Genetic Algorithm Used to Produce New Local Rice Varieties

Su Myo Swe*

Abstract

Genetics can be defined as the study of heredity mechanisms, by which characteristics are passed from generation to generation. Genetic Algorithm (GA) is randomized searching and optimization techniques guided by the principles of evolution and natural genetic. Genetic Algorithm handles a population of possible solutions represented by a chromosome and a chromosome is a sequence of genes. Rice is being a self-pollinated plant, cultivating only one rice variety in one season and that is used as in next season for parents may be reduce the quality and quantity of rice. Accordingly, rice growing countries are greatly interested in new rice variety. New rice variety is any genealogy of rice produced by cross-breeding different kind of rice. New rice variety gives high-yielding inbred rice varieties it can produce up to 30% more rice. This paper introduced the local rice varieties operate with genetic algorithm to produce new rice variety.

Key word: Genetic Algorithm, Local Rice Variety

Introduction

Organisms (animals or plants) produce a number of offspring which are almost, but not entirely, like themselves. Variation may be due to mutation (random changes) and may be due to sexual reproduction (offspring have some characteristics from each parent). Some of these offspring may survive to produce offspring of their own and some are not. The "better adapted" offspring are more likely to survive in their environment. Over time, later generations become better and better in adaption to their life. Genetic algorithms use this same process to "evolve" the generations. In this paper, local rice varieties are used in genetic algorithm to evolve new local rice variety.

In the organisms, genetic material is saved into the chromosomes. Parents' genetic material is mixed during reproduction and the offspring has genes of both parents. The flow chart of the genetic algorithm shows in Figure 1and the structure of genetic algorithm shows in Figure 2.

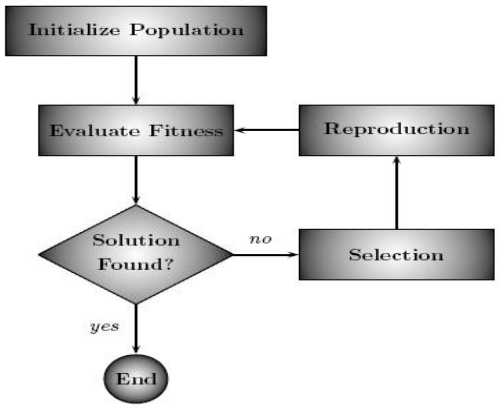


Figure 1. Flow chart of basic genetic algorithm iteration

produce an initial population of individuals
evaluate the fitness of all individuals
while termination condition not met **do**
select fitter individuals for reproduction
mutate individuals
evaluate the fitness of the modified individuals
generate a new population
end while

Figure 2. Structure of the genetic algorithm

Implementation of Genetic Algorithm with Local Rice Varieties

To develop the propose system, Microsoft SQL Server 2005 is used as database, Visual Studio 2008 is used as the programming tools and 26 Local Rice Varieties are used as applied data.

*Dr., Assistant Lecturer, Department of Computer Studies, Dagon University

Populate the Local Rice Varieties in the Genetic Algorithm

To produce new local rice variety, 26 local rice varieties are populated; pure rice varieties are 13 and ordinary rice varieties are 13. Table 1 shows the list of rice varieties in local name.

Table 1. Local Names of Different Varieties of Rice

(a) Pure Rice Varieties		(b) Ordinary Rice Varieties	
No.	**Ordinary Rice**	**No.**	**Pure Rice**
1	Khun Ni	1	Taun Bjan
2	Shwebo Khun Ni	2	Bei Kyar Taun Bjan
3	Bay Gyar Gyi	3	Pyapon Pohsan Hmwei
4	Shwe War Tun	4	Ayeyarwady Pohsan Hmwei
5	Manow Thukha	5	Shwebo Pohsan Hmwei
6	Hsin Thukha	6	Ayeyarwady Poung Hsan
7	Hsin New Yin	7	Shan Hsan
8	Thee Htet Yin	8	Paw Kjwe
9	Yatanr Toe	9	Nga Kjwe
10	Zeayar	10	Hnjin Hsan
11	Hsin Ma	11	Pathein Pohsan Hmwei
12	Hsin Shwe Li	12	Kjwe Thwa
13	Nga Sein	13	Taun Paw Kjwe

Data Representation for Rice Varieties

Rice' chromosome number is 24, contains 12 pairs of chromosomes. Each chromosome has its own properties and functions. Each chromosome is assigned an index within a binary string and if the chromosome is placed in the rice genomic structure, it is represented by a one, if not, a zero. Data representation of rice varieties are show in Table 2.

