Frame Modelling of Dynamic Ecosystems

Mark Sherwood Quadling

Degree awarded with distinction 6 May 1993

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science

Johannesburg 1992

Abstract: Frame Modelling of Dynamic Ecosystems

This thesis develops the theoretical basis of the qualitative frame based modelling technique, a paradigm recently proposed by Starfield for the modelling of ecosystems with a multiplicity of stable states. This technique is a refinement of the **State-and-Transition** conceptual model of Westoby *et al* which involves the division of the ecosystem dynamics into a catelog of stable 'states' and a suite of transitions between these states. The frame models of Starfield associate with each stable configuration of the ecosystem a qualitative rule based model for the **key processes** in that stable configuration.

The aims of this thesis are the following.

1. A rigorous definition of frame modelling of dynamic ecosystems is proposed, and this theoretical foundation is used to demonstrate that qualitative frame models may be used to model dynamic ecosystems to an arbitrary accuracy.

2. The development of implementation software. A qualitative rule based frame modelling environment is presented, and a specification for an improved environment is proposed based on the theoretical work.

Declaratio.

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any other degree or examination in any other University.

Mark Sherwood Quadling

Date.

Table of Contents

List of Figures

List of Tables

Preface

When starting out on this project, the aim was to investigate a new modelling technique, **Frame-based Modelling**, by building software to implement several frame models and to investigate these frame models in an empirical manner. It soon became clear, however, that there was much more to frame modelling, and slowly but surely more effort was put into the investigation of the theoretical aspects. The result is that this thesis is concerned almost entirely with the formal frame theory, and the frame modelling software which was to form the basis of the work has really just become a bonus, useful for demonstrating frame models. I am thankful to my supervisor, Tony Starfield, for allowing me to shift the emphasis in this way.

I am particularly grateful to Tony Starfield for suggesting such a fascinating project. Thanks are also due to Dave Cumming who has been building frame models (and has suffered greatly from early buggy versions of the software), and John Field and Ricky Taylor for their useful suggestions. The Brachystegia frame model presented in **Chapter 3** was written by Tony Starfield, Dave Cumming and Russell Taylor, and is currently being modified and extended.

Theorems from the standard texts which have been referred to have not been proved in this thesis: A reference for the proof has been given instead. The theorems which have been proved are those stated and proved for the first time in this thesis.

This work was supported by the South African Nature Foundation (SANF), the Multispecies Animal Production Systems project of the World Wide Fund for Nature (WWF) in Harare and Senate Research at the University of the Witwatersrand.

Errata

1. Page 11, third line from the bottom: Replace *an* with *a*.
2. Page 40, fifth last line: Delete *system*.
3. Page 50, third last paragraph: Replace *popular* with *been popular*.
4. Page 54, second line in the definition: Replace *nonnegative* with *a nonnegative*.
5. Page 97, last paragraph: Replace *english* with *English*.
6. Page 98, in definition: Insert $v \in V$.
7. Page 100, second line: Replace *system values* with *system variables*.
8. Page 103, top line: Replace *proof* with *prove*.
9. Page 104, third last line: Delete *for*.
10. Page 106, last paragraph: Delete *presented*.
11. Page 107, second paragraph: Replace *possibly* with *possible*.

Chapter 1

1.1 Introduction

Ecological systems are notoriously difficult to model; there are few examples of systems where a satisfactory model has been built which may be used to predict future behaviour or explore different hypotheses with an acceptable degree of accuracy. (An 'acceptable degree of accuracy' in ecology is much less demanding than in more precise sciences such as physics). Their great complexity obviously presents problems but is by no means the only reason for this. The very existence of highly accurate models for complicated systems in physics, chemistry and engineering suggests that this is not the case. Perhaps the most important reason for this is the fact that the observables of the systems being studied are so difficult to measure accurately, because of the difficulty in carrying out the actual measurement as well as the variable and seemingly random fluctuations in the observables. These problems confound modellers in many real life systems other than ecology (although perhaps to a lesser extent), notable examples being in economics or climate simulation, and the gap in the development of useful models between these areas and those of fields such as physics is quite considerable.

Many of the modelling approaches which have been attempted appear to have been inspired by successful paradigms in the mathematical and physical sciences. In addition to the large scale borrowing of these modelling techniques, much of the modelling in the 60s and 70s was characterized by the dictum 'Use all available data in the model'. Generally, this modelling philosophy has not been successful in the sense that it has not provided ecologists with useful tools which may be used confidently and indeed has led to a certain suspicion on the part of many ecologists towards the efforts of modellers. Perhaps the greatest flaw in this philosophy is the way it excludes the majority of ecologists (the experts themselves) by turning the study of ecological systems into a mathematical curiosity and testbed for numerous modelling techniques from unrelated fields. This research concerns a modelling technique which has been born from the ideas and suggestions of ecologists themselves. This research represents an attempt to develop an appropriate way of modelling ecosystems and **not** the appropriate way of using existing techniques from other unrelated fields. The formalization and mathematical tools have been moulded to fit the problem, and not the reverse.

In particular, this thesis outlines a technique proposed by Starfield[1] for the modelling of ecosystem dynamics where the quality of available data is poor. This technique is based on the **frame** and on the concept of **qualitative modelling.** A **qualitative model** is a model whose variables take only a finite number of values each of which may be identified with a qualitative label[2,3]. The concept of a **frame** has been borrowed from Artificial Intelligence and was first used in connection with computer vision[4]. Frames are useful for simulating natural commonsense phenomena, an area notoriously difficult to master on computers. In the frame-based modelling of ecosystems, each stable state of the ecosystem is associated with a frame. In the case of an African savanna, two such frames could be a woodland frame where trees dominate and a grassland frame where the trees have been effectively removed and grasses dominate. The model dynamics are then present at two different levels: Since each frame is associated with a unique state of the system, each frame has its own model simulating the most important processes present for that particular state of the ecosystem. Typically these models are predominately qualitative. The second level of dynamics concerns the switching between frames and refers to the processes which perturb the ecosystem sufficiently to move it from one neighbourhood of attraction to another. A complete frame model is presented in **Chapter 3**.

This type of modelling is very tempting to pursue because of the ease with which ecologists can set up simple models of those systems which have eluded the development of satisfactory 'conventional' models. Frame based models are also easy to expand because of their nature. On the surface, frame models may appear to be too simple to lead ultimately to useful predictive models, but frame modelling is only deceptively simple: it embodies a deep appreciation of the system being studied. Consequently, there are two main themes in this thesis as summarized by **Figure 1.1**:

