Information Theory: Learning Through Example

Alexander R. May, *Student Member, IEEE*

Optoelectronics Research Centre
University of Southampton
Southampton, United Kingdom
arm103@orc.soton.ac.uk

Rallis C. Papademetriou, *Member, IEEE*

School of Engineering
University of Portsmouth
Portsmouth, United Kingdom
rallis.papademetriou@port.ac.uk

Abstract—In this paper, the perceived difficulty for students learning purely through auditory and visual means, such as via the standard lecture format, is highlighted through a literature review of learning styles and methods, and a solution based upon interactive computer case-based learning through worksheets is proposed. A brief history of Information Theory is discussed and a review of the various software packages available is made and Mathematica selected. The contents of the worksheets are described and conclusions made.

Index Terms—Information Theory, Education, Mathematica, Worksheets

I. INTRODUCTION

THIS paper describes the development of a student-centered teaching and learning methodology within the field of Information Theory, with particular emphasis on Information Channels. It is important that learners acquire knowledge and experience in the contextualized use of Information Channels since these have a broad specification. It is essential that the learner appreciates the fact that Information Channels encompass more than just the physical medium through which communication is made, but also the nature of the flow of Information and the factors affecting it, such as the kind of transmission signal used and the receiver.

Existing textbooks have a tendency to discuss the properties of Information Channels, but without the required contextualization with the result that whilst the student may understand the properties of the channels, they have little idea how they might be applied in a real situation. Furthermore, theories to do with different styles of learning are very well rehearsed in this day and age and the development of interactive packages to assist students with their learning is highly relevant when these are considered. As described by [1], it is important to keep in mind that learners broadly learn via an auditory, visual or kinesthetic style of learning. Whilst it is the case that most learners may use an amalgamation of all three styles, there is nevertheless much evidence to suggest that learners do tend to have a preferred style of learning.

Traditional lectures may capture the interest of an auditory-based learner, albeit with some limitations based upon the length of time an individual can concentrate. The addition of graphics and visual aids such as 'Powerpoint' can be used to help assist the visually stimulated learner. However, it is generally the case that unless an activity is very practical and/or laboratory based, many learning sessions do not provide the kinesthetic learner with many opportunities.

This report seeks to address the problem learners have in accessing, understanding and using material in the field of Information Theory by developing a learner-centered, interactive, computer-based package. The interactive package described here integrates different styles of learning and gives the student every opportunity to progress at his or her own rate. Allowing students to proceed at their own rate is highly important as argued by [2]. Learners tend to acquire knowledge and confidence in material at different rates and the interactive package described in this report fully recognizes this important fact.

II. A BRIEF REVIEW OF INFORMATION THEORY

October 1998 was the fiftieth anniversary of the birth of Information Theory as we know it. This was commemorated in the IEEE Transactions of Information Theory with an article by [3] providing a brief description of the historical development of the field. Here we follow that author in providing an introduction to the concept of 'Information' and, in particular, the concept of an 'Information Channel'.

In October 1948 Shannon published a remarkable paper [4] that for the first time formalized the fundamental laws of data compression and transmission using the language of probability theory. Prior to this [5] had discussed the need to introduce a measure of Information content in order to determine whether a method of transmission of data was feasible or not over a 'channel'. It was realized from observations of RLC circuits that the capacity of a channel was in some way proportional to the bandwidth of the channel, but this was not investigated in a probabilistic manner.

When Shannon was an undergraduate he came across the paper by [5] and soon developed an abstraction of a communication system in terms of a mean-square fidelity criterion. After completing a Master's Thesis on the application of Boolean Algebra to switching circuits, he switched fields to Mathematics partly on the advice of his supervisor [6] and began a PhD in the field of population dynamics. During this work he continued to research into

Communication Systems in his spare time and, later, on return to Bell Laboratories he published his monumental work [4]. Here Shannon introduced formally the concept of 'Information' and a way of quantifying what a 'message' is. Prior to this it was understood how to transmit, receive and process a given waveform, but there was no understanding of how to turn a 'message' into a waveform.

Shannon began by explaining based upon previous work [5] that a 'message' should be thought of as a choice between alternatives with this number of alternatives increasing exponentially with the duration of the transmission from a message source. As a result the idea of Information loosely as a logarithm appeared. A fundamental idea was that it is the choice between alternatives that was important, not the representation of the choice. It was at this point that he introduced the concept of the 'channel' as a device which accepts a sequence of letters at its input and maps this to a noise-corrupted version of these letters at the output. Vital to this development was the characterization of a channel completely by its capacity, and the realization that provided the information rate of a source is less than the channel capacity, the data may be transmitted with arbitrarily small error though a choice of appropriate processing.

