Using Genetic Algorithm for classification of flower plants

by
Mohammed Abbas Kadhum
AL-Qadisiyia University-College of science
Computer Dept.

Abstract

The aim of this research is use adaptive search (genetic Algorithm) for classification of flower plants depend on its features, these features are extracted from outline shape of these flower's plants. The classification of flower plants are a well known problem in plants science and encountered almost of researchers in that domain.

In this paper we produce a Simple Genetic Algorithm (SGA) for classifier system, the classifier system is an application mode because it have sufficiency learning is accomplished, the proposed system receive a massage (features) from environment (that features may be overlapping) , the SGA responsible for take that features to arrive final appropriate class, the proposed system implemented using visual basic programming language and applied on P4 processor

I-Introduction

The classification of flower plants is a well known problem in plants science, in this paper we process this problem by Genetic Algorithm (GA) using classifier system. The genetic Algorithm is an optimization and search technique based on the principle of genetics and natural selection, a GA allows a population composed of many individuals to evolve under specific selection rule to a state that maximizes the fitness (i.e. minimizes the cost function)[3].While the classifier system is a machine learning system that learns syntactically simple string rules, a classifier system derives its name from its ability to learn to classify massages from the environment in to general sets and is similar to control system in many respects[7,2], inside the classifier system is a set of condition action rules[6] or conditional statement whose constituents are words drawn from the ternary alphabet (0,1,#). It has one or more words or conditions as antecedent, an action statement as the consequent, the symbol "#" acts don't care in the condition matching either 0 or 1 [7].

The classifier performs as traditional production system, the environment sends a message to classifier system detectors, this event is decoded and placed as a pattern on the internal message list, the working memory for the production system match the condition patterns of the classifier rule, the selection of classifier is determine by a biding scheme, where a bid is a function of both the accumulated fitness of the classifier and the quality of the match between the input stimulus and its condition pattern add messages (the action of fired rules) to working memory [4,2].

This paper is organized as follow, in the next section we introduce our classification problem followed by the proposed system that include two approaches the first one is extraction of classifier rules and second is using that classifier rules and combination with GA to produce classifier system using SGA, finally we produce conclusions in the last section.

II- Classification of Flower plants

The taxonomy strives for putting live vegetations into groups which reflects the real image of the genetic relationships which connect it to each other. And when the agreement on the shelves which gives greater importance becomes not simple than other in the classification and on a manner of the evolutionary connection explanation between these shelf and by the nature of the case, there are thousands of prosperities which may be the alive being will contain it. On it for an eternity of a choice of limiting number of it which is sufficient to the scientific hedgy. From here the problem of shelves selection becomes liquefied vital. Some shelves with it's nature are specific either it is one of two cases only such as the existence of generosity bracts or nonexistence of it or multi cases like the kind of the fruit for example it may be amygdaloidal or a box or a craw and all of these shelves don't create between it the medial cases. Or the second type of

shelves which have quantities properties with continuous variations while measurement makes it permitted to itself with one numerical meanses such as the vegetable piece of leaf long or counted of the petaloid leaves. Generally the classification leans on the properties of the form for the aims of plant or animal classification to the appearances of the form and it's discriminators more than any other properties in the a live body being [1].Where the importance of the form shelf is being measured with the extent of it's front stead fastness the environmental stakes where, whenever it is fixed and an inherited so it carried higher taxonomic value for that it is possible to depend on the assembling like the flower and fruit because indeed it is distinguished by the fixed properties and in the following the most important used assembling of the semblances:-

1- Corolla: divided the vegetations with filqaten lean according to properties of corolla to three groups and it is either without corolla or possess or of separate petaloid leaves or the conjunctive petaloid leaves.
2- Letting fall the pink fractions (position of ovary): In the categories the florescence is the nisleia fractions which occurs under the ovary or some times the fractions are peripheral while in other flowers occurs in higher forms.
3- The nature of pink envelope (covers syphilis): the possessors of one device for holding the legs of the delinquent during the bastinado divided according to the nature of it's pink envelope then in me some of her is being formed for like the ace of color green mutual rings and in other categories it may be shorten to squames or pinnulaes.
4- Stamens: which the stamens become one of the important fundamentals which concentrates on classification operation it is distinguished by various and numerous properties from it the number; the length; the support and the cohesion of a florescence [5], in addition to the above appearances there are other semblances can be depend on it in the proposed system.

