degree of similarity $S(\overline{A}, \overline{B})$ between the fuzzy sets A and B can be calculated as follows (Chen, 1988):

$$S(\overline{A}, \overline{B}) = \frac{\overline{A} \cdot \overline{B}}{Max(\overline{A} \cdot \overline{A},\ \overline{B} \cdot \overline{B})}, \tag{8}$$

where $S(\overline{A}, \overline{B}) \in [0, 1]$. The larger the value of $S(\overline{A}, \overline{B})$, the higher the similarity between the fuzzy sets A and B.
Biswas (1995) presented a "fuzzy evaluation method" (*fem*) for evaluating students' answerscripts, based on the matching function S. He used five fuzzy linguistic hedges, called Standard Fuzzy Sets (SFS), for students' answerscripts evaluation, i.e., E (excellent), V (very good), G (good), S (satisfactory) and U (unsatisfactory), where

X = {0%, 20%, 40%, 60%, 80%, 100%},
E = {(0%, 0), (20%, 0), (40%, 0.8), (60%, 0.9), (80%, 1), (100%, 1)},
V = {(0%, 0), (20%, 0), (40%, 0.8), (60%, 0.9), (80%, 0.9), (100%, 0.8)},
G = {(0%, 0), (20%, 0.1), (40%, 0.8), (60%, 0.9), (80%, 0.4), (100%, 0.2)},
S = {(0%, 0.4), (20%, 0.4), (40%, 0.9), (60%, 0.6), (80%, 0.2), (100%, 0)},
U = {(0%, 1), (20%, 1), (40%, 0.4), (60%, 0.2), (80%, 0), (100%, 0)}.

He used the vector representation method to represent the fuzzy sets E, V, G, S and U by the vectors $\overline{E}$, $\overline{V}$, $\overline{G}$, $\overline{S}$ and $\overline{U}$, respectively, where

$\overline{E}$ = <0, 0, 0.8, 0.9, 1, 1>,
$\overline{V}$ = <0, 0, 0.8, 0.9, 0.9, 0.8>,
$\overline{G}$ = <0, 0.1, 0.8, 0.9, 0.4, 0.2>,
$\overline{S}$ = <0.4, 0.4, 0.9, 0.6, 0.2, 0>,
$\overline{U}$ = <1, 1, 0.4, 0.2, 0, 0>.

Biswas pointed out that "***A***", "***B***", "***C***", "***D***" and "***E***" are letter grades, where $0 \leq \boldsymbol{E} < 30$, $30 \leq \boldsymbol{D} < 50$, $50 \leq \boldsymbol{C} < 70$, $70 \leq \boldsymbol{B} < 90$ and $90 \leq \boldsymbol{A} \leq 100$. Furthermore, he presented the concept of "mid-grade-points", where the mid-grade-points of the letter grades ***A***, ***B***, ***C***, ***D*** and ***E*** are *P(A)*, *P(**B**)*, *P(**C**)*, *P(**D**)* and *P(**E**)*, respectively, *P(A)* **=** 95, *P(**B**)* = 80, *P*(C) = 60, *P(**D**)* = 40 and *P(E)* = 15. Assume that an evaluator evaluates the first question (i.e., *Q.1*) of the answerscript of a student using a fuzzy grade sheet as shown in Table 2.

Table 2. A fuzzy grade sheet (Biswas, 1995)

Question No.	Fuzzy mark						Grade
Q.1	0.1	0.2	0.3	0.6	0.8	0.9	
Q.2							
Q.3							
:	:	:	:	:	:	:	:
					Total mark =		

In the second row of Table 2, the fuzzy marks 0.1, 0.2, 0.3, 0.6, 0.8 and 0.9, awarded to the answer of question *Q.1*, indicate that the degrees of the evaluator's satisfaction for that answer are 0%, 20%, 40%, 60%, 80% and 100%, respectively.
In the following, we briefly review Biswas' method (1995) for students' answerscript evaluation as follows:
Step 1: For each question in the answerscript repeatedly perform the following tasks:
(1) The evaluator awards a fuzzy mark F_i to each question $Q.i$ and fills up each cell of the ith row for the first seven

columns shown in Table 2, where $1 \le i \le$ n. Let $\overline{F_i}$ be the vector representation of F_i, where $1 \le i \le n$.

(2) Based on Eq. (8), calculate the degrees of similarity $S(\overline{E}, \overline{F_i})$, $S(\overline{V}, \overline{F_i})$, $S(\overline{G}, \overline{F_i})$, $S(\overline{S}, \overline{F_i})$ and $S(\overline{U}, \overline{F_i})$, respectively, where $\overline{E}$, $\overline{V}$, $\overline{G}$, $\overline{S}$ and $\overline{U}$ are the vector representations of the standard fuzzy sets E (excellent), V (very good), G (good), S (satisfactory) and U (unsatisfactory), respectively, and $1 \le i \le n$.

