GENETIC ALGORITHMS (GAs) IN THE ROLE OF INTELLIGENT REGIONAL ADAPTATION AGENTS FOR AGRICULTURAL DECISION SUPPORT SYSTEMS

Gianni Jacucci, Mark Foy, and Carl Uhrik

Laboratorio di Ingegneria Informatica (LII); Università degli Studi di Trento
Via Fortunato Zeni 8; I-38068 Rovereto (TN) Italy

ABSTRACT

An AI adaptation methodology designed to assist in transporting agricultural models between regions is presented. A methodology to perform model adaptation (viz. localization) is frequently necessary when models are transported because models developed in one region often do not produce valid results when used in a different region. In this methodology, a GA plays the role of the adaptation agent. By linking a GA to an agricultural model, the model become more robust because it is able to adapt to the region in which the model is being used. This methodology has been implemented within a decision support system (DSS) developed within an EC project called SYBIL. The DSS helps farmers predict when diseases and funguses will attack their plants, so they can make intelligent decisions on preventing these attacks. Preliminary testing within this environment indicates this adaptation methodology has the ability to allow agricultural models developed in one area to be effectively utilized in other regions.

INTRODUCTION

In the domain of agriculture, the utilization of already developed models in a broad area is often hindered by one or more factors. One frequent factor which impedes transportation is model inaccuracy. For example, when models that perform well in one region, are transported to be used in a different region, they often do not give accurate output (such as, recommendations, results, and/or indicators) in their new environment (i.e., when they are run in a new region).

This is one of the major difficulties of *model technology transfer*. To address this difficulty, an artificial intelligence (AI) methodology is proposed. In this methodology, a GA plays the role of an intelligent adaptation agent by adjusting agricultural risk assessment models to particular regions. The general component created by this combinational methodology is called an 'Agricultural Model-GA' or an AGMOD-GA.

A decision support system (DSS) (which addresses grape and apple management) developed under an EC project called SYBIL has utilized this AI adaptation methodology. The models included within this DSS are designed to help growers determine when it is necessary to spray against certain pests. One of the models in this DSS is called the P.R.O. model, and the methodology has been applied to this model to create a PRO-GA component.

The following sections will give an overview of the adaptation methodology, discuss under what conditions this methodology can be applied, describe the methodology's elements, and then give details about linking a GA to a model under this scenario. Finally, conclusions will be given.

OVERVIEW OF THE AI ADAPTATION METHODOLOGY

The theory of this AI adaptation methodology is that by utilizing historical data from a particular region, a model's parameter settings can be adapted so that the new parameters allow the model to work well in the particular region. This adaptation is done by trying to match the model parameter settings (which should be easy to change) to the particular region. That is, new model parameter settings are found which allow this model to give accurate output values (such as, recommendations, results, and/or indicators) when run in the particular region. To find matching model parameter settings, intelligent search is performed which utilizes historical outcome data as part of the objective function (i.e., the search attempts to fit the model parameter settings such that when used in the model, the model gives output near to the given historical outcome data).

For example, suppose a model has been developed in **RegionA** that gives an assessment of the risk that a certain plant will be damaged by a certain fungus. This model uses an instantiation of model parameter setting values (such as values for thresholds and coefficients, etc.) called **ModelParmsA**. The model has proved quite useful in **RegionA**, and now agriculturalists in **RegionB** are interested in using this model. Upon first running this model in **RegionB** (with the original model parameter settings, i.e., **ModelParmsA**), the model gives output values (recommendations, results, and/or indicators) which are inaccurate. As an attempt to make this model work accurately in **RegionB**, new model parameter settings (which are different from **ModelParmsA**) could be derived. By applying the methodology discussed here (which requires historical data and an intelligent search method to be described below), it is possible to locate new model parameter settings, and once this is done, this new instantiation of settings can be called **ModelParmsB**. After applying this methodology, the model should produce accurate output values in **RegionB** when using **ModelParmsB**.

