Mathematical & Computational Applications, Vol. 1, No. 2, pp 1-6, 1996

A NEW TECHNIQUE TO PROCESS AND RECOGNIZE BARCODES USING INDUCTION

M.S. AKSOY* and M.BAYRAM**

ABSTRACT

In this paper, a new technique to recognize and process Barcodes is introduced. The technique employs Inductive Learning. It is suitable to use, for example, in a factory to control the workers, staff, stock etc. In this technique only vertical lines are considered while the spaces in between are ignored. This results faster processing. Each Barcode is considered to represent an item. For each Barcode a rule is extracted from the necessary information using Inductive Learning. So the unnecessary information is eliminated. This causes faster processing time and less amount of memory. In order to use this technique no special hardware is required. Only a PC and a Barcode reader is enough.

Keywords: Artificial Intelligence, Inductive Learning, Barcode, Expert Systems

* Sakarya Univ., Engineering Faculty, Industrial Eng. Dept., Esentepe, Adapazarı-Turkey
** TÜVASAŞ Adapazarı-Turkey

1. INTRODUCTION

Barcodes are made of some thin and bold bars, some spaces in between and some meaningful numbers[1]. They can be used to recognize some items such as the type, price etc. of a product; the names and other necessary information of staffs in a factory; the names, subjects, author's names, year and other information of a book or publication in a library and so on. For example in a factory the worker's and staff's comings and goings can be controlled and for example payroll can be designed using barcodes. In order to do these the barcode must be read and processed. Normally a barcode is read by a reader and recognized using a special hardware which employs a special technique[2].

In this paper a new technique to process and recognize Barcodes is introduced. The technique employs Inductive Learning. In the paper RULES-3 inductive learning algorithm is introduced. For a number of randomly generated barcodes, how the necessary rules are extracted and using the extracted rules how barcodes are recognized are explained.

2. RULES-3 INDUCTIVE LEARNING ALGORITHM

In recent years, there has been a growing amount of research on inductive learning. In its broadest sense, induction (or inductive inference) is *a method of moving from the particular to the general - from specific examples to general rules*. Induction can be considered the process of generalizing a procedural description from presented or observed examples [3,4,5]. These examples may be specified by an expert as a good tutorial set, or may come from some neutral source such as an archive. The induction process will attempt to find a method of classifying an example expressed as a function of the attributes, that explains the training examples and that may also be used to classify previously unseen cases.

RULES-3 [6] is a simple algorithm for extracting a set of classification rules from a collection of examples for objects belonging to one of a number of known classes. An object must be described in terms of a fixed set of attributes, each with its own range of possible values which could be nominal or numerical. For example, attribute "length" might have nominal values {short, medium, long} or numerical values in the range {-10, 10}.

An attribute-value pair constitutes a condition in a rule. If the number of attributes is N_a , a rule may contain between one and N_a conditions. Only conjunction of conditions is permitted in a rule and therefore the attributes must all be different if the rule comprises more than one condition.

RULES-3 extracts rules by considering one example at a time. It forms an array consisting of all attribute-value pairs associated with the object in that example, the total number of elements in the array being equal to the number of attributes of the object. The rule forming procedure may require at most N_a iterations per example. In the first iteration, rules may be produced with one element from the array. Each element is examined in turn to see if, for the complete example collection, it appears only in objects belonging to one class. If so, a candidate rule is obtained with that element as the condition. In either case, the next element is taken and the examination repeated until all elements in the array

have been considered. At this stage, if no rules have been formed, the second iteration begins with two elements of the array being examined at a time. Rules formed in the second iteration therefore have two conditions. The procedure continues until an iteration when one or more candidate rules can be extracted or the maximum number of iterations for the example is reached. In the latter case, the example itself is adopted as the rule. If more than one candidate rule is formed for an example, the rule which classifies the highest number of examples, is selected and used to classify objects in the collection of examples. Examples of which objects are classified by the selected rule are removed from the collection. The next example remaining in the collection is then taken and rule extraction is carried out for that example. This procedure continues until there are no examples left in the collection and all objects have been classified. This algorithm can be summarizes as follow:

Step 1. Define ranges for the attributes which have numerical values and assign labels to those ranges
Step 2. Set the minimum number of conditions (N_{cmin}) for each rule
Step 3. Take an unclassified example
Step 4. $N_c = N_{cmin} - 1$
Step 5. If $N_c < N_a$ then $N_c = N_c + 1$
Step 6. Take all values or labels contained in the example
Step 7. Form objects which are combinations of N_c values or labels taken from the values or labels obtained in Step 6
Step 8. If at least one of the objects belongs to a unique class then form rules with those objects;
ELSE go to Step 5
Step 9. Select the rule which classifies the highest number of examples
Step 10. Remove examples classified by the selected rule
Step 11. If there are no more unclassified examples then STOP;
ELSE go to Step 3