Rice Varieties and Its Chromosome Characters Used in the Genetic Algorithm

In the rice varieties, there are two types; pure rice varieties and ordinary rice variety. Pure rice varieties are Pathein Pohsan Hmwei and Nga Kjwe and ordinary rice varieties are Khun Ni and Zeayar. Table 3 shows the pure rice varieties lists and its properties of chromosome used in the system. Table 4 shows the ordinary rice varieties lists and its properties of chromosome used in the system.

Fitness of Local Rice Varieties

In this paper, there are phenotypic fitness and genotypic fitness. For the phenotypic fitness; among the chromosome 3,5,6,8 and 10, chromosome locus 4 and above must be binary value 1. For the genotypic fitness; among the chromosome 1,9,11 and 12, chromosome locus 3 and above must be binary value 1.

Define the Termination Condition

According the genetic algorithm, the termination condition must be defined. In this paper, the generation and run number per generation will be defined. **One generation** means **one growing season** and **run number per generation** means **plot of the growing rice varieties**.

Table 2. Properties or Function of Rice Chromosome and Binary Value

Chromo -some	Properties / Function	Binary Value
C1	Several gene families that are dispersed or arranged in tandem repeats. (determining parentage) Bacterial blast resistance.	0 = non-resistance to bacteria 1 = resistance to bacteria
C2	Reveals a breakdown of rice.	0 = breakdown 1 = non-breakdown
C3	The availability of sequence and annotation will provide the function for trait identification and genome evolutionary studies.	0 = bad-texture 1 = good-texture
C4	Six representative rice organ types were examined, chromosome organization and regulation of chromosome level.	0 = short grain length after cooking 1 = long grain length after cooking
C5	Distinguish between two cultivated rice sup-species (Japonia and Indica) (Lingage study of DNA)	0 = Indica 1 = Japonia
C6	Waxy in rice.	0 = non-waxy 1 = waxy
C7	Diverse panel germplasm of rice widely used parental lines in genetic analysis or breeding.	0 = translucence 1 = opacity
C8	Aroma, major gene for fragrance.	0 = bad-fragrance 1 = good-fragrance
C9	Resistance to the rice blast fungal pathogen.	0 = non-resistance to fungus 1 = resistance to fungus
C10	Heterochromation, improvement of rice production varieties, a number of disease resistance genes present.	0 = short grain 1 = long grain
C11	Rich in disease resistance genes.	0 = not rich in disease resistance 1 = rich in disease resistance
C12	Recent gene duplication, new gene for bacterial blight resistance.	0 = do not duplicate genes of bacteria resistance 1= duplicate genes of bacteria resistance

Table 3. Pure Rice Varieties and their Chromosome Characters

No.	**Rice Varieties**	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**C11**	**C12**
1	Taun Bjan	0	0	0	0	0	0	1	1	0	1	1	1
2	Bei Kyar Taun Bjan	1	1	1	1	0	1	1	1	1	0	0	0
3	Pyapon Pohsan Hmwei	1	0	0	0	0	1	1	1	0	0	1	0
4	Ayeyarwady Pohsan Hmwei	0	1	1	1	0	1	1	1	0	1	0	0
5	Shwebo Pohsan Hmwei	1	1	1	1	0	1	1	1	1	1	0	0
6	Ayeyarwady Poung Hsan	0	1	1	1	1	1	1	1	0	1	1	1
7	Shan Hsan	0	1	0	0	1	0	1	1	0	0	0	0
8	Paw Kjwe	1	0	0	0	0	1	1	0	0	1	1	1
9	Nga Kjwe	0	0	0	0	0	0	1	0	1	1	1	1
10	Hnjin Hsan	0	0	0	1	1	0	0	0	0	0	1	1
11	Pathein Pohsan Hmwei	0	0	1	0	0	1	1	1	1	1	0	1
12	Kjwe Thwa	1	1	0	1	1	0	0	0	0	0	1	1
13	Taun Paw Kjwe	1	0	0	1	1	1	0	0	1	0	0	1

Table 4. Ordinary Rice Varieties and their Chromosome Characters

No.	Rice Varieties	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1	Khun Ni	0	1	0	0	0	1	1	1	0	1	1	1
2	Shwebo Khun Ni	1	0	1	0	0	1	1	1	1	1	0	0
3	Bay Gyar Gyi	0	0	1	1	1	1	1	0	1	1	0	1
4	Shwe War Tun	0	1	0	0	0	0	1	0	0	0	0	0
5	Manow Thukha	1	0	1	1	0	0	1	0	0	1	1	1
6	Hsin Thukha	0	0	0	1	0	1	1	0	0	1	1	0
7	Hsin New Yin	0	0	1	0	0	0	0	0	0	1	0	0
8	Thee Htet Yin	0	1	0	0	0	0	1	0	1	0	0	0
9	Yatanr Toe	1	0	0	0	0	0	1	0	0	0	1	0
10	Zeayar	0	1	0	1	0	0	1	0	1	0	0	0
11	Hsin Ma	0	1	0	0	0	0	0	0	1	1	0	0
12	Hsin Shwe Li	1	0	0	1	0	1	1	1	0	1	1	0
13	Nga Sein	1	1	0	0	0	0	0	1	1	0	1	1