Overall, by following this methodology, a component will be created which can search for good model parameter settings such that when the given model is applied and run at the location in question, the output values given will be consistent with the historical outcome data. Moreover it is hoped that this will also allow the model to be generally used in this region, producing accurate output values on data which it has not seen. This component that performs this search/adaptation can be called an expert system component since it modifies and adjusts a model to work in a new location in the same way an expert would modify and adjust a model.

Additionally, it should be emphasized that this methodology is particularly appealing because it is not a strictly empirical or analytical, but both. That is, this methodology does not perform a search to fit the historical data from a particular location into an empirical algorithm; rather it performs the search in a larger context, fitting the model parameter settings to a particular location. Therefore, the resulting instantiation of the adapted/localized agricultural model (with the new parameter settings inside) is as good (or as bad) as the original model. Consequently, if the model is biologically significant (e.g., if it simulates biological events), then this feature is not lost by this adaptation methodology since the model is used in the same form (i.e., the structure of the model is left intact), only the model parameter settings are changed.

Before describing the elements of this methodology, a discussion of the limitations of the methodology will be given

APPLICABILITY OF THE METHODOLOGY

The difficulty of transporting models from one location to another is not limited to being explain in one way (i.e., one or more different reasons could explain model inaccuracy in a new location). Likewise, the methodology described here is not applicable to all models that have difficulty being transported; therefore, not all models that have difficulty being transported will necessary benefit by following this methodology.

Distinguishing models that can be treated by this methodology from those that can not, agricultural models can be divided into two cases:
(Case1) models that are not transportable because they are too location-specific and not robust enough to allow different conditions to enter into the model and
(Case2) models that are transportable, but the model parameter settings need to be altered to allow the model to give accurate output values for the new region.

In **Case1**, the model is just too specific to operate outside one area. This situation is difficult to address and solve if there is a strong desire to transfer the model. This case will not be addressed here. On the other hand, **Case2** can be addressed by the adaptation methodology described here which determines location-specific model parameter settings so that the model functions accurately in a new location.

ELEMENTS OF THE METHODOLOGY

Generally stated, this methodology prescribes the utilization of:
(i) historical situation data,
(ii) historical outcome data,
(iii) the model, and
(iv) an intelligent search method (in this case, a GA).

Historical situation data is the basic data required by the model in question. In the domain of agricultural models, this often includes meteorological data since this is frequently an important input to the model. In this methodology, the more historical situation data that is available, the better.

The other type of historical data needed is **historical outcome data**. The methodology assumes that for every piece of historical situation data, there is a corresponding outcome. Both the situation and the outcome are observable in the real world. The model accepts situations and computes outcomes. For risk assessment models, historical outcome data regards the occurrence of fungus or pest problems in past years (epidemiological data); or for crop growth models, historical outcome data regards crop yield in past years. Thus, historical outcome data is a particularly notable part of this methodology because it is used to *fit* the model parameter settings to the new region in question. Therefore, when constructing a component using the methodology described here, it must be possible to match model outputs to some combination of outcomes and/or events in the real-world. For example, with a risk assessment model, if the only historical epidemiological outcome data available is a set of historical primary infection dates, then the model should be able to produce a primary infection date (albeit a guess of the primary infection date) as one of it's outputs. In this way, it will be possible for the model to be adapted such that the primary infection date produced by the model nearly matches a set of actual

primary infection dates (i.e., dates specified by the historical outcome data).

In this methodology, the agricultural **model** is fundamental because it will be used to evaluate how well particular model parameter settings work in the given region (i.e., with the given historical data). In particular, the intelligent search method will repeatedly call upon the model as it constructs new model parameter settings that need to be evaluated.

This methodology has the capability to address many types of agricultural models: risk assessment models, damage prediction models, crop growth simulation models, and other models that use meteorological data.

