ODL-086-2012

ENTROPIC MEASURES FOR STUDENT ENROLMENTS AT UNISA

M Ilunga
Department of Civil and Chemical Engineering
University of South Africa
e-mail: ilungm@unisa.ac.za

Abstract

Entropy originated from physics, was introduced by Boltzman in 1872 and shows a degree of uncertainty or chaos for a system/subsystem. The versatility of the entropy concept was proven through several fields of science: chemical engineering, electrical engineering, metallurgical engineering, communication theory, theoretic entropy, water resources, and student performance in an open distance learning context, etc. In the current study, theoretic entropy is extended to quantify the amount of information/uncertainty of student enrolment associated with graduations and academic staff head counts (HC) at Unisa. Entropic computations revealed preliminarily the degree of uncertainty of enrolment; hence the information contained in enrolments did not vary sensibly from college to college. Similar results were also obtained for graduations as well as staff (HC). Around 16 to 17% degree of uncertainty, ignorance or chaos can be accounted for for enrolments, graduations and staff HC. To reduce the level of uncertainty, a reconfiguration/integration of Unisa as a system in terms of planning of its different subsystems was suggested: registration-enrolment plan, learner support, academic staff, ICT capabilities. Hence a schematic diagram has been derived for system integration. Further work needs to be done for the integration of different elements of Unisa in order to reduce the level of uncertainty in the system.

1. Introduction

Enrolment is an important element in any academic institution, particularly an open distance education institution. The importance of distance education (DE) in the world has been recognised by many scholars (Peters 2010). For instance Unisa and many other DE institutions have played a major role in the global world by giving all classes of people the opportunity to further their studies. The openness of DE institutions led to the industrialisation of institutions of higher learning (Peters 2010), hence massification. Enrolments should go hand in hand with organisational resources to ensure service delivery and learner-centredness. For example, openness in admissions at the Open University of the United Kingdom (OU UK), that is adequate organisational infrastructure (Haughey 2010:50), and enhanced learner success (Subotzky & Prinsloo 2011). Unisa currently handles more than 300 000 students as a mega-university at the same level as OU UK and many others. The Vice-Chancellor and other members of the Unisa Senate have repeatedly raised concerns about the escalating enrolments versus available resources of the institution. This is likely to have an impact on the service delivery and workload of academic staff. Therefore the situation can become chaotic, uncertain or stochastic.
The current study introduces an entropy approach (i.e. theoretic entropy from communication theory) to quantify the degree of uncertainty (disorder, ignorance or

chaos) associated mainly with student enrolments. Since enrolments should target the success of the learners through support from academics, entropic measures will also be expanded to graduations, and academic staff HC. The literature of entropy theory in an open distance learning environment as approached in this paper is almost nonexistent: except that this concept was used for the first time for student numbers per qualification type at the University of South Africa (Ilunga 2007) and subsequently to measure the student performance in the Statistics I module (Ilunga & Muchengetwa 2011). It is argued that the entropy approach offers certain advantages as opposed to the traditional statistical approach (Harmancioglu, Alpaslan & Singh 1994). The current study will preliminarily evaluate the degree of uncertainty associated with enrolments, and subsequently with graduations and staff HC. In the end, a simplified diagram depicting the entropy level of a system will be given. Five colleges will be considered: College of Agriculture & Environmental Sciences (CAES), College of Economic & Management Sciences (CEMS), College of Education (CEDU), College of Human Sciences (CHS), College of Law (CLAW) and College of Science, Engineering & Technology (CSET).

2. Open distance education and entropy

The openness of DE institutions is of interest in a global world. Peters (2010) argued that "[o]pen DE is the most industrialised form of education" as it could be compared to industrial mass production and at the same time reduce costs of service delivery and maintain an acceptable academic standard. Student enrolment versus the available resources (learner support, ICT capabilities, etc.) needs to be managed properly, otherwise service delivery can be jeopardised. This has been raised by the Vice-Chancellor, Prof Makhanya, and Senate members in several meetings. It is believed that the move at Unisa to online teaching and learning may have an impact on service delivery, considering the massive number of students. The same move has been done by other DE institutions, for instance OU UK (Haughey 2010).