(3) Find the maximum value among the values of $S(\overline{E}, \overline{F_i})$, $S(\overline{V}, \overline{F_i})$, $S(\overline{G}, \overline{F_i})$, $S(\overline{S}, \overline{F_i})$ and $S(\overline{U}, \overline{F_i})$. Assume that $S(\overline{V}, \overline{F_i})$ is the maximum value among the values of $S(\overline{E}, \overline{F_i})$, $S(\overline{V}, \overline{F_i})$, $S(\overline{G}, \overline{F_i})$, $S(\overline{S}, \overline{F_i})$ and $S(\overline{U}, \overline{F_i})$, then award the letter grade "***B***" to the question *Q.i* due to the fact that the letter grade "***B***" corresponds to the standard fuzzy set V (very good). If $S(\overline{E}, \overline{F_i}) = S(\overline{V}, \overline{F_i})$ is the maximum value among the values of $S(\overline{E}, \overline{F_i})$, $S(\overline{V}, \overline{F_i})$, $S(\overline{G}, \overline{F_i})$, $S(\overline{S}, \overline{F_i})$ and $S(\overline{U}, \overline{F_i})$, then award the letter grade "***A***" to the question *Q.i* due to the fact that the letter grade "***A***" corresponds to the standard fuzzy set E (excellent).

Step 2: Calculate the total mark of the student as follows:

$$\textit{Total Mark} = \frac{1}{100} \times \sum_{i=1}^{n} [T(Q.i) \times P(g_i)], \qquad (9)$$

where $T(Q.i)$ denotes the mark allotted to *Q.i* in the question paper, g_i denotes the grade awarded to *Q.i* by Step 1 of the algorithm, $P(g_i)$ denotes the mid-grade-point of g_i, and $1 \le i \le n$. Put this total score in the appropriate box at the bottom of the fuzzy grade sheet.

Biswas (1995) also presented a generalized fuzzy evaluation method (*gfem*) for students' answerscripts evaluation, where a generalized fuzzy grade sheet shown in Table 3 is used to evaluate the students' answerscripts.

Table 3. A generalized fuzzy grade sheet (Biswas, 1995)

Question No.	Fuzzy mark	Derived letter grade	Mark
	F_{11}	g_{11}	
	F_{12}	g_{12}	
	F_{13}	g_{13}	
Q.1	F_{14}	g_{14}	m_1
	F_{21}	g_{21}	
	F_{22}	g_{22}	
	F_{23}	g_{23}	
Q.2	F_{24}	g_{24}	m_2
…	…	…	…
…	…	…	…
…	…	…	…
		Total mark =	

In the generalized fuzzy grade sheet shown in Table 3, for all j = 1, 2, 3, 4 and for all i, g_{ij} denotes the derived letter grade by the fuzzy evaluation method *fem* for the awarded fuzzy mark F_{ij} and m_i denotes the derived mark awarded to the question *Q.i*, where

$$m_i = \frac{1}{400} \times T(\text{Q}.i) \times \sum_{j=1}^{4} P(g_{ij}), \qquad (10)$$

and the *Total Mark* $= \sum_{i=1}^{n} m_i$.

A New Method for Evaluating Students' Answerscripts Based on the Similarity Measure between Vague Sets

In this section, we present a new method for evaluating students' answerscripts based on the similarity measure between vague sets. Let X be the universe of discourse. We use five fuzzy linguistic hedges, called Standard Vague Sets (*SVS*), for students' answerscripts evaluation, i.e., $\widetilde{E}$ (excellent), $\widetilde{V}$ (very good), $\widetilde{G}$ (good), $\widetilde{S}$ (satisfactory) and $\widetilde{U}$ (unsatisfactory), where

X = {0%, 20%, 40%, 60%, 80%, 100%},

$\widetilde{E}$ = [0, 0]/0% + [0, 0]/20% + [0, 0]/40% + [0.4, 0.5]/60% + [0.8, 0.9]/80% + [1, 1]/100%,

$\widetilde{V}$ = [0, 0]/0% + [0, 0]/20% + [0, 0]/40% + [0.4, 0.5]/60% + [1, 1]/80% + [0.7, 0.8]/100%,

$\widetilde{G}$ = [0, 0]/0% + [0, 0]/20% + [0.4, 0.5]/40% + [1, 1]/60% + [0.8, 0.9]/80% + [0.4, 0.5]/100%,

$\widetilde{S}$ = [0, 0]/0% + [0.4, 0.5]/20% + [1, 1]/40% + [0.8, 0.9]/60% + [0.4, 0.5]/80% + [0, 0]/100%,

$\widetilde{U}$ = [1, 1]/0% + [1, 1]/20% + [0.4, 0.5]/40% + [0.2, 0.3]/60% + [0, 0]/80% + [0, 0]/100%.

Assume that "***A***", "***B***", "***C***", "***D***" and "***E***" are letter grades, where $0 \le \boldsymbol{E} < 30$, $30 \le \boldsymbol{D} < 50$, $50 \le \boldsymbol{C} < 70$, $70 \le \boldsymbol{B} < 90$ and $90 \le \boldsymbol{A} \le 100$. Assume that an evaluator evaluates the first question (i.e., *Q.1*) of a student's answerscript, using a vague grade sheet as shown in Table 4.

