Article

Temporal support in the identification of e-learning efficacy: an example of object classification in the presence of ignorance

Paul Jones[1] and Malcolm J. Beynon[2]

(1) Department of Enterprise and Economic Development, Business School, University of Glamorgan, UK
E-mail: wpjones1@glam.ac.uk
(2) Cardiff Business School, Cardiff University, UK
E-mail: BeynonMJ@cardiff.ac.uk

Abstract: *This study investigates the efficacy of an in-place e-learning facility, towards the performance of students on a university module, through the utilization of an information system based on CaRBS (Classification and Ranking Belief Simplex). It attempts to discern the final indifferent and good performance classifications of students based on their activity during the module (online page views). The ongoing assessment of the students every two weeks during the module is modelled with separate classification results reported as the number of two-week periods builds up. Students who completed the course as well as those who withdrew are considered. The CaRBS system has its emphasis on the visual representation of results (using simplex plots), as well as its operations in the presence of ignorance (using the Dempster–Shafer theory of evidence). In the analysis of the module considered, the activity over certain weeks is consistently shown to be important to discern student performance. This study is of interest to practitioners and theorists, with the results shown offering a benchmark to the type of findings pertinent to expositing e-learning efficacy.*

Keywords: CaRBS, e-learning, efficacy, performance, assessment

1. Introduction

Electronic learning (e-learning) is an expanding phenomenon, implicit in the education of individuals at universities and colleges (Alexander, 2001), its development constrained by the power and range of the available computing facilities (Gunasekaran *et al.*, 2002). For those associated with e-learning, its success is advantageous for a number of reasons, including cost savings for the provider and employee (Van Dam, 2001; Williams, 2002) and flexibility facilitated to the learner (Frank *et al.*, 2002; Piskurich, 2004). McVay-Lynch (2002) identifies that e-learning respects the differences in learning style and pace, fostering a greater degree of communication among students and e-moderators (Jolliffe *et al.*, 2001). Further advantages include consistent learning material (Voci & Young, 2001), accessible location of module content (McVay-Lynch, 2002) and easily updatable materials (Barabash *et al.*, 2003).

The quantification of the success of an e-learning facility is through the positive relationship between the online activity of the students and their concomitant module success. The module considered (an online undergraduate degree offered by a UK university), like other e-learning initiatives, has suffered from a significant withdrawal rate (Rovai, 2003; Jones *et al.*, 2004), with a key strategy to increase retention being improved student monitoring. This study considers the configuration of a concomitant information system (IS), including its ability to elucidate the intermediate progress of students with respect to their online activity, which may advocate supportive action by the university, possibly mitigating the decision to withdraw by certain students.

The module success of students here is through their classification to good or indifferent performance (discerning those achieving a module mark of at least a UK graded 2–1 class or not, or non-UK equivalent). The e-learning module considered has ongoing assessment every two weeks during the 12-week duration of the module. Hence, six individual analyses are undertaken with configured ISs, which elucidate the performances of the students based on their levels of online activity over the two, four, six, eight, ten and 12 weeks of the module. The analysis of students who withdrew from the module is also reported, demonstrating the future utilization of the presented ISs.

The Classification and Ranking Belief Simplex (CaRBS) (Beynon, 2005a), utilized here to form the basis of an IS, is considered to operate its classification of students in the presence of ignorance (its mathematical foundation is based around the Dempster–Shafer theory of evidence; see Dempster (1968) and Shafer (1976)). A second motivation for this study is an exposition of the CaRBS system as an appropriate IS to aid the elucidation of a student's performance. The inclusion of the simplex plot approach to data representation in the CaRBS system offers the opportunity for a visual interpretation of the analysis presented (Beynon & Buchanan, 2004; Beynon, 2005b). This visual emphasis is in tune with views presented in Breiman (2001), highlighting the need for new techniques to be developed where the ability to fully interpret results is inherent.

Bhatt and Zaveri (2002) acknowledge that artificial intelligence is being embedded in many organizational learning based applications, enhancing their supportive capabilities. The issue of ISs in the education field has been considered previously. White (1987) suggests that they will be used more and more in forthcoming years; it is hoped this study supports their maxim. Here, we adhere to Keen and Scott-Morton (1978) who posit that the use of an IS can allow organizational decision-makers to effectively process data to identify trends and meaningful patterns. Importantly, the simplex plot facet of the CaRBS system is the standard domain for the representation of results, which, while novel, needs to be considered in terms of future familiarity.

The structure of the rest of the paper is as follows. In Section 2 the notion of e-learning and a relevant online university module are discussed, and in Section 3 the CaRBS system and the process of its optimum configuration are described. An initial CaRBS-based analysis is made on the module–student data set in Section 4, considering only the first six weeks' activity of students towards their final performance. In Section 5, further CaRBS analyses are made on different intermediate weeks of the module, and in Section 6 performance analyses of withdrawn students are elucidated. Conclusions are given in Section 7, as well as directions for future research.

2. E-learning and the module data set

E-learning implies education by means of digital media such as computers, Web pages, video conferencing systems and CD-Roms and learning enabled via the Internet (Keller & Cernerud, 2002; Jones *et al.*, 2004). The module in question is part of an undergraduate degree run by a UK university, developed specifically for online delivery. The module is supported by the 'Blackboard' software, utilizing synchronous and asynchronous communication mechanisms including discussion boards, e-mail and virtual classrooms, with module materials all held within individual online pages. The students are able to access relevant pages on a week-by-week basis, completing associated tasks and discussing the content with a tutor and fellow students; hence it is in their interests to participate in each week's activities.

This study analyses student activity on their initial module (Entrepreneurial Competencies) using the number of online pages they viewed each week. These details are recorded by a virtual learning environment system and thereafter reported to the module tutor who has the opportunity to contact individual inactive students. A total of 70 students were analysed, who fully completed the module. A number of students were not included, owing to their withdrawal. While not used in the performance analysis, these withdrawn students are used to exposit the concurrent information of the configured CaRBS system.

To utilize the CaRBS system in this study the pre-preparation of the data set is simply the standardization of the weekly activity values of the students, which removes any inherent scale effects of between-week activity levels (also beneficial when evaluating the necessary control variables, see later). The standardization process simply infers the subtraction of the associated mean and subsequent division by the associated standard deviation values for each week's activity; see Table 1.

The mean values in Table 1 show different levels of activity of the online pages accessed over the different weeks, as do the standard deviation values. Also reported are the minimum and maximum values associated with each week's activity, again supporting the variation in activity. Indeed, it is this inherent variation in the utilization of the e-learning facilities by the students over the different weeks that is a motivation for this study, as well as the exposition of the CaRBS system for performance analysis.

Table 1: *Descriptive details of weekly student 'activity' levels*

Week	*Mean*	*Min*	*Max*	σ
c_1	285.44	0	1114	244.59
c_2	273.64	0	1427	293.76
c_3	227.03	0	838	172.77
c_4	112.23	0	358	100.50
c_5	68.93	0	466	98.09
c_6	107.60	0	505	110.76
c_7	335.17	0	1519	254.03
c_8	189.36	0	894	168.81
c_9	129.11	0	575	120.38
c_{10}	112.74	0	561	118.98
c_{11}	172.40	0	825	164.31
c_{12}	204.43	0	791	189.60

The performance classification of the students considered here is based on their final module mark, with the threshold defined between 'less than 60%' and 'greater than or equal to 60%', described exclusively here as indifferent and good performance, respectively (in summary it is based on students 'not attaining' and 'attaining' a UK 2–1 class or above). Of the 70 students who completed the module (s_i, $i = 1, \ldots, 70$), 23 and 47 students are classified as indifferent and good, respectively (marks are spread over a continuous domain from a minimum of 32.8% to a maximum of 78.2%).