Currently, the many students registered at Unisa impacts heavily on the workload of staff, in particular academics. For instance, the turnaround time of assignments easily extends beyond two weeks, unlike the figure suggested by Bates (2011). In principle students want feedback on their assignments as soon as possible.

Due to the complexity introduced by openness in DE institutions, there is a degree of uncertainty associated for instance with massive numbers of students that could not be serviced properly otherwise. It has been mentioned previously that the literature of theoretic entropy in open distance studies is scant, except for a few articles by Ilunga (2007), and Ilunga and Muchengetwa (2011). These two studies concluded that there was a need to reduce the level of uncertainty associated with student numbers per qualification type at the University of South Africa and student performance respectively, by enhancing the learner support system.

It is argued in this paper that if student enrolments still increase uncontrollably without integrating (re-engineering) the institution, the system may degenerate to a status of chaos, uncertainty, ignorance, such that service delivery (Unisa 2005) and student-centredness will be affected. Hence in what follows the theoretical background of theoretic entropy is briefly explained.

3. Theoretic entropy

Although other forms of entropy exist (e.g. thermodynamics), Shannon and Weaver (1949) are the pioneers of theoretic entropy. The current study will deal with the theoretic entropy form, a term borrowed from communication theory. The concept of entropy has a wide range of applications from physics, communication, engineering and other fields and now also to teaching and learning. Hence entropy depicts its versatility. Considering a continuous random variable X, the theoretic entropy is given by the following equation (Singh 1998):

$$H(X) = -\int_a^b f(x) \log f(x) dx \qquad (1)$$

Under the following normality condition

$$\int_a^b f(x) dx = 1 \qquad (2)$$

where $f(x)$ is the probability distribution function of the random variable *x* and $H(x)$ is the marginal entropy of the random variable which is associated with the information content of *x*. Given that entropy is an indication of uncertainty represented by the probability distribution, equation (1) can be written as follows:

$$H(X) = -K \sum_{i=1}^{n} p_i \log_2 p_i \qquad (3)$$

where K is a function of the base used or the scale factor (bits for base 2, Napiers for base e, decibels for base 10). So, this definition holds only numbers of outcomes, which are countable and equal to some integer.

Since the reduction of the uncertainty by means of making observations is equal to the amount of information gained, the entropy criterion indirectly measures the information content of a given series of data (Harmancioglu & Yevjevich 1987). Hence, for independent variables, it is evident that the total uncertainty or chaos of independent variables is the summation of marginal entropy values of each variable (Ilunga & Muchengetwa 2011).

4. Methodology and data availability

To reach the main objective of the study, i.e. evaluating the degree of uncertainty and indirectly the information contained in the student enrolment, the methodology is explained below.

External factors that might impact on student enrolment are ignored and hence are not taken into consideration. The current study assumes enrolment (yearly) per college as occurring at random in an open DE context. Hence student enrolment per college is considered to be an independent variable from any other college. The abovementioned variable reflects some randomness in its occurrence. The uncertainty associated with each variable, i.e. enrolment at each college, is calculated by using equation (2) as it is regarded as a discrete variable rather than a continuous variable. Frequency of each variable is approximated to be equal to its probability of occurrence. The total entropy associated with all colleges is computed by summing up uncertainties related to each college. The above procedure could be extended to student graduations as well as staff HC. Similar to that presented by Ilunga and Muchengetwa (2011), a simplified diagram will finally be presented to show the states of system: from highest entropy to lowest entropy. For the sustainability of the system, change in entropy should be negative or decreasing.

Data used were obtained mainly from the report by Van Zyl and Barnes (2012): student enrolments (HC), student graduations by college and staff (HC) by college. Staff are limited to academic staff. The following colleges will be considered: CAES, CEDU, CHS, CLAW and CSET.

5. Results and discussion

The results presented in Tables 1, 2 and 3 show the frequency, probability of occurrence and marginal entropy values for the following variables: student enrolments (HC), student graduations by college and staff (HC) by college. Frequency values for each variable are assumed to be probability values.