3. OBTAINING THE SET OF BARCODE EXAMPLES

For this application it is assumed that there are 20 attributes. Each attribute is considered to represent the thickness of a vertical line. The vertical lines in a barcode are numbered from left to right for each barcode. Each line represents an attribute. For example the first line represents Attribute-1 (or A1 in short), the second A2 etc. The value for each attribute can vary from 1 to 5 pixels. In this work the set of examples consists of 30 randomly generated barcodes. Only the thickness of lines are considered while the spaces in between are ignored. Each example consists of 20 values for each barcode. For instance the following example represents Barcode-1:

4,3,3,2,2,4,1,4,5,4,1,3,5,4,2,5,5,1,5,2,Barcode1

It means the thickness for the first line in the Barcode-1 is 4 pixels, for the second it is 3 pixels and so on. The whole set of randomly generated examples is given in Table 1.

Table 1. The set of examples for 30 Barcodes.

No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	BARCODE
1	4	3	3	2	2	4	1	4	5	4	1	3	5	4	2	5	5	1	5	2	Barcode1
2	3	4	1	3	3	2	4	4	2	2	5	5	3	5	5	2	4	5	2	3	Barcode2
3	1	5	4	1	3	1	1	4	2	1	2	2	2	5	5	3	2	1	1	4	Barcode3
4	3	3	4	2	4	2	1	3	1	3	5	2	4	2	2	5	4	4	3	1	Barcode4
5	3	4	5	5	1	3	5	3	4	3	3	3	2	3	2	1	2	5	1	2	Barcode5
6	2	3	1	3	2	4	3	1	5	4	3	2	1	4	3	4	3	5	1	2	Barcode6
7	1	1	2	1	1	3	4	3	5	1	1	4	3	2	1	4	5	3	1	4	Barcode7
8	3	3	3	2	2	1	3	1	5	1	3	2	5	4	2	4	2	1	1	2	Barcode8
9	4	2	2	3	2	2	1	3	2	5	3	4	5	2	3	1	4	3	5	1	Barcode9
10	5	4	2	1	3	2	5	1	1	2	3	5	3	3	5	5	4	4	5	2	Barcode10
11	3	3	4	1	3	3	5	3	3	3	2	4	3	4	5	2	2	2	1	3	Barcode11
12	2	3	2	5	3	1	4	4	4	4	1	5	2	5	1	1	4	2	3	1	Barcode12
13	1	1	1	4	3	3	2	3	4	4	2	5	4	1	4	4	1	3	3	2	Barcode13
14	5	5	4	3	4	2	2	3	5	1	4	2	1	1	1	3	5	3	3	5	Barcode14
15	2	2	2	2	3	1	3	5	3	5	4	3	4	1	4	4	3	1	2	2	Barcode15
16	4	1	3	4	5	2	5	5	4	5	5	3	1	5	4	4	3	1	1	1	Barcode16
17	3	3	1	2	4	5	3	1	5	2	2	4	2	3	2	5	4	2	2	5	Barcode17
18	3	5	3	2	5	2	2	1	2	1	4	5	2	4	3	5	4	3	1	3	Barcode18
19	2	4	2	4	1	3	5	2	5	1	5	4	3	5	3	2	5	1	3	4	Barcode19
20	2	2	2	3	2	3	2	1	5	1	4	5	2	5	5	3	3	4	5	2	Barcode20
21	1	2	3	1	4	1	2	4	3	1	5	3	1	3	2	4	5	5	1	1	Barcode21
22	1	1	2	1	5	5	3	4	1	2	1	5	4	4	2	3	3	5	3	2	Barcode22
23	2	2	3	2	3	3	3	1	3	2	5	5	5	5	3	5	5	2	1	4	Barcode23
24	1	2	4	1	1	2	2	3	2	2	5	2	2	3	5	3	4	4	3	5	Barcode24
25	4	3	2	3	4	3	5	2	4	5	2	2	1	2	5	4	3	4	2	3	Barcode25
26	4	5	1	3	3	1	3	4	3	3	5	5	3	3	4	1	1	2	1	1	Barcode26
27	3	3	3	4	1	3	1	2	4	3	3	2	5	4	3	3	4	1	1	2	Barcode27
28	4	5	2	4	4	1	4	3	2	4	5	5	4	1	4	5	1	5	1	2	Barcode28
29	3	5	1	1	3	2	3	3	4	3	2	1	3	3	4	3	4	5	3	4	Barcode29
30	4	4	4	1	4	1	2	5	5	2	1	2	5	2	2	4	4	4	4	2	Barcode30

4. FORMING THE KNOWLEDGE BASE

Using RULES-3, 30 rules can be extracted from the set of examples given in Table1. The rule set (knowledge base) is given in Table 2. As can be seen from Table 2, none of the rules contains more than two conditions and even some of them have only one condition. It shows that out of 20 possible conditions only one or two of them are enough to represent and recognize each barcode while the rest are not necessary for this application. This helps to spent less effort and time to store, recognize and process a barcode. The unnecessary information is eliminated by means of inductive learning.