3. Technical description of the CaRBS system

This section briefly describes the main classification technique used here, namely the CaRBS system; for a more in-depth discussion see Beynon (2005a, 2005b). When used as a classification tool, it undertakes the predicted classification of objects (students) based on a number of characteristics (activity levels). The rudiments of CaRBS are based on Dempster–Shafer theory (Dempster, 1968; Shafer, 1976), which itself considers a finite set of p elements $\Theta = \{o_1, o_2, \ldots, o_p\}$, called a frame of discernment. A *mass value* is a function m: $2^\Theta \rightarrow [0, 1]$ such that $m(\emptyset) = 0$ ($\emptyset$ is the empty set) and $\sum_{s\in 2^\Theta} m(s) = 1$ (2^Θ is the power set of Θ). Any proper subset s of the frame of discernment Θ for which $m(s)$ is non-zero is called a focal element and the $m(s)$ value represents the exact belief in the proposition depicted by s.

Within CaRBS, the information from a characteristic value is quantified in a body of evidence (BOE) denoted by $m(\cdot)$, where all assigned mass values sum to unity and there is no belief in the empty set. Moreover, for a student s_j ($1 \le j \le n_O$) with ith week's activity c_i ($1 \le i \le n_C$), an *activity* BOE defined as $m_{j,i}(\cdot)$ has mass values $m_{j,i}(\{x\})$ and $m_{j,i}(\{\neg x\})$, which here denote levels of exact belief in the classification of a student to a hypothesis x (good performance) and not the hypothesis $\neg x$ (indifferent performance), and $m_{j,i}(\{x, \neg x\})$ is the level of concomitant ignorance. Following

Safranek *et al.* (1990), they are given by

$$m_{j,i}(\{x\}) = \frac{B_i}{1 - A_i}\mathrm{cf}_i(v) - \frac{A_i B_i}{1 - A_i}$$

$$m_{j,i}(\{\neg x\}) = \frac{-B_i}{1 - A_i}\mathrm{cf}_i(v) + B_i$$

$$m_{j,i}(\{x, \neg x\}) = 1 - m_{j,i}(\{x\}) - m_{j,i}(\{\neg x\})$$

where

$$\mathrm{cf}_i(v) = \frac{1}{1 + \mathrm{e}^{-k_i(v - \theta_i)}}$$

and k_i, θ_i, A_i and B_i are incumbent control variables. Importantly, if either $m_{j,i}(\{x\})$ or $m_{j,i}(\{\neg x\})$ is negative they are set to zero, and the respective $m_{j,i}(\{x, \neg x\})$ is then calculated. Figure 1 presents the progression from a value v to an activity BOE and its representation as a single simplex coordinate in a simplex plot.

In Figure 1, an activity level v is first transformed into a confidence value (1(a)), from which it is de-constructed into its activity BOE (1(b)), made up of a triplet of mass values $m_{j,i}(\{x\})$, $m_{j,i}(\{\neg x\})$ and $m_{j,i}(\{x, \neg x\})$. The notion of ignorance here is a part of the ambiguity between where there is more certainty in the evidence supporting more $\{x\}$ or $\{\neg x\}$. Stage 1(c) shows a BOE $m_{j,i}(\cdot)$; $m_{j,i}(\{x\}) = v_{j,i,1}$, $m_{j,i}(\{\neg x\}) = v_{j,i,2}$ and $m_{j,i}(\{x, \neg x\}) = v_{j,i,3}$ can be represented as a simplex coordinate ($p_{j,i,v}$) in a simplex plot (equilateral triangle). That is, a point $p_{j,i,v}$ exists within an equilateral triangle such that the least distances from $p_{j,i,v}$ to each of the sides of the equilateral triangle are in the same proportion (ratio) as the values $v_{j,i,1}$, $v_{j,i,2}$ and $v_{j,i,3}$. In Figure 1(c), a number of BOEs are exhibited as points in the simplex plot, which can be used to demonstrate the relationship between BOEs and representation in a simplex plot.

The set of activity BOEs $\{m_{j,i}(\cdot), i = 1, \ldots, n_C\}$ associated with the student s_j can be combined using Dempster's combination rule into a *student* BOE, defined as $m_j(\cdot)$. Moreover, using $m_{j,i}(\cdot)$ and $m_{j,k}(\cdot)$ as two independent activity BOEs, $[m_{j,i} \oplus m_{j,k}](\cdot)$ defines their combination, given by

$$[m_{j,i} \oplus m_{j,k}](\{x\}) = \frac{m_{j,i}(\{x\})m_{j,k}(\{x\}) + m_{j,k}(\{x\})m_{j,i}(\{x, \neg x\}) + m_{j,i}(\{x\})m_{j,k}(\{x, \neg x\})}{1 - [m_{j,i}(\{\neg x\})m_{j,k}(\{x\}) + m_{j,i}(\{x\})m_{j,k}(\{\neg x\})]}$$

$$[m_{j,i} \oplus m_{j,k}](\{\neg x\}) = \frac{m_{j,i}(\{\neg x\})m_{j,k}(\{\neg x\}) + m_{j,k}(\{x, \neg x\})m_{j,i}(\{\neg x\}) + m_{j,k}(\{\neg x\})m_{j,i}(\{x, \neg x\})}{1 - [m_{j,i}(\{\neg x\})m_{j,k}(\{x\}) + m_{j,i}(\{x\})m_{j,k}(\{\neg x\})]}$$

$$[m_{j,i} \oplus m_{j,k}](\{x, \neg x\}) = 1 - [m_{j,i} \oplus m_{j,k}](\{x\}) - [m_{j,i} \oplus m_{j,k}](\{\neg x\})$$

This process is then used iteratively to combine all the activity BOEs describing a student into the student BOE. For a student s_j, the student BOE contains the information necessary for his/her final classification.

To illustrate the method of combination employed here, the two example BOEs, $m_1(\cdot)$ and $m_2(\cdot)$, are further considered, with their combination to a BOE denoted $m_C(\cdot)$, and evaluated to be $m_C(\{x\}) = 0.467$, $m_C(\{\neg x\}) = 0.224$ and