III. A REVIEW OF LEARNING STYLES

The use of interactive methods of teaching and learning in the field of undergraduate mathematics is not new. As the availability of computers became more common in the mid-20th century, the field of mathematics began to embrace them in its teaching and learning. The studies of [7] and [8] predicted that computers would become increasingly integrates into the classroom. However, it has been argued by authors such as [9] that the incorporation of computers has been relatively slow and a variety of reasons for this have been put forward. Recent studies such as by [10] suggest that it is the beliefs and attitudes of teachers, as well as cultural attitudes, that have most influenced integration. It is certainly true to say that the younger generations have much more familiarity with and confidence in the use of Information Technology in their learning than do older people.

Anecdotal evidence suggests that most university mathematics departments do utilize some form of computer algebra system (CAS) in their courses. A variety of studies, such as those by [11] and [12] have examined teaching practices using CAS. Looking beyond the field of mathematics and into that of Information Theory, it is our belief that interactive and contextualized teaching and learning is the way forward.

There are many ways in which a review of interactive computer-based learning and teaching solutions could be started, but as a way of indicating quite how long the introduction of computers in learning has been considered, it was felt that starting with the work of [13] would be appropriate. In that work, it was argued that by the year 2000 the majority of teaching at all levels would be performed through the interactive use of computers. However, it is also argued that teachers tend to be the highly successful products of the existing educational system and therefore start with preconceived views of how education should happen. These may then create a barrier for the learners.

Throughout human history education in communities was, and still is, in the case of many communities in the developing world, achieved through individuals' interaction with the environment. This is a key point since formal education based on 'learning situations', classrooms and similar, is indeed a relatively new development in human history. Natural learning was an active process with the child or adult engaging in learning through necessity for survival, the use of games, or for fun. Each individual would have different and unique experiences of interacting with their environment. In addition, this individual experience allowed for learning at a pace that felt comfortable to the individual.

It was only in the Greek times that a new form of learning began in the form of Socratic dialogues such as seen in the works of Plato. It was key here that the student learnt through exercising thought and responding to carefully formulated questions provided by the teacher. Although Plato did not totally subscribe to the idea himself, it was during this period and the idea of the lecture as a method of educational delivery began. Somewhat later the development of the textbook also began, although it is important to note that this occurred almost 200 years after the development of the printing press itself.

These relatively recent innovations in education were not necessarily due to their superiority but purely as a result of the need to education and increasingly large number of people. While the idea of one-to-one tuition is still highly respected, and can still be found in education establishments such as Oxford and Cambridge, its use is not generally feasible considering the millions of students worldwide in schools and universities.

However, in the Sciences, in particular, it is posited that emphasis should be made not just on the learning of facts, but on the methodologies involved in obtaining them. There would appear to be reasonable evidence to suggest that the lecture conveys information relatively poorly. The idea of simply giving the student the solution to a problem in a magical fashion does not appear to achieve very much since when the time comes for students to solve a problem of their own, they simply do not know how to go about the process. They do not possess skills to transfer knowledge and practice from one situation to another. It is therefore essential that some form of 'play', or interaction, is necessary such that the student actually understands how to solve problems him or herself.

However, in their experiences of teaching electrostatics through the use of interactive experiments it was noted by [13] and colleagues that it was only the brighter students who made the most use of these particular experiences. Many other students had better success when there was at least some form

of guidance or structure to the learning experience such as by the provision of additional written material or a set of questions for students to demonstrate through a quiz that they understood the material.

Moving more to the present day, [14] discusses interactive learning systems with respect to modern theories of education. One of the first attempts at describing interactive learning was that put forward by [15] consisting of the computer as a tutor, tool and tutee. As a tutor a computer can instruct a learner through predefined interactive lessons. As a tool it helps educate though the use of tools such as word processors, spreadsheets or database programs. In the case of the computer acting as a tutee, the student teaches the computer to perform some task through the use of a user friendly computer language. A contemporary update of this theory has been developed by [16] and consists of five different teaching architectures,

(1) SIMULATION-BASED LEARNING BY DOING

(2) INCIDENTAL LEARNING

(3) LEARNING BY REFLECTION

(4) CASE-BASED TEACHING

(5) LEARNING BY EXPLORING

In this work, focus is made on the 1st and 4th of these teaching architectures. [16] argue that Simulation-Based Learning by Doing is critical when the subject matter is experiential at heart. Emphasis is made on learning by doing and preparing the student for the challenges that might be faced with real world problems. Case-Based Teaching is described as being a useful addition to that of learning by doing – sometimes students will realize that in order to progress in their 'doing', they need further knowledge. This can be achieved through the use of Case Studies. While isolated facts may be difficult to remember, if knowledge is presented in the form of a story or particular case, it is generally integrated better into memory.