For enrolment at each college, it can be observed that the entropy values vary between 0.67 and 0.70 decibels (Dcb). This suggests that all colleges are affected more or less in the same way by the numbers of students enrolled. The total entropy (total amount of uncertainty) associated with student enrolment for all colleges was calculated as 4.18 Dcb. In the last column of Table 1, it is shown that the relative entropy of enrolment for each college is the ratio between the marginal entropy for student enrolment at each college divided by the total entropy of all colleges. Values calculated were very close and varied between 16 and 17%. This could be explained again by the fact that since colleges have similar problems due to student enrolments, there is a need to reduce this proportion of uncertainty by improving the overall system. For instance by looking at integration of different subsystems within Unisa, ICT support, finances, human resources, learner support and many others to reach a system with fewer problems. On the other hand, a move to a new mode of delivery (e.g. online teaching and learning) should be done cautiously and in such a way to ensure that new problems are not introduced. In Table 2, student graduations show the same characteristics as student enrolments in terms of entropic calculations, which suggests that the uncertainty in enrolments may lead to uncertainty in graduations. In other words, if enrolments are not properly managed with adequate resources, the system is likely to reflect complexity, degree of chaos or disorder. Hence entropy will increase. From Table 3, staff (HC) depict similar characteristics to student enrolment, in terms of entropy calculations. Enrolments are very likely to affect staff as well: heavy workload, reduced responsiveness, etc. Hence issues associated with the enrolment plan due to openness can affect graduations and staff, as well as less uniformly all the colleges across the institution. A simplified diagram depicts the change in entropy of the system (Figure 1). The system relates to the variables under study. Since $H2<H1$ (i.e. system 2), the change in entropy from system 1 to system 2 is negative. In other words, the reduction in entropy of the system corresponds to the information inferred about the system (Ilunga & Muchengetwa 2011). Through system integration (re-engineering), the move from 1 to 2 is possible. Hence the uncertainty can be reduced. In case $H1<H2$, the system 1 is associated with higher uncertainty due to open enrolment that may impact on graduations and the workload for academics.

6. Conclusion and suggestions

This study showed again that entropy is a versatile tool. In this case, it quantifies uncertainty/information contained in enrolments, graduations and staff HC. Entropic computations revealed preliminarily that the degree of uncertainty of enrolment, hence the information contained in enrolments, did not vary sensibly from college to college. Similar results were also obtained for graduations and staff HC. Due to openness, enrolment is likely to affect graduation and staff workload. Around 16 to 17% degree of uncertainty, ignorance or chaos can be accounted for for enrolments, graduations and staff HC. To reduce the level of uncertainty, it is suggested that there be an integration (re-engineering) of Unisa as a system: registration-enrolment plan, learner support, academic staff, ICT capabilities, etc. Hence a schematic diagram has been derived for system integration. Further work needs to be done for the integration of different elements and resources of Unisa in order to reduce the level of uncertainty in the system.

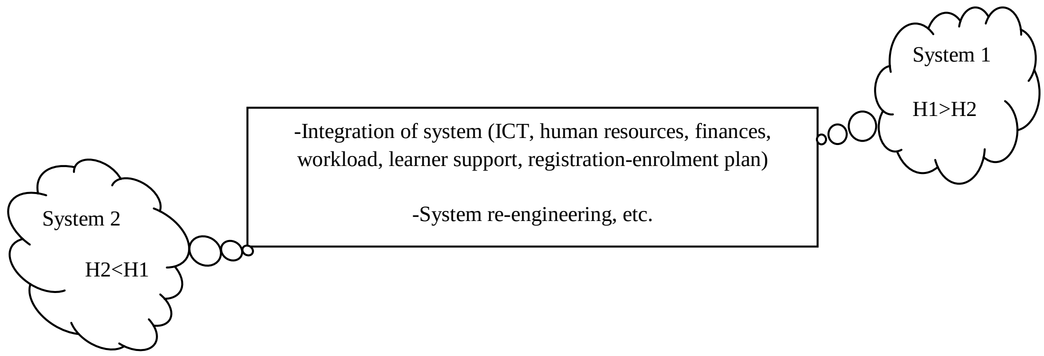

Figure 1: Simplified diagram showing the change in entropy of the system

Table 1: Entropy calculations for student enrolments

	Marginal entropy (uncertainty per college) in decibels						Relative entropy (%)
	2006	2007	2008	2009	2010		
CAES	0.12	0.13	0.14	0.14	0.15	0.69	17
CEDU	0.12	0.12	0.14	0.15	0.15	0.68	16
CEMS	0.13	0.14	0.14	0.14	0.14	0.70	17
CHS	0.13	0.13	0.14	0.14	0.15	0.70	17
CLAW	0.14	0.14	0.14	0.14	0.14	0.70	17
CSET	0.13	0.14	0.14	0.14	0.15	0.70	17
Total entropy						4.16	