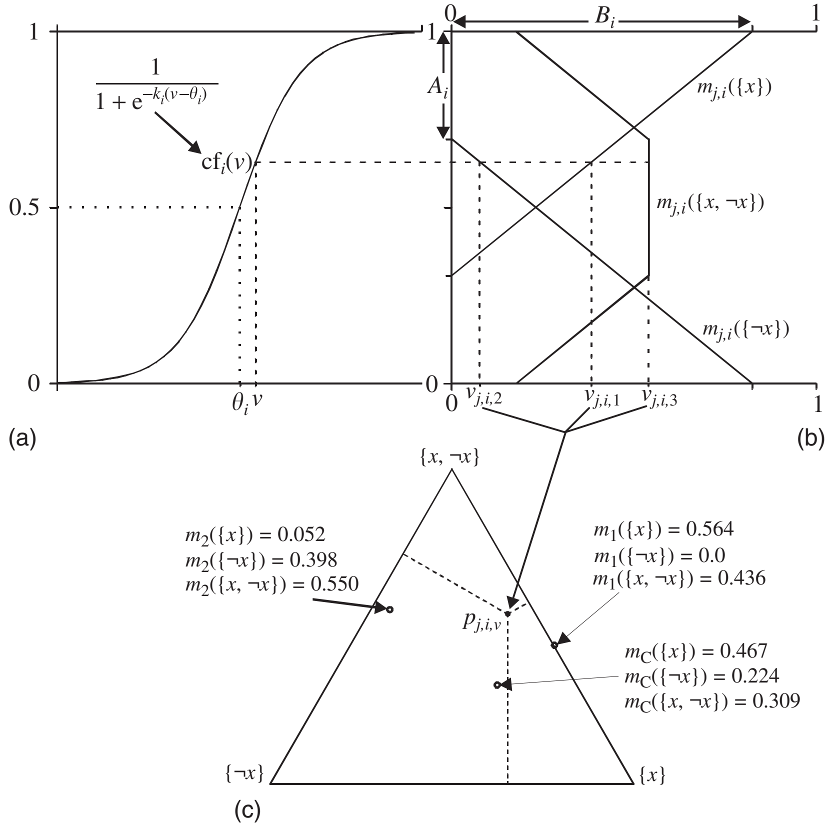

Figure 1: *Graphical representation of stages in CaRBS for a single value.*

$m_C(\{x, \neg x\}) = 0.309$. The combination process is graphically shown with the simplex coordinate representation of the combined BOE $m_C(\cdot)$ in Figure 1(c). In this case, $m_1(\cdot)$ offers more evidential support to the combined BOE $m_C(\cdot)$ than $m_2(\cdot)$, since the ignorance in $m_2(\cdot)$ is more than that associated with $m_1(\cdot)$. In the limit, a final object BOE will have a lower level of ignorance than that associated with the individual variable BOEs.

The configuration of a CaRBS system depends on the assignment of values to the incumbent control variables (k_i, θ_i, A_i and B_i, $i = 1, \ldots, n_C$). With the weekly activity levels standardized, the domains of the control variables are set as $0 \le k_i \le 2$ (positive relationship between activity level and improved performance), $-1 \le \theta_i \le 1$, $0 \le A_i < 1$ and $B_i = 0.3$ (see Beynon, 2005b). With closed domains of the control variables this becomes a constrained optimization problem, solved here using an evolutionary algorithm called trigonometric differential evolution (Fan & Lampinen, 2003) with the following operation parameters: amplification control $F = 0.99$, crossover constant $\mathrm{CR} = 0.85$, trigonometric mutation probability $M_t = 0.05$ and number of parameter vectors $\mathrm{NP} = 10 \times$ number of control variables $= 360$.

Associated with any evolutionary algorithm is an objective function (OB), here a positive function that measures the misclassification of students from their known performance classification. The equivalence classes $E(x)$ and $E(\neg x)$ are sets of students known to be classified to $\{x\}$ and $\{\neg x\}$, respectively. For objects in $E(x)$ and $E(\neg x)$ the optimum solution is to maximize the weighted difference values $m_j(\{x\}) - m_j(\{\neg x\})$ and $m_j(\{\neg x\}) - m_j(\{x\})$, respectively.

The subsequent OB is given by

$$\frac{1}{4}\left(\frac{1}{|E(x)|}\sum_{o_j\in E(x)}[1-m_j(\{x\})+m_j(\{\neg x\})]\right.$$

$$\left.+\frac{1}{|E(\neg x)|}\sum_{o_j\in E(\neg x)}[1+m_j(\{x\})-m_j(\{\neg x\})]\right)$$

In the limit, each difference value can attain -1 and 1; hence $0 \le \mathrm{OB} \le 1$. Maximizing a difference value such as $m_j(\{x\}) - m_j(\{\neg x\})$ only indirectly affects the associated ignorance, rather than making it a direct issue, since the OB does not incorporate the respective $m_j(\{x, \neg x\})$ mass values. The division of elements of OB by $|E(\cdot)|$ takes account of unbalanced data sets, in this case with different numbers of indifferent and good students.

An indication of the evidential support offered by each week's activity to the known indifferent and good students is made with the evaluation of average activity BOEs. More formally, partitioning the students into the equivalence classes $E(x)$ and $E(\neg x)$, then the average activity BOEs, defined as $\mathrm{am}_{i,x}(\cdot)$ and $\mathrm{am}_{i,\neg x}(\cdot)$ respectively, are given by

$$\mathrm{am}_{i,x}(\{x\}) = \sum_{s_j\in E(x)}\frac{m_{j,i}(\{x\})}{|E(x)|}$$

$$\mathrm{am}_{i,x}(\{\neg x\}) = \sum_{s_j\in E(x)}\frac{m_{j,i}(\{\neg x\})}{|E(x)|}$$

$$\mathrm{am}_{i,x}(\{x,\neg x\}) = \sum_{s_j\in E(x)}\frac{m_{j,i}(\{x,\neg x\})}{|E(x)|}$$

$$\mathrm{am}_{i,\neg x}(\{x\}) = \sum_{s_j\in E(\neg x)}\frac{m_{j,i}(\{x\})}{|E(\neg x)|}$$

$$\mathrm{am}_{i,\neg x}(\{\neg x\}) = \sum_{s_j\in E(\neg x)}\frac{m_{j,i}(\{\neg x\})}{|E(\neg x)|}$$

$$\mathrm{am}_{i,\neg x}(\{x,\neg x) = \sum_{s_j\in E(\neg x)}\frac{m_{j,i}(\{x,\neg x\})}{|E(\neg x)|}$$

where s_j is a student. As BOEs they can be represented as simplex coordinates in a simplex plot describing the evidential support of each week's activities to the performance classification of the students.

4. CaRBS performance classification of students based on the first six weeks' activity

The analysis here on the module–student data set, using the CaRBS system, utilizes only the first six weeks' activity of the students on the module (first three two-week periods). As such the findings could be used to check on the performance of the students at the halfway stage of the module (different multiples of two-week time periods could be analysed, see later).

Following the description of the CaRBS system in the previous section, its optimum configuration, through the defined objective function, attempts to minimize the level of ambiguity in each student's classification but not the concomitant level of ignorance. The first technical results presented are the incumbent control variable values in the configured CaRBS system, found using the trigonometric differential evolution algorithm when only the first six weeks' activity of the students on the module are considered; see Table 2.

Considering these control variables, their effect on the predicted classification of an individual is presented next, namely the student s_{14}. This starts with the construction of the activity BOE $m_{14,1}(\cdot)$, describing the evidence from week 1's activity. First the associated $\mathrm{cf}_{14,1}(\cdot)$ is calculated, using the week 1's standardized activity level $v = -0.7009$ (see Table 3):

$$\mathrm{cf}_{14,1}(-0.7009) = \frac{1}{1+\mathrm{e}^{-2.0000(-0.7009-0.0384)}}$$

$$= \frac{1}{1+4.3872} = 0.1856$$

This value is used in the expressions making up the mass values in the activity BOE $m_{14,1}(\cdot)$, namely $m_{14,1}(\{x\})$, $m_{14,1}(\{\neg x\})$ and

Table 2: *Control variable values associated with six-week 'e-learning activity' problem*

Variables	*Week 1 (c_1)*	*Week 2 (c_2)*	*Week 3 (c_3)*	*Week 4 (c_4)*	*Week 5 (c_5)*	*Week 6 (c_6)*
k_i	2.0000	0.2144	2.0000	1.9988	2.0000	1.9959
θ_i	0.0384	0.7169	–0.0029	0.2610	–0.5142	–0.2085
A_i	0.2144	0.8933	0.9840	0.9775	0.4068	0.9723

Table 3: *Activity BOEs $m_{14,i}(\cdot)$, $i = 1, \ldots, 6$, and student BOE $m_{14}(\cdot)$ for student s_{14}*

	c_1	c_2	c_3	c_4	c_5	c_6	*Student BOE*
	114	*10*	*50*	*148*	*337*	*300*	
Standardized values	–0.7009	–0.8975	–1.0247	0.3559	2.7330	1.7371	
$m_{14,i}(\{x\})$	0.0000	0.0000	0.0000	0.0000	0.2991	0.0814	0.2991
$m_{14,i}(\{\neg x\})$	0.2291	0.0000	0.0000	0.0000	0.0000	0.0000	0.1606
$m_{14,i}(\{x, \neg x\})$	0.7709	1.0000	1.0000	1.0000	0.7009	0.9186	0.5403