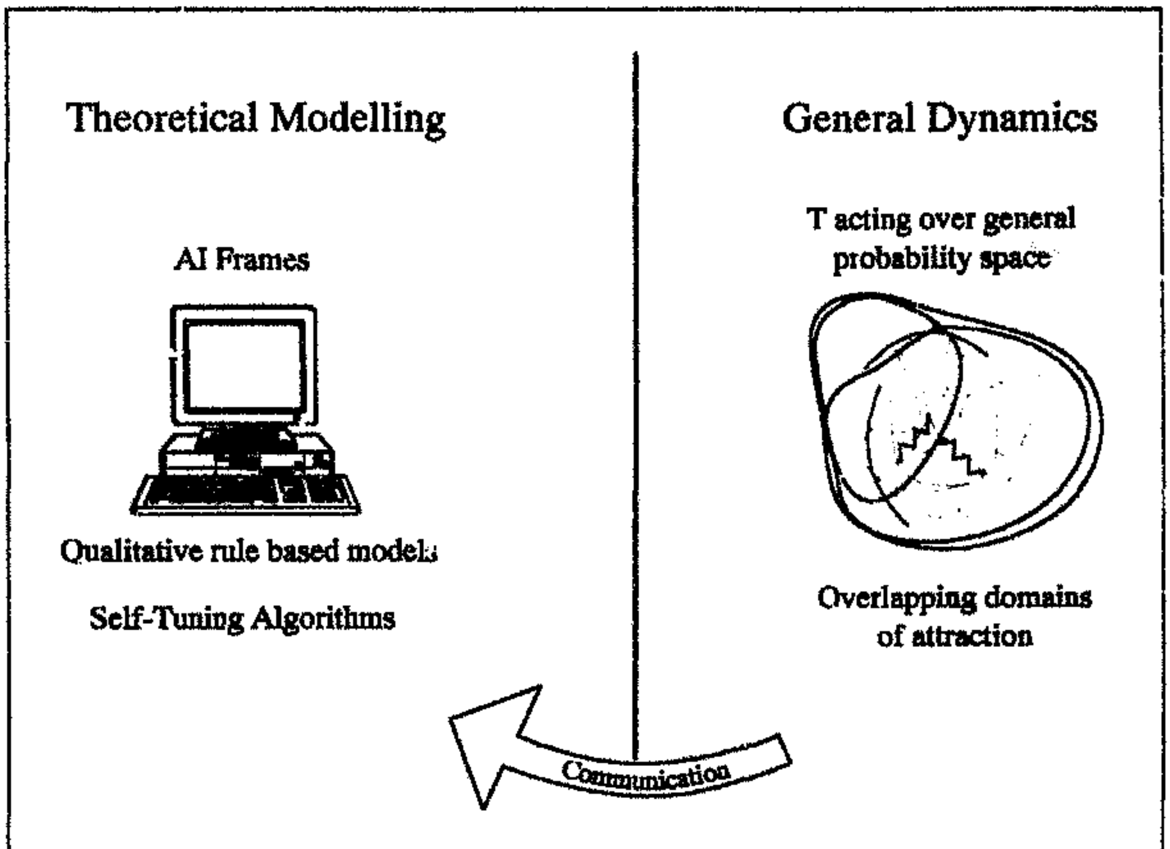


Figure 1.1: The two aspects of frame modelling.

The **Theoretical Modelling** component concerns the way qualitative frame models are built, and the development of tools to assist ecologists in this task. This component will be studied in detail in **Chapter 5**. By contrast, the **General Dynamics** component examines a hypothetical dynamical function **T** acting over a general probability space formed from the phase space of the ecosystem variables. This dynamical view will be presented in **Chapter 4**. The function **T** represents the true dynamics of the ecosystem, *but the ultimate aim is **not** to determine **T***. The link between the two will be demonstrated by considering the information flow from the general dynamical space to the observer. One of the aims of this thesis, therefore, is to show that even by assuming so little about the system and allowing it to be distinguished by non-determinism in its purest form, the frame-based modelling technique permits a model to be constructed which is as close to the hypothetical true dynamical function **T** as the available information will allow. Such a demonstration constitutes the ultimate justification for the use of frame-based modelling. It should be noted that the

General Dynamics component is considered only for theoretical purposes and does not form part of the 'final product': the modelling methodology and software tools being developed for ecological modelling. Its sole use is to demonstrate the feasibility of using frame-modelling and for considering issues of theoretical interest such as the effect of reversible processes on the dynamical description.

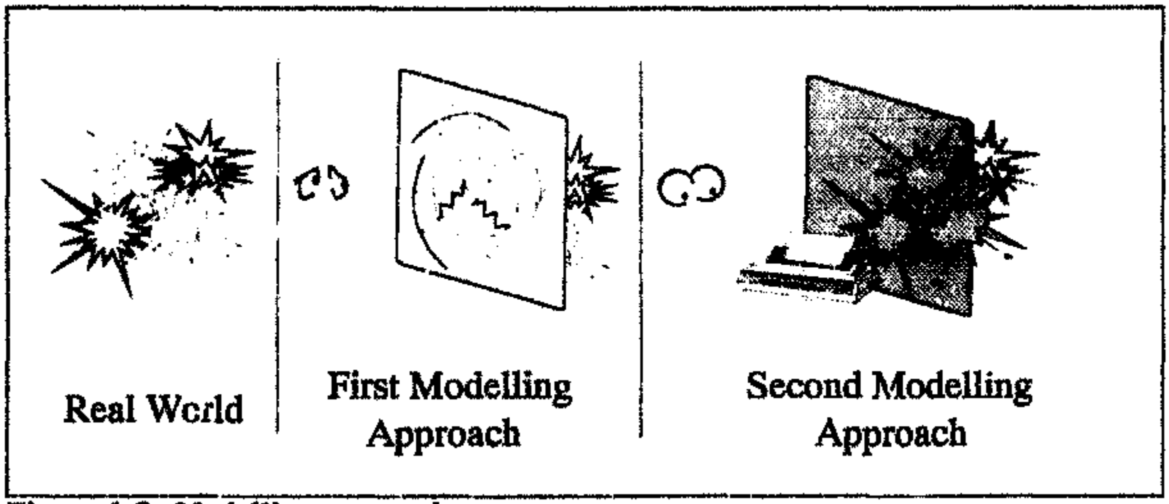


Figure 1.2: Modelling approaches.

Another way to look at this is as follows: Consider a scientist observing a complex dynamical system through a window which is not perfectly clear. It is not necessary to elaborate on reasons why the window is not perfectly clear. it may be because of something as simple as an incomplete data set or a more fundamental obstacle such as measuring difficulty or even an innate property of the system under observation. (An example of the latter in quantum systems is the Heisenberg Uncertainty Principle which puts an upper limit on the accuracy of the simultaneous measurement of position and momentum). In trying to build models which can emulate the system being studied, the traditional approach has been to try to find a formal description of the dynamics as perceived by the observer. This approach has been depicted in the middle in **Figure 1.2**. Often this formal description is imposed on the system, it is assumed that if the window were perfectly clear this description would be very close or identical to the true dynamics. In frame modelling, an entirely different approach is taken as is depicted on the right hand side of **Figure 1.2**. No attempt is made whatsoever to find a formal description of the dynamical system (to find **T** in the preceding remarks), rather we concentrate on extracting the maximum amount of information about the system (symbolized by a computer printout) through the appreciation of the deficiencies of the window as a communication channel.

Before looking at frame modelling in detail, the properties perculiar to ecosystems and modelling approaches generally used are examined in the next chapter in more detail.

Chapter 2: Perspectives on Ecosystem Modelling

2.1 Conceptual models in ecology

An aspect of considerable importance in ecology is that of the management of ecosystems, and much effort has been put into the development of conceptual models which seek to identify the most important processes and provide a framework for representing the knowledge about the system. A successful conceptual model of a particular ecosystem will focus research so that a balanced store of knowledge (information on how and why the system functions as well as information from the field which is sufficiently accurate) may be assembled. Conceptual models are therefore primarily organisational tools and do not necessarily lead directly to predictive system models able to simulate the system being studied with an acceptable level of accuracy. Since a successful ecological model must be able to do just that, one may be tempted to dismiss conceptual models as not useful to the modelling process. But conceptual models are very important in that they represent an operational description of the system as perceived by those who have the greatest understanding of it and consequently are vitally important for any subsequent attempt to build predictive models. The modelling process may be greatly speeded up by first building conceptual models and their related simple predictive models. Conceptual models of particular importance to frame-based modelling will be discussed in detail in **Chapter 4**.

