PUMS99:1

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS@

JUDUL: APPLICATION OF MATHEMATICS IN EAN 13 AND EAN 8 WITH PROGRAMMING

IJAZAH: SARJANA MUDA SAINS DENGAN KEPUJIAN (MATEMATIK DENGAN EKONOMI)

SAYA SHALINI A/P NAGARATNAM (HURUF BESAR) SESI PENGAJIAN: 2003/2004

Mengaku membenarkan tesis (~~LPSM~~/Sarjana/~~Doktor Falsafah~~) ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis adalah hakmilik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institutsi pengajian tinggi.
4. Sila tandakan (/)

[ ]	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau Kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
[✓]	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
[ ]	TIDAK TERHAD	

Disahkan Oleh

Shalini

(TANDATANGAN PENULIS)

(TANDATANGAN PUSTAKAWAN)

Alamat Tetap: NO. 23, JALAN ERULING 40, KAWASAN 3, TAMAN KLANG JAYA, 41200 KLANG.

Tarikh: 2/5/06

MR. TIONG KUNG MING

Nama Penyelia

Tarikh: 2/5/06

CATATAN:- *Potong yang tidak berkenaan.

**Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa /organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT dan TERHAD.

@Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan atau disertai bagi pengajian secara kerja kursus dan Laporan Projek Sarjana Muda (LPSM).

APPLICATION OF MATHEMATICS IN EAN 13 AND EAN 8 WITH PROGRAMMING

SHALINI A/P NAGARATNAM

THIS DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF SCIENCE WITH HONOURS

MATHEMATICS WITH ECONOMICS PROGRAM
SCHOOL OF SCIENCE AND TECHNOLOGY
UNIVERSITI MALAYSIA SABAH

April 2006

DECLARATION

I declare that this thesis is the result of my work, except the quotations and summaries, each of which the source has been mentioned.

25th April 2006

SHALINI A/P NAGARATNAM

HS2003-3121

CONFIRMATION

Signature

1. **Supervisor**

    Mr. Tiong Kung Ming ____________________

2. **Examiner 1**

    Dr. Jumat Sulaiman ____________________

3. **Examiner 2**

    Cik Suzelawati Zenian ____________________

4. **Dean**

    Supt. Prof. Madya Dr. Shariff A. K. Omang ____________________

ACKNOWLEDGMENT

Special acknowledgment is heartily granted to my supervisor Mr. Tiong Kung Ming, who has given me guidance and suggestions from the preliminary until the end of my dissertation. I would like to convey my appreciation to my parents, Mr.&Mrs. Nagaratnam, who has been supporting and motivating me through these times. My gratitude also goes to Mr. Anandraj Theseira and Mr. Rajan Supramaniam for a great deal of help, guidance and moral support. Finally, I thank Mr. Lim Keng Poh for patiently assisting me with my C language programming.

ABSTRACT

EAN 8 and EAN 13 are the most common barcodes used in retail products. EAN 8 consists 8 digits while EAN 13 consists 13 digits. The last digit for both is known as the check digit. A check digit serves to verify the order of the preceding digits or the validity of a barcode which is attained through mathematical calculations using mod 10 and weight (1, 3). The understanding of Congruence Modulo *n* and Modular Arithmetic is required. After validating EAN 8 and EAN 13, using binary digits, the symbol of EAN 13 and EAN 8 is formed. The understanding of probability, permutation, binary and base two system are required to understand the formation of the symbol. To assist the validity check of EAN 8 and EAN 13, a C-language program was encoded. Another C-language program was developed to assist the detection of the possible errors in the barcode string of digits. In conclusion, Mod 10 is the most suitable mod for attaining check digit. However using the weight (1, 3) does not allow the detection of a Jump Transposition Error. For further research, analysis on different weight for EAN 13 and EAN 8 can be conducted as well as improving the C-program for Error Detection in EAN 13 and EAN 8.