Selection of Local Rice Varieties for Reproduction

In this paper, there are **two local rice varieties** are selected as the parents for reproduction.

Recombination or Crossover Methods of Local Rice Varieties

It is the cross-combination of two chromosomes. It selects genes from parent chromosomes and creates a new offspring. The simplest way to do this is to choose randomly some crossover point and everything before this point is copied from the first parent and then everything after a crossover point is copied from the second parent. A single point crossover is illustrated in Figure 3, (| is the crossover point).

Chromosome A : 11111 | 00100110110
Chromosome B : 10011 | 11000011110
Offspring A : 11111 | 11000011110
Offspring A : 10011 | 00100110110

Figure 3. A Single Point Crossover Operation

Mutation of the Chromosomes of Each Rice Variety

This is the chance that a bit within a chromosome will be flipped (0 becomes 1, 1 becomes 0). In mutation, flip every bit with a specified probability. Usually one or only few bits should be mutated. Figure 4 shows a simple mutation operation.

Chromosome A : 11**01**1110000011110
Chromosome B : 11011**00**100110110
Offspring A : 11**00**1110000011110
Offspring B : 11011**01**100110110

Figure 4. A Simple Mutation Operation

Evaluation Procedure

The given fitness function is used to evaluate the population of string.

Case Study

The two Ordinary Rice Varieties; **Bay GyarGyi** and **HsinShwe Li**are selected as parent 1 and parent 2 and crossover with their **phenotype**.

Crossover Operations in Step by Step Generation

The crossover operation is step by step generation and run number per generation (i.e. generation 1-run number per generation 1, generation 1-run number per generation 2, …….., generation 1-run number per generation 100 to generation 10-run number per generation 1, generation 10-run number per generation 2, ……., generation 10-run number per generation 100). If the iteration of pre-defined generation and run number per generation is complete produce five possible new rice result or five possible new rice results and one expected result and if the pre-defined generation and run number per generation is stop before complete the iteration produce exact new rice result. First, maximum number of generation 1 to 10 and run number per generation 1 to 100 are operated as a case study. There are all together 1000 times are operated.

Results of the Crossover Operations

After crossover **Bay Gyar Gyi** and **Hsin Shwe Li** (as parent 1 and parent 2) with their **phenotype,** the operation results are shown in Table 5.

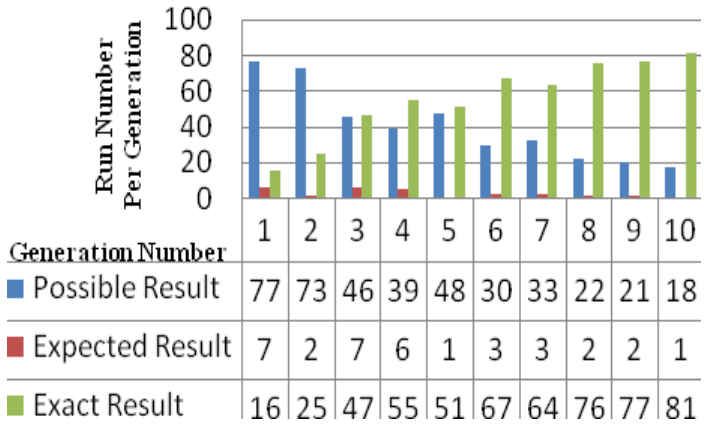
Table 5. Operation Results of Step by Step on Both Generation and Run Number per Generation

Generation	Possible Result	Expected Result	Exact Result	Total Operation Time
1	77	7	16	100
2	73	2	25	100
3	46	7	47	100
4	39	6	55	100
5	48	1	51	100
6	30	3	67	100
7	33	3	64	100
8	22	2	76	100
9	21	2	77	100
10	18	1	81	100

- Possible Result – Full iteration of pre-defined generation and run number per generation
- Expected Result – Full iteration of pre-defined generation and run number per generation
- Exact Result – Partial iteration of pre-defined generation and run number per generation

For the generation 1, run number per generation is 1 to 100 operated. According to the results, the possible results are more than the exact results. For the generation 2, 3, 4 and 5, the possible results are gradually reduced and the exact results are more and more. For the generation 6 and above, the exact results are more and more. In the generation 10, the new rice variety reaches the pure line. Therefore the operating the program is stop.