The **intelligent search method** is an important part of this component because, in this particular domain of agricultural models, knowledge of the domain is often hard to codify (i.e., 'rules of thumb' are vague and difficult to construct), and the selection of an intelligent search method can help to alleviate this difficulty. This is due to the fact that intelligent search methods do not rely on 'rules of thumb'.

The selection of the actual intelligent search method to be employed was made among the following possible methods: hill-climbing, simulated annealing, and GAs. In the end, GAs were selected because of their unbiased, effective search capabilities and their applicability to real-world problems [1] [2] [3] [4].

UTILIZING THE AI ADAPTATION METHODOLOGY - THE AGRICULTURAL MODEL-GA (AGMOD-GA)

Linking an Agricultural Model to a GA

To allow a GA to search the space of an agricultural model's parameters, the GA uses the model as the evaluation function. Furthermore, the model uses the historical situation data (as this is necessary to run the model in the given historical years), and the GA additionally uses the historical outcome data (discussed earlier) in combination with the output of the model. Whenever the GA wants to evaluate one instance of model parameter settings, the agricultural model is called, and the final outcome is returned through an objective function to the GA so that a "fitness" can be computed. This resulting general component is called an 'Agricultural Model-GA' or an AGMOD-GA.

Figure 1 illustrates an agricultural model and a GA linked to form an AGMOD-GA.

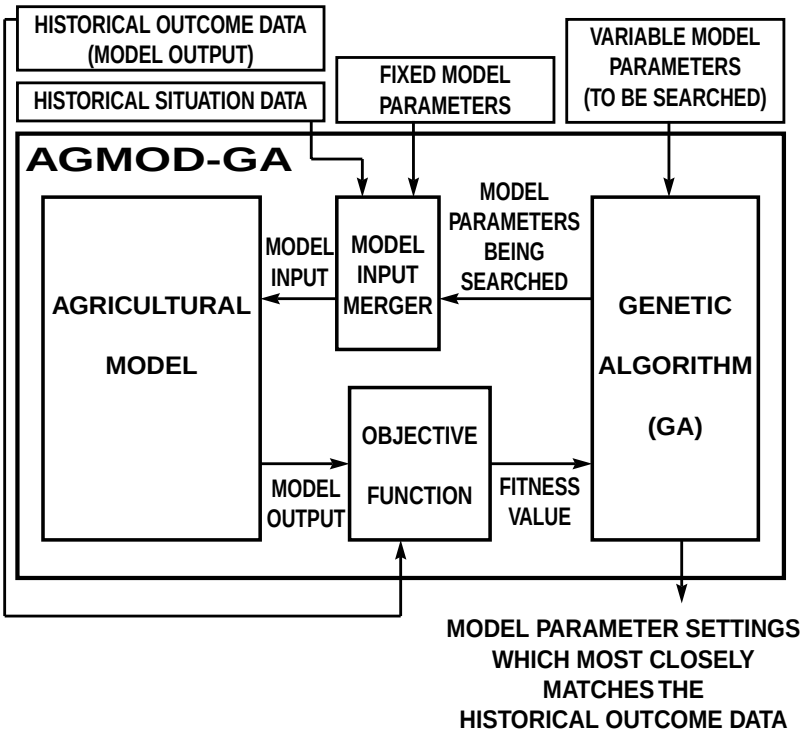

Fig. 1. Structure of an AGMOD-GA

The function of the AGMOD-GA is to find near-optimal model parameter settings for the given desired behavior (i.e., matching the given historical outcome data). There are three main steps involved in the execution of a typical AGMOD-GA. First, the agricultural model and the GA are initialized. The second step in a typical AGMOD-GA is the fitness computation. This involves taking each GA population member and executing one or more model simulations using the model parameter settings represented by this member. This fitness evaluation step is executed many times because new population members are continually being generated by the GA. Fitness evaluation is usually continued until the GA has converged on a suitable optimal or quasi-optimal solution. The last step is the evolution of the GA population. This involves applying operations to the population members. The three operators used in a typical GAs are reproduction, crossover, and mutation. They act by treating the GA bit strings (which represent model parameters) in a way analogous to the evolution of chromosomes in genetics [1].