III-Proposed system

The proposed system include two stages as follow:

1-Extraction of classifier rules

The expert in classification of flower plants able to classify plant by observation of flower's plant, the problem here is how we can extraction that features of flower's plant which depend by experts to classify that plants, the first step is interview with domain expert to extract features of flower which depend by experts to specify the class of plant for that features, the features or conditions represent antecedent and the class of plant represent consequent in a rule that rule called classifier rule. The second step is satisfied that rules by comparing with rules are extracted from other experts, books, or scientific

researches, for example of that classifier rule which used in proposed system as follow :

Class1 if featur1, feature2, feature3,…

after that, classifier rules are coding to binary encoding for the above rule as:

101 if 101#1#0…

In the condition part of classifier rule 1 denoted find that feature, 0 denoted not find that feature, and # symbol denoted don't care that feature, figure (1) illustrate main step of extraction of classifier rules from expert.

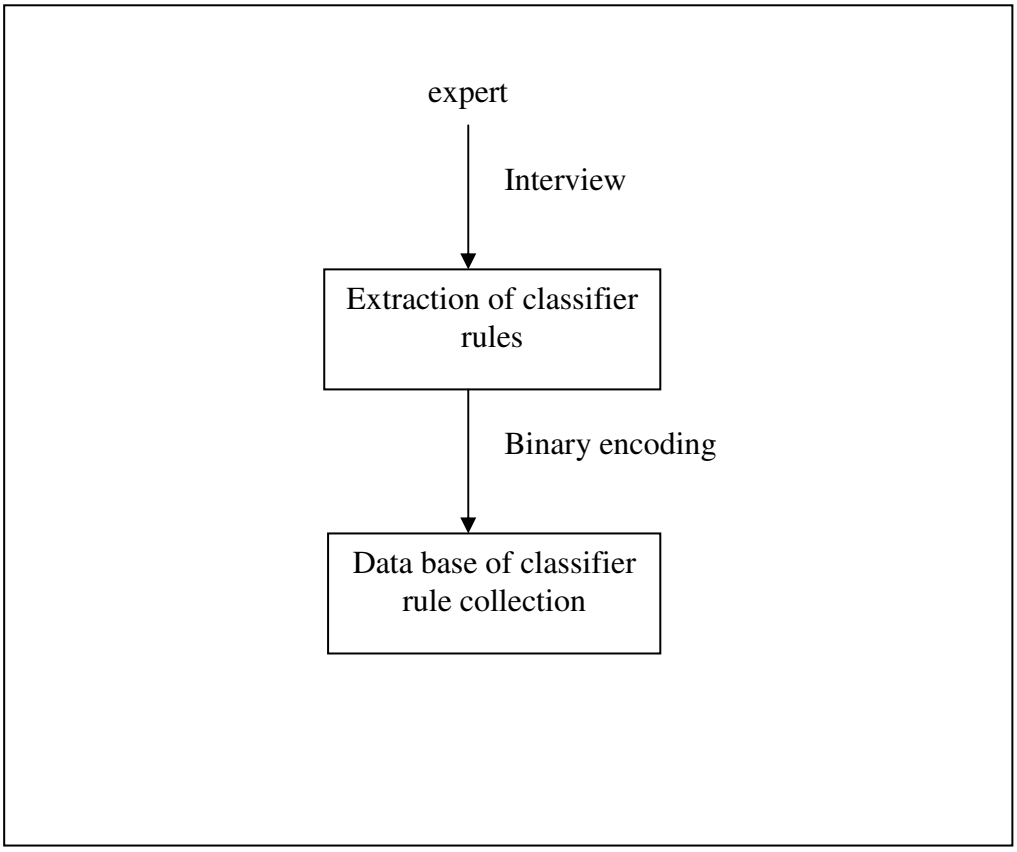


Figure (1) steps of classifier rule extraction

2-genetic Algorithm elements

For our problem of determining the parameters for flower plants classification, each individual is defined by a vector V=(P1,P2….P15) each gene corresponding to one of fifteen parameters of the problem and its value (1,0 or #), the proposed system is initialized by interact with environment through input massage from that environment to proposed system, the massage is features of plant which need to classify, that features are coding by binary encoding, after that genetic Algorithm is work, the over all Algorithm is described by the flowchart in figure (2), we can show the following the genetic Algorithm elements in proposed system:

A-Binary encoding

The coding method used in proposed system is binary encoding because this method is appropriate to our problem, the features is coding to 1 it is found in that rule or 0 if it is not found and # if it is don't care.