Table 4. A vague grade sheet

Question No.	Vague mark						Derived letter grade
Q.1	[0, 0]	[0.1, 0.2]	[0.3, 0.4]	[0.6, 0.7]	[0.7, 0.8]	[1, 1]	
Q.2							
Q.3							
:	:	:	:	:	:	:	:
Q.n							
					Total mark =		

In the second row of the vague grade sheet shown in Table 4, the vague marks [0, 0], [0.1, 0.2], [0.3, 0.4], [0.6, 0.7], [0.7, 0.8] and [1, 1], awarded to the answer of question *Q.1*, indicate that the degrees of the evaluator's satisfaction for that answer are 0%, 20%, 40%, 60%, 80% and 100%, respectively. Let the vague mark of the answer of question *Q.1* be denoted by $\widetilde{F}_1$. Then, we can see that $\widetilde{F}_1$ is a vague set of the universe of discourse X, where

X = {0%, 20%, 40%, 60%, 80%, 100%},

$\widetilde{F}_1$ = [0, 0]/0% + [0.1, 0.2]/20% + [0.3, 0.4]/40% + [0.6, 0.7]/60% + [0.7, 0.8]/80% + [1, 1]/100%.

The proposed vague evaluation method (*VEM*) for students' answerscripts evaluation is presented as follows:

Step 1: For each question in the answerscript repeatedly perform the following tasks:

(1) The evaluator awards a vague mark $\widetilde{F}_i$ represented by a vague set to each question *Q.i* by his/her judgment and fills up each cell of the *i*th row for the first seven columns shown in Table 4, where $1 \le i \le n$.

(2) Based on Eq. (7), calculate the degrees of similarity $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$, respectively, where $\widetilde{E}$ (excellent), $\widetilde{V}$ (very good), $\widetilde{G}$ (good), $\widetilde{S}$ (satisfactory) and $\widetilde{U}$ (unsatisfactory) are standard vague sets.

(3) Find the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$. If $H(\widetilde{W}, \widetilde{F}_i)$ is the largest value among the values of $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$, where $\widetilde{W} \in \{\widetilde{E}, \widetilde{V}, \widetilde{G}, \widetilde{S}, \widetilde{U}\}$, then translate the standard vague set $\widetilde{W}$ into the corresponding letter grade, where the standard vague set $\widetilde{E}$ is translated into the letter grade "***A***", the standard vague set $\widetilde{V}$ is translated into the letter grade "***B***", the standard vague set $\widetilde{G}$ is translated into the letter grade "***C***", the standard vague set $\widetilde{S}$ is translated into the letter grade "***D***", and the standard vague set $\widetilde{U}$ is translated into the letter grade "***E***". For example, assume that $H(\widetilde{V}, \widetilde{F}_i)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$, then award grade "***B***" to the question *Q.i* due to the fact that the letter grade "***B***" corresponds to the standard vague set $\widetilde{V}$ (very good). If $H(\widetilde{E}, \widetilde{F}_i) = H(\widetilde{V}, \widetilde{F}_i)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$, then award the letter grade "***A***" to the question *Q.i* due to the fact that the letter grade "***A***" corresponds to the standard vague set $\widetilde{E}$ (excellent).

Step 2: Calculate the total mark of the student as follows:

$$\textit{Total Mark} = \frac{1}{100} \times \sum_{i=1}^{n} [T(Q.i) \times K(g_i) \times H(\widetilde{w}, \widetilde{F}_i)], \tag{11}$$

where $T(Q.i)$ denotes the mark allotted to the question *Q.i* in the question paper, g_i denotes the letter grade awarded to *Q.i* by Step 1, $K(g_i)$ denotes the derived grade-point of the letter grade g_i based on the index of optimism λ determined by the evaluator, where $\lambda \in [0, 1]$, $H(\widetilde{W}, \widetilde{F}_i)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_i)$, $H(\widetilde{V}, \widetilde{F}_i)$, $H(\widetilde{G}, \widetilde{F}_i)$, $H(\widetilde{S}, \widetilde{F}_i)$ and $H(\widetilde{U}, \widetilde{F}_i)$, $\widetilde{W} \in \{\widetilde{E}, \widetilde{V}, \widetilde{G}, \widetilde{S}, \widetilde{U}\}$, such that the derived letter grade awarded to the question *Q.i* is g_i, and $1 \le i \le n$. If $0 \le \lambda < 0.5$, then the evaluator is a pessimistic evaluator. If $\lambda = 0.5$, then the evaluator is a normal evaluator. If $0.5 < \lambda \le 1.0$, then the evaluator is an optimistic evaluator. Assume that the derived letter grade obtained in Step 1 with respect to the question *Q.i* is g_i, where $g_i \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}\}$ and $0 \le y_1 \le g_i \le y_2 \le 100$, then the derived grade-point $K(g_i)$ shown in Eq. (8) is calculated as follows:

$$K(g_i) = (1 - \lambda) \times y_1 + \lambda \times y_2, \tag{12}$$

where λ is the index of optimism determined by the evaluator, $\lambda \in [0, 1]$, and $0 \le y_1 \le K(g_i) \le y_2 \le 100$. Put the derived total mark in the appropriate box at the bottom of the vague grade sheet.

Example 2: Consider a student's answerscript to an examination of 100 marks. Assume that in total there are four questions to be answered:

TOTAL MARKS = 100,
Q.1 carries 30 marks,
Q.2 carries 30 marks,
Q.3 carries 20 marks,
Q.4 carries 20 marks.