Table 2. Extracted set of rules for 30 Barcodes

1	IF A1=4 and A4=2 then it is BARCODE1
2	IF A1=3 and A4=3 then it is BARCODE2
3	IF A1=1 and A2=5 then it is BARCODE3
4	IF A1=3 and A9=1 then it is BARCODE4
5	IF A3=5 then it is BARCODE5
6	IF A1=2 and A3=1 then it is BARCODE6
7	IF A1=4 and A7=4 then it is BARCODE7
8	IF A1=3 and A5=2 then it is BARCODE8
9	IF A1=4 and A2=2 then it is BARCODE9
10	IF A1=5 and A2=4 then it is BARCODE10
11	IF A1=3 and A9=3 then it is BARCODE11
12	IF A1=2 and A4=5 then it is BARCODE12
13	IF A1=1 and A3=1 then it is BARCODE13
14	IF A1=5 and A2=5 then it is BARCODE14
15	IF A1=2 and A8=5 then it is BARCODE15
16	IF A1=4 and A2=1 then it is BARCODE16
17	IF A1=3 and A6=5 then it is BARCODE17
18	IF A1=3 and A5=5 then it is BARCODE18
19	IF A1=2 and A2=4 then it is BARCODE19
20	IF A1=2 and A7=2 then it is BARCODE20
21	IF A1=1 and A3=3 then it is BARCODE21
22	IF A1=1 and A5=5 then it is BARCODE22
23	IF A1=2 and A3=3 then it is BARCODE23
24	IF A1=1 and A6=2 then it is BARCODE24
25	IF A1=4 and A6=3 then it is BARCODE25
26	IF A1=4 and A3=1 then it is BARCODE26
27	IF A1=3 and A4=4 then it is BARCODE27
28	IF A1=4 and A7=4 then it is BARCODE28
29	IF A1=1 then it is BARCODE29
30	IF A9=4 then it is BARCODE30

5. BARCODES RECOGNITION

In order to recognize a barcode, the rules obtained by RULES-3 and given in Table 2 are used. For this, first of all the barcode must be read via a reader and it must be transferred to a PC. Afterwards using a very simple software the thickness of each line is obtained. The extracted rules are examined whether

they match with this object or not. When a rule is satisfied then the class (barcode) is obtained. As an example in this work each barcode was considered to represent a book. Each barcode was used for the name, the authors names, date, place and some other necessary information of a book. It was realized that 30 different barcodes were correctly recognized. When the barcode is recognized, the related information can be obtained from a data-base. No special hardware is required for this process, only a reader and a PC will do the job. Since the number of conditions is usually small for extracted rules, the process will not take long time provided that the number of rules is not too big. The number of extractable rules depends on the number of examples. So if the number of examples is too big then this technique may not be suitable.

6. RESULTS AND DISCUSSION

In this paper a new technique to process and recognize barcodes is presented. The necessary set of examples which contains 30 randomly generated examples was used. Using RULES-3 inductive learning algorithm 30 rules was extracted. Each rule has one or two conditions. The rest of unnecessary information was eliminated by means of inductive learning. The technique only the thickness of lines are considered while the spaces in between are ignored. Using extracted rules 30 Barcodes were correctly recognized. This technique is especially suitable to control the staff, stock etc. in, for example, a factory. Each barcode was used for an item. The necessary information about each barcode can be stored in a data-base. When the barcode is recognized this information can be reached. For this technique no special hardware is required because only a reader and a PC will do the job. Since the number of conditions is usually small for extracted rules, the process will not take long time. If the number of examples is too big it means the number of rules will be too big and in that case this technique may not be suitable.

REFERENCES

[1] İşmen M., "Çizgikodun teknik özellikleri, master film hazırlanması ve baskıya uygulanması", TOBB EAN İhtisas Toplantısı, Ankara, Turkey, 1988,.

[2] TS 10147, "Barkod (çubuk kodlama) sembolleri, EAN-13 ve EAN-8 Genel Kurallar", UDK 681.327.12.003.2(083.73), Birinci baskı, Ankara, Turkey, 1992

[3] Quinlan J.R. "Induction, knowledge and expert systems", in Artificial Intelligence Developments and Applications, Eds: J.S. Gero and R. Stanton, Amsterdam, Nort-Holland, pp.239-266, 1988.

[4] Forsyth R. "Machine learning principles and techniques", Ed: R. Forsyth, Chapman and Hall, London, 1989.

[5] Hancox P.J., Mills W.J. and Reid B.J. "Artificial intelligence/expert systems", Ergosyst Associates, Lawrence, Kansas, 1990.

[6] Pham D.T. and Aksoy M.S., "A new algorithm for inductive learning", Journal of Syst. Eng. , 5, pp. 115-122, London, 1995.