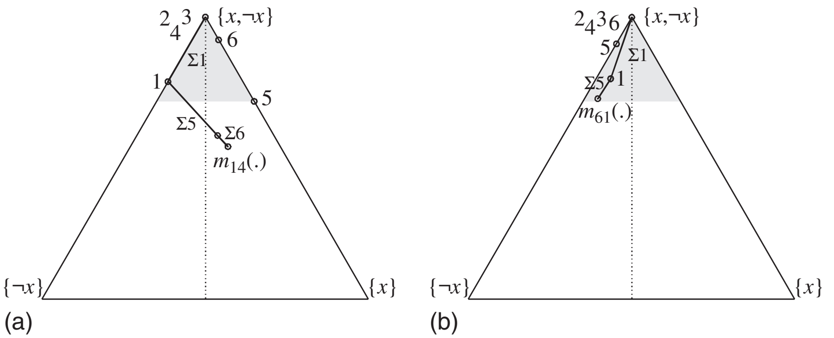

Figure 2: *Simplex coordinates of activity, p_student and student BOEs of (a) s_{14} and (b) s_{61}.*

$m_{14,1}(\{x, \neg x\})$:

$$m_{14,1}(\{x\}) = \frac{0.3}{1 - 0.2144}0.1856 - \frac{0.2144 \times 0.3}{1 - 0.2144}$$
$$= 0.0709 - 0.0819$$
$$= -0.0110 < 0 \text{ so } 0.0000$$

$$m_{14,1}(\{\neg x\}) = \frac{-0.3}{1 - 0.2144}0.1856 + 0.3$$
$$= -0.0709 + 0.3$$
$$= 0.2291$$

$$m_{14,1}(\{x, \neg x\}) = 1 - 0.0000 - 0.2291 = 0.7709$$

This BOE is representative of the activity BOEs $m_{14,i}(\cdot)$, $i = 1, \ldots, 6$, presented in Table 3.

The activity BOEs reported in Table 3 show large levels of ignorance associated with their evidence towards the final predicted classification of the student s_{14}. Indeed, the activity-based evidence from weeks 2, 3 and 4 are all total ignorance towards the performance classification. This collection of evidence is then combined to form the respective student BOE $m_{14}(\cdot)$, using Dempster's combination rule: $m_{14}(\{x\}) = 0.2991$, $m_{14}(\{\neg x\}) = 0.1606$ and $m_{14}(\{x, \neg x\}) = 0.5403$. These activity and student BOEs can be represented as simplex coordinates in a simplex plot (see Figure 2).

In Figure 2(a), the details of the performance classification of the student s_{14} are presented in a simplex plot (known to have a good performance (x) with final mark 62.6%). At the top of the equilateral triangle in the shaded region are circles (simplex coordinates) that represent the activity BOEs ($m_{14,i}(\cdot)$, $i = 1, \ldots, 6$, in this case) associated with each week's activity contribution to the final performance (labelled $i = 1, \ldots, 6$). The positions of the activity BOEs on either side of the vertical broken line identify whether they confer correct (right) or incorrect (left) support for the classification of this student (known good performance x). Here correct supporting evidence comes from weeks 5 and 6, and incorrect evidence from week 1, with weeks 2, 3 and 4 offering no evidence (only ignorance).

The emphasis in this study is on the graphical representation of the results regarding the predicted performance classification of the individual students. However, more linguistic interpretation can be gained from the graphs; in the case of student s_{14}'s performance classification, it can be stated as

> Based on the first six weeks' activity in the module, the level of activity in week 1 suggested a final indifferent performance, the activity in weeks 2, 3 and 4 contributed nothing to the performance classification, but the activities in weeks 5 and 6 were adequate to support a good performance – culminating in an overall good performance with a noticeable ignorance associated with it.

The configured CaRBS system allows the ongoing performance activity of a student during the first six weeks of the module considered. This is undertaken by combining the evidence from the successive weeks' activity levels (using Dempster's combination rule iteratively). That is, the concurrent temporal evidence is a result of the combination of the weekly activity BOEs up to that week. This is shown in Figure 1(a); moving down the simplex plot are further simplex coordinates which represent the *progressive student* BOEs (p_student BOEs), defined as $m_{j,\Sigma i}(\cdot)$. These are BOEs constructed from the combination of the activity BOEs for the student from week 1 up to and including the ith week, labelled Σi (not included are those weeks when a week's activity confers total ignorance).

For the student s_{14}, the $\Sigma 1$, $\Sigma 5$ and $\Sigma 6$ labels are shown near the lines adjoining the simplex coordinates of the respective p_student BOEs. For example, the p_student BOE $m_{14,\Sigma 5}(\cdot)$ is constructed from the combination of the five activity BOEs $m_{14,1}(\cdot)$, $m_{14,2}(\cdot)$, $m_{14,3}(\cdot)$, $m_{14,4}(\cdot)$ and $m_{14,5}(\cdot)$, even though only $m_{14,1}(\cdot)$ and $m_{14,5}(\cdot)$ contribute specific evidence. The successive p_student BOEs' positions progressively down the simplex plot highlight that as BOEs are combined the p_student BOE produced has less ignorance. With six weeks' activity, the p_student BOE $m_{14,\Sigma 6}(\cdot)$ is the final student BOE for this student, also labelled $m_{14}(\cdot)$, which is to the right of the vertical broken line and suggests good performance classification (correct classification in this case).

In Figure 2(b), a similar analysis is given on the student s_{61}, known to have an indifferent performance on the module. The simplex coordinate of the student BOE $m_{61}(\cdot)$ is to the left of the vertical broken line which again confers correct indifferent performance classification. Contributing evidence is from weeks 1 and 5, which both offer evidence to the indifferent performance. These two examples of the results from the configured CaRBS system highlight the standard domain of the simplex plot and in future applications its familiarity will allow full interpretation to be available. For all the 70 students considered, their classification details, the simplex coordinates of their associated final student BOEs, are reported in Figure 3.

The two simplex plots in Figure 3 separate the presentation of the classification results of the 70 students with indifferent (3(a)) and good (3(b)) performances. For the correct classification of the students the circles presented should be to the left (indifferent) and right (good) of the vertical broken lines in Figures 3(a) and 3(b), respectively. The simplex coordinates are spread across the upper part of the simplex plot domains, at different heights indicating varying levels of ignorance associated with all the students' performance classifications. The number of simplex coordinates advocating correct classification for students shows that 68.57% were correctly classified (48 out of 70; 15 out of 23 indifferent (2(a)) and 33 out of 47 good (2(b)) students). These results are comparable with a random 50% accuracy since the objective function utilized takes into account unbalanced data