2.2 Stable configurations within ecosystems

Suppose an ecosystem is in a certain configuration which is labelled by P. If the ecosystem is then slightly perturbed (for example by fire or disease) away from P and eventually returns back to P, then P is said to be a stable configuration or stable state of the ecosystem. The property of ecosystems being able to 'recover' from a small perturbation from the stable state is known as their resilience, a measure of the resilience being indicated by the size of the area of attraction around P.

Historically the management of ecosystems has been channelled into maintaining a particular stable configuration which is perceived as desirable, usually the state the system was in when first observed. This view amounted to an implicit assumption that ecosystems have only one stable configuration or that any other possible stable configurations are undesirable and to be avoided at all costs. This is best illustrated by means of example, an interesting one being given by the Serengeti-Mara ecosystem in East Africa[5]. Woodlands in this ecosystem have declined from their 1930-1950 levels with the greatest decline during the 1960s to the extent that by the 1980s, the area had become grassland. It is recognized that elephants and fire were primarily responsible for the decline. However, management of the elephants and burning has failed to return the Serengeti to its state of the first half of the century. In fact, the period prior to the 1960 decline may have been highly unusual in the history of the Serengeti-Mara: During the l[illegible] century, a great rinderpest epidemic swept through Africa which in combination with heavy ivory hunting presumably reduced the human and browser [illegible] such a level as to allow massive woodland regeneration in the early grasslands. [illegible] then, that the Serengeti-Mara has at least two stable states: the first dom[illegible] [illegible]lands and the second by woodlands. The large elephant population and fire of the 1970s [illegible] the system out of the area of attraction associated with the woodland state and into the grassland state area of attraction.

In terms of the ideas to be introduced in this thesis, we would represent each stable state by a **frame**. The intuitive notion of a **frame** is that of a unit of knowledge associated with each stable configuration of the ecosystem under observation. Frames originate in the field of Artificial Intelligence where they were first used in connection with computer vision, but find wider application now. Perhaps because of their wide range of use, there is no absolute definition of a frame. They were proposed by Minsky[4] as a means for the large scale organisation of knowledge, and in all implementations a frame consists of a number of slots, each of which may contain further knowledge in the form of data, frames or procedures (active knowledge). Generally, the frame structure does not prescribe any temporal or causal relations between the slots (such a property being reserved for **scripts** which may be considered to be time-ordered sequences of frames). By associating a frame to every stable configuration, we are also making the implicit assumption that available knowledge is most effectively categorized into compartments enumerated by the stable configurations of the system (the definition of a formal frame to be given in **Chapter 4** will be proved to do just that). In **Chapter 4**, this abstract computer science theoretic structure will be adapted to relate the available knowledge possessed about the system with the dynamics of the system, but it is important to appreciate the intuitive concept of a frame as a descriptor of a stable configuration.

2.3 Gross structure of ecosystems

While the purpose of this thesis is not to explore the detailed structure of ecosystems, an appreciation of the gross structure and unique properties of ecosystems is essential for the design of successful modelling techniques. From a modelling viewpoint, the following are proposed as bei[illegible]e most important characteristics of ecosystems.

a. Trophic levels
b. Niches
c. Complexity
d. Large measurement error (*)
e. Multiple stable configurations (*)

(The last two are especially important for frame-based modelling).

2.3.1 Trophic levels in ecosystems

Trophic levels are one of the most fundamental concepts in the study of ecosystems and indeed are often used to define an ecosystem. Every self-contained closed ecosystem contains a stack of trophic levels, a hierarchy of subsystems in which each level extracts energy from those below and contributes to those above. Each of these subsystems is vital to the functioning of the whole, but they are not tightly coupled in the sense that each trophic level must give up most of its energy to higher levels. An energy transfer efficiency of around ten percent has been observed in lakes, indicating that in these systems approximately one tenth of the energy is taken up by the next level[5]. Such a generalization has not been possible for terrestrial systems where estimates give a lower efficiency. Various reasons have been forwarded for this[6]. The appreciation of this hierarchical structure to ecosystems has led to various modelling methodologies such as the hierarchical models proposed by O'Neill *et al*[35]. It is also true that in many modelling scenarios, the dynamics within one trophic level (and perhaps including the interactions with a lower and higher level) are of prime importance. The resolution of a model[7] is consequently often related to this hierarchical structure of ecosystems.

2.3.2 Niches and competition within ecosystems

Along with the presence of trophic levels within ecosystems, the existence of niches is a localization property of ecosystems where the localization applies both spatially and functionally. The term niche was first applied by Grinell[8,9] to denote the ultimate distributional unit and later defined by Elton[10] as the functional role and position of the species in its community. These two interpretations stress respectively the spatial and ethological aspects of niches, and indicate the wide variety of ways in which niches have been viewed by ecologists. Following various empirical studies[11], the view of niches came to be linked with competition to the point where it became ecological dogma that one and only species is related to every niche. Put another way, an important premise of ecology is the **Competitive Exclusion Principle** prohibiting equivalent species from a stable coexistence[12] (two or more species being equivalent if their functional roles are indistinguishable). Looking at ecosystems from an abstract point of view, it may be argued that the development of niches within ecosystems is essential to the optimal utilization of resources and the promotion of global stability. For an ecosystem to extract as much energy as possible from its environment, diversity is encouraged to benefit from the many different forms of energy available. As will be discussed later, increased complexity in general systems may lead to decreased stability.

But the existence of niches allows complexity to be reduced without decreasing diversity. From this abstract point of view, developing such localization properties as niches and trophic layers is beneficial to the global stability of ecosystems.

2.3.3 Complexity of ecosystems

Complexity of ecosystems refers to both diversity as well as the number and type of interactions between the species. It is a fact that all ecosystems are complex in both these ways. To get some kind of idea of what happens when complexity is increased, consider a general system S_1 with a finite number of state variables and a finite number of relationships between these variables. Suppose also that P is an attractor of the system. If another system S_2 is formed by adding to S_1 an additional relationship, then it is clear that the domain of attraction around P in S_2 is a subset of the domain of attraction around P in S_1 since points within this region must satisfy an additional relationship. We may therefore argue that by increasing the complexity within ecosystems, the regions of stability are shrunk (at most they would be the same as before) which means that the global stability and resilience of ecosystems would be decreased. But as is outlined by May[12], it may be argued that many ecosystems of great complexity (such as tropical rainforests) exhibit greater stability than many ecosystems of lesser complexity (such as arctic systems). This apparent paradox can be resolved by considering the following: Firstly, the informal proof given above does **not** claim that more complex ecosystems are always less stable than less complex ecosystems, it only states that adding complexity to an existing ecosystem cannot result in increased stability. The stability may in fact be reduced. Secondly, the structural properties of ecosystems are such that interactions are localized by niches and trophic levels so that the web complexity is reduced somewhat even in cases of a great number of variables.

