

A hybrid machine model of rice blast fungus, *Magnaporthe grisea*

Mike Holcombe^{a,*}, Lucy Holcombe^b, Marian Gheorghe^a, Nick Talbot^b

^a Department of Computer Science, University of Sheffield, Sheffield S1 4DP, UK

^b Department of Biological Sciences, University of Exeter, Exeter EX4 4QJ, UK

Abstract

The fungus, *Magnaporthe grisea* (Rice blast fungus) is a major agricultural problem affecting rice and related food crops. The way that the fungus invades the host plant and propagates itself is a very important scientific problem and recent advances in research into the genetic basis of these processes can be used to build a simple partial model using hybrid computational modelling techniques. The possible potential benefits of doing this include the use of computer simulation and automated analysis through techniques such as model checking to understand the complex behaviour of such systems. The example is a metaphor for the process of trying to integrate and understand much of the vast amounts of genomic and other data that is being produced in current molecular biology research.

© 2002 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Computational model; State machine; X-machine; Agent; Hybrid model; Fungal infection; Genomics; Bioinformatics

1. Introduction

Computational models have been of interest in biology for many years and have represented a particular approach to trying to understand biological processes and phenomena from a systems point of view (Holcombe, 1991; Krohn et al., 1967). Much of the early work was rather abstract and high level and probably seemed to many to be of more philosophical than practical value. There have, however, been some advances in the development of more realistic models and the current

state of computer science research provides us with new opportunities both through the emergence of model types that can model seriously complex systems but also the support that modern software can give to the modelling process.

2. Modelling continuous state-based phenomena

Finite state machines and their generalisations, such as X-machines (Holcombe, 1982; Eilenberg, 1976; Holcombe and Ipate, 1998; Balanescu et al., 1999), are examples of discrete computational models that operate in finite environments (finite input sets, finite output sets and finite memory variables). They are suitable for modelling many types of system. However, they can only model

* Corresponding author. Tel.: +44-114-222-1802; fax: +44-114-222-1810

E-mail address: m.holcombe@des.shef.ac.uk (M. Holcombe).

instantaneous processing and only finite discrete data is processed. Continuous functions and real valued data cannot be incorporated into traditional finite state machine models. Such systems are problematic when trying to deal with the complexities of some biological models and the hybrid X-machine (Duan et al., 2000), overcomes some of these problems.

A hybrid machine has states and transitions as usual and responds to discrete events and performs discrete actions, which are observable. The internal memory consists of:

- a set of discrete variable; and
- a set of continuous variables.

The way the model works is that we identify a number of significant states of the system, each of which describes some aspect of the system's behaviour. At certain moments the system changes state and moves to another state where different activities occur. The events and conditions that prompt state changes are explicitly identified as are the sorts of processing that goes on in each state. All of these make use of the identified variables, both continuous and discrete and are expressed in terms of mathematical equations and properties.

A key aspect of the hybrid machine is the memory, a collection of variables, continuous and discrete which help to model the cellular metabolism and structure.

An important component of the memory is a description of the internal structure of the cells and one way that this might be achieved is through the use of a *membrane system* or a *p-system* (Paun, 2000). We hope to adapt the idea of a membrane system to our hybrid context in a later paper. For the present the memory is not structured in any specific way.

When the system is in a given state there are sets of equations that apply to the system's continuous variables and all the while it is in that state, with time progressing, these variables change according to these equations.

When either an appropriate external event occurs or a leaving condition is met (e.g. a set

point) the system moves to its next state where a different set of equations takes over.

This sort of generic model can model many types of biological phenomena but needs refining to deal with some important issues.

The overall system is then described by using a modified state transition diagram, see Fig. 1, which describes the main states of the system; the transitions between these states; the equations pertaining to each state; the events which cause a state change—this could be either an external event or an internal leaving condition; and the results of the transition which will effect either internal discrete variables or external properties of the system. In some cases the transitions will be prompted by signals derived either internally or by the communication between one part or component of the system, or by some external signal perceived in interaction with the system's environment—other cells in the organism or signals from outside the organism.

Some simple examples that can be modelled this way include:

- Ion flow through voltage gated channels.
- Antigen–antibody interaction (Duan et al., 2000).

The continuous variables can exhibit complex behaviour as shown in Fig. 2.

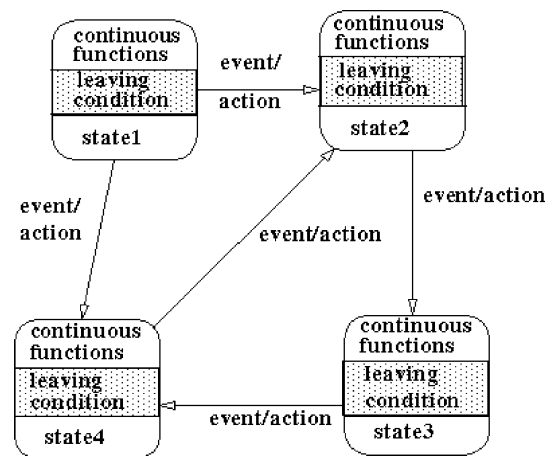


Fig. 1. A hybrid machine.