IV. A REVIEW OF COMPUTER ALGEBRA SYSTEMS

So, moving on to reviewing the more interactive ways of teaching and learning, in the case of CAS (Computer Algebra Systems) in the educational environment, the choice of Computer Algebra System packages generally falls between three products – MAPLE [17], Mathematica [18] and MATLAB [19]. As a result of this, [20] published an article reviewing the three packages from the perspective of teaching, the ease of use of a CAS package being a key issue.

Each of these packages has a different user interface that finds itself suited to a particular type of user. Of these packages MAPLE is the only one that was rooted in

Education, the other two being developed for Scientific Computing and Engineering respectively and it is important to discuss the characteristics of each.

MAPLE has a history of being used in the teaching of Calculus and this author has personal experience of it. In a similar way Mathematica has also been used for similar purposes, and in both cases using each package gives access to a considerable community of educators. It is believed that the main reason MAPLE and Mathematica have been used with such success in teaching Calculus is because of their interactive worksheets. In each case, learning is achieved by following a teaching narrative with linked symbolic algebra or calculations for the student to 'play with', to interact with. By varying parameters as they go along, they develop an understanding of the concepts. However, it was found that for the purposes of this project there was more scope for use and more interactivity with Mathematica when compared with MAPLE.

While MATLAB has become significant in Engineering, it is felt to be less suited to education because of its simple command line interface and lack of a worksheet. While this straightforward interface is a powerful benefit for a practicing Engineer or Scientist, it does not allow for easy manipulation by the student as will be discussed further below.

The next step in determining what form a set of interactive notes and tutorials will take comes in identifying what previous work has been done in the field of Information Theory, with comparisons between MATLAB and Mathematica.

One such MATLAB example which this particular author purchased previously is that by [21]. While the content of the book is exemplary, the presentation of the interactive MATLAB code is felt to be less than optimal. MATLAB code, some of it of considerable length, is simply listed for the student to copy as necessary. The result of this is that the student spends much time copying the code rather than actually understanding its contents. This is felt to be a considerable disadvantage. An example of this is illustrated in Fig 1 below. This illustrates the disadvantage of MATLAB in an educational setting – code is not truly interactive as in the case of a worksheet.

```
function [sqnr,a_quan,code]=u_pcm(a,n)
%U_PCM      uniform PCM encoding of a sequence
%           [SQNR,A_QUAN,CODE]=U_PCM(A,N)
%           a=input sequence.
%           n=number of quantization levels (even).
%           sqnr=output SQNR (in dB).
%       a_quan=quantized output before encoding.
%       code=the encoded output.

amax=max(abs(a));
a_quan=a/amax;
b_quan=a_quan;
d=2/n;
q=d.*[0:n-1];
q=q-((n-1)/2)*d;
for i=1:n
  a_quan(find((q(i)-d/2 <= a_quan) & (a_quan <= q(i)+d/2)))=...
  q(i).*ones(1,length(find((q(i)-d/2 <= a_quan) & (a_quan <= q(i)+d/2))));
  b_quan(find( a_quan==q(i) ))=(i-1).*ones(1,length(find( a_quan==q(i) )));
end
a_quan=a_quan*amax;
nu=ceil(log2(n));
code=zeros(length(a),nu);
for i=1:length(a)
  for j=nu:-1:0
    if ( fix(b_quan(i)/(2^j)) == 1)
    code(i,(nu-j)) = 1;
    b_quan(i) = b_quan(i) - 2^j;
    end
  end
end
sqnr=20*log10(norm(a)/norm(a-a_quan));
```

Fig. 1. An example of MATLAB code presented in [21]

