Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

BioSystems 85 (2006) 46–54

BioSystems

www.elsevier.com/locate/biosystems

The Paton individual-based model legacy

R. Gregory[a]*, J.R. Saunders[b], V.A. Saunders[c]

[a] *Department of Computer Science, University of Liverpool, Chadwick Building, Peach Street, Liverpool L69 7ZF, United Kingdom*
[b] *Microbiology and Genomics Division, School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, United Kingdom*
[c] *School of Biomolecular Sciences, Liverpool John Moores University, Liverpool L3 3AF, United Kingdom*

Received 16 February 2006; accepted 17 February 2006

Abstract

Ray Paton oversaw the creation of a long lineage of Individual-based Models (IbMs) and this paper discusses the five most successful. All of these concern the development of adaptation, covering both evolutionary time and organism lifetime (somatic time). Of the five models discussed here, the first is based on a plant–herbivore model, the other four are based on a substrate–bacteria model, with the option of antibiotics.

Keywords: Evolution; Individual-based model; Liverpool history; Bacteria; Classifier ; State machine

---

1. Introduction

This paper reviews the long history of Individual-based Models (IbMs) developed at Liverpool under the guidance of Ray Paton. It is unclear precisely what introduced Ray to the IbM approach. However, with his biological background and early papers in statistics (Paton, 1990, 1991a,b; Meyer and Paton, 1992), it was presumably intuitive to him that statistics, logic and differential equations (DEs) are usually inadequate at describing the processes found in living organisms. Ray was not afraid to go against the established norms and only in hindsight is it clear that he was correct in his view that IbMs were the obvious path to have taken.

Logic and differential equation-based models have been entrenched in the minds of modellers. Not all modellers were so restricted and workers in this area could well have taken inspiration from the early simulations of Conrad and Pattee (1970). These were extremely crude and abstract, but encompassed some of the essential ideas, namely individuality and ecology. The Individual-based Model is a philosophy that embraces the uniqueness of the individuals in a system. Simulations are composed of multiple individuals, with each individual having its own set of state variables and parameters. Looking in state space, the population is akin to clouds of individuals with similar behaviour, and other clouds, amounting to separate species. Fundamentally, this allows for species to exist, speciation to occur and permits extinction. Of course, the modeller decides at what level the philosophy is to be applied and, for practical reasons of computability, a given model need not be entirely individual-based. This is especially true when the model covers many scales of biological process, under which circumstances the power of the computer seems somewhat inadequate.

The overwhelming advantage of the IbM is that evolution becomes open-ended. IbM moves away from traditional techniques that start by assuming *x* many species exist (looking at the macro level) and then observing population shifts. IbM brings together the macro and the

---

* Corresponding author. Tel.: +44 7796955729; fax: +44 666666666.
E-mail addresses: greg@csc.liv.ac.uk (R. Gregory), jrs@liverpool.ac.uk (J.R. Saunders), v.a.saunders@ljmu.ac.uk (V.A. Saunders).

0303-2647/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.biosystems.2006.02.009

micro scale. Thus, as a simulation runs, one can visualise the emergence and death of a given virtual "species", and can decide at the time what exactly defines a separate "species". Individuals are only considered together when analysed, if it is relevant so to do. Moreover, the individuals in the simulation can be analysed and tracked as individuals. An individual bacterium can be monitored so that its movements, interaction with other bacteria and any infection events (for example by a virus) are known. The reasons for the behaviour of that bacterium can then be determined. Average behaviour can normally be seen in population-based data, but only individual-based data can capture unusual events.

The IbM also has the characteristic that dissimilar processes can be coupled together. Dynamic numbers of time-based events and dynamic numbers of analogue parameter changes will work together with stochastic and crisp process control logic. This makes IbM ideal for biological simulations, as knowledge may be patchy, unreliable and explained with mutually incompatible metaphors. In turn, knowledge integration is possible, though not necessarily trivial. The resultant behaviour from the IbM rules can often lead to complex behaviour, even when the rules appear simple. The rules themselves need only be elementary if–then statements, but in an IbM there would normally be implicit feedback and interaction between rules. Of course there are problems with the IbM approach, computability being a frequent obstacle. Individuality can be added at any point to a model, but each individual bears a potentially large set of state variables, some of which could themselves be based on individuals. The required storage space can then suddenly explode and require impractical amounts of RAM and potentially impractically large amounts of permanent storage space. As a result, there should be a perverse tendency not to use IbM in an IbM model unless it is expected to be useful.

The IbM is an inclusive philosophy that allows incompatible modelling processes to be brought together and considered. The problem comes when those processes interact and generate incorrect behaviour. Identifying interaction is non-trivial as the whole system is a feedback loop, where all processes have had some (possibly small) influence on the next iteration. This is related to the problem of testing in general. The simulation is heterogeneous in terms of processes and interactions and so is described using many lines of program code. The chances for error are then higher and the testing must be much more rigorous.

In the following sections, we introduce some of the IbM projects developed by Paton and his co-workers. All of these models concerned populations of organisms and their adaptation to the environment, over a combination of evolutionary time and organism lifetime. As shown in Fig. 1, the first in a line of models was HERBY (Devine et al., 1996), which was an ecological simulation where agents eat and move around in a discrete grid world. The underlying control system in this case was a classifier system (Bull, 2004), which gave the agents the ability to co-evolve survival strategies. The COSMIC project (Gregory et al., 2004a) started as HERBY was coming to an end. COSMIC was more ambitious, as it modelled bacterial evolution in a continuous world. Gene transcription cascades were used to generate the control logic for chemotaxis (movement of organisms towards or away from chemical stimuli). Here the IbM philosophy was manifested on two levels of scale, as both the bacteria and the gene products were individuals.

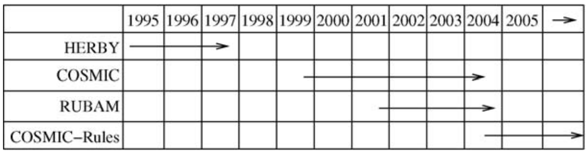

Fig. 1. Time line of the IbM projects developed at Liverpool.

While the COSMIC project was still running, RUBAM (Vlachos et al., 2004) was created to remove some of the complexity of COSMIC, but still used a bacterial theme. RUBAM modelled bacterial evolution in a continuous world with multiple substances that could be either nutrient or toxic, e.g. antibiotics. The control logic was a fuzzy classifier, with the goal of learning to climb substrate gradients and avoid concentrations of toxic substances. The lessons learnt with both RUBAM and COSMIC were then employed to derive COSMIC-Rules. This returned to the complexity of COSMIC, but acknowledged the computability issue, and also contained more abstraction than COSMIC. Worlds were continuous and contained multiple substances. Each bacterium contained a genome that gave it affinity for one substance, whilst at the same time conferring susceptibility to the lethal effects of another substance, thus deriving a "multi-species" model. Importantly, the genome was a functional abstraction rather than an abstraction of a DNA string. This allowed the assignment of bit strings to genes and the mutation of gene function while still being computable. The result was a mix of bit matching bit strings, matching functional genes to other functional genes, and Gaussian noise. Using these mechanisms it has been possible to add plasmid transfer and bacteriophage infection, and a variety of other events mediated by mobile genetic elements.

COSMIC-Rules was started just months before Ray Paton passed away. Although he was therefore unable