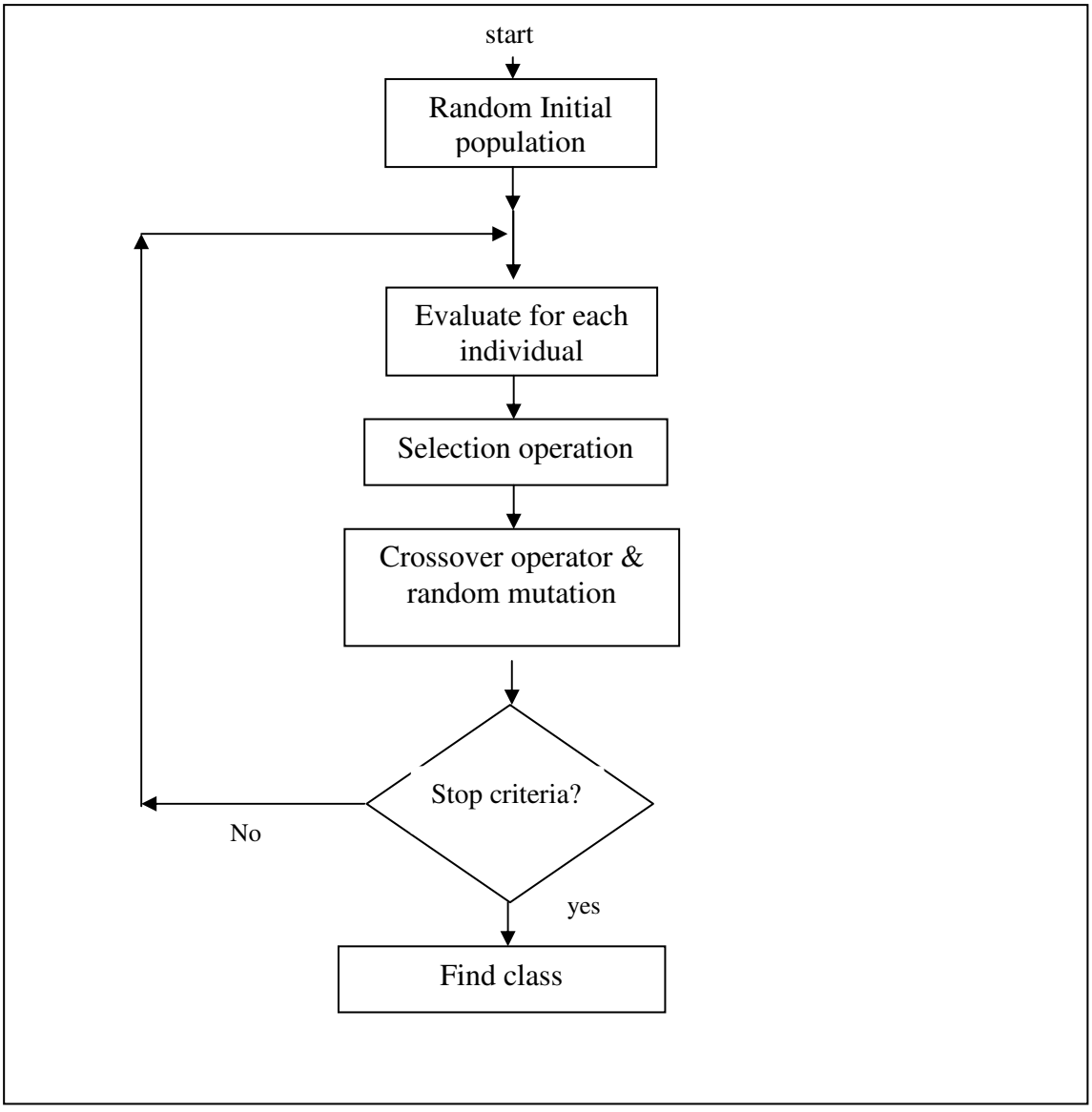


Figure (2): A GA for classification system of flower plant

B-Initial population

The initial population is strings of {0,1,#} each string is individual (chromosome)in that population the chromosome length is fifteen genes each gene represent feature (1 if found, 0 if not found, # don't care) the population size is 10 individuals for example of initial population as figure (3)