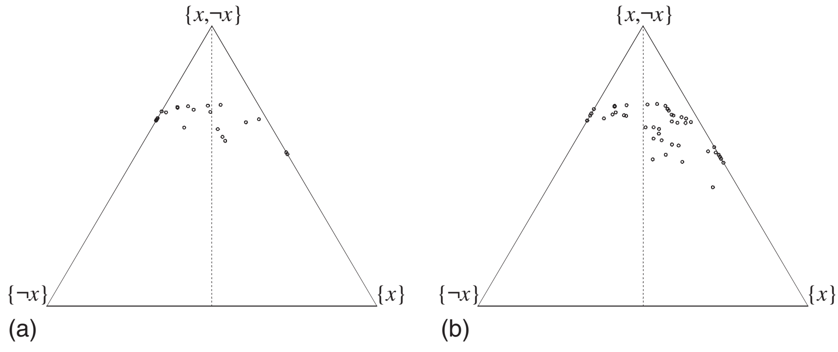

Figure 3: *Simplex coordinates of all student BOEs: (a) indifferent; (b) good.*

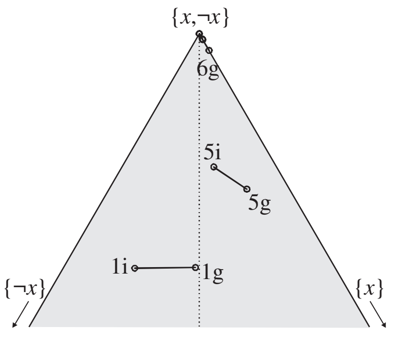

Figure 4: *Simplex coordinates of* $\mathrm{am}_{i,\neg x}(\cdot)$ *'i' and* $\mathrm{am}_{i,x}(\cdot)$ *'g' average activity BOEs.*

sets. A limited benchmark to this result is a traditional multivariate discriminant analysis (MDA), which offered a similar 68.57% classification accuracy (a Shapiro–Wilk test for normality identified that all these weeks' activity levels could be rejected against having a normal distribution (1% significance), hence mitigating the MDA results).

The CaRBS system also aids the education facilitators (module tutors) in the evidential contribution of each week's activity to discern between indifferent and good students. The weekly *average activity* BOEs are calculated separately for those students known to be indifferent ($\neg x$) and good (x) students (defined as $\mathrm{am}_{i,\neg x}(\cdot)$ and $\mathrm{am}_{i,x}(\cdot)$; see Section 3). Since they are themselves BOEs they can be represented as simplex coordinates in a simplex plot (see Figure 4).

In Figure 4, the simplex coordinates labelled 'i' and 'g' in the presented sub-domain of the simplex plot (shaded region only shown) denote the positions of the average activity BOEs $\mathrm{am}_{i,\neg x}(\cdot)$ and $\mathrm{am}_{i,x}(\cdot)$, respectively (numbers indicate the week of activity in question). It is noticeable that the simplex coordinates associated with weeks 2, 3 and 4 are all very near or at the $\{x, \neg x\}$ vertex, indicating that they contribute little or no evidence to the students' final performance classifications (see later). Most noticeable results concern weeks 1 and 5 which are considerably far away from the $\{x, \neg x\}$ vertex and therefore have less ignorance associated with the evidence they confer.

The module tutors can utilize the findings in Figure 4 to identify which of the first six weeks' activities are crucial (indicative) to the students' final performances; noticeably week 1's activity embraces least ignorance ('1i' and '1g' are furthest down the simplex plot). For this module, the activity in week 1 involved a face-to-face induction, whereby students were encouraged to access and experiment within the virtual learning environment. Interestingly, this week was instruction driven, and hence it could be considered that particularly proactive work was undertaken. However, it is possible that a combination of the activities of weeks 1 and 5 is the main contributing factor – so interpretation of individual weeks is limited.

The contribution of each week's activity level is further exposited to show the direct relationship between the level of student activity, online pages viewed, and the mass values in the activity BOE subsequently constructed; see Figure 5

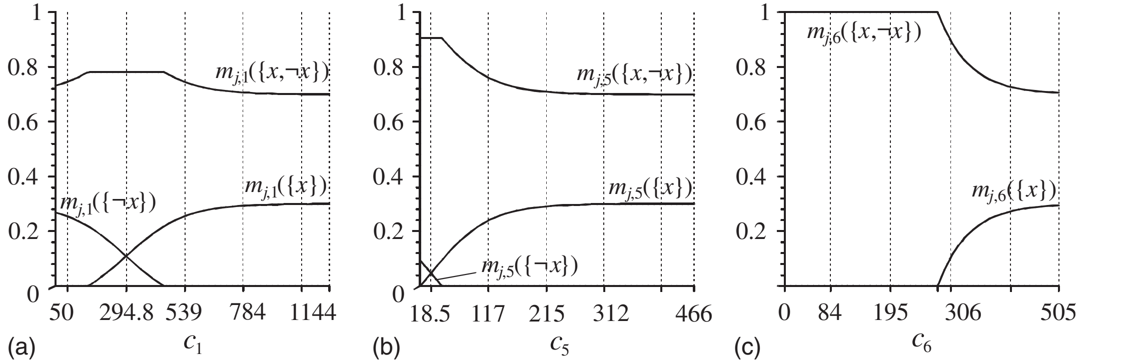

Figure 5: *Graphs of mass values in the activity BOEs from weeks 1 (c_1), 5 (c_5) and 6 (c_6).*

where weeks 1, 5 and 6 are considered (those shown to contribute to the constructed student BOEs, see Figure 4). These graphs are special cases of that reported in Figure 1 (a combination of stages 1(a) and 1(b)); in each graph the 'online pages viewed' domains shown are between zero and the largest number present amongst the students considered.

In Figure 5(a), the graph shows that a week 1 activity above 294.8 online pages viewed is what is needed to confer more evidence to achieving a good performance ($m_{j,1}(\{x\})$) than an indifferent one ($m_{j,1}(\{\neg x\})$) (its simplex coordinate would be to the right of the vertical broken line in a simplex plot). The dominance of the $m_{j,i}(\{x, \neg x\})$ values in each graph is a consequence of the need to have a large level of ignorance associated with the individual activity BOEs. The vertical broken lines indicate the changes in the activity BOEs if they are an integer number of standard deviations away from the $\theta_1 = 294.8$ value (see previously).

The graph in Figure 5(c) supports the limited evidence that the activity in week 6 offered in the performance classification of the students (see Figure 4). That is, it suggests that only those students who viewed over about 282 online pages would be undertaking activity that confers more evidence to them achieving a good performance. Below this level the evidence is total ignorance and not supporting evidence to achieving an indifferent performance. The utilization of these graphs can be demonstrated with reference to the student s_{14}. The evaluation of the activity BOEs is shown in Table 3 including $m_{14,1}(\cdot)$; with $c_1 = 114$ the resultant activity BOE is $m_{14,1}(\{x\}) = 0.0000$, $m_{14,1}(\{\neg x\}) = 0.2291$ and $m_{14,1}(\{x, \neg x\}) = 0.7709$ (see Figure 5(a)).

A further elucidation of what evidence an activity BOE confers is further elucidated here with a brief comparison with what happens in a technique such as linear regression, where a single activity value is one dimensional and can only offer a negative or positive value. An activity BOE with its triplet of mass values includes the extreme values to some hypothesis, not the hypothesis and the intermediate ignorance. With respect to the e-learning problem the ignorance term is an uncertainty as to whether the number of online pages viewed is enough to add to their chances of achieving a good performance or not enough, instead suggesting an indifferent performance. In the week 6 case, the evidence from between zero and around 282 online pages viewed is totally uncertain – hence the activity BOE would only confer total ignorance.

This section concludes with a further utilization of the graphs given in Figure 5. Moreover, they can be used to identify where further activity in a single week would offer limited gain towards achieving a good performance. To illustrate, student s_{43} had the highest activity in week 1, with 1144 online pages viewed (see Figure 5(a)), and consistently achieved near highest activity in all of the 12 weeks of the module. However, this student withdrew from the course during the next module stating as a reason that they could not maintain their perceived level of commitment (activity). This perception could be detrimental to their benefit, as the results in Figure 5 show, anecdotally, that there is marginal gain from over-activity in a single week of a module. For a student in week 1 it could be argued that activity above around 784 online pages viewed offers limited additional evidence to their achieving a good performance classification. Instead, in the notion of pastoral care, it would be to the benefit of the student and module tutor(s) that this over-activity is highlighted. This is a novel contribution that the CaRBS system can offer to the efficacy of a module on this course with respect to potential impact on student retention.