In step by step generation, there are totally 1000 times operated. Among them, the results are so many changes between 1 to 500 times (generation 1 to 5). Many changes occur because of the mutation contains in Genetic Algorithm. Figure 5 shows the changes of operation results in generation by generation. Above 500 times, New Rice Results are gradually closed to the pure line and stable. Therefore, if pre-defined generation is operated in above generation 5, the new rice variety will close to the pure line.



Generation Number	1	2	3	4	5	6	7	8	9	10
Possible Result	77	73	46	39	48	30	33	22	21	18
Expected Result	7	2	7	6	1	3	3	2	2	1
Exact Result	16	25	47	55	51	67	64	76	77	81

Figure 5. Three operational results of crossover operation

Discussion

After operation, there are three kinds of operation results. They are possible result, exact result and expected result. There are so many possible results, therefore these are do not record.

Exact Results of Step by Step Generation Operation

Parent 1 (Bay Gyar Gyi) 001111101101 and parent 2 (Hsin Shwe Li) 100101110110 produced exact new result 110101100011. Both parent 1(Bay Gyar Gyi) and parent 2 (Hsin Shwe Li) are bad (0) in chromosome 2 (Reveals a Breakdown of Rice). But in exact new result this chromosome 2 is good (1). Thus, ordinary rice (Bay Gyar Gyi and Hsin Shwe Li) are good for not breakdown of rice in New Rice. Figure 6 shows two parents' chromosomes and new rice chromosome.

Parent 1	001111101101
Parent 2	100101110110
Exact Result	110101100011

Figure 6. Two Parents' Chromosomes and New Rice Chromosome

Expected Results of Step by Step Generation Operation

In step by step generation, there are totally 1000 times operated. Among them, there are 34 expected results. By the comparison of the expected results with the exact results, it may be better than or worse or equal (exact result is 110101100011). Table 6 shows all expected results of step by step operation.

Table 6. Expected Result of Step by Step Operation

Generation	Run No per Generation	Expected Result	Generation	Run No per Generation	Expected Result
1	4	110111011101	4	14	110111111111
1	9	110111010111	4	17	110111111101
1	12	110111110111	4	23	110111111101
1	44	110101100011	4	68	110111111101
1	70	110111011110	4	91	111001011110
1	88	110111011111	5	5	110111011111
1	90	110111010111	6	1	110111010110
2	40	110111010111	6	8	110111011111
2	55	111001011100	6	19	110101100011
3	11	110111111111	7	9	110111111111
3	38	110111010101	7	28	110111010100
3	50	110111010100	7	34	110111010101
3	51	110111110101	8	14	110111010111
3	55	110101100011	8	64	110111011110
3	85	110111111101	9	4	110111111111
3	87	110111011100	9	14	110111111101
4	8	110111110101	10	11	110111110111

Conclusion

Rice is the most important food source feeds for half of the world's population. International Rice Research Institute (IRRI) from Philippine has made both laboratory research and field research for good quality and yield for new rice. The main food source of our country (Myanmar) is Rice. Therefore, Ministry of Agriculture tries to pay guidelines to the farmers for good quality and yield. Myanmar Rice Research Department performs field research from China and Philippine new rice varieties. A Genetic Algorithm is a search technique which emulates natural evolution, attempting to search for an optimal solution. Genetic Algorithm tries to evolve better solutions by selecting the best performing individuals (through the use of a fitness function), allowing only the best individuals to survive into the next generation through reproduction. This paper aimed by using genetic algorithm to produce the new rice varieties in the country.

References

Abraham, Ajith et al., (2009), "Evolutionary Computation: from Genetic Algorithms to Genetic Programming", *American J. of Engineering and Applied Sciences*.

Kumar, Anit, (2013), "Encoding Schemes in Genetic Algorithm", *International Journal of Advanced Research in IT and Engineering*, **2**(3).

Mathew, Tom V., Genetic Algorithm.

Myint, MyoKhin et al., (2012), "A PCR-based Marker for a Locus Conferring the Aroma in Myanmar Rice (*Oryza sativa* L.)", *Theoretical and Applied Genetics*, **125**.

Obitko, Marek, (1998), Introduction to Genetic Algorithms.

Online Materials

www.patentlens.com, 5.12.2013

www.pubmed.org

www.wikipedia.org