PENGGUNAAN MATEMATIK DALAM EAN 13 DAN EAN 8 DENGAN ATURCARA KOMPUTER

ABSTRAK

EAN 8 dan EAN 13 merupakan kod bar yang paling biasa digunakan untuk produk yang dipasarkan. EAN 8 mempunyai lapan digit manakala EAN 13 mempunyai 13 digit. Digit yang terakhir untuk kedua-duanya dikenali sebagai angka kawalan. Angka kawalan memainkan peranan untuk menentukan aturan digit-digit sebelumnya atau untuk mengesahkan suatu kod bar yang diperoleh melalui pengiraan dengan menggunakan mod 10 dan pemberat (1, 3). Pemahaman kongruen modulo n dan modular aritmetik diperlukan. Selepas mengesahkan suatu EAN 13 dan EAN 8 melalui penggunaan digit-digit binary, maka suatu kod bar EAN 13 dan EAN 8 dapat dibentuk. Kefasihan dalam kebarangkalian, gabungan dan aturan, binari dan sistem asas dua diperlukan untuk memahami pembentukan simbol kod bar. Untuk membantu pengesahan kod bar EAN 13 dan EAN 8, suatu aturcara C telah dikodkan. Satu lagi aturcara C telah direka untuk mengesan kesilapan yang mungkin berlaku dalam digit-digit kod bar. Kesimpulannya, mod 10 merupakan mod yang paling sesuai untuk memperoleh angka kawalan. Walau bagaimanapun, penggunaan pemberat (1, 3) dalam pengiraan angka kawalan tidak membenarkan pengesanan kesilapan transposisi loncatan. Untuk penyelidikan lanjut, analisis terhadap pemberat yang berbeza untuk EAN 13 dan EAN 8 boleh dijalankan, dan kajian secara mendalam untuk memperbaiki aturcara untuk mengesan kesilapan mengkodkan digit-digit kod bar EAN 13 dan EAN 8.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\in$	Element of
$\equiv$	Congruence equivalence
$\not\equiv$	Congruence inequivalence
$\sim$	Equivalence relation to
$\{\ \}$	Set of
ξ	Sample space
$\bullet$	Dot product

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Barcodes are found practically on all items that are being produced these days. It is more common to consumers as it is on all the products that we purchase, let it be a small bar of chocolate to a carton of soft drinks. Barcodes are those rectangular-like symbol on the item produced or purchased, consisting dark lines and white spaces with numerical digits below. Each time we purchase an item, the store identifies a price that appears on the cash register when the products are scanned.

This generates our curiosity to understand barcodes and how it works. Questions that often arise are:

i. Could we create our own barcode by drawing some lines and placing a few digits below?
ii. Do the lines and spaces in a barcode represent the digits below it?
iii. Why are some lines thicker and some spaces thinner compared to the rest?

iv. How do we know if a barcode is valid?

The answers to all these unuttered questions are basically some mathematical algorithms and mathematical properties. Reading these mathematical contents in a barcode can be easily done through the use of technology.

Learning and understanding more about the barcodes not only allows us to increase our knowledge about barcodes, but also increase our interest in applied mathematics and computer technology. Each time we see a barcode, perhaps we will tend to appreciate it more compared to other consumers. Perhaps this would also foster our interest in cracking barcodes to gain self satisfaction. However, this must be done for only knowledge purposes to see the little wonders of mathematics in our daily life and not for other gains.

1.2 GENERAL VIEW OF BARCODES AND THE CHECK DIGIT

Barcode has existed for a long time, evolving from some few digits to a longer string of digits. However, these barcodes are very similar in nature. Each barcode are made up of lines and spaces representing the digits that are printed below it. Below are examples of barcodes that are commonly used:

Figure 1.1 Universal Product Code Version A

Figure 1.2 European Article Numbering 8 Barcode (EAN 8)

Figure 1.3 European Article Numbering 13 Barcode (EAN 13)

The digits in a barcode represent certain information of a product. For Universal Product Code Version A, the first digit is a prefixed digit which identifies the product category. The next five digits, assigned directly by Uniform Code Council (UCC), identify the manufacturer. The following five digits are the product numbers allocated and controlled by the manufacturer to its products. The last digit is the check digit.

For the EAN 8 and EAN 13, the first three digits identify the country in which the product has been numbered. The three digits are allocated by IANA (International Article Numbering Association) to the national numbering organization in the country. For Malaysia, the Federation of Malaysian Manufacturers (FMM) is the designated organization which manages and promotes the EAN system under the name of the Malaysian Product Numbering (MPN) system. FMM has been given the number 955 by the IANA (FMM, 1996). The next four digits for the EAN 13 represent the manufacturer's number allocated by the national numbering organization to the individual companies while for the EAN 8, the next four digits are the article numbers determined by the national numbering organization to a particular product. The following five digits for the EAN 13 represent the product code which is allocated by the manufacturer to his products. The last digit is the check digit.

Figure 1.4 Code 39

Code 39 is the most popular symbology in the non retail world and is used in manufacturing, military and health applications. Code 39 optionally allows for a (modulo 43) check character in cases where data security is important. The health care industry has adopted the use of this check character for health care applications.

Figure 1.5 RSS-14

RSS was developed by Uniform Code Council (UCC). The purpose of this code is to increase the amount of information that can be put in a barcode while decreasing the overall area of the code. Industries such as pharmaceutical, healthcare, logistics and transportation, and supermarkets in North America have recently begun utilizing this new symbol (Taltech, 2005).