5. Performance classification of students over varying numbers of weeks of activity

This section presents similar results to those presented earlier, considering the evidence from the students' weekly activity over certain numbers of weeks in the 12-week module. Following the ongoing assessments on the students every two weeks, the separate CaRBS analyses of their performance presented here are based on activity in the two, four, six, eight, ten and 12 weeks. In each study the optimization of the students to their known performance classification is attempted. Figure 6 reports the results associated with student s_{14} (considered earlier).

The six graphs in Figure 6 represent the performance classification of student s_{14} over six different sets of activity weeks. Working through the graphs in order, the first two graphs (6(a) and 6(b)) similarly show that only the activity in week 1 contributed evidence using the CaRBS system and the evidence was of the student s_{14} achieving only an indifference performance at the end of the module. After six weeks (6(c)), the activities in weeks 5 and 6 both offer evidence to achieving a good performance – hence the overall performance classification is suggesting a good performance. In Figures 6(d), 6(e) and 6(f) the performance classifications of

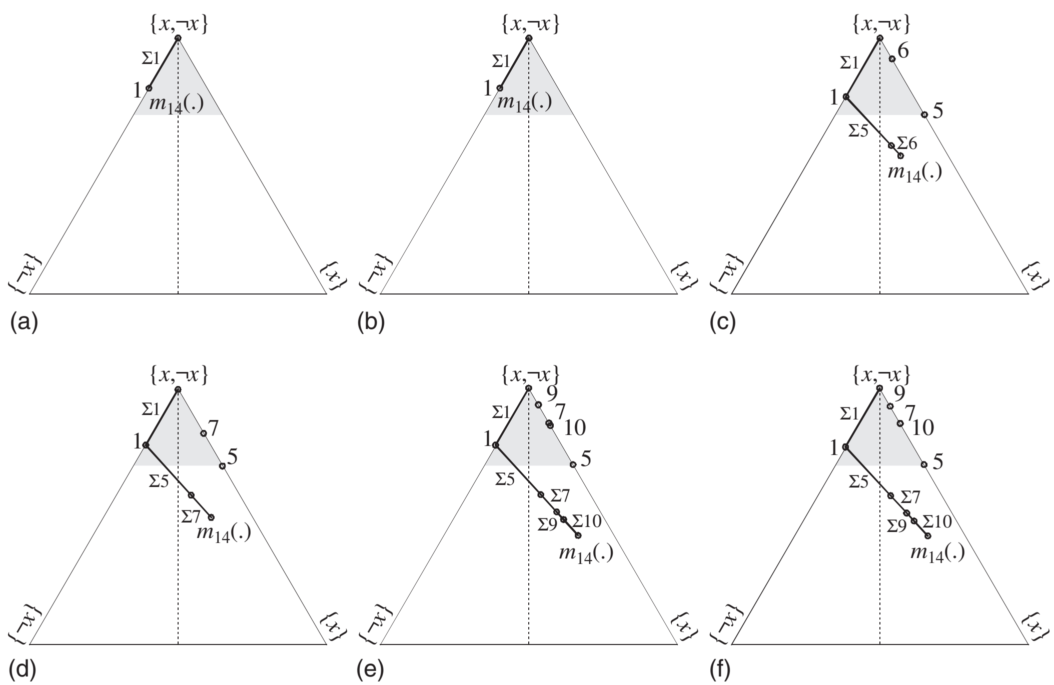

Figure 6: *Simplex coordinates of activity, p_student and student BOEs of s_{14} using different numbers of weeks of activity.*

s_{14} over the first eight, ten and 12 weeks of the module show a consistent 'good' classification with less contribution from week 6 but further contributions from weeks 7 (6(d)) and 7, 9 and 10 (6(e) and 6(f)).

The next set of results concerns the final performance classification of all the 70 students considered. For brevity, only numerical accuracy results are presented (graphs similar to Figure 3 could be produced). For the sets of weeks 2 to 12, the correct classification results are as follows: week 2, 39/70 (55.71%); week 4, 41/70 (58.57%); week 6, 48/70 (68/57%); week 8, 48/70 (68/57%); week 10, 51/70 (72.86%); and week 12, 51/70 (72.86%), respectively. This progressive improving correct classification accuracy is understandable since there are an increasing number of activity weeks to help discern between the indifferent and good performances of the students. The less than 100% values identify the limited results, signifying that the activities over the different weeks are not the only antecedents to their success (left for future research). The contributions of the different activity weeks in the separate CaRBS analyses are considered next (see Figure 7).

Throughout the graphs reported in Figure 7 the relatively small movement of the 1i ($\mathrm{am}_{i,\neg x}(\cdot)$) and 1g ($\mathrm{am}_{i,x}(\cdot)$) labels show the consistent influence in discerning the indifferent and good performance classification of the students. The case of the activity in week 4 is interesting in that its influence changes when considering the evidence from the first four (7(b)) and six (7(c)) weeks. From Figure 7(c) onwards the contribution of the activity in week 5 is noticeable with the positions of 5i and 5g consistently down from the $\{x, \neg x\}$ vertex. In the latter figures the limited contributions of weeks 7, 8, 9, 10 and 11 are also noticeable.

6. Intermediate performance of withdrawn students

This section utilizes the information in the CaRBS system on those students who have

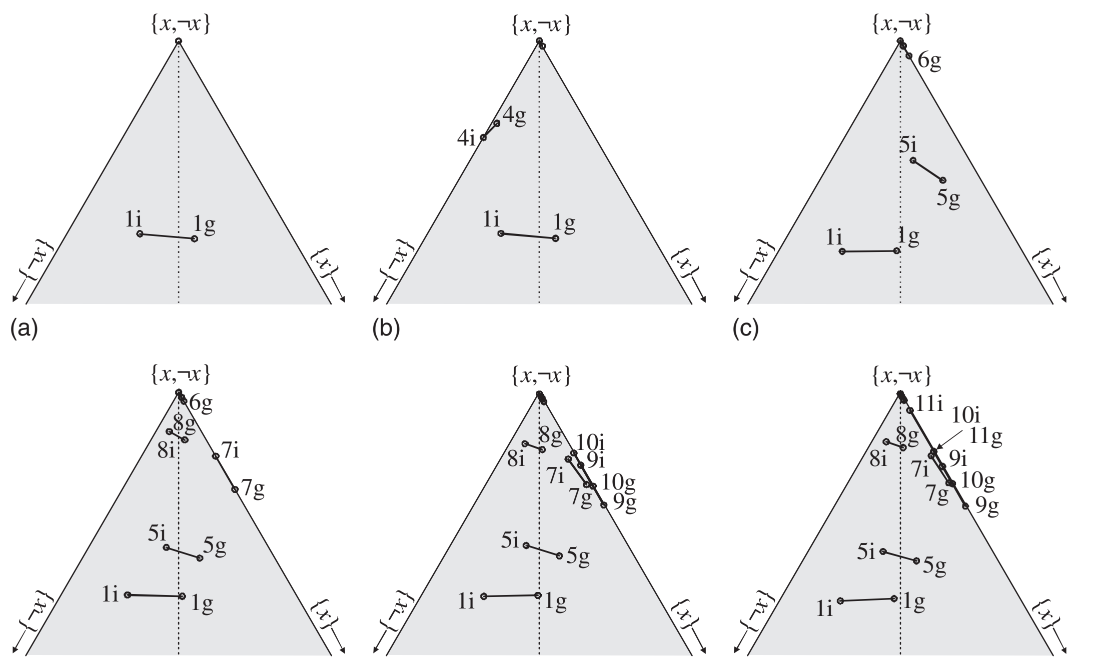

Figure 7: *Simplex coordinates of* $\mathrm{am}_{i,\neg x}(\cdot)$ *'i' and* $\mathrm{am}_{i,x}(\cdot)$ *'g' average activity BOEs: (a) two weeks; (b) four weeks; (c) six weeks; (d) eight weeks; (e) ten weeks; (f) 12 weeks.*

withdrawn during the module. More specifically, 15 students withdrew at various stages, their reasons including health, pressures of work, personal circumstances and incompatibility with e-learning. Suggested strategies to overcome student withdrawal include effective module flexibility, student monitoring and support, assessment strategy and pedagogical design in addition to effective recruitment and induction policies (Rovai, 2003). Unlike the previous completing students, the activity levels of the withdrawn students could be incomplete (over the 12 weeks of the module). Hence only those analysis results which include activity weeks that they were present for are included. Here the evidence of three withdrawn students are considered (s_{wd1}, s_{wd2} and s_{wd3}) (see Figures 8, 9 and 10).