Assume that an evaluator awards the student's answerscript using the vague grade sheet shown in Table 5, where the index of optimism λ determined by the evaluator is 0.60, i.e., $\lambda = 0.60$. Assume that "***A***", "***B***", "***C***", "***D***" and "***E***" are letter grades, where $0 \le \boldsymbol{E} < 30$, $30 \le \boldsymbol{D} < 50$, $50 \le \boldsymbol{C} < 70$, $70 \le \boldsymbol{B} < 90$ and $90 \le \boldsymbol{A} \le 100$.

Table 5. Vague grade sheet of Example 2

Question No.	Vague mark						Derived letter grade
Q.1	[0, 0]	[0, 0]	[0, 0]	[0.4, 0.5]	[1, 1]	[0.5, 0.6]	
Q.2	[0, 0]	[0, 0]	[0, 0]	[0.4, 0.5]	[0.8, 0.9]	[1, 1]	
Q.3	[0, 0]	[0.4, 0.5]	[1, 1]	[0.6, 0.7]	[0.4, 0.5]	[0, 0]	
Q.4	[0.8, 0.9]	[0.5, 0.6]	[0.2, 0.3]	[0, 0]	[0, 0]	[0, 0]	
					Total mark =		

From Table 5, we can see that the vague marks of the questions *Q.1*, *Q.2*, *Q.3* and *Q.4* represented by vague sets are $\widetilde{F}_1$, $\widetilde{F}_2$, $\widetilde{F}_3$ and $\widetilde{F}_4$, respectively, where

$\widetilde{F}_1$ = [0, 0]/0% + [0, 0]/20% + [0, 0]/40% + [0.4, 0.5]/60% + [1, 1]/80% + [0.5, 0.6]/100%,

$\widetilde{F}_2$ = [0, 0]/0% + [0, 0]/20% + [0, 0]/40% + [0.4, 0.5]/60% + [0.8, 0.9]/80% + [1, 1]/100%,

$\widetilde{F}_3$ = [0, 0]/0% + [0.4, 0.5]/20% + [1, 1]/40% + [0.6, 0.7]/60% + [0.4, 0.5]/80% + [0, 0]/100%,

$\widetilde{F}_4$ = [0.8, 0.9]/0% + [0.5, 0.6]/20% + [0.2, 0.3]/40% + [0, 0]/60% + [0, 0]/80% + [0, 0]/100%.

[Step 1] According to the standard vague sets $\widetilde{E}$, $\widetilde{V}$, $\widetilde{G}$, $\widetilde{S}$, $\widetilde{U}$ and the vague marks $\widetilde{F}_1, \widetilde{F}_2, \widetilde{F}_3, \widetilde{F}_4$, we can get the vague values, as shown in Table 6.

Table 6. Vague values of Example 2

t	$V_{\widetilde{E}}(t)$	$V_{\widetilde{V}}(t)$	$V_{\widetilde{G}}(t)$	$V_{\widetilde{S}}(t)$	$V_{\widetilde{U}}(t)$	$V_{\widetilde{F}_1}(t)$	$V_{\widetilde{F}_2}(t)$	$V_{\widetilde{F}_3}(t)$	$V_{\widetilde{F}_4}(t)$
0 %	[0, 0]	[0, 0]	[0, 0]	[0, 0]	[1, 1]	[0, 0]	[0, 0]	[0, 0]	[0.8, 0.9]
20 %	[0, 0]	[0, 0]	[0, 0]	[0.4, 0.5]	[1, 1]	[0, 0]	[0, 0]	[0.4, 0.5]	[0.5, 0.6]
40 %	[0, 0]	[0, 0]	[0.4, 0.5]	[1, 1]	[0.4, 0.5]	[0, 0]	[0, 0]	[1, 1]	[0.2, 0.3]
60 %	[0.4, 0.5]	[0.4, 0.5]	[1, 1]	[0.8, 0.9]	[0.2, 0.3]	[0.4, 0.5]	[0.4, 0.5]	[0.6, 0.7]	[0, 0]
80 %	[0.8, 0.9]	[1, 1]	[0.8, 0.9]	[0.4, 0.5]	[0, 0]	[1, 1]	[0.8, 0.9]	[0.4, 0.5]	[0, 0]
100 %	[1, 1]	[0.7, 0.8]	[0.4, 0.5]	[0, 0]	[0, 0]	[0.5, 0.6]	[1, 1]	[0, 0]	[0, 0]

By applying Eq. (5), we can get scores of the vague values, as shown in Table 7.