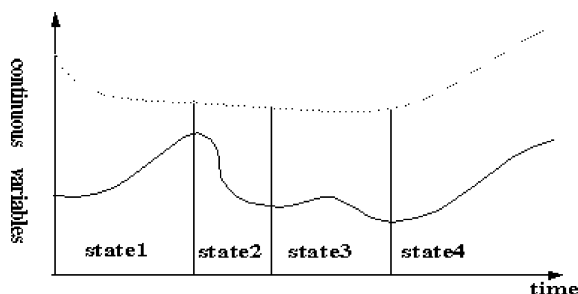


Fig. 2. Continuous variable in a hybrid machine.

The equations are often composed of relatively simple functions compared with the equations that try to describe the complete functions over all states. This is an advantage of a state-based approach and can be exploited in a variety of ways. Furthermore, these functions might be decomposable into lower level hybrid machines thus providing a mechanism for dealing with the undoubted complexity that such systems exhibit.

3. Hybrid machines and hybrid logic

The basic parameters of a model include a finite set, M of macrostates which describe the way that the model is decomposed into coherent phases of activity during which some continuous processing is carried out, governed by some suitable set of equations that model the way that the variables change over time under the specific conditions associated with the state, $m \in M$.

These equations will involve a number of continuous variables, we regard them as part of the memory of the machine, along with a number of discrete variables.

The logic is a symbolic language that enables precise mathematical formulae to be expressed.

$$p = o_c(x = x_0 e^{kt}) \wedge (y = y_0) \wedge \text{halt}_c(t = L)$$

which is read as: all the time that the system is in this state then the equation $x' = x_0 e^{kt}$ holds for the variable x until the halt condition occurs when variable t reaches the value L .

This would describe, for example, the operation of a very simple model of an immune system where x represents the population of an antigen intro-

duced at time $t = 0$ and L represents a time delay before a response sets in.

This would then trigger a state change to another state in which antibody responses start to operate. A new set of equation models this.

$$q = o_c(x = x_0 e^{kt} - y) \wedge (y = y_0 e^{rt}) \wedge \text{halt}_c(x = 0)$$

Which translates as: all the time that the system is in this state the equations:

$$x = |x_0 e^{kt} - y|$$

$$y = y_0 e^{rt}$$

hold until the halting condition described by the removal of all the antigen, $x = 0$ (k and r are parameters, $k < r$).

A typical property that one might want to prove is that the condition $O_c(y' = 0 \Rightarrow x = 0)$ i.e. at the next state if there is no antibody growth then the infection is over.

4. A case study of a hybrid machine to model fungal infection of plants

4.1. *Magnaporthe grisea* (Rice blast fungus), (Talbot and Foster, 2001)

Rice is one of the world's most important food crop. This fungus can destroy up to 40% of crops. The way that this fungus infests rice plants is a major area of research, which has made significant advances in understanding the genetic basis of this process. This is a partial model of the infection stage.

It is a hybrid machine and uses some of the most recent information about the genetic basis of the behaviour of the fungus.

The spore or *conidium* is a three celled structure, Fig. 3 which is present in the atmosphere in affected areas. These alight on the surface of rice leaves, normally contained within a dew drop, and attach themselves to the surface, this is possible despite the fact that the leaf surface is highly hydrophobic and is achieved by the conidium releasing from its tip a powerful adhesive stimulated by wetting. The spore then germinates and

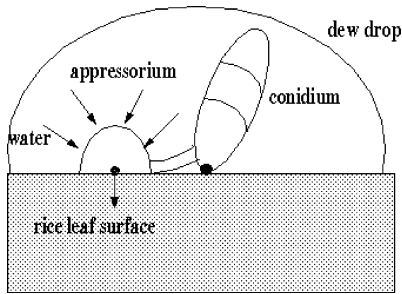


Fig. 3. The Magnaporthe infection process (figure after Talbot, 1995).

produces a germ tube from one of the terminal cells of the conidium which then forms a hook and adheres to the leaf. An *appressorium* of a roughly hemispherical shape then develops at this point of contact. The penetration of the rice leaf surface is carried out by the build up of pressure within the appressorium. The pressures generated are as high as 8 MPa (or 40 times the pressure in a car tyre). As a result of this enormous pressure a penetration peg, formed at the part of the appressorium where it joins the leaf surface, is forced to (mechanically?) penetrate the leaf surface. Once penetration has been achieved the fungus then forms cylindrical cells called hyphae which initially spread into plant cells without causing damage or overt disease symptoms, but later produce toxic compounds and degradative enzymes which cause plant cell death. The fungus causes necrotic disease lesions on leaves, which can be seen as dark oval spots, each represents the point of a single appressorium-mediated infection and when they coalesce in heavy infection, whole leaves or entire seedlings can die (Talbot and Foster, 2001). The fungus produces conidia from disease lesions which propagate the fungus to new plants (Talbot, 1995).