In Figure 8 the four simplex plots presented are due to the known withdrawal of the student s_{wd1} after the eighth week; hence details up to week 8 are presented. The results in Figures 8(a) and 8(b) show progressive indifferent performance, but week 5 showed noticeable improvement with week 8 again relapsing. This student's reasons for withdrawing included incompatibility with e-learning as a pedagogical style and low prior information technology skills. This meant limited activity in the early weeks, but they were then contacted to enquire about their activity which resulted in increased activity during weeks 5 and 6 (see Figures 8(c) and 8(d)). However, further limited activity in the next weeks contributed to their decision to withdraw. The results presented in Figure 8 offer some evidence to support this student's case. The case study presented and the results show that just because limited or no activity may actually take place this may not be as critical in their final performance for certain weeks, simply because, when optimizing the performance of students based on their activity, some weeks were not as important as others. A similar series of results are presented for the second withdrawn student s_{wd2} in Figure 9.

The four simplex plots in Figure 9 again respect the issue that the student withdrew after

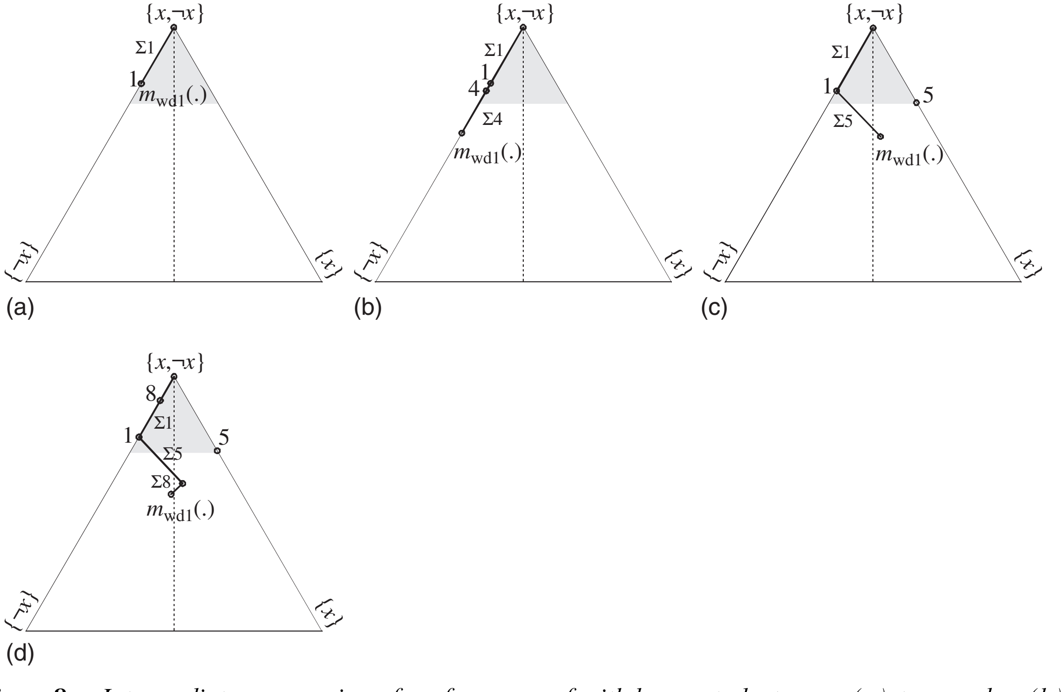

Figure 8: *Intermediate progression of performance of withdrawn student s_{wd1}: (a) two weeks; (b) four weeks; (c) six weeks; (d) eight weeks.*

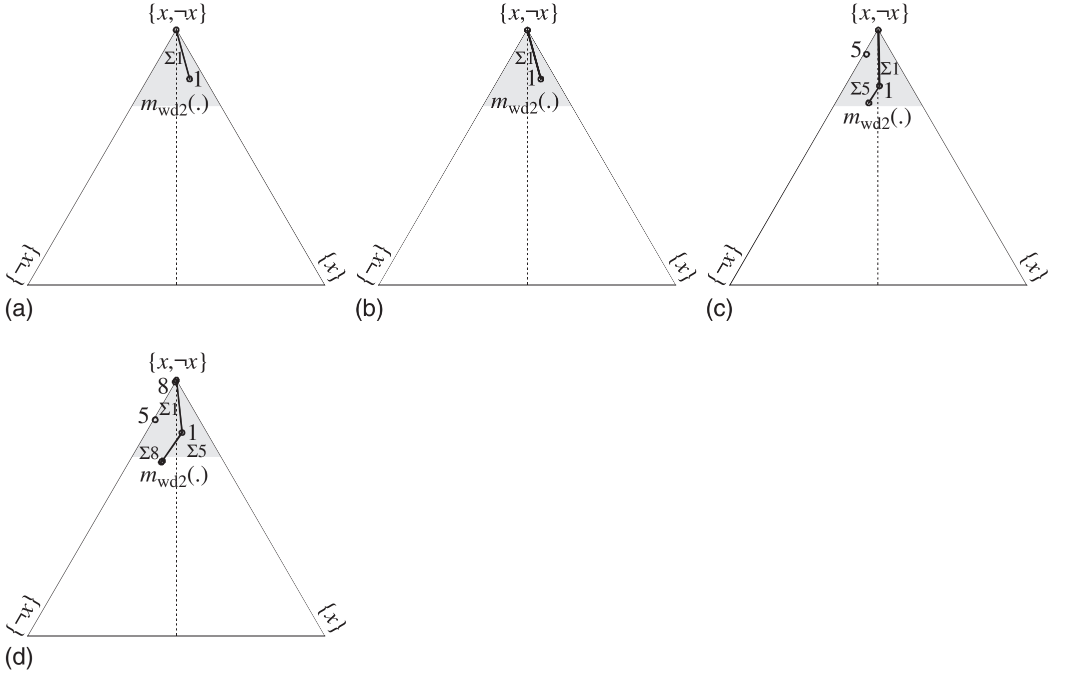

Figure 9: *Intermediate progression of performance of withdrawn student s_{wd2}: (a) two weeks; (b) four weeks; (c) six weeks; (d) eight weeks.*

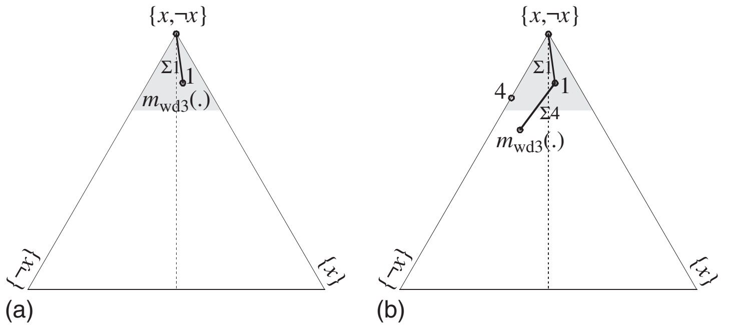

Figure 10: *Intermediate progression of performance of withdrawn student s_{wd3}: (a) two weeks; (b) four weeks.*

the eighth week of the module. The results shown over the different weeks identify the good contribution from week 1 but indifferent contribution from weeks 5 and 8 (marginally). This student withdrew from the course after week 8, citing a change of work circumstances as the main reason. He had tried to keep up with the course, but as the results in Figures 9(c) and 9(d) show, the activity particularly in weeks 5 and 8 show an indifferent contribution; week 1 was a good contribution. The results on the final withdrawn student s_{wd3} are reported in Figure 10.