2.3.4 Measurement errors

It is ironic that subatomic phenomena can be observed and measured so accurately while measurements in biological systems in general are notoriously inaccurate. The basic constituents of ecological systems appear to be so observable, but in many cases a measurement error of up to 50% is considered good. This does not reflect on a deficiency on the part of ecologists in taking measurements, since physical systems cannot be compared to ecological systems. The constraints acting on the basic components of ecosystems are very weak compared to those acting on particles in many physical systems (consider for example the constraints on the dynamics of a charged particle in an electric field as opposed to the relative freedom enjoyed by a herbivore in an ecosystem). These weak constraints translate into a system which has a complex behaviour very difficult to predict, and which has large random fluctuations. Modelling techniques used for ecosystems will be discussed in **Section 2.5**, but it should be noted that most of these do not adequately take into consideration this vital attribute of ecosystems.

2.4 Modelling of ecosystems

Perhaps more so than in any other discipline, ecologists are faced with an enormously wide range of problems which require an equally extensive range of approaches to their solution or attempted solution. These techniques range from those which have been successful in other fields and are adapted for use in modelling ecosystems, to those which are tailor-made for the structures peculiar to ecological systems. Because of this, many different classifications or means of categorizing modelling techniques have been proposed, each very much dependent on the class of problems being considered. Of course there is no single 'correct' classification, since there are infinitely many ways of partitioning modelling approaches, but for any particular class of problem being studied there is possibly an useful classification which assists in the critical analysis of the various approaches. One such a classification proposed by Holling[13,14] may be into the **tactical** and **strategic** groups. A tactical model is an attempt to provide as detailed a description of the system as possible, while a strategic one attempts rather to capture the most general features of the system. The purpose of a strategic model therefore is only to test broad hypotheses and to encourage discussion on the weaker issues. Another classification proposed by Holling[15] is summarized by the diagram in **Figure 2.1**, with the models being classified according to the quality of the data and of the understanding of the basic processes.

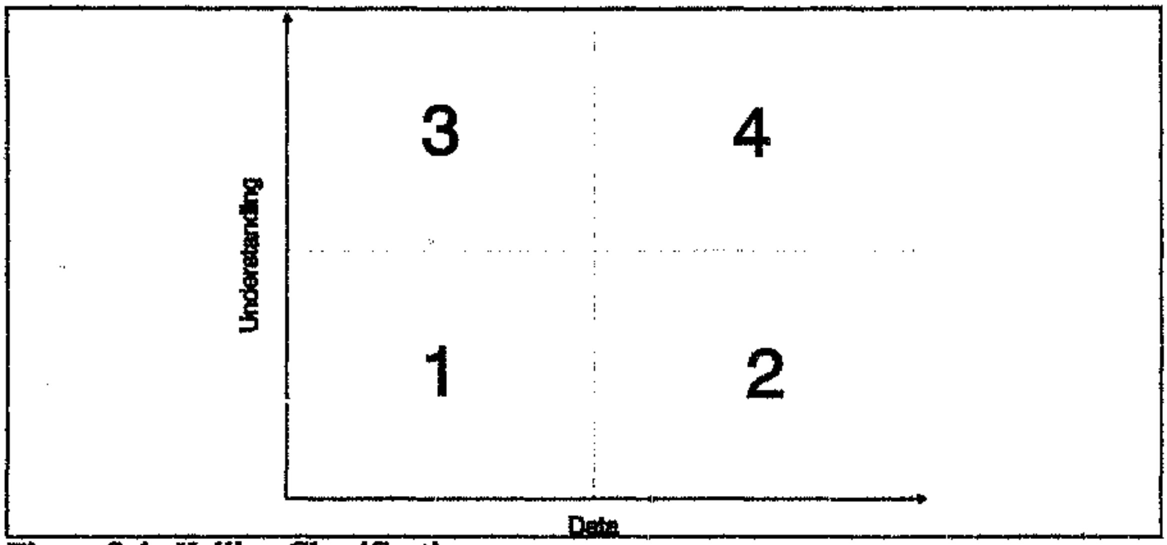


Figure 2.1: Holling Classification

Levins[16,17] suggested a classification based on the concepts of **Realism**, **Precision** and **Generality**. Referring more to the mathematical details of a model, other classifications can be **Time Dependent/Time Independent** or **Deterministic/Non-deterministic** or a combination of any two or more of the above[7]. The list appears to be endless which may prompt one to doubt the usefulness of considering model classifications at all. But much can be learned about precisely what is characteristic of the problem under consideration by looking at the type of model required, since the form of the model may be dictated to by the managerial requirements of the model and other non-ecological constraints posed by the hardware, expertise or time factors of implementation. In what follows, a classification is proposed which is useful for analysing the essential features of frame-based modelling.

2.4.1 Mechanistic vs Non-mechanistic models

One of the coarsest and most general classification of modelling approaches is into the categories of **Mechanistic** and **Non-mechanistic**, roughly the good understanding and poor understanding halves of the Holling diagram respectively. Generally if an ecologist knows enough about the system being investigated to be able to suggest a basic form (a formal description) either deterministic or non-deterministic, then he or she is following a mechanistic approach. If on the other hand, the ecological modeller has a poor understanding of why the system behaves as it does, a statistical or non-mechanistic approach may be necessary to derive this information through the observation of trends and other statistical analyses. Usually if there is simultaneously poor understanding (in the sense of being able to quantify the important processes) and poor data available for the system being studied, ecologists have had more faith (perhaps unjustified) in pursuing a statistical approach. In this fuzzy border region there have been attempts to develop mechanistic models (Starfield has done much pioneering work in this field[2.1.7]), but there is a prevailing viewpoint that the statistical approach is more reliable. It may be argued, however, that since the information about any particular system is more than just the measurable data, a statistical approach cannot use all the available information and consequently does not represent the best approach to solving the problem. Others may argue that a mechanistic approach is nevertheless futile because the lack of precise knowledge about the dynamics of the system will **not** allow the building of realistic and useful models.

The purpose of this thesis is to investigate a mechanistic approach to the modelling and solution of problems in this border region which ***does*** *allow a useful description of the system under observation.*

2.4.2 Exact versus phenomenological mechanistic models

Corresponding to the two approaches to modelling ecosystems depicted in **Figure 1.2**, consider a further subdivision of the mechanistic models into the classes **Exact** and **Phenomenological**:

2.4.3 Exact class of models

If a modelling approach is an attempt to find the exact form of the dynamical behaviour of the system, then that approach belongs in the **Exact** class. Consequently, each modelling approach within the **Exact** class represents a claim that the form of the model is believed to be either exact or very close to the exact form. In this class the modeller is effectively following one of two possible routes:

In the first route, the modeller knows enough about the system to be able to give an exact description of the dynamical behaviour, which is usually the case in the fields of engineering and physics. This exact description need not be entirely deterministic: The postulates of quantum physics are believed to be an exact description of the universe and yet do not lead to a system which is entirely deterministic.

The second route is prescribed for those systems whose dynamics are not as easily derived as above but for which the modeller is confident that the experience gained in examining type 1 problems may be applied. That is, the modeller assumes that the underlying dynamics of the system are very similar to the dynamics of another system whose behaviour may be described exactly.

Both these routes have been highly successful in many of the applied sciences, but their application within ecology has not been as successful as hoped.

2.4.4 Differential equation form of exact models

It may be argued that the first successful mathemetical description and modelling of a real physical system was the mechanics developed by Newton. One of the reasons why it was so successful is that Newtonian mechanics may be built up from a precise set of simple relations such as the differential equation given by his second law

$$F=\frac{dP}{dt} \tag{2.1}$$

In fact, the necessity of these relations in the description of Newtonian mechanics resulted in the Calculus being invented (independently by Leibnitz). Another reason for the success of this approach is that using mathematical analysis, these problems could be solved analytically which was an essential requirement in the days before the existence of computers. Because of the huge store of study and literature on the analytical approach to modelling it is natural that a similar approach would be tried for more complex systems. This is the most obvious example of the second route within the exact modelling class. The idea behind the differential equation form of exact modelling is simple, but it is instructive to identify the assumptions being made[12]:

Assumption 1: The system being studied may be described by a finite set of differentiable functions N_i where i=1,...,m. In the case of a multispecies population model, the populations numbers would be given by N_i.

