Empirical Modelling: The Computer as a Modelling Medium
==========================================

Steve Russ, Department of Computer Science, University of Warwick

****************{sub-heading / brief abstract}
A research group at the University of Warwick is developing a novel style of modelling based on observation, dependency and agency. This 'empirical modelling' brings computer-based models closer to the user by combining intuitive computing methods (e.g. spreadsheet principles) with an emphasis on the way a user experiences a given real-world domain (e.g. through interactions with observables).

****************

****************{definitions-in-a-box }

*************************************************************
* empirical: based on observations or experiment;
*
* referent: the object or application that is being modelled,
* whether already existing or only imagined;
*
* artefact : a computer-based model of a referent;
*
* agent: anything (human or otherwise) capable
* of changing the state of an artefact or referent;
*
* state: the collection of all observables regarded from some viewpoint
* as essential to the identity and function of an artefact or referent;
*
* observable: any feature of a phenomenon that can be reliably perceived,
* identified, and compared with similar features;
*
* environment: computer-based collection of artefacts with a context
* (e.g. model of furniture in a room, with gravity and other agents
* such as occupant, architect, or builder made explicit) .
*
*************************************************************

*******{main text begins here: headings are logical, not to be included literally}

{Constraints of conventional modelling}
=============================

The use of models proliferates everywhere—in business as much as science, in management as much as research. In large measure this is due to the spread of powerful desktop computers. But the role of the computer is usually limited to the manipulation of the model (e.g. fast mathematics and convenient interface). We are advocating a novel approach to modelling in which the power of the computer, through interaction with users, plays a significant part in the initial conception of the model.

While there is an essential role for the physical, analogue or scale model (e.g. the wind tunnel, wave tank and architect's mock-up), the majority of explicit scientific or business models are probably ones which have a mathematical or theoretical basis, and are therefore highly amenable to being 'put on' a computer.

A feature common to both physical and mathematical models is the need for selection, in advance of any modelling process, of those properties of the original which are regarded as essential to preserve in the model. The choice of these properties will often depend on the purpose of the model. This forms a commitment on the part of the modeller which it may not be easy to revise at a later stage in the modelling process (e.g. because of physical properties of materials, or the methods of finding a solution space). Thus a theoretical model will typically allow exploration of the variation of a given set of parameters, but will not so easily allow an exploration of alternative sets of parameters.

In a mathematical model where physical relationships may be represented by functions, differential equations, inequalities etc. the possible outcomes are completely circumscribed. There is no question of experimental interaction to imitate changes in an environment unless that has been explicitly incorporated in advance. To this extent a mathematical model is an artificial and closed world. The space of all behaviours is preconceived.

For a mathematical model to be practically useful we need a high degree of knowledge about the domain. As our modelling ambitions increase and we move further from those scientific areas for which we have good theories it becomes all the more needful to make allowance for ignorance, and make revisability a natural property of our models. In such areas we may wish to build models simply in order to aid our understanding: any more specific purpose may be unknown, or provisional, and it is then only an impediment to make early commitments to certain properties we wish to preserve in the model. Again, the need for openness and flexibility in the model-making process is paramount.

{Empirical modelling: motives and analogy with mental modelling}
==========================================

Two major motives of our observation-based approach are to bring computer modelling closer to the user, and to preserve in computer artefacts the openness that is characteristic of our real-world experience. Both these features are related to the essential role which interaction plays in the development of our models.

A familiar method of modelling, and one in which the computer *does* play an important part in the model development, including its initial conception, is the spreadsheet. In our empirical modelling the computer builds a virtual machine which resembles a large distributed spreadsheet with several users. This 'spreadsheet' can include analogue quantities of the kind we observe or can measure - such as the width of a doorway, the speed of a vehicle, or the force acting on a sail. Corresponding to a spreadsheet formula, a key construct in notations for empirical modelling is that of a *definition*. This expresses a dependency of one variable, typically an observable feature of a domain, on the values of other observables. The dependency is expressed in terms of appropriate operators (arithmetic, geometric, physical etc). The simplest dependency is the assignment of an explicit value to a variable. Sequences of definitions represent chains of dependencies with the only restriction that they be acyclic. (See Figure 1.) Dependencies are automatically maintained—this is not a control mechanism in the programming sense, but an indivisible state change propagated so as to reflect change in the domain. A script of definitions corresponds to a single state in the artefact. The user makes state transitions by re-definition, that is, by modifying existing definitions, or by defining new variables. In such an environment revision has equal status with construction. Nothing in such a model is preconceived—no-one knows what re-definitions an agent may make next.

{Figure 1 = example of definitions of a clock face beside picture of clock face}
{ there is no caption provided for Fig. 1}

Why do most people find spreadsheets so appealing? Perhaps the 'what-if' exploratory nature of a spreadsheet has deep roots in the way we learn to structure experience; after all, for the young child the 'what-if' experiment is almost a way of life. Making sense of new experience is related to the way we learn to represent parts of the world in our minds. It is not only as children we need to do this: arriving in a country that is new to us, entering an unfamiliar room in darkness, greeting a long-lost relative. A similar need to accommodate the new in terms of the old arises in any context of creation, growth, discovery or surprise. We make sense of the unfamiliar by successfully assimilating it into a representation (a mental model) of the familiar.

Much of human life is possible only because our minds are able to represent parts of the world, and to render these private representations public. How do we do it? In constructing computer models we have been guided by the way people seem to perform this remarkable feat of mental modelling.