AGMOD-GA Performance

When particular instantiations of this adaptation methodology are employed, AGMOD-GA search runs provide the user with model parameter settings that the operator can use to ameliorate the original model. Instantiations of the AGMOD-GA perform similarly to most GAs, in that the average fitness of a generation, over time, will increase. That is, the members of later populations will converge on maximum fitness members of the space (in this case, on optimal sets of model parameter settings). This performance of a typical AGMOD-GA is shown in Figure 2. The two curves ("POPULATION MAXIMUM FITNESS" and "POPULATION AVERAGE FITNESS") illustrate the GA converging on high fitness members. This type of performance has in fact been achieved on an instantiation of the AGMOD-GA called the PRO-GA. This PRO-GA will be discussed in the next section.

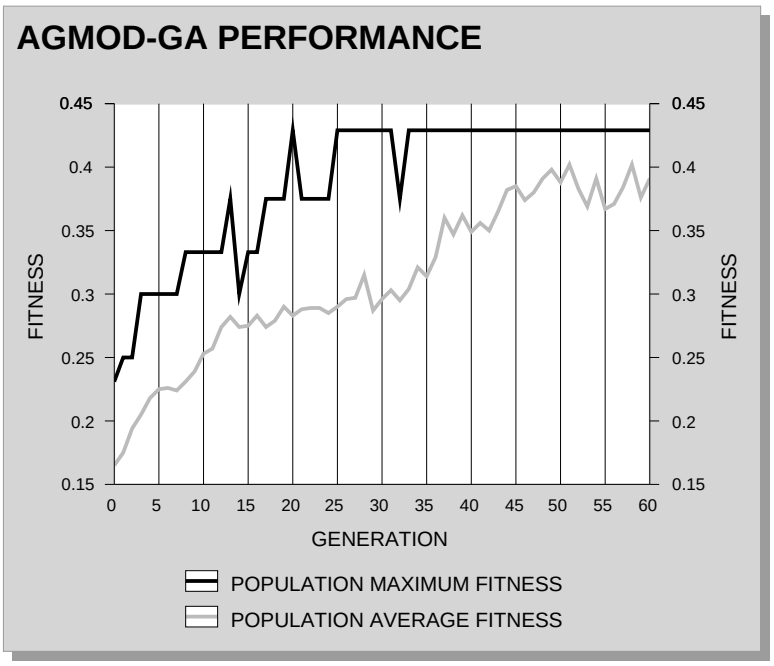

Fig. 2. Typical AGMOD-GA Performance
(Note: The values plotted are from a PRO-GA run)

An Example: The PRO-GA

The AI adaptation methodology described has been employed within a DSS developed under an EC project called SYBIL. In particular, this methodology has been utilized within one of the developed DSSes to adapt the P.R.O. model (i.e., find localized model parameter settings for the P.R.O. model). The P.R.O. model is a biological life cycle model that simulates the growth of downy mildew (viz. Plasmopara viticola, also called peronospora) on grape vines. This model is designed to help growers determine when it is necessary to spray against downy mildew. The model has been used in Germany to help grape growers reduce the amount of fungicide needed to control this fungus [5] [6] [7] [8] [9] [10]. The AI component intelligently searches the space of P.R.O. model parameter settings and locates settings that match local conditions. Preliminary testing with the P.R.O. model indicates that this adaptation methodology has the capacity to allow the parameter values of regional models to be effectively adapted to regions other than where they were developed [11].

CONCLUSION

This AI methodology addresses *model technology transfer* (i.e., the moving of agricultural models that are developed in one location to a new location so they can be used in this new location). In particular, it addresses one of the major difficulties within this area, namely, model accuracy; i.e., it addresses the instance when a useful model is transported from a region where it is functioning accurately (viz. producing accurate recommendations, results, and/or indicators) to a new region where it subsequently does not function accurately.