Table 7. Scores of the vague values of Example 2

t	$S(V_{\widetilde{E}}(t))$	$S(V_{\widetilde{V}}(t))$	$S(V_{\widetilde{G}}(t))$	$S(V_{\widetilde{S}}(t))$	$S(V_{\widetilde{U}}(t))$	$S(V_{\widetilde{F}_1}(t))$	$S(V_{\widetilde{F}_2}(t))$	$S(V_{\widetilde{F}_3}(t))$	$S(V_{\widetilde{F}_4}(t))$
0 %	-1	-1	-1	-1	1	-1	-1	-1	0.7
20 %	-1	-1	-1	-0.1	1	-1	-1	-0.1	0.1
40 %	-1	-1	-0.1	1	-0.1	-1	-1	1	-0.5
60 %	-0.1	-0.1	1	0.7	-0.5	-0.1	-0.1	0.3	-1
80 %	0.7	1	0.7	-0.1	-1	1	0.7	-0.1	-1
100 %	1	0.5	-0.1	-1	-1	0.1	1	-1	-1

Table 8. The degrees of similarity between the vague sets

H(X, Y) *Y* / *X*	$\widetilde{F}_1$	$\widetilde{F}_2$	$\widetilde{F}_3$	$\widetilde{F}_4$
$\widetilde{E}$	0.900	1.000	0.492	0.342
$\widetilde{V}$	0.967	0.942	0.508	0.358
$\widetilde{G}$	0.792	0.742	0.633	0.350
$\widetilde{S}$	0.508	0.458	0.967	0.425
$\widetilde{U}$	0.300	0.250	0.508	0.825

By applying Eq. (7), we can get the degree of similarity $H(X, Y)$ between the vague values X and Y, where $X \in \{\widetilde{E}, \widetilde{V}, \widetilde{G}, \widetilde{S}, \widetilde{U}\}$ and $Y \in \{\widetilde{F}_1, \widetilde{F}_2, \widetilde{F}_3, \widetilde{F}_3\}$, as shown in Table 8.

Because $H(\widetilde{V}, \widetilde{F}_1)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_1)$, $H(\widetilde{V}, \widetilde{F}_1)$, $H(\widetilde{G}, \widetilde{F}_1)$, $H(\widetilde{S}, \widetilde{F}_1)$ and $H(\widetilde{U}, \widetilde{F}_1)$, we award grade "***B***" to the question *Q.1* due to the fact that the letter grade "***B***" corresponds to the standard vague set $\widetilde{V}$ (very good).

Because $H(\widetilde{E}, \widetilde{F}_2)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_2)$, $H(\widetilde{V}, \widetilde{F}_2)$, $H(\widetilde{G}, \widetilde{F}_2)$, $H(\widetilde{S}, \widetilde{F}_2)$ and $H(\widetilde{U}, \widetilde{F}_2)$, we award grade "***A***" to the question *Q.2* due to the fact that the letter grade "***A***" corresponds to the standard vague set $\widetilde{E}$ (excellent).

Because $H(\widetilde{S}, \widetilde{F}_3)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_3)$, $H(\widetilde{V}, \widetilde{F}_3)$, $H(\widetilde{G}, \widetilde{F}_3)$, $H(\widetilde{S}, \widetilde{F}_3)$ and $H(\widetilde{U}, \widetilde{F}_3)$, we award grade "***D***" to the question *Q.3* due to the fact that the letter grade "***D***" corresponds to the standard vague set $\widetilde{S}$ (satisfactory).

Because $H(\widetilde{U}, \widetilde{F}_4)$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_4)$, $H(\widetilde{V}, \widetilde{F}_4)$, $H(\widetilde{G}, \widetilde{F}_4)$, $H(\widetilde{S}, \widetilde{F}_4)$ and $H(\widetilde{U}, \widetilde{F}_4)$, we award grade "***E***" to the question *Q.4* due to the fact that the letter grade "***E***" corresponds to the standard vague set $\widetilde{U}$ (unsatisfactory).

[Step 2] Because $90 \le \boldsymbol{A} \le 100$, $70 \le \boldsymbol{B} < 90$, $30 \le \boldsymbol{D} < 50$ and $0 \le \boldsymbol{E} < 30$, where "***A***", "***B***", "***D***" and "***E***" are letter grades, and the index of optimism λ determined by the evaluator is 0.60 (i.e., $\lambda = 0.60$), based on Eq. (12), we can get the following results:

$K(\boldsymbol{A}) = (1 - 0.60) \times 90 + 0.60 \times 100 = 96,$
$K(\boldsymbol{B}) = (1 - 0.60) \times 70 + 0.60 \times 90 = 82,$
$K(\boldsymbol{D}) = (1 - 0.60) \times 30 + 0.60 \times 50 = 42,$
$K(\boldsymbol{E}) = (1 - 0.60) \times 0 + 0.60 \times 30 = 18.$

Because the questions *Q.1*, *Q.2*, *Q.3* and *Q.4* carry 30 marks, 30 marks, 20 marks and 20 marks, respectively, and because $H(\widetilde{V}, \widetilde{F}_1) = 0.967$, $H(\widetilde{E}, \widetilde{F}_2) = 1.000$, $H(\widetilde{S}, \widetilde{F}_3) = 0.967$ and $H(\widetilde{U}, \widetilde{F}_4) = 0.825$, based on Eq. (11), the total mark of the student is evaluated as follows:

$$\frac{1}{100}(30 \times 82 \times 0.967 + 30 \times 96 \times 1.000 + 20 \times 42 \times 0.967 + 20 \times 18 \times 0.825)$$
$$= \frac{1}{100}(2378.82 + 2880 + 812.28 + 297)$$
$$= 63.681$$
$= 64$ (assuming that no half mark is given in the total mark).

A Generalized Method for Evaluating Students' Answerscripts Based on the Similarity Measure between Vague Sets

In this section, we present a generalized vague evaluation method (*GVEM*) for students' answerscripts evaluation based on the similarity measure between vague sets, where a generalized vague grade sheet shown in Table 9 is used to evaluate the students' answerscripts.