Table 2: Entropy calculations for student graduations

	Marginal entropy (uncertainty per college) in decibels						Relative entropy (%)
	2006	2007	2008	2009	2010		
CAES	0.10	0.12	0.14	0.15	0.15	0.67	16
CEDU	0.12	0.11	0.14	0.15	0.16	0.67	16
CEMS	0.12	0.13	0.14	0.14	0.15	0.69	17
CHS	0.13	0.13	0.14	0.15	0.15	0.69	17
CLAW	0.09	0.14	0.14	0.15	0.15	0.67	16
CSET	0.09	0.14	0.15	0.14	0.15	0.68	17
Total entropy						4.08	

Table 3: Entropy calculations for student graduations

	Marginal entropy (uncertainty per college) in decibels						Relative entropy (%)
	2006	2007	2008	2009	2010		
CAES	0.12	0.14	0.14	0.15	0.15	0.69	17
CEDU	0.14	0.14	0.14	0.14	0.14	0.70	17
CEMS	0.13	0.14	0.14	0.14	0.15	0.70	17
CHS	0.14	0.14	0.14	0.14	0.14	0.70	17
CLAW	0.14	0.14	0.14	0.14	0.14	0.70	17
CSET	0.14	0.14	0.14	0.14	0.14	0.70	17
Total entropy						4.18	

References

Bates, AW. (2011, November) *The second wave of distance education and history of the Open University United Kingdom* [online video]. http://vimeo.com/32292234 (Transcript: http://www.box.com/s/cvygk4334sub0i6atrn8)

Harmancioglu, NB & Yevjevich, V. (1987) Transfer of hydrologic information among rivers points. *Journal of Hydrology* 93: 103–118.

Harmancioglu, NB, Alpaslan, N & Singh, VP. (1994) Assessment of the entropy principle as applied to water quality monitoring network design, in Hipel, KW et al. (eds), *Stochastic and statistical methods in hydrology and environmental engineering* 3: 135–148. The Netherlands: Kluwer Academic Publishers.

Haughey, M. (2010) Teaching and learning in distance education before the digital age. In MF Cleveland-Innes & DR Garrison, *An introduction to distance education: understanding teaching and learning in a new era* (pp. 46–66). New York and London: Routledge.

Ilunga, M. (2007) Entropic measures for student numbers per qualification type at the University of South Africa (Unisa). Poster presentation, HELTASA Conference 2007, Bloemfontein, South Africa.

Ilunga, M & Muchengetwa, S. (2011) Towards using e-portfolios in an ODL development state. *Progressio, South African Journal for Open and Distance Learning Practice*, Under review.

Oldenburg, Germany: BIS-Verlag der Carl von Ossietzky Universität Oldenburg. http://www.box.com/shared/ktx7ipccetotqrr11mct

Peters, O. (2010) The theory of the "most industrialized education", in O. Peters *Distance education in transition: developments and issues.* 5th edition. pp. 15–19.

Shannon, CE & Weaver, W. (1949) *The mathematical theory of communication*. Urbana, Chicago, London: Univeristy of Illinois Press.

Singh, VP. (1998) Entropy as a decision tool in environmental and water resources, *The Indian Association of Hydrologists*, *Hydrology Journal* 21(4): 1–12.

Singh, VP. (1998) Entropy-based modeling in hydrology and water resources. Water resources Program, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, USA.

Subotzky, G & Prinsloo, P. (2011) Turning the tide: a socio-critical model and framework for improving student success in open distance learning at the University of South Africa. *Distance Education* 32(2): 177–193.

Unisa. (2005) 2015 strategic plan: an agenda for transformation. http://staff .unisa.ac.za/cmsys/staff/docs//unisa_2015_strategicplan_nov_final.pdf

Van Zyl, D & Barnes, D. (2012) *An institutional profile of Unisa: Unisa facts & figures.* Department of Institutional Statistics & Analysis. http://staff.unisa.ac.za.