The methodology employs four main elements, with a GA at the center. By employing this AI component in conjunction with the engine of an agricultural model and historical data, model parameter settings can be adapted to new locations, allowing the model to give accurate results when run in the new location. Specifically, the module created by this methodology can be applied to localize models by deriving new model parameter settings that can be employed in the particular location to give good suggestions/decision support.

With the assumption that model technology transfer is an advantageous action (refer to [11] for an elaboration of advantages and disadvantages in transporting models between regions), this AI methodology has been found to efficiently addresses this issue and improves the current state-of-the-art in model technology transfer.

This is briefly discussed here through an example which describes the utilization of this methodology with the P.R.O. model which is contained within one of the decision support systems (DSSes) developed under an EC project called SYBIL. The resulting PRO-GA component has been used to produce new model parameter settings, and testing with these settings has shows that this methodology has great potential to localize model parameter settings, and this should assist in achieving the goal of making sound models more widely available.

REFERENCES

[1] Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA, USA.

[2] Grefenstette, J.J. 1985, ed. Proc. of an International Conference on Genetic Algorithms and Their Applications. Carnegie-Mellon University, Pittsburgh, PA, USA, July 24-26, 1985. U.S. Navy Center for Applied Research in Artificial Intelligence, USA.

[3] Grefenstette, J.J. 1987, ed. Proc. of the Second International Conference on Genetic Algorithms. Massachusetts Instituted of Technology, Cambridge, MA, USA, July 28-31, 1987. L. Erlbaum Associates, USA.

[4] Schaffer, J.D. 1989, ed. Proceedings of the Third International Conference on Genetic Algorithms. George Mason University, Fairfax, VA, USA, June 4-7, 1989. Morgan Kaufmann Publishers, Inc., USA.

[5] Hill, G.K. 1989. Mit P.R.O. System Spritzungen Sparen. Die Weinwirtschaft, 125, pp. 28-31.

[6] Hill, G.K. 1990. Das Peronospora-Riskoprognosemodell Oppenheim (P.R.O.) im Praxistest. Der Deutsche Weinbau, 44 (12/27 April 1990), pp. 514-517.

[7] Hill, G.K. 1990. Studies on Plasmopara Viticola Epidemics in the Vineyard. Proc. Atti del Convengo su Modelli Euristici ed Operativi in Agricoltora, Caserta, Italy, 27-29 Sept. 1990, Societa Italiana de Fitoiatria ed., pp. 266-273.

[8] Hill, G.K. 1990. Plasmopara Risk Oppenheim - A Deterministic Computer Model for the Viticultural Extension. Proc. Atti del Convengo su Modelli Euristici ed Operativi in Agricoltora, Caserta, Italy, 27-29 Sept. 1990, Societa Italiana de Fitoiatria ed., pp. 182-194.

[9] Hill, G.K. 1993. Severe Primary Infections of Plasmopara Viticola at bloom-Causal Climatic Conditions and Threshold of Damage on Bunches. Proc. OILB Meeting on Integrated Control in Viticulture, Bordeaux, France, 2-5 March, 1993.

[10] Hill, G.K., Breth, K., Spies, S. 1993. The Application of the P.R.O. - Simulator for Minimizing of Plasmopara Sprays in the Frame of an Integrated Control Project in Rheinhessen/Germany. Vitic. Enol. Sci., 48, pp. 176-183

[11] Jacucci, G., Foy, M., and Uhrik, C. 1994. SYBIL DSS:Localization of Agricultural Risk Assessment Models. Proceedings of Decision Support 2001 (DSS2001) (The 17th Annual Geographic Information Seminar & The Resource Technology ’94 Symposium), Toronto, Canada, 12-16 Sept. 1994, American Society of Photogrammetry and Remote Sensing.