After a review of the relevant literature, the work of [22] was found. The beauty of this particular work is that it is presented in the form of a clickable presentation with a narrative and well presented graphical illustration. This is shown in Fig 2.
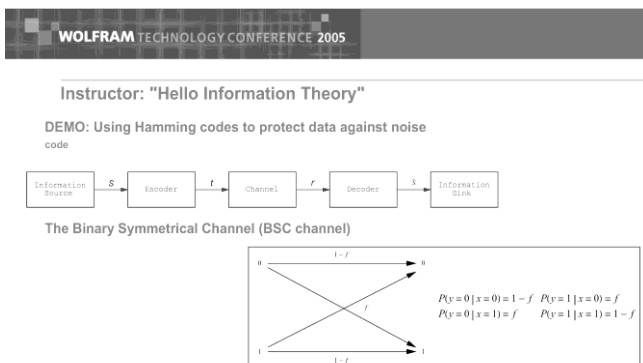

Fig. 2. The layout of the Mathematica worksheet from [22]

As can be seen from Fig 2, Mathematica allows for a nice presentation style, with graphical representations formed using hidden code that can, if necessary, be viewed by clicking on 'code'. This allows for student manipulation should he or she want to create a worksheet of his or her own. As an illustration of the effect of channel noise, the author uses comic book images and subjects them to transmission over a Binary Symmetric Channel. At the same time as presenting the result of such a transmission, the code that was used to perform the operation is shown to the student. This is shown in Fig 3.

Fig. 3. An illustration of the transmission of an image over a Binary Symmetric Channel, with necessary code shown above. From [22]

The above worksheets shown in Fig 2 and 3 are in stark contrast to the typically available textbooks on the topic of Signal Processing and Information Theory.

V. THE PACKAGE AIMS

As indicated previously, the aim of this project was to expand upon existing methodologies for the teaching and learning of Information Theory at the University of Portsmouth. While this field is traditionally 'feared' by students for its mathematical content, it was our belief that this fear is unfounded. It is the authors' experience that the use of Computer Algebra Systems has been highly beneficial to the teaching of university mathematics in general, from the basics of calculus to the manipulation of tensors in General Relativity.

Following on from these experiences, it is believed that Information Theory, too, can be made more accessible to the student. Consequently, the project aimed to develop a series of student-centered, interactive, computer-based tutorials to support learning in the field of Information Theory. These tutorials are accompanied by a clear set of notes and instructions regarding the use of the online material. The expected outcome was that learners would become much more familiar with Information Channels and how they operate, and gain more confidence in applying this knowledge to real life situations.

VI. THE PACKAGE ACHIEVEMENTS

While [22] approached their worksheets from the point of view of a 'research orientated activity' and Mathematica coding exercise, it is clear from their student feedback that this was considered a significant challenge. In addition, it is our view that the range of material, which even includes coverage of Cellular and Graph Automata, could be off-putting from the start. The package described here was specifically designed with clarity and simplicity of layout in mind, since increasing students' confidence, and thereby their knowledge and understanding, was of prime importance. As has been discussed in [23], too much material can overwhelm students, and our aim was to increase hands on experience of existing lecture notes which have been developed and optimized through years of experience in the Higher Education environment.

To summarize their content, the worksheets began with an introduction to the concept of self information as being a quantifiable measure of the unexpectedness of an event (see Fig 4). This was then reinforced by considering examples of unexpectedness and surprisal, such as with the lottery, where a lottery win with its low probability of occurrence is associated with a high amount of unexpectedness or surprisal. The concept of entropy associated with a random variable was then

The choice of base a gives rise to various units:

for a=2 → bits (binary units)
for a=e → nats (natural units)
for a=10 → hartleys (or decits)

$\log_a x = \log_b x \times \log_a b \rightarrow 1$ hartley $= 3.32$ bits, 1 nat $= 1.44$ bits

A particularly interesting property of base 2 for those familiar with BITS is that if $P(E) = \frac{1}{2} \rightarrow I(E) = 1$ bit. That is, ONE BIT is the amount of information we obtain when one of the TWO possible EQUALLY LIKELY alternatives is specified (e.g flipping a coin or examining the output of a binary communication system)

Formalising our intial 'hand waving' description of information being a way of quantifying unexpectedness, we may say better say that:

Interpretation: I(E) can be interpreted as:

measure of how unexpected the event was

OR

measure of the information yieled by the event

We now consolidate this with some examples. At each stage please TAKE PART in the problem by evaluating the answers using the keys SHIFT AND ENTER

Fig. 4. An illustration of the format of the worksheets produced during this project

introduced as being an average measure of unexpectedness for such a variable. Entropy was then interpreted as being the amount of information gained upon knowledge of the outcome of an experiment. Once the above concepts were introduced more formal derivations took place where properties of such a measure were discussed.