Assumption 2: The dynamics of the system are given by the set of m differential equations

$$\frac{dN_i}{dt}=F_i(N_1,...,N_m) \tag{2.2}$$

where each F_i is an arbitrary function.

Assumption 3: There exists a set of constants N^*_i so that for all i

$$F_i(N_1^*, N_2^*, \ldots, N_m^*) = 0 \tag{2.3}$$

N^*_i are called the equilibrium populations.

An example of such a system is given by the widely used Lotka-Volterra-Gause equations:

$$\frac{dN_i}{dt} = N_i\left(\epsilon_i - \sum_{j=1}^{m} \alpha_{ij} N_j\right) \tag{2.4}$$

where i=1,..,n and ε_i and α_{ij} are constants (see **Section 2.4.6** for description of α_{ij}). For a discussion of other typical models used in ecology, see Starfield[7].

The set of values $(N_1(t),..,N_m(t))$ may be viewed as a surface in an m+1 dimensional space, a visualization which is particularly meaningful in the two dimensional case.

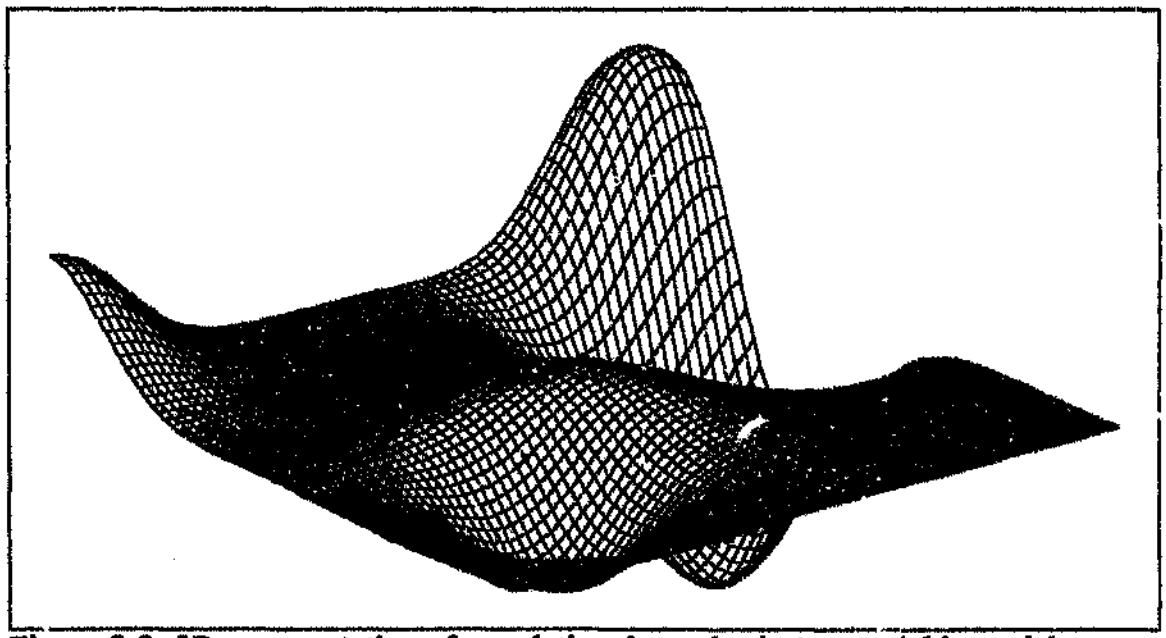


Figure 2.2: 3D representation of population dynamics in two variable model.

In **Figure 2.2** we have a possible plot of F for a system with two population variables, $N_1(t)$ being the parametrized variable along the x-axis and $N_2(t)$ being the parametrized variable along the y-axis. Since N_1 and N_2 are continuous, the surface will always be characterized by dips and bumps (extrema), ridges and valleys (saddle points). The equilibrium points then correspond to those parts of the surface which are flat (where the derivative vanishes). By the definition of an equilibrium point, an infinitesimal displacement from the equilibrium point in any direction must result in movement either towards the equilibrium point, or away from it. The first is an example of a stable equilibrium point (the dips in the 3D representation), while the second one of an unstable equilibrium point (the bumps in the 3D representation). Traditionally stability was a yes-no question, either a point was stable or unstable. But when it comes to identifying stability in real systems, it is not as straightforward as presented above. If the response surface in the neighbourhood of a point is fairly flat, then during the time that the system is observed any perturbation from this point may not appear to result in the system deviating significantly from the point. It is clear, therefore, that there must be more to stability than simply a yes-no affair. A possibility is to express the degrees of stability in terms of the slope of the response surface[12]. Another possibility is suggested by studying systems which have no points of equilibrium as defined in **Assumption 3**, or have no stable points of equilibrium but which nevertheless do exhibit stable behaviour. A classic example is given by the one-prey-one-predator model:

$$\frac{dx}{dt} = x f_1(x,y)$$
$$\frac{dy}{dt} = y f_2(x,y) \qquad (2.5)$$

for which Kolmogorov[18] showed that there is either a stable point of equilibrium or a stable limit cycle, as is illustrated in **figure 2.3**.

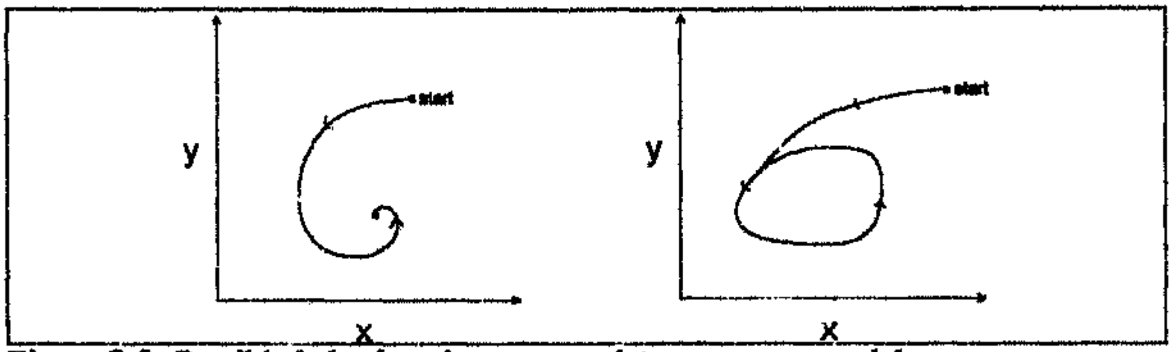


Figure 2.3: Possible behaviour in a one-predator-one-prey model

By considering attractors (as in chaotic systems) instead of simplistic points of equilibrium, stable phenomena such as the above are far more easily treated. Another more complex example is given by **figure 2.4**.