10#10#001#010#1
011#1101##01#0#
:
101###0101#101#

figure (3) sample of initial population

c-Evaluation function

The quality of the individual is assessed with fitness function the result is a value for each individuals will survive and are allowed to produce new individuals in the proposed system the initial population must be evaluated by fitness function. The fitness function here is percentage value between the input massage (after coding) with all the individuals in that population that mean we will compare each gene in input massage with each gene in individual of that population as following equation:

$$\text{Fitness value} = \frac{\text{sum of similar gene + sum of \# symbol } *0.5}{\text{length of individual}}$$

D-Selection operator

Selection is the process of choosing parents for reproduction (usually based on fitness value)[3] after the evaluate of each individual in that population we will use Whitly method selection. This method is one of binary tournament selection, in this method we select two individuals randomly, then adding the individual that have higher fitness value to mating pool (we select the individual that have higher fitness value because of our problem is maximize problem) this procedure is repeated until we arrive inform number, the following relation illustrate that:

$$\text{Select}_n = \begin{cases} \text{Individual}_i & \text{fitness } i < \text{fitness } j \\ \text{Individual}_j & \text{otherwise} \end{cases}$$

Where :-

select_n :selected individual

i,j a random numbers € {1,2,…N},i≠j.

fitness i: fitness for individual i.

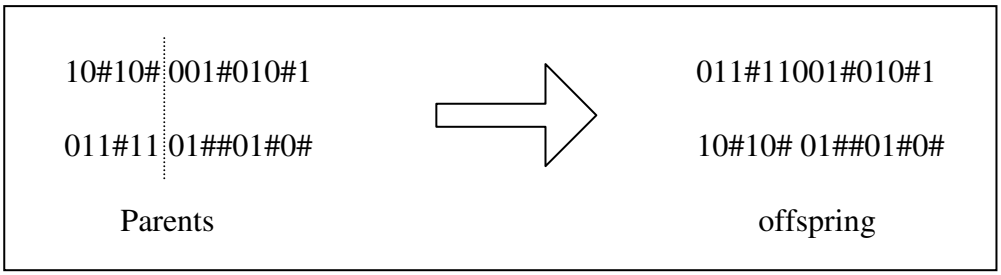
fitness j: fitness for individual j.

N: population size.

E-crossover operator

An operator that forms a new chromosomes from two parent chromosomes by combining part of the information from each[3] the crossover take two individuals to produce two new individuals we use in proposed system

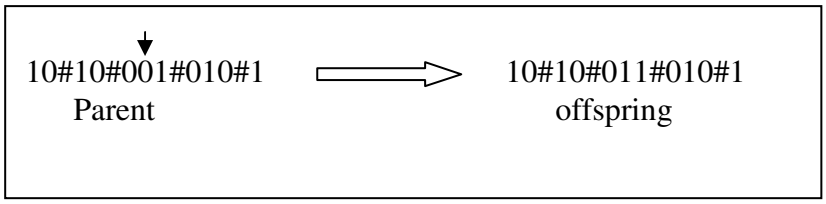
a single point method of crossover and select cross point randomly figure (4) illustrate that operation



Figure(4):crossover operation

F- mutation operator

Mutation is a reproduction operator that randomly alters the values of genes in a parent chromosome in our problem we alter randomly one or more of chromosome components as in example in figure (5). The aim of mutation operator is verity of population and try to capture individual have highest fitness value