Having introduced the concept of information associated with a random variable, the concept of mutual information was then introduced, though the use of a generalized information channel, as being a measure of how much information one random variable gives about another through the use of conditional probabilities. Given an initial uncertainty in a random variable such as the input of a channel, and a final uncertainty upon observation of the channel output, the mutual information was described as being a gain in Information obtained in this process. If entropy can be considered the average self-information of a random variable, then as a consequence, average mutual information was introduced. Once again, with the concepts introduced, more formal derivations took place where properties such as symmetry were discussed.

Given the use of conditional probabilities in the introduction of mutual information, the concept of conditional and joint entropies were then discussed and useful relationships obtained, such as the definition of mutual information in terms of the entropy of a source and a conditional entropy.

The general concept of an information channel was then introduced and its position between an information source and information destination illustrated. With conditional

probabilities being fundamental to the discussion of information channels, a brief review of such a representation with matrices and associated properties was then made.

Nine basic channels were then introduced and their capacities derived. These channels consisted of the,

(1) LOSSLESS CHANNEL

(2) DETERMINISTIC CHANNEL

(3) IDEAL CHANNEL

(4) UNIFORM CHANNEL

(5) BINARY SYMMETRIC CHANNEL

(6) BINARY ERASURE CHANNEL

(7) USELESS CHANNEL

(8) GENERAL BINARY CHANNEL

(9) PACKET ERASURE CHANNEL

Of these channels the General Binary Channel was considered to be the most complex requiring a detailed derivation of properties, but emphasis was made through the use of examples of determining the capacity of such a channel using the 'method of curves' as a graphical method.

With the above set of basic channels introduced, the concept of a cascade of channels was then discussed and properties determined. Examples were used to consolidate learning.

Entropy was extended to the case of a continuous random variable where differential entropy was introduced as a relative measure. Probability density functions which result in entropy maximizations were then determined under constraints upon a random variable. In each case the student was asked to determine the probability density function for a given case and to determine its entropy. The fact that the Gaussian distribution had the highest possible entropy was then noted.

The concept of optimal decision making was discussed with respect to backward and forward probabilities. The ideal observer and the maximum likelihood decision schemes were introduced and examples studied.

Finally a brief discussion of more advanced channels was made. The Additive White Gaussian Noise (AWGN) channel was discussed and the Shannon Capacity derived from an Information Theoretic point of view. The capacity of such a channel with a given signal-to-noise ratio was then shown interactively to the student. The Gaussian Multiple Access Channel and Broadcast channels were then discussed and the

Shannon Capacity relationships used to introduce the student to the more complex Network Information Theory.

VII. DISCUSSION

Having completed the worksheets it was felt that a step forward had been made in giving students a 'Simulation-Based Learning by Doing' and 'Case-Based Teaching' opportunity to experience the topic of Information Theory as in the learning architectures by [16].

Throughout the worksheets emphasis was made to highlight important points in red text. Key points were also emphasized through the use of 'fact boxes' which summarize properties. Both are shown in Fig 4. While formal derivations were performed, practical examples were provided for the student to interact with.

It would be possible, potentially advisable, to further the use of this package, for more examples to be developed, some of which could require students to complete them wholly on their own to demonstrate a solid understanding of the material.

This method of providing tutorial examples and then detailed solutions the following week is standard practice in the course on Information Theory at the University of Portsmouth and it is felt that this works well. It is often the case that students only truly understand material when they are able successfully to answer questions with little in the way of notes. This author, a student himself, has often found that it is only when you are forced to use formulae in answering questions, such as in the case of an exam paper, that you find out whether you are able to use them, let alone have an understanding. One particular example of this is in the determination of the capacity of a Uniform Channel where the result is independent of the row index of the channel matrix. Students may well remember the capacity relationship from memory, but when it comes time to apply it, they may suddenly realize that they are unsure of which row index to use, when in reality it does not matter.

It is felt that Mathematica as a package has allowed for the writing of lectures notes and tutorial examples with relative ease while also supplying a sophisticated computational backend, which this study has only touched on, due to the beauty of the simply closed form solutions involved in this material. There is also still room for expansion of the notes to utilize the more sophisticated graphical user interface, such a through the use of 'sliders' and interactive plotting.