Fig. 4 provides a top level hybrid machine model of the life history of the fungus. The internal variables (or memory) provide information about the status of various internal aspects of the fungus, for example the initial concentration of glycogen, glycerol etc. have to be modelled as variables. The build up of glycerol is thought to be responsible for the generation of appressorial turgor. One potential source of this glycerol is from the breakdown of glycogen. So we need to describe

the internal turgor pressure of the appressorium generated by the concentration of glycerol.

Each state has either a leaving condition, some condition that has to be satisfied by some internal parameter in order for a state change to occur, or there is some external event that triggers the state change.

When the spore lands on the leaf, the release of its glue to attach to the surface is a passive process caused simply by the presence of water. This triggers germination and germ tube development. The break of dormancy is a genetic signal, but derives from a passive external effect (presence of water). The dormancy breaking signal would trigger gene expression involved with polar growth and further adhesion.

Some state changes will be triggered by the recognition of some intracellular signal and the switching on of some specific genes which then become active in the new state and control the metabolism until the next state transition. In the diagram we indicate this thus:

signal4 / *MAGB*, *MAC1*, *PMK1*

where *signal4* is some internal signal indicating that the hook formation process has completed and the genes *MAGB*, *MAC1*, *PMK1* become more active, that is the protein products of these genes that become active, not the genes themselves. These genes are probably always switched on at low levels, but the proteins react to and transduce an inductive signal to the cell's nucleus leading to new gene expression. This would trigger genes required for movement to stages 4 through to 7.

A state transition involving a leaving condition might be state 8 to state 9 where the leaving condition is the internal appressorium pressure reaching a set point sufficient to cause the puncturing of the host leaf surface.

Whilst in state 8 this pressure is undergoing change mediated by some suitable set of equations determined by the inflow of water through the appressorium surface.

If this surface is regarded as a hemisphere and the porosity of the surface is constant throughout then the rate of increase in mass, the pressure is given by the linear equation:

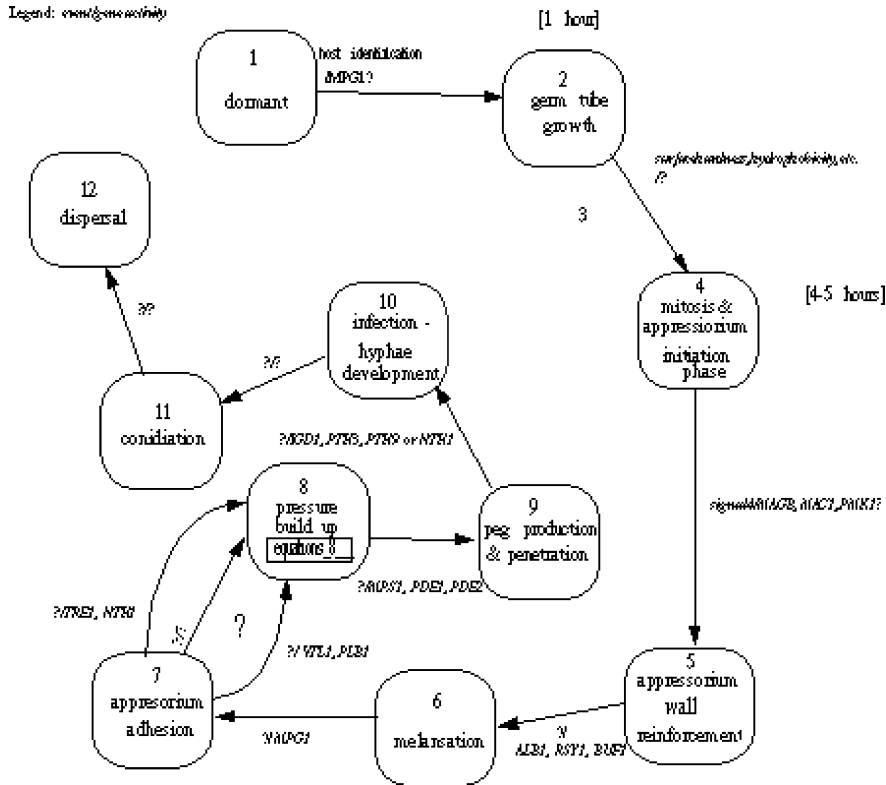


Fig. 4. A possible Magnaporthe hybrid machine.