The process of representing part of the world presupposes experience of some phenomenon. We therefore begin with the simple constituents of experience: the observables we find significant in a domain or application; these observables directly correspond to names (or variables) in our model, or artefact. Changes in the phenomenon, particularly those stimulated by our interaction, then form the basis of 'experiments' from which we make hypotheses about likely dependencies between observables, and about the state-changing agents in the environment (other people or devices, gravity, etc). Until we have made sense of, or understood the phenomenon, this cycle of observation, experiment and conjectures about dependencies and agents continues in an open-ended manner. It is a personal and subjective process leading to a provisional representation in the mind, or a provisional artefact in a computer.

The style of computer modelling we have developed is exploratory, unpredictable and 'hand-driven': human interaction is an essential part of the process. There is no prescription in advance of the tasks or problems that may be posed in the application domain, nor of their solutions (if any). Useful by-products of the process are: gaining a better personal understanding of a domain, explaining the mechanisms present, and communicating this knowledge.

At a practical level we have developed a general purpose tool for the maintenance of dependencies, and the representation of agent view-points, as well as a number of special purpose notations for window management, line drawing, domain analysis, etc. Simple models have been developed using these notations for a wide range of applications, such as a vehicle cruise control simulation, a sailing boat, a digital watch, various board games and sports, a railway animation, a jigsaw-puzzle assistant, classroom interaction, and many others.

{Empirical Modelling: novelty, principles and characteristics}
=======================================

When modelling via conventional programming, or with physical media, there is an irreducible separation between the user and the modelling medium which calls for deliberate, self-conscious planning. Even while sketching a diagram with paper and pencil, for example, revision is always possible but the effort involved keeps the thinking process distinct from its representation.

In contrast many people have had the experience that a word processor does not simply make writing easier to change but leads to a new, closer relationship between writing and thinking. When words can be manipulated almost as easily as thoughts, the think-write-think-revise cycle merges into a single malleable process. The way we use a computer in empirical modelling for representing states and state-changing mechanisms has a similar plastic quality: the immediate feedback and revisability makes for an intimate, symbiotic relation between the artefact and the user's mental model.

With empirical modelling we aspire to do for modelling in general, what the spreadsheet has done for accounting, or what parametric modelling has done for CAD, that is, to bring the modelling process closer to the user's own mental representation and to introduce more flexibility and openness.

The principles running through our approach can be summarised (dogmatically!) as follows. For a thorough understanding of a system or phenomenon we need to start from the notion of state in terms of possible observations. State is a matter of viewpoint, and so a notion of agency (viewer, and potential changer of state) is essential. We adopt as a fundamental relation (between states) the notion of a one-way dependency, which can be expressed and maintained in a definition (like a spreadsheet formula). The interpretation of states is fundamentally a product of interaction between a user and a model.

Four features of empirical modelling may help to characterise it :

* Empirical elements are explicitly incorporated into the model. Many variables are directly displayed - and in principle experienced in any appropriate way that devices allow. The user can make comparisons between states in the artefact and states in the referent in terms of direct experience (e.g. observations and measurements). This is rather striking and gives our models a quality akin to 'electronic sculpture'.

* The state-changing mechanisms ensure that perceived dependencies in the referent are automatically, and indivisibly, maintained in the artefact. These are 'chunks of change' corresponding to naturally interpretable empirical elements. Such chunks of change can also reflect different viewpoints on the domain, for example in the vehicle cruise control the views of the driver, designer and engineer are all represented. ? simulation

* The rich quality of interaction with the artefact during the initial modelling process contrasts with the limited, preconceived, interaction inevitable in an automated or conventionally programmed model. An analogy is the contrast between a car driver, needing constant open-ended interaction with the controls in response to the environment, and a car passenger who has little concern for the mechanism of progress, only for appropriate entry and exit from the vehicle.

* The fundamental nature of the primitive concepts (observation, dependency and agency), and the affinity with human modelling, make for a broad framework within which to describe and use both empirical, and more conventional, modelling methods.

****************************

Figure 2 Picture of digital watch and statechart.... for visual interest and evidence of significant modelling. See separate caption text below.

*****************************

In the light of these characteristics it is not too fanciful to suggest that empirical models are rather like an extension of our own mental representations of a system, and that creating models this way is like working with a 'thought processor'. But being computer-based, rather than brain-based, we can reliably record, elaborate, communicate, and manipulate these models in powerful ways. ?

Some applications lend themselves particularly to empirical modelling: such as those with a high degree of interaction, or where multiple viewpoints are important; others will be best treated by conventional methods. Our approach probably has most to offer for the early, creative and provisional stage in the modelling of a complex real-world system.

*********************{final box}

The author is indebted to the whole Empirical Modelling research group at Warwick for assistance with this article. Full details of who we are, a bibliography and details of access to a PC version of some of our tools are available from our Web page: http://www.dcs.warwick.ac.uk/pub/research/modelling/

We welcome enquiries on any aspect of the work described here.
e-mail: sbr@dcs.warwick.ac.uk

*********************

************{basic caption for Figure 2}

Figure 2 is a screenshot from a model that combines simulated digital and analogue watches with an associated statechart to represent the display functions of the digital watch, as decribed by David Harel in his book *Algorithmics*. The state of the entire model is recorded using a script of definitions, and animated in such a way that the digital and analogue watches run in real time, the buttons on the digital watch activate the watch display, and the statechart is updated appropriately.

*************{more interesting comment on Figure 2!}

This model was jointly developed by two members of the Empirical Modelling group at Warwick in two separate phases of independent work in 1992 and 1995, and has been the focus of subsequent development and extension in several ways. For instance, the model has been incorporated into a chess clock simulation without any modification to the original script of definitions. The entire model is based on a software tool that has been specially developed for empirical modelling: the tkeden interpreter. A software tool for the interactive construction and animation of statecharts was developed using tkeden as a final year undergraduate project in 1995-6. This tool can be used in the tkeden environment for the rapid development of models such as that depicted in Figure 2, which integrate statecharts with their associated animations.

**************

7