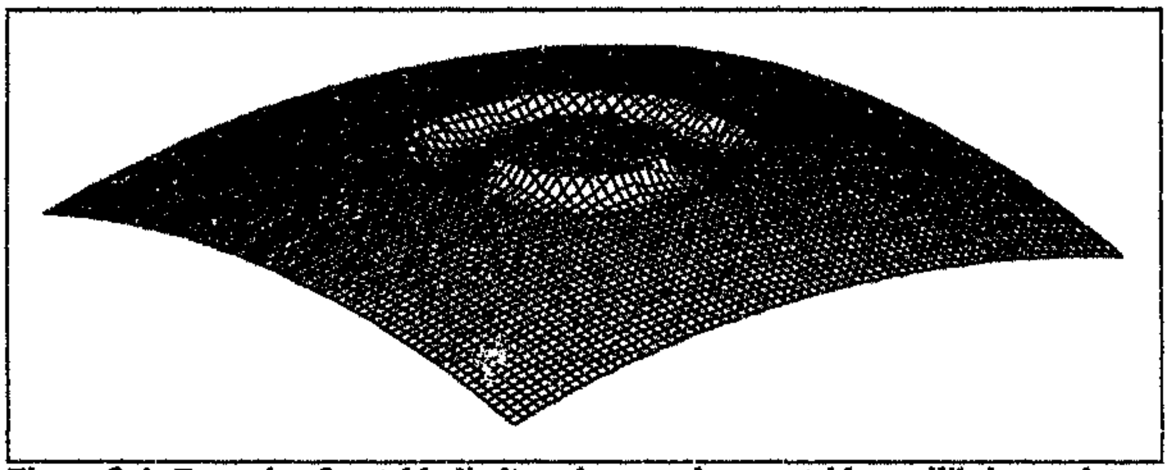


Figure 2.4: Example of a stable limit cycle around an unstable equilibrium point

The system represented by the figure has only one point of equilibrium P at the centre of the surface. But although P is clearly not a stable point, it is also not quite true to characterize it as unstable. The situation becomes even more complicated when real ecosystem problems such as measurement errors are factored in. In the diagram, the system behaves in a stable manner only if it is in the stable limit cycle. While in this cycle, it is a fixed distance away from P. What if this distance is less than the smallest measurable error? In this case, the system would appear to be stable at P and any attempt to formulate an exact model based on that assumption would lead to an incorrect form of the system dynamics. In **Chapter 4**, a more general notion of stability in ecosystems will be proposed which takes the variable nature of measurements in ecosystems into account.

2.4.5 Non-deterministic exact models

The above approach may be altered to provide a non-deterministic modelling technique[12]. The most intuitively simple way to do this is to retain all the assumptions, but interpret the space over which they apply differently. Instead of $(x_1,\ldots,x_m)$ representing a point in phase space, it instead represents the probability that $N_1=x_1,\ldots, N_m=x_m$. Referring to the two dimensional case in **Figure 2.2**, the surface no longer refers to the exact values of the variables but rather to the probability of them achieving a particular value. It is emphasized that although convenient, these assumptions are **not** sufficient for general non-deterministic systems since the probability space they operate over is far too specialized.

2.4.6 Linearizing the differential equation form of exact models

In any mathematical model if the dynamics can be linearized through approximation or under special circumstances, the advantages are enormous because of the great ease in dealing with linear systems. In the differential equation form of exact models, a further approximation is often made to aid the analysis of the system being modelled. This is done by assuming that the system being studied is close to an equilibrium point. If this equilibrium point is $(N_1^*,\ldots,N_m^*)$ and a small perturbation is given by $(x_1,\ldots,x_m)$ then **(2.2)** becomes

$$\frac{dx_i(t)}{dt}=F_i(N_1^*+x_1(t),\ldots,N_m^*+x_m) \qquad \textbf{(2.6)}$$

which when Taylor expanded around $(N_1^*,\ldots,N_m^*)$ in terms of x_i and retaining only those terms linear in x_i gives

$$\frac{dx_i(t)}{dt}=\vec{x}\cdot\nabla_{\vec{x}}F_i \tag{2.7}$$

where

$$\vec{x}=(x_1,\ldots,x_m) \tag{2.8}$$

and where we have used the fact that

$$F_i(N_1^*,\ldots,N_m^*)=0 \tag{2.9}$$

The usefulness of (2.7) is made apparent by rewriting it as a matrix equation

$$\frac{d\vec{x}}{dt}=A\vec{x} \tag{2.10}$$

where A is an mxm matrix often referred to as the community matrix[12]. A conventional linear stability analysis may then be performed. While convenient, it is very difficult to justify these approximations. The constraint that the ecosystem be near an equilibrium point is also far too restrictive in the study of dynamics ecosystems, and is questionable in the vicinity of more general attractors. In **Chapter 4**, the conditions under which a general dynamical system may be linearizable will be studied in detail.

2.5 Phenomenological class of models

The second class of models which we consider corresponds to the modelling approach on the right hand side of **Figure 1.2** and contains all those approaches which do not represent attempts to find the exact form of the model, or even a close approximation. The phenomenological approaches are traditionally used in those fields where the understanding of the basic processes is poor but the data good, or where both the understanding and the data are of a poor quality[7]. It may be argued that even for systems for which there is enough data and sufficient understanding to build an exact class model, a more realistic model could conceivably be obtained by a phenomenological approach. The central concept in the phenomenological class of models is the utilization of as much of the information about the system being investigated, and the most efficient way to use the available data. The modelling technique to be investigated in this thesis is in this class, but the recognition that we are attempting to extract as much information about the system is explicit in the technique

Chapter 3: Overview of Frame Modelling

3.1 Introduction

In **Chapter 1**, an informal description of frame modelling was given through the reference to the existence of multiple stable configurations in ecosystems. In the frame modelling technique proposed by Starfield[1], a **frame** is used to represent each possible stable configuration of the ecosystem together with a model for the **key processes** associated with that configuration. These models are **qualitative** and **rule-based**, so that the dynamics are described in terms of rules in much the same way as conventional dynamical models are described in terms of equations. **Figure 3.1** presents the basic structure of a frame model diagrammatically:

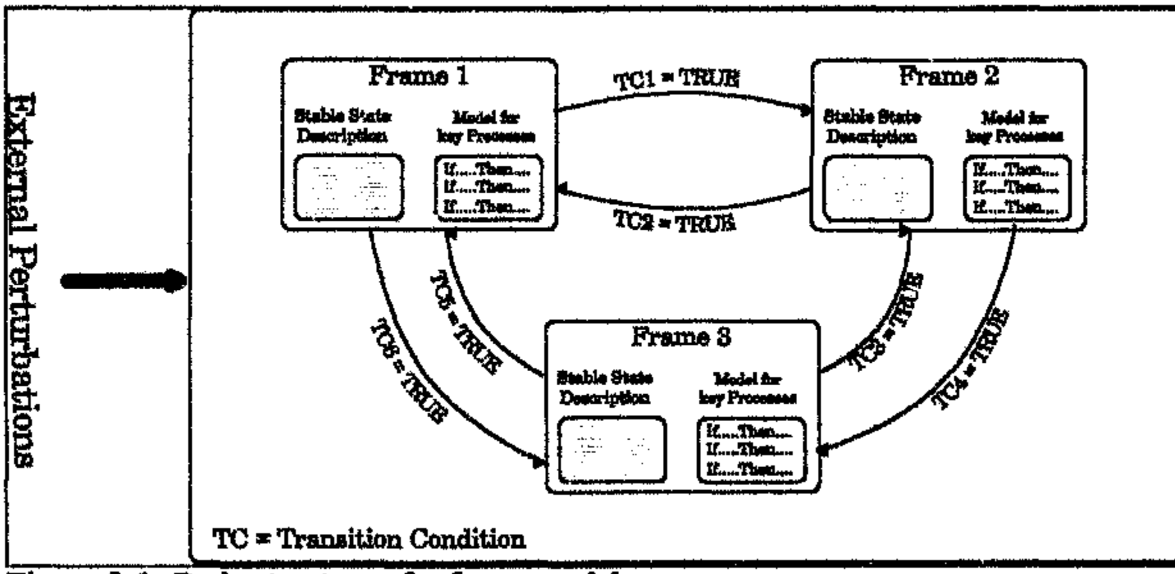


Figure 3.1: Basic structure of a frame model

As is discussed in depth by Starfield[2,3], the use of qualitative and rule based models is very useful in the building of models for systems which are not well understood, or which are well understood but difficult to measure: Qualitative and rule based models are ingenious tools allowing the building of models using the type of qualitative information known about the system which conventional modelling techniques have such difficulty in dealing with. A model demonstrating all the essential features of frame modelling is given below: This model was developed by Starfield, Cumming and Taylor and is taken directly from a draft paper by them.

3.2 A frame model of a Brackystegia boehmii woodland

A frame-based model is constructed in six steps:

1. First we clarify the objectives of the model.
2. Next we identify the driving (or input) variables.
3. Then we choose the frames and determine the pathways between frames.
4. Next, we identify the key variables and processes in each of the frames.
5. Then we construct the rules for switching from one frame to another.
6. Finally we build models for the key processes within each frame.

3.2.1 Objectives

To test our understanding of the system dynamics and explore alternative management options over a time-scale of up to 100 years.

3.2.2 Driving variables

Annual rainfall (low or high); inception of a fire, or "match" (on or off); time of burning (early or late in the dry season); elephant density (0, 1, 2 or 3 where 0 means no or very few elephants, 1 corresponds to densities of about 0.25 per sq km, 2 to densities of about 0.50 per sq km and 3 to densities of 0.75 per sq km or higher). These driving variables can either be input directly to the model or can be calculated via a random number generator from specified probabilities. For example, the probability of a "match" might be specified as 0.5, implying that there is a 50% chance in any one year that an attempt will be made to ignite a fire.

3.2.3 Frames

(1) A miombo Woodland frame dominated by Brachystegia boehmii woodland which is stable in the absence of elephants;

(2) A tall Grassland frame with a mosaic of Brachystegia shrubs;

(3) A Bushland frame of Combretum and other woody species.

The system can switch from Woodland to Grassland and back to Woodland or from Grassland to Bushland and back to Grassland. There is a possibility that the Bushland could return directly to a Woodland, without going through the Grassland frame, as Brachystegia shrubs re-invade and eventually dominate the canopy, but this could take 50 years or more; we therefore use the objectives of our model to avoid this pathway. Instead, we assume that the system will be irreversibly "trapped' in the Bushland frame after 7 years.

3.2.4 Key variables

A "match" implies that an attempt is made to ignite a fire, but it does not necessarily follow that there is sufficient fuel to sustain it. In all three frames, therefore, we introduce the variable "fuel-load" on a scale of 0 to 6 (or very low to very high). The type of fire that occurs is also important and will depend on the fuel-load. In all three frames we introduce the variable "fire" (cool or hot) which can occur in the early or late dry season. These are the only global variables (i.e. variables which apply to all frames).

In the Woodland frame: We track the state of the mature Brachystegia trees using a "canopy" variable on a logarithmic scale as in **Table 3.1**.

Canopy category	Representing (% canopy)
1	0 to 4
2	4 to 12
3	12 to 28
4	28 to 60
5	> 60

Table 3.1: Canopy categories in the Woodland frame.

In the Grassland frame: We represent the Brachystegia shrubs by the variable "height" on a scale of 0 to 3.

In the Bushland frame: We introduce the variable "bushtime" to count the number of years since switching into this frame.

3.2.5 Rules for switching from one frame to another

The switch from Woodland to Grassland occurs when mature trees have been removed from the canopy. The rule is: switch to Grassland definitely in a year in which the canopy category is 1 and with probability 0.3 if the category is 2.

There are three conditions that will trigger a switch from Grassland to Woodland: (i) if the shrub height has been continuously in category 3 (the highest category) for 5 years; (ii) if the annual rainfall for the current year is high and the shrub height has been continuously in category 3 for at least 3 years; (iii) irrespective of the rainfall, with probability 0.5 if the height has been in category 3 for 3 years and probability 0.75 in the fourth year.

The switch from Grassland to Bushland is associated with a depletion of Brachystegia boehmii seedlings and shrubs and the encroachment of other shrubs and emergent tree species. The rule is: switch if the height variable has been continuously less than 1 for at least 5 years and there have been no hot fires for the past two years.

The switch from Bushland is somewhat tentative. The encroaching shrubs are more fire resistant and could eventually out-compete the Brachystegia, but they will remain fire-susceptible until they have escaped from the grass layer. The system could, therefore, switch back into a Grassland, but only if a hot fire occurs soon after the switch to Bushland. The rule for switching back to Grassland is: switch definitely if there is a hot fire and bushtime < 3, with probability 0.75 if bushtime is 3 or 4, and with probability 0.5 if bushtime is 5 or 6.

If the system does not switch back to Grassland within 7 years, then it remains trapped in Bushland for the rest of the simulation.

3.2.6 Modeling the dynamics within frames

It is easiest to describe the within-frame models working backwards from the simplest frame (Bushland) to the most complex (Woodland).

3.2.6.1 The Bushland frame

All that matters in the Bushland frame is whether or not there is a hot fire during the first six years. The variable bushtime keeps track of the time and when it reaches 7 years the simulation ends with the message "trapped in Bushland". If bushtime is less than 7, the following rules apply:

The variable fuel-load is reset to zero after any fire, is increased by one if rainfall is low, and by two if the rainfall is high. A fire can only occur if there is a "match" (ignition). It will always be cool if the ignition is early or if the fuel-load is 1. If ignition is late, the fire will be hot if the fuel-load is 3 or greater, and can be either hot or cool (with equal probability) if the fuel-load is 2.

3.2.6.2 The Grassland frame

The Grassland frame has identical rules for fuel-load and fire. In addition, it has a set of rules for increasing or decreasing the height of Brachystegia shrubs. This variable is set equal to 1.0 whenever the system switches into the Grassland frame. Thereafter, low rainfall adds 0.2 to the height, while high rainfall adds 0.25. **Table 3.2** is used to determine how much to subtract from the height variable, depending on the category of elephants and on the height itself. (The height category is just the truncated integer value of the height variable.) It is assumed that elephants do not find bushes in height category 0 because they are hidden in the grass layer.

		Brachystegia height category		
		1	2	3
	1	0.075	0.05	0.03
Elephant	2	0.15	0.10	0.06
density	3	0.25	0.20	0.10
	4	0.35	0.35	0.35

Table 3.2: Reduction in height due to herbivory, as a function of elephant and height categories.

Shrub height can also be reduced by hot fires. If the height category is 0, 1 or 2 then we subtract 1.0 if the fuel-load is less than 3; for higher fuel-loads we subtract 2.0. Taller shrubs (height category 3) will only some times be affected by hot fires; the probability of this happening is a function of fuel-load as shown in **Table 3.3** and we subtract 2.0 when it does happen.