$$p(t) = \frac{\rho t}{r} + \text{initial pressure.} \quad (1)$$

During state 8, then, the turgor pressure is increasing linearly, assuming that there is sufficient water surrounding the appressorium, until the set point is reached at which time the transition to state 9 occurs, thus triggering further gene activation for the next phase of development. In state 8 there are a number of important processes which determine the production of glycerol. Genes involved in the breakdown of lipids, glycogen and trehalose (a fungal storage carbohydrate) are all thought likely to be involved in the production of glycerol. For example, the genes *GPH1* which encodes glycogen phosphorylase and *AGL1* which encodes an amyloglucosidase are involved in breaking down glycogen into glucose units which

can then be used to make glycerol. The equations operating during state 8 (Eq. (1)) will describe the metabolic activity relating to glycerol production. Stage 8 may require a different signalling pathway, perhaps also involving cAMP. *GPH1* and *AGL1* are targets of this pathway, rather than effectors. They bring about the change in metabolic state, however, acting in concert with a number of other enzymes that leads to state 8.

There are a number of gaps in this model which will be filled as further research into the genetic and molecular basis of the disease is carried out. It is essentially a high level model which needs to be developed in a hierarchical way so that the individual transitions actually involve complex hybrid submachines and these, themselves, will also break down into lower level structures until we reach an appropriate representational level.

5. Conclusions and further work

One fundamental problem in modelling such complex systems as biological systems will be trying to understand the complex interactions between many subsystems and the vastly complicated molecular and genetic activity that exists. We might be able to build these models but will we be able to understand and analyse them? It is likely that we will only be able to do this if we simplify them greatly. As an alternative approach we developed the Hybrid Projection Temporal Logic (HPTL) (Duan et al., 2000) specifically for hybrid machines. This logic allows us to define such a machine in a precise formal logic which is the first step towards using automated reasoning techniques. The basic process involves trying to establish properties about the model, now represented as a logical formula in HPTL. There are two, related, ways of doing this. First, we could try to prove theorems about the system by using theorem proving engines, this is probably impractical since the success of automated theorem provers in dealing with extremely complex systems is limited. An alternative approach is the use of model checking techniques, (Clarke et al., 1986) either alone or in combination with theorem proving. This is potentially feasible and would allow us to ask ‘what if’ questions and query whether the system could ever get into a state with a given property holding etc. This is more feasible since model checking technology can handle models with very large state spaces. However, the technology needs to be substantially extended to cope with hybrid machines of this type. It does, however, offer a potentially rewarding direction for research.

We could, in the mean time, use simulation, in virtue, to run these models and derive some useful information about the system from it. Software

already exists that could be used to simulate these models.

Once we have such a model we can then try to use machine reasoning—model checking—to automatically discover properties and answer questions. Admittedly this is long term research—we may be able to do something realistic in 5 years, definitely by 10. It is a framework for modelling and analysis that is needed if we are to manage the complexity of the models that we are going to have to deal with.

References

- Balanescu, T., Holcombe, M., Cowling, A.J., Georgescu, H., Gheorghe, M., Vertan, C., 1999. Communicating stream X-machines systems are no more than X-machines. *J. Universal Comput. Sci.* 5 (9), 494–507.
- Clarke, E., Emerson, E., Sistla, A., 1986. Automatic verification of finite state concurrent systems using temporal logic specifications. *ACM Trans. Prog. Lang. Syst.* 8, 244–263.
- Duan, Z., Holcombe, M., Bell, A., 2000. A logic for biological systems. *Biosystems* 55, 93–105.
- Eilenberg, S., 1976. *Automata, Languages and Machines*, vol. B. Academic Press.
- Holcombe, M., 1982. *Algebraic Automata Theory*. Cambridge University Press.
- Holcombe, M., 1991. Mathematical models of biochemistry. In: Ji, S. (Ed.), *Molecular Theories of Cell Life and Death*. Rutgers University Press, pp. 250–263.
- Holcombe, M., Ipate, F., 1998. *Correct Systems—Building a Business Process Solution*. Springer.
- Krohn, K., Langer, R., Rhodes, J., 1967. Algebraic principles for the analysis of a biochemical system. *J. Comput. Syst. Sci.* 1, 119–136.
- Paun, G., 2000. Computing with Membranes. *J. Comput. Syst. Sci.* 61 (1), 108–143.
- Talbot, N.J., 1995. *Trends Microbiol.* 3 (1), 9–16.
- Talbot, N.J., Foster, A.J., 2001. Genetics and genomics of the rice blast fungus *Magnaporthe grisea*: developing an experimental model for understanding fungal diseases of cereals. *Adv. Bot. Res.* 34, 263–287.