However, from the point of view of a 'first iteration', it is felt that the project has been a success. The authors have been able to implement their notes into Mathematica and develop additional tutorial problems as well as provide a small expansion into the topic of Network Information Theory. The synopsis in plain language of the background to Information Theory and the explanation of the different types of channels was written with actual students on an MSc program in mind, since these students are not only dealing with a potentially complex field but may be doing so in a language other than their own mother tongue. It is therefore hoped that this project will be of some practical use to such learners.

It would be interesting to consult with other Mathematica worksheet authors, even from different topics, to see what improvements could be made given the vast array of such worksheets on the internet.

Given the availability of similar course notes on the topic of Signal Processing at the University of Portsmouth by the same author, it is recommended that not only further tutorial examples on the topic of Information Theory be provided and implemented into Mathematica, but also that these former notes be considered for interactive learning also.

REFERENCES

[1] B. Prashnig, *The power of diversity: new ways of learning and teaching through Learning Styles*. Stafford: Network Educational Press Ltd, 2004.

[2] P. D. Ely, Plomp, T., *Classic writings on instructional technology* vol. 1. Englewood, CO, USA: Libraries Unlimited, Inc, 1996.

[3] S. Verdu, "Fifty Years of Shannon Theory," *IEEE Transactions on Information Theory,* vol. 44, pp. 2057-2078, 1998.

[4] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.,* vol. 27, pp. 379-423, 623-656, 1948.

[5] R. V. L. Hartley, "Transmission of Information," *Bell Syst. Tech. J.,* pp. 535-563, 1928.

[6] R. G. Gallager, "Claude E. Shannon: A Retrospective on His Life, Work, and Impact," *IEEE Transactions on Information Theory,* vol. 47, pp. 2681-2695, 2001.

[7] J. J. Kaput, *Technology and mathematics education*, 1992.

[8] L. A. Steen, "The science of patterns," *Science,* vol. 240, pp. 611-616, 1988.

[9] L. Cuban, Kirkpatrick, H., and Peck, C., "High access and low use of technologies in high school classrooms: Explaining the apparent paradox," *American Educational Research Journal* vol. 38, pp. 813-834, 2001.

[10] S. Hennessy, Ruthven, K. and Brindley, S. , "Teacher perspectives on integrating ICT into subject teaching: commitment, constraints, caution and change," *Journal of Curriculum Studies,* vol. 37, pp. 155-192, 2005.

[11] A. Franco, Franco, P., Garcia, A., Garcia, F., Gonzalez, F.J., Hoya, S., et al., "Learning Calculus of Several Variables with New Technologies," *The International Journal of Computer Algebra in Mathematics Education* vol. 7, pp. 295-309, 2000.

[12] M. Thomas, and Holton, D. , *Technology as a tool for teaching undergraduate mathematics* vol. 1: Fluwer Academic Publishers, 2003.

[13] A. Bork, "Interactive Learning," *American Journal of Physics,* vol. 47, 1979.

[14] T. C. Reeves, and Hedberg, J.G., *Interactive learning systems evaluation*. New Jersey, USA: Educational Technology Publications, 2003.

[15] R. Taylor, *The Computer in the School: Tutor, Tool, Tutee*. New York: Teachers College Press, 1980.

[16] R. C. Schank, and Cleary, C., *Engines for education*. New Jersey: Lawrence Erlbaum Associates, Inc, 1995.

[17] Maplesoft. (2011, 27/02/2011). *Math Software for Engineers, Educators and Students*. Available: http://www.maplesoft.com

[18] Wolfram. (2011, 27/02/2011). *Wolfram Research: Mathematica, Technical and Scientific Software*. Available: http://www.wolfram.com

[19]   Mathworks. (2011, 27/02/2011). *Mathworks - MATLAB and Simulink for Technical Computing*. Available: http://www.mathworks.com

[20]   N. Chonacky, and Winch, D., "3Ms for Instruction: Reviews of Maple, Mathematica, and Matlab," *IEEE Journal of Computing in Science & Engineering,* vol. 7, pp. 7-13, 2005.

[21]   J. G. Proakis, Salehi, M., and Bauch, G., *Contemporary communication systems using MATLAB and Simulink*. Mason, Ohio, USA: Cengage Learning, 2004.

[22]   Y. Ben-Shimol, "Teaching Information Theory with Mathematica," presented at the 2005 Wolfram Technology Conference, Champaign IL, USA, 2005.

[23]   P. Ramsden, *Learning to teach in Higher Education*. London: RoutledgeFalmer, 2003.