Table 9. A generalized vague grade sheet

Question No.	Sub-questions	Vague mark	Derived letter grade	Mark
	Q.11	$\widetilde{F}_{11}$	g_{11}	
Q.1	*Q.12*	$\widetilde{F}_{12}$	g_{12}	m_1

	Q.13	$\widetilde{F}_{13}$	g_{13}	
	Q.14	$\widetilde{F}_{14}$	g_{14}	
	Q.21	$\widetilde{F}_{21}$	g_{21}	
	Q.22	$\widetilde{F}_{22}$	g_{22}	
	Q.23	$\widetilde{F}_{23}$	g_{23}	
Q.2	*Q.24*	$\widetilde{F}_{24}$	g_{24}	m_2
:	:	:	:	:
	Q.n1	$\widetilde{F}_{n1}$	g_{n1}	
	Q.n2	$\widetilde{F}_{n2}$	g_{n2}	
	Q.n3	$\widetilde{F}_{n3}$	g_{n3}	
Q.n	*Q.n4*	$\widetilde{F}_{n4}$	g_{n4}	m_n
			Total mark =	

In the generalized vague grade sheet shown in Table 9, each question *Q.i* consists of four sub-questions, i.e., *Q.i1*, *Q.i2*, *Q.i3* and Q.*i4*. For all j = 1, 2, 3, 4 and for all i, g_{ij} is the derived letter grade by the proposed vague evaluation method *VEM* of the awarded vague mark $\widetilde{F}_{ij}$ with respect to the sub-question *Q.ij*, and m_i is the derived mark awarded to the question *Q.i*,

$$m_i = \frac{1}{400} \times T(Q.i) \times \sum_{j=1}^{4} [K(g_{ij}) \times H(\widetilde{w}, F_{ij})], \tag{13}$$

and

$$Total\ Mark = \sum_{i=1}^{n} m_i .$$

where $T(Q.i)$ denotes the mark allotted to *Q.i* in the question paper, g_{ij} denotes the derived letter grade awarded to *Q.i*, and $K(g_{ij})$ denotes the derived grade-point of the letter grade g_{ij} based on the index of optimism λ determined by the evaluator, where $\lambda \in [0, 1]$, $H(\widetilde{W}, \widetilde{F}_{ij})$ is the maximum value among the values of $H(\widetilde{E}, \widetilde{F}_{ij})$, $H(\widetilde{V}, \widetilde{F}_{ij})$, $H(\widetilde{G}, \widetilde{F}_{ij})$, $H(\widetilde{S}, \widetilde{F}_{ij})$ and $H(\widetilde{U}, \widetilde{F}_{ij})$, $\widetilde{W} \in \{\widetilde{E}, \widetilde{V}, \widetilde{G}, \widetilde{S}, \widetilde{U}\}$, such that the derived letter grade awarded to the question *Q.ij* is g_{ij}, $1 \le j \le 4$, and $1 \le i \le n$. If $0 \le \lambda < 0.5$, then the evaluator is a pessimistic evaluator. If $\lambda = 0.5$, then the evaluator is a normal evaluator. If $0.5 < \lambda \le 1.0$, then the evaluator is an optimistic evaluator. Assume that the derived letter grade with respect to the sub-question *Q.ij* is g_{ij}, where $g_{ij} \in \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}\}$ and $0 \le y_1 \le g_{ij} \le y_2 \le 100$, then the derived grade-point $K(g_{ij})$ shown in Eq. (13) is calculated as follows:

$$K(g_{ij}) = (1 - \lambda) \times y_1 + \lambda \times y_2, \tag{14}$$

where λ is the index of optimism determined by the evaluator, $\lambda \in [0, 1]$, and $0 \le y_1 \le K(g_{ij}) \le y_2 \le 100$. Put the derived total mark in the appropriate box at the bottom of the generalized vague grade sheet.

Experimental Results

We have made an experiment to compare the evaluating results of the proposed method with the Biswas' method (1995) for different days. In our experiment, there are four questions to be answered in a student's answerscript, where

TOTAL MARKS = 100,
Q.1 carries 20 marks,

Q.2 carries 25 marks,
Q.3 carries 25 marks,
Q.4 carries 30 marks.

Assume that the index of optimism λ of the evaluator is 0.60 (i.e., λ = 0.60). The evaluator uses Biswas' method (1995) and the proposed method to evaluate the student's answerscript on different days, respectively. The results are shown in Fig. 2 and Fig. 3, respectively. A comparison of the evaluating results of the student's answerscript is shown in Table 10. From Table 10, we can see that the proposed method is more stable to evaluate students' answerscripts than Biswas' method (1995). It can evaluate students' answerscripts in a more flexible and more intelligent manner.