Figure 1.6 Code 93

Code 93 was developed as an advancement to the Code 39 symbology by providing a slightly higher character density than Code 39. Code 93 also attains two check digits as an added measure of security (Taltech, 2005).

Figure 1.7 POSTNET

POSTNET (POSTal Numeric Encoding Technique) is a bar code symbology used by the U.S. Postal Service to encode ZIP Code information for automatic mail sorting by zip code (Taltech, 2005). Data is encoded in the height of the bars instead of in the widths of the bars and spaces. POSTNET has a fixed dimension symbology meaning that the height, width and spacing of all bars must fit within exact tolerances (Taltech, 2005). The U. S. Postal money orders, divides the identification numbers by 9 (modulo 9), the remainder is the check digit to ensure validity (Gallian & Winters, 1988).

Figure 1.8 BPO 4 State Code (British Post Office, Royal Mail Code)

BPO (British Post Office) 4 State Code is a postal bar code symbology that has been developed by the British Post office for encoding European postcode data similar to the way the U.S. PostNET symbology is used for encoding Zip Code data. The goal of BPO 4 State Code is to provide European countries with a simple and efficient postal bar coding scheme. BPO 4 State Code is a fixed dimension symbology meaning that the height, width and spacing of all bars must fit within exact tolerances (Taltech, 2005).

Figure 1.9 PDF 417

PDF417 is a high density 2 dimensional bar code symbology that consists of a stacked set of smaller bar codes. The symbology was published by Symbol Technologies, Inc. to fulfill the need for higher density bar codes. The low level structure of a PDF417 symbol consists of an array of code words (small bar and space patterns) that are grouped together and stacked on top of each other to produce the complete printed symbol (Taltech, 2005).

Figure 1.10 Data Matrix

Data Matrix is a high density 2 dimensional matrix style bar code symbology that can encode up to 3116 characters from the entire 256 byte ASCII character set (Taltech, 2005). The security level of this barcode symbology is comparatively high. The symbol is built on a square grid arranged with a finder pattern around the perimeter of the bar code symbol (Taltech, 2005).

Figure 1.11 Maxicode

MaxiCode is a fixed size matrix style symbology which is made up of offset rows of hexagonal modules arranged around a unique bulls-eye finder pattern (Taltech, 2005). Each MaxiCode symbol has 884 hexagonal arranged in 33 rows. The symbology was designed by United Parcel Service for package tracking applications. The design of the MaxiCode symbology was chosen because it is well suited to high speed, orientation independent scanning.

1.2.1 Check Digit

Most barcodes append an extra digit, called the check digit. In our discussion above, we notice that each barcode in the example has a check digit. The check digit is the last digit encoded in the barcode and it serves to verify the validity of the order of the other digits. Thus we cannot simply create a string of numbers for a barcode. As humans, we cannot avoid making mistakes so it is important for us to detect any errors in a barcode. The check digit also notifies us the possibility of an error while encoding the digits for a barcode. There are several types of errors in transmitting the digits, namely, a single digit error, transposition-of-adjacent-digit errors, jump-transposition, twin error, phonetic error, and jump twin error. In the following, we discuss four of the major types of errors based on study presented by Verhoeff (Kirtland, 2001).

1.2.2 Single Digit Error

A single digit error occurs when one of the digits in the number changes to a different value. 79.1% of all transmission errors that occur are single digit errors.

Example: Actual number: 12<u>3</u>456

Transmitted error: 12<u>8</u>456

1.2.3 Transposition-of-adjacent-digits Error

Transposition-of-adjacent-digits error occurs when two different side-by-side digits change places. 10.2% of all transmission errors that occur are transposition-of-adjacent-digits errors.

Example: Actual number: 12<u>34</u>56

Transmitted error: 12<u>43</u>56

1.2.4 Jump-transposition Error

Jump-transposition error occurs when two different digits, separated by a third digit between them, changes places. 0.8% of all transmission errors that occur are jump-transposition error.

Example: Actual number: 12<u>345</u>6

Transmitted error: 12<u>543</u>6

1.2.5 Twin error

Twin error occurs when two identical side by side digits change to a different pair of identical digits. 0.5% of all transmission errors that occur are twin errors.

Example: Actual number: 1234<u>55</u>

Transmitted error: 1234<u>33</u>

1.3 RESEARCH OBJECTIVES

The objectives of this research are:

i. To understand the check digit system for the barcode
ii. To develop a C program to verify barcodes and to correct errors in barcodes
iii. To understand the connection of the barcode digits with the lines and spaces in a barcode

1.4 RESEARCH SCOPE

The research scopes this research are:

i. EAN 13, and

ii. EAN 8.