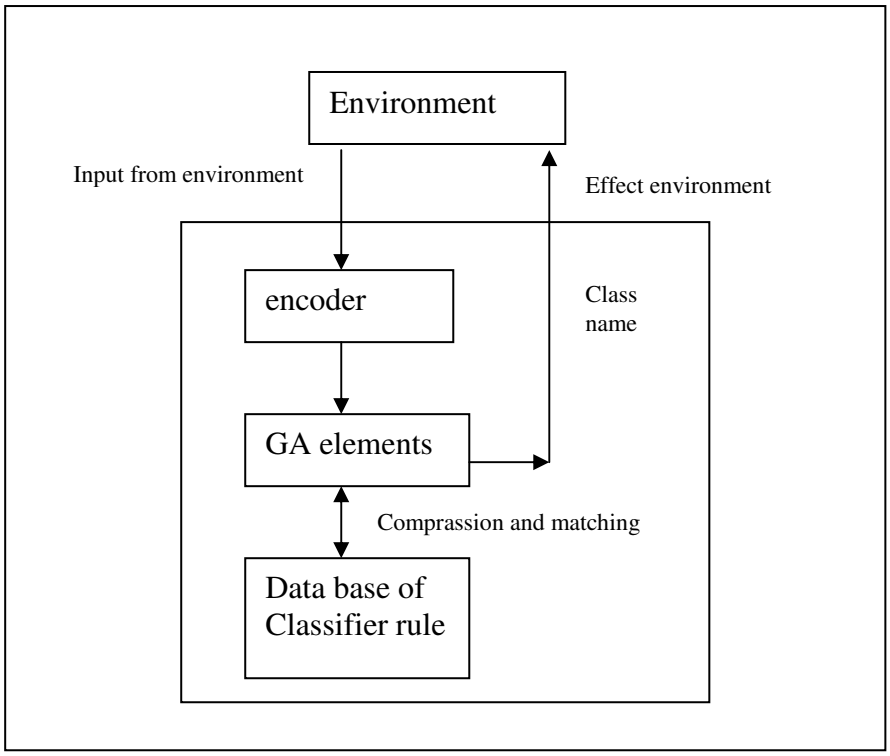
Figure(5): mutation operator

G-stop criteria

The stop condition is used to determine the end of the Algorithm in proposed system for each generation we will matching between the chromosome which have a highest fitness value with the data base of classifier rules which extracted in first step of proposed system if there is a matching with one of classifier rule then stopped else continue with GA cycle. The following figure(6) illustrate the interactive between environment and proposed classifier system.

IV-Case study

In proposed system we have fifteen features for each class that feature may be found(1), not found(0)or don't care(#), the number of classes which can be distinguished are eight classes of flower plants depend on that features, figure(7) illustrate a dialogue between proposed system and user to produce appropriate class



Figure(6):interactive between environment and proposed system

Leaves mostly parallel – veined; flower mostly trimerous (monocots)
Yes No don't care

Leaves mostly reticulate-veined; flowers tetramerous or pentamerous (dicots)
yes No don't care

Perianth absent
Yes No don't care

Stem jointed ; leaves scaly
Yes No don't care

The name of class is CASUARINACEAE

Figure(7) an example of dialogue between user and proposed system

Conclusions

In this paper we have discussed how we can build classifier system using GA. after implemented the proposed system we can conclude the following:

1-We can use the system to help the biological department students to classify flower plants in libratory

2-In extraction of classifier rule we produce a method to restructure the features of flower plant and classes

3-In proposed system we produce a combination between GA and classifier system depend on production system

References

[1]-Alkatab, Yusaf Mansor , "Classification of plant's seeds " , House book publish comp., Mousal University, 2nd edition, 2000.

[2]-Cho,siu-yeumg and Zheu Chi, "genetic evolution processing of data structures for Image classification", IEEE transactions on knowledge and data engineering, Vol. 17, No. 2, February 2005.

[3]-Haupt, Randy L. and sue Ellen H. "Practical Genetic Algorithms" Wiley Interscience published comp. ,2nd edition,, 2004.

[4]-Luger, George F. and William A. Stubblefild , " Artificial Intelligence ", Addison Wesley Published comp.,2nd edition ,1998.

[5]-Naik,V.,N., " Taxonomy of Angiosperms" , McGraw Hill Published comp. , Ney Delhi 1984.

[6]-Nicolas, Pech-Gourg and Hao Jin-kao, "A genetic Algorithm for the classification of natural corks" ,www.imener.com,1999.

[7]-Richards, Robert A., "Classifier system & Genetic Algorithm" , www.dca.fee.uncamp,1995.