July 1, 2006

Question No.	Satisfaction Levels						Grade
Q.1	0	0	0	0.6	0.9	0.8	
Q.2	0	0	0.6	0.9	0.8	0	
Q.3	0	0	0	0.6	0.8	0.9	
Q.4	0	0.6	0.9	0.8	0.2	0	
					Total Mark =		

July 2, 2006

Question No.	Satisfaction Levels						Grade
Q.1	0	0	0	0.8	0.9	1	
Q.2	0	0	0.7	0.8	0.9	0	
Q.3	0	0	0	0.7	0.9	0.8	
Q.4	0	0.5	0.8	0.7	0	0	
					Total Mark =		

July 3, 2006

Question No.	Satisfaction Levels						Grade
Q.1	0	0	0	0.6	0.9	0.7	
Q.2	0	0	0.6	0.8	0.7	0	
Q.3	0	0	0	0.5	0.7	0.9	
Q.4	0	0.5	0.8	0.6	0	0	
					Total Mark =		

July 4, 2006

Question No.	Satisfaction Levels						Grade
Q.1	0	0	0	0.6	0.8	0.7	
Q.2	0	0	0.5	0.9	0.7	0	
Q.3	0	0	0	0.7	0.9	0.8	
Q.4	0	0.6	0.9	0.7	0	0	
					Total Mark =		

Figure 2. Evaluating the student's answerscript at different days using Biswas' method (1995)

July 1, 2006

Question No.	Vague marks						Grade
Q.1	[0, 0]	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.8, 0.9]	
Q.2	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.8, 0.9]	[0, 0]	
Q.3	[0, 0]	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.8, 0.9]	
Q.4	[0, 0]	[0.5, 0.6]	[0.8, 0.9]	[0.7, 0.8]	[0.1, 0.2]	[0, 0]	
					Total mark =		

July 2, 2006

Question No.	Vague marks						Grade

Q.1	[0, 0]	[0, 0]	[0, 0]	[0.7, 0.8]	[0.8, 0.9]	[0.9, 1.0]	
Q.2	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.8, 0.9]	[0, 0]	
Q.3	[0, 0]	[0, 0]	[0, 0]	[0.7, 0.8]	[0.8, 0.9]	[0.8, 0.9]	
Q.4	[0, 0]	[0.5, 0.6]	[0.8, 0.9]	[0.7, 0.8]	[0, 0]	[0, 0]	
					Total mark =		

July 3, 2006

Question No.	Vague marks						Grade
Q.1	[0, 0]	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.7, 0.8]	
Q.2	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.7, 0.8]	[0, 0]	
Q.3	[0, 0]	[0, 0]	[0, 0]	[0.5, 0.6]	[0.7, 0.8]	[0.8, 0.9]	
Q.4	[0, 0]	[0.5, 0.6]	[0.8, 0.9]	[0.6, 0.7]	[0, 0]	[0, 0]	
					Total mark =		

July 4, 2006

Question No.	Vague marks						Grade
Q.1	[0, 0]	[0, 0]	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.8, 0.9]	
Q.2	[0, 0]	[0, 0]	[0.5, 0.6]	[0.8, 0.9]	[0.7, 0.8]	[0, 0]	
Q.3	[0, 0]	[0, 0]	[0, 0]	[0.7, 0.8]	[0.8, 0.9]	[0.8, 0.9]	
Q.4	[0, 0]	[0.6, 0.7]	[0.8, 0.9]	[0.7, 0.8]	[0, 0]	[0, 0]	
					Total mark =		

Figure 3. Evaluating the student's answerscript at different days using the proposed method

Table 10. A comparison of the evaluating results for different methods

Methods / Total mark / Days	Biswas' method (1995)	The proposed method
July 1, 2006	69	68
July 2, 2006	72	68
July 3, 2006	55	68
July 4, 2006	55	68

The Merits of the Proposed Methods

The proposed methods have the following advantages:

(1) The proposed methods are more flexible and more intelligent than Biswas' methods (1995) due to the fact that we use vague sets rather than fuzzy sets to represent the vague mark of each question, where the evaluator can use vague values to indicate the degree of the evaluator's satisfaction for each question. Especially, the proposed methods are particularly useful when the assessment involves subjective evaluation.

(2) The proposed methods are more stable to evaluate students' answerscripts than Biswas' methods (1995). They can evaluate students' answerscripts in a more flexible and more intelligent manner.

Conclusions

In this paper, we have presented two new methods for evaluating students' answerscripts based on the similarity measure between vague sets. The vague marks awarded to the answers in the students' answerscripts are represented by vague sets, where each element belonging to a vague set is represented by a vague value. An index of optimism λ determined by the evaluator is used to indicate the degree of optimism of the evaluator, where $\lambda \in [0, 1]$. Because the proposed methods use vague sets to evaluate students' answerscripts rather than fuzzy sets, they can evaluate students' answerscripts in a more flexible and more intelligent manner. The experimental results show that the proposed methods can evaluate students' answerscripts more stable than Biswas' methods (1995).

Acknowledgements

The authors would like to thank Professor Jason Chiyu Chan, Department of Education, National Chengchi University, Taipei, Taiwan, Republic of China, for providing very helpful comments and suggestions. This work was supported in part by the National Science Council, Republic of China, under Grant NSC 95-2221-E-011-117-MY2.

References

Biswas, R. (1995). An application of fuzzy sets in students' evaluation. *Fuzzy Sets and Systems*, *74* (2), 187-194.