The two simplex plots in Figure 10 respect the fact that the student withdrew after the fourth week of the module. This student felt the course materials were not meeting his business needs, and hence withdrew. The results in Figure 10 identify a good contribution in week 1 but a noticeable downturn in week 4, which correlates with his withdrawal at the end of week 4.

7. Conclusions

It is acknowledged that universities are 'buying in' to e-learning, one reason being that it opens up a new cohort of students to whom they can offer university education. Here, a single university module is considered, with the weekly activity details of the students' access to online pages used, within a configured IS, to predict a known binary classification of their performance. A second motivation for this study was the elucidation of a novel classification technique, namely the CaRBS system, as a practical IS in this area. A number of CaRBS analyses are developed, which use different two-week sets of the total activity (duration of the module) of the students to classify them to a final indifferent or good performance classification (defined as a constrained optimization problem). The utilized objective function attempts to minimize ambiguity but not the inherent ignorance in the classification results.

The CaRBS system is shown to be an effective IS in this study, where, based on the simplex plot method of visual data representation, the support from each week's activity to a student classification is elucidated. For the module tutor the simplex plot offers a standard domain within which to judge the temporal progress of different students as well as to view the importance/relevance of the individual characteristics (activity in each week). This is extended to the intermediate stages of the module, whereby the progression of a student to achieving the indifference or good performance classification can be identified. Indeed, this is most beneficial in future years, where, with stability of the module content, the CaRBS system found from previous years' data can be used to check on the progression of students in future years.

In light of the dearth of literature connecting e-learning with ISs, the study here enables concomitant module tutors to closely monitor student performance and inform the respective pedagogical design. Such activity monitoring should improve student interaction and retention, and ultimately performance. In conclusion, we believe that this study contributes a novel opportunity to monitor student behaviour. Future development of the utilization of the CaRBS system can inform on the relationship between the student and the incumbent e-learning community.

References

ALEXANDER, S. (2001) E-learning developments and experiences, *Education + Training*, **43**, 240–248.

BARABASH, M., R. GUBERMAN-GLEBOV and R. BARUCH (2003) Decision-making in construction of courses combining classroom based and Internet based learning and teaching strategies in mathematics teachers' education, *Journal of Educational Media*, **28**, 147–163.

BEYNON, M.J. (2005a) A novel technique of object ranking and classification under ignorance: an application to the corporate failure risk problem, *European Journal of Operational Research*, **167**, 493–517.

BEYNON, M.J. (2005b) Optimizing object classification under ambiguity/ignorance: application to the credit rating problem, *International Journal of Intelligent Systems in Accounting, Finance and Management*, **13**, 113–130.

BEYNON, M.J. and K. BUCHANAN (2004) A novel approach to gender classification under ignorance: the case of the European barn swallow (*Hirundo rustica*), *Expert Systems with Applications*, **27**, 403–415.

BHATT, G.D. and J. ZAVERI (2002) The enabling role of decision support systems in organisational learning, *Decision Support Systems*, **32**, 297–309.

BREIMAN, L. (2001) Statistical modelling: the two cultures, *Statistical Science*, **16**, 199–231.

DEMPSTER, A.P. (1968) A generalization of Bayesian inference (with discussion), *Journal of the Royal Statistical Society, Series B*, **30**, 205–247.

FAN, H.-Y. and J. LAMPINEN (2003) A trigonometric mutation operation to differential evolution, *Journal of Global Optimization*, **27**, 105–129.

FRANK, M., G. KURTZ and N. LEVIN (2002) Implications of presenting pre-university courses using the blended e-learning approach, *Educational Technology and Society*, **5**, 1–16.

GUNASEKARAN, A., D. MCNEIL and D. SHAUL (2002) E-learning: research and applications, *Industrial and Commercial Training*, **34**, 44–53.

JOLLIFFE, A., J. RITTER and D. STEVENS (2001) *The Online Learning Handbook – Developing and Using Web Based Learning*, London: Kogan-Page.

JONES, P., G. PACKHAM, C. MILLER and A. JONES (2004) An initial evaluation of student withdrawals within an e-learning environment: the case of e-College Wales, *Electronic Journal of E-learning*, **2**, 113–120.

Keen, P.G.W. and M.S. Scott-Morton (1978) *Decision Support Systems: An Organizational Perspective*, Reading, MA: Addison-Wesley.
Keller, C. and L. Cernerud (2002) Student perceptions of e-learning in university education, *Journal of Educational Media*, **27**, 55–67.
McVay-Lynch, M. (2002) *The Online Educator – A Guide to Creating the Virtual Classroom*, London: Routledge Falmer.
Piskurich, G.M. (2004) *Getting the Most from Online Learning – a Learner's Guide*, San Francisco, CA: Pfeiffer.
Rovai, A. (2003) In search of higher persistence rates in distance education online programs, *The Internet and Higher Education*, **6**, 1–16.
Safranek, R.J., S. Gottschlich and A.C. Kak (1990) Evidence accumulation using binary frames of discernment for verification vision, *IEEE Transactions on Robotics and Automation*, **6**, 405–417.
Shafer, G.A. (1976) *Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press.
Van Dam, N. (2001) Where is the future of learning? Blended learning is the ultimate answer, *E-learning*, November, 160.
Voci, E. and K. Young (2001) Blended learning working in a leadership development programme, *Industrial and Commercial Training*, **33**, 157–160.
White, G.P. (1987) A survey of recent management sciences applications in higher administration, *Interfaces*, **17**, 97–108.
Williams, C. (2002) Learning on-line: a review of recent literature in a rapidly expanding field, *Journal of Further and Higher Education*, **26**, 263–272.

The authors

Paul Jones

Paul Jones is a Senior Lecturer in the Department for Enterprise and Economic Development within the University of Glamorgan Business School. He gained a BA Business Studies degree with Swansea Institute of Higher Education and an MSc with Cardiff University. He is currently in the process of submitting his PhD investigating e-business growth with the University of Glamorgan. His research areas, in which he has published over 40 articles, include e-business development, e-learning and enterprise education. He is a member of the British Computer Society, the Institute of Small Business Entrepreneurship, the European Council for Small Business and Entrepreneurship and the UK Association for Information Systems.

Malcolm J. Beynon

Malcolm J. Beynon is a Reader in Cardiff Business School at Cardiff University (UK). He gained his BSc and PhD in pure mathematics and computational mathematics, respectively, at Cardiff University. His research areas include the theory and application of uncertain reasoning methodologies, including rough set theory, Dempster–Shafer theory and fuzzy set theory; also the introduction and development of multicriteria-based decision-making and classification techniques, including the Classification and Ranking Belief Simplex. He has published over 90 research articles. He is a member of the International Operations Research Society and the International Multicriteria Decision Making Society.