European Article Numbering 13 (EAN 13) is a string of 13 digits including a check digit. It is extensively used for retail production (Taltech, 2005). EAN 13 is an enhancement of Universal Product Code (UPC) Version A. It is common to us as it is on every product we purchase.

European Article Numbering 8 (EAN 8) is a shorter version of EAN 13 consisting 8 digits including a check digit. EAN 8 is an enhancement of Universal Product Code (UPC) Version E. It is also used for retail product but only in special cases such as round containers (soft drink can), small items and magazine (Gallian, 1996).

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

The inventory departments of supermarkets before the inventions of barcodes had spent a long time to update inventory stock check to keep track of the items that are being checked out. Grocers knew they desperately needed something to assist them in this problem. Punch cards, first developed for the 1890 U. S. Census, seemed to offer some hope (Seideman, 2000). In 1932, a business man named Wallace Flint, a Harvard University graduate from the school of Business Administration, wrote a masters degree thesis in which he envisioned a supermarket where customers would perforate cards to mark their selections, at the checkout counter they would insert them into a reader which would activate machinery to bring the purchases to them on conveyer belt (Seideman, 2000). Therefore, store management would have a record of what was being bought. However, the card reading equipment was bulky, expensive and was economically unfeasible to create the system.

REFERENCES

Adams, R., 2005. *Bar Code History.* Retrieved 12th August 2005, http://www.adams1.com/pub/russadam/history.html

Billstien, R., Libeskind, S, and Lott, J.W., 2004. *A Problem Solving Approach to Mathematic for Elementary School Teachers.* 8th Ed. Pearson Addison Wesley, United States of America.

Carlson, B. and Thorne, B., 2003. *Statistics for Business and Economics.* 5th Ed. Prentice Hall, New Jersey.

Crossfield, R. T., Taylor, J., Dale B.G., and Plunkett J.J., 1988. The use of Bar Coding in the goods receiving function in order to improve and better control information flow: a pilot study. In: *Proceeding of the 1st International Conference Tracking, Identification and Control,* November 1988, UK, 103-113.

Federation of Malaysian Manufacturers, 1996. Barcode It Right, A Step by Step Guide on Bar-coding of Products. *Workshop on the Malaysian Product Numbering System,* Kuala Lumpur.

Gallian, J. A., 1996. Error Detection Methods. *ACM Computing Surveys.* Volume 28, Issue 3, 504-517.

Gallian, J. A. and Winters, S., 1988. Modular Arithmetic in Marketplace. *Amer. Math. Monthly* ***95***, 548-551.

Harmon, C. and Adams, R., 1989. *Reading Between Lines.* Helmers Publishing. NH

Ho, C.M., Amran Ahmed, Rozaini Roslan, 2004. *Lecture Notes Real Analysis.* Universiti Sabah Malaysia, Kota Kinabalu, Sabah.

Kirtland, J., 2001. Identification Numbers and Check Digit Schemes. In: Verhoeff, J., *Error Detecting Decimal Codes.* The Mathematical Association of America, United States of America, 4-5.

Long, C.T., and DeTemple, 1995. *Mathematical Reasoning For Elementary Teachers.* Harper Collins Publisher, New York.

Palazollo, M., 1994. The use of EAN in Non-Retail Environments. *FMM Seminar on EAN International using Bar Code Technology to Enhance Business Efficiency.* Kuala Lumpur, 6-7.

Rainman, 2005. Drawing EAN 13. Retrieved 4 August 2005, http://www.codeproject.com

Sato, 1999. Barcode History. Retrieved 12 August 2005, http://www.sato_europe.html

Seideman, T., 2000. Barcodes Sweep the World. In: *Wonders of Modern Technology.* Retrieved 22 August 2005, http://www.barcoding.com/information /Barcode_history.shtml.

Taltech Instrumental Software Solutions, 2005. Introduction to Barcode Symbologies. Retrieved 8 March 2005, http://www.taltech.com/TALtech_web/resources/intro_to_bc/bcsymbol.htm.

Van Lennap, R., 1994. EAN – A Global Business Solution. *FMM Seminar on EAN International using Bar Code Technology to Enhance Business Efficiency.* Kuala Lumpur, 2-4.

Wikipedia, 2005a. *Binary Numeral System.* Retrieved 4 September 2005, http://en.wikipedia.org/wiki/Binary_numeral_system

Wikipedia, 2005b. *Congruence*. Retrieved 4 September 2005, http://en.wikipedia.org/wiki/Congruence

Williams, P., 1996. *Congruence Modulo n*. Retrieved 4 September 2005, http://planetmath.org/encyclopedia/Congruences.html