Chang, D. F., & Sun, C. M. (1993). Fuzzy assessment of learning performance of junior high school students. *Paper presented at the First National Symposium on Fuzzy Theory and Applications*, June 25-26, 1993, Hsinchu, Taiwan.

Chen, S. M. (1988). A new approach to handling fuzzy decisionmaking problems. *IEEE Transactions on Systems, Man, and Cybernetics*, *18* (6), 1012-1016.

Chen, S. M. (1995a). Arithmetic operations between vague sets. *Paper presented at the International Joint Conference of CFSA/IFIS/SOFT'95 on Fuzzy Theory and Applications*, December 7-9, 1995, Taipei, Taiwan.

Chen, S. M. (1995b). Measures of similarity between vague sets. *Fuzzy Sets and Systems*, *74* (2), 217-223.

Chen, S. M. (1997). Similarity measures between vague sets and between elements. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, *27* (1), 153-158.

Chen, S. M. (1999). Evaluating the rate of aggregative risk in software development using fuzzy set theory. *Cybernetics and Systems*, *30* (1), 57-75.

Chen, S. M., & Lee, C. H. (1999). New methods for students' evaluating using fuzzy sets. *Fuzzy Sets and Systems*, *104* (2), 209-218.

Chen, S. M., & Wang, J. Y. (1995). Document retrieval using knowledge-based fuzzy information retrieval techniques. *IEEE Transactions on Systems, Man, and Cybernetics*, *25* (5), 793-803.

Cheng, C. H., & Yang, K. L. (1998). Using fuzzy sets in education grading system. *Journal of Chinese Fuzzy Systems Association*, *4* (2), 81-89.

Chiang, T .T., & Lin C. M. (1994). Application of fuzzy theory to teaching assessment. *Paper presented at the 1994 Second National Conference on Fuzzy Theory and Applications*, September 15-17, 1994, Taipei, Taiwan.

Frair, L. (1995). Student peer evaluations using the analytic hierarchy process method. *Paper presented at the Frontiers in Education Conference*, November 1-4, 1995, Atlanta, GA, USA.

Gau, W. L., & Buehrer, D. J. (1993). Vague sets. *IEEE Transactions on Systems, Man, and Cybernetics*, *23* (2), 610-614.

Echauz, J. R., & Vachtsevanos, G. J. (1995). Fuzzy grading system. *IEEE Transactions on Education*, *38* (2), 158-165.

Hwang, G. J., Lin, B. M. T., & Lin, T. L. (2006). An effective approach for test-sheet composition with large-scale item banks, *Computers & Education*, *46* (2), 122-139.

Kaburlasos, V. G., Marinagi, C. C., & Tsoukalas, V. T. (2004). PARES: A software tool for computer-based testing and evaluation used in the Greek higher education system. *Paper presented at the 2004 IEEE International Conference on Advanced Learning Technologies*, August 30 – September 1, 2004, Joensuu, Finland.

Law, C. K. (1996). Using fuzzy numbers in education grading system. *Fuzzy Sets and Systems*, *83* (3), 311-323.

Liu, C. L. (2005). Using mutual information for adaptive item comparison and student assessment. *Educational Technology & Society, 8* (4), 100-119.

Ma, J., & Zhou, D. (2000). Fuzzy set approach to the assessment of student-centered learning. *IEEE Transactions on Education, 43* (2), 237-241.

McMartin, F., Mckenna, A., & Youssefi, K. (2000). Scenario assignments as assessment tools for undergraduate engineering education. *IEEE Transactions on Education*, *43* (2), 111-119.

Nykanen, O. (2006). Inducing fuzzy models for student classification. *Educational Technology & Society, 9* (2), 223-234.

Pears, A., Daniels, M., Berglund, A., & Erickson, C. (2001). Student evaluation in an international collaborative project course. *Paper presented at the First International Workshop on Internet-Supported Education*, January 08 - 12, 2001, San Diego, CA, USA.

Wang, H. Y., & Chen, S. M. (2006a). New methods for evaluating the answerscripts of students using fuzzy sets. *Paper presented at the 19th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems*, June 27-30, 2006, Annecy, France.

Wang, H. Y., & Chen, S. M. (2006b). New methods for evaluating students' answerscripts using fuzzy numbers associated with degrees of confidence. *Paper presented at the 2006 IEEE International Conference on Fuzzy Systems*, July 16-21, 2006, Vancouver, BC, Canada.

Wang, H. Y., & Chen, S. M. (2006c). New methods for evaluating students' answerscripts using vague values. *Paper presented at the 9th Joint Conference on Information Sciences*, October 8-11, 2006, Kaohsiung, Taiwan.

Wang, H. Y., & Chen, S. M. (2006d). Evaluating students' answerscripts based on the similarity measure between vague sets. *Paper presented at the 11th Conference on Artificial Intelligence and Applications*, December 15-16, 2006, Kaohsiung, Taiwan.

Weon, S., & Kim, J. (2001). Learning achievement evaluation strategy using fuzzy membership function. *Paper presented at the 31st ASEE/IEEE Frontier in Education Conference*, October 10-13, 2001, Reno, NV, USA.

Wu, M. H. (2003). *A research on applying fuzzy set theory and item response theory to evaluate learning performance*, Master Thesis, Department of Information Management, Chaoyang University of Technology, Taiwan.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338-353.