	Fuel-load				
	2	3	4	5	6
Probability	0.00	0.05	0.10	0.20	0.40

Table 3.3: Probability that a tall shrub will be affected by a hot fire.

3.2.6.3 The Woodland frame

The fuel-load model is slightly more complex here than in the other two frames since fuel-load builds up more slowly as the canopy increases. **Table 3.4** shows the increment in fuel-load as a function of annual rainfall and tree canopy.

	Canopy category				
	1	2	3	4	5
Rainfall					
low	1.00	1.00	0.87	0.75	0.50
high	2.00	2.00	1.74	1.50	1.00

Table 3.4: The increase in fuel-load in the Woodland frame.

Fuel-load is reset to zero whenever a fire occurs and the rest of the fire model is the same as in the other two frames.

The canopy variable is randomly initialized to a value between 3.5 and 4.5 whenever there is a switch to the Grassland frame We add 0.11 to the canopy variable in a dry year and 0.22 in a wet year. The amount to be subtracted each year because of elephants is shown in **Table 3.5**.

		Canopy category				
		1	2	3	4	5
Elephant density	1	0.16	0.14	0.12	0.10	0.075
	2	0.35	0.30	0.25	0.20	0.15
	3	0.70	0.60	0.50	0.50	0.50

Table 3.5: Reduction in canopy due to elephants.

Finally, we also subtract an amount from the canopy variable if there is a hot fire. **Table 3.6** shows how this amount depends on the fuel-load.

	Fuel-load				
	2	3	4	5	6
Loss of canopy	0.00	0.10	0.20	0.35	0.45

Table 3.6: Loss of canopy in a hot fire.

3.3 Frame modelling environment

Because the purpose of qualitative and rule-based models is to assist potential modellers in capturing qualitative knowledge of the system in the form of a frame model, a modelling environment was developed with the objective of aiding users in setting up frame models. Because of this, a great deal of effort was put into designing a Graphical User Interface (GUI) which is both easy to use and also assumes an active role in the model building process. Since the model is built up out of rules which have a predictable structure, the user of this environment *is not required to know the exact syntax used in the writing of the rules*. Precisely what is meant by this will be demonstrated through example in this section. To illustrate the use of this environment, the Brachystegia woodland model given above has been entered into the environment, and is given below:

3.3.1 Global variables and ruleset

The global variables are those variables which are common to each frame, and the global ruleset applies no matter which frame the system is in The global variables are summarized in **Table 3.7**:

Name	Type	Format	Levels
Rainfall	Driving	Qualitative	low, high
Elephant	Driving	Qualitative	none, low, medium, high
Burn Time	Driving	Qualitative	early, late
Match	State	Qualitative	off, on
Fire Type	State	Qualitative	none, cool, hot
Fuel load	State	Qualitative	0,..,6

Table 3.7: Global Variables

The global rules are:

***IF** random 0.5 **THEN** match: on **ELSE** match: off*

***IF** match off **THEN** fire type: none*

***IF** match on **AND** (burn time early **OR** fuel load 1) **THEN** fire type: cool*

***IF** match on **AND** burn time late **AND** fuel load 3 **THEN** fire type: cool*

***IF** match on **AND** burn time late **AND** fuel load 2 **AND** random 0.5*

***THEN** fire type: hot*

***ELSE IF** match on **AND** burn time late **AND** fuel load 2 **THEN** fire type: cool*

***IF** fire type none **THEN** fuel load: 0*

***IF** fire type none **THEN** fuel load. 0*

3.3.2 Woodland Frame

The only variable distinct to the woodland frame is **canopy** which is defined in **Table 3.8**:

Name	Type	Format	Levels
canopy	state	qualitative	0% to 4% 4% to 12% 12% to 28% 28% to 60% over 60%

Table 3.8: Definition of *canopy* in woodland frame.

The rules distinct to the woodland frame are:

IF (canopy = #x) AND (rainfall = #y) THEN fuel load:= fuel load + #z [Using Table 3.4]

IF rainfall = high THEN canopy:= canopy + 0.22 ELSE canopy:= canopy + 0.11

IF (canopy = #x) AND (elephant = #y) THEN canopy:= canopy-#z [Using Table 3.5]

IF fire type = hot AND fuel load = 3 THEN canopy:= canopy-0.1

IF fire type = hot AND fuel load = 4 THEN canopy:= canopy-0.2

IF fire type = hot AND fuel load = 5 THEN canopy:= canopy-0.35

IF fire type = hot AND fuel load = 6 THEN canopy:= canopy-0.45

IF fire type = hot AND fuel load 6 THEN canopy:= canopy-0.45

3.3.3 Grassland Frame

The variable unique to the grassland frame is **height** which refers to the height of the Brachystegia shrubs, and is defined in **Table 3.9**:

Name	**Type**	**Format**	**Levels**
height	state	qualitative	zero, low, medium, high

Table 3.9: Definition of *height* in grassland frame.

The rules for the grassland frame are:

IF rainfall low THEN height: height 0.2 ELSE height: height 0.25

IF (height #x) AND (elephant #y) THEN height: height-#z ***[Using Table 3.2]***

IF height medium AND fuel load 3 AND fire type hot THEN height: height-1

IF height medium AND fuel load 3 AND fire type hot THEN height: height-2

IF height high AND fire type hot AND fuel load 3 AND random 0.05
THEN fuel load: fuel load-2

IF height high AND fire type hot AND fuel load 3 AND random 0.1
THEN fuel load: fuel load-2

IF height high AND fire type hot AND fuel load 3 AND random 0.2
THEN fuel load: fuel load-2

IF height high AND fire type hot AND fuel load 3 AND random 0.4
THEN fuel load: fuel load-2

IF height high AND fire type hot AND fuel load 3 AND random 0.4
THEN fuel load: fuel load-2

3.3.4 Bushland Frame

The only variable unique to the bushland frame is **bushtime**, which counts the number of years that the system is in the bushland frame. It is defined in **Table 3.10**:

Name	**Type**	**Format**	**Levels**
bushtime	state	quantitative	N/A

Table 3.10: Definition of *bushtime* in bushland frame.

> ***IF** rainfall = low **THEN** fuel load: = fuel load + 1; bushtime: = bushtime + 1*
> ***ELSE** fuel load: = fuel load + 2; bushtime: = bushtime + 1*

3.3.5 Transition rules

The transitions indicated below will occur if the particular transition rule is true:

From **woodland** to **grassland**:
*canopy 0% to 4% **OR** (canopy 0% to 4% **AND** random 0.3).*

From **grassland** to **woodland**:
*(height high **FOR** 5 time units) **OR** (rainfall high **AND** height high **FOR** 3 time units)*
***OR** (height high **FOR** 3 time units **AND** random 0.5)*
***OR** (height high **FOR** 4 time units **AND** random 0.75)*

From **woodland** to **bushland**:
Not possible.

From **bushland** to **woodland**:

Not possible.

From **grassland** to **bushland**:

height zero FOR 5 time units AND fire type hot FOR 2 time units.

From **bushland** to **grassland**:

fire type hot AND (bushtime 3 OR (bushtime 3 OR bushtime 4) AND random 0.75 OR (bushtime 5 OR bushtime 6) AND random 0.5).

3.3.6 Sample screens from modelling environment

Figures 3.2 to **3.8** are screen dumps taken from the first prototype of the modelling environment during the process of implementing the Brachystegia model. This prototype environment was developed for MS-DOS machines with VGA graphics and mouse; for clarity the screen dumps have been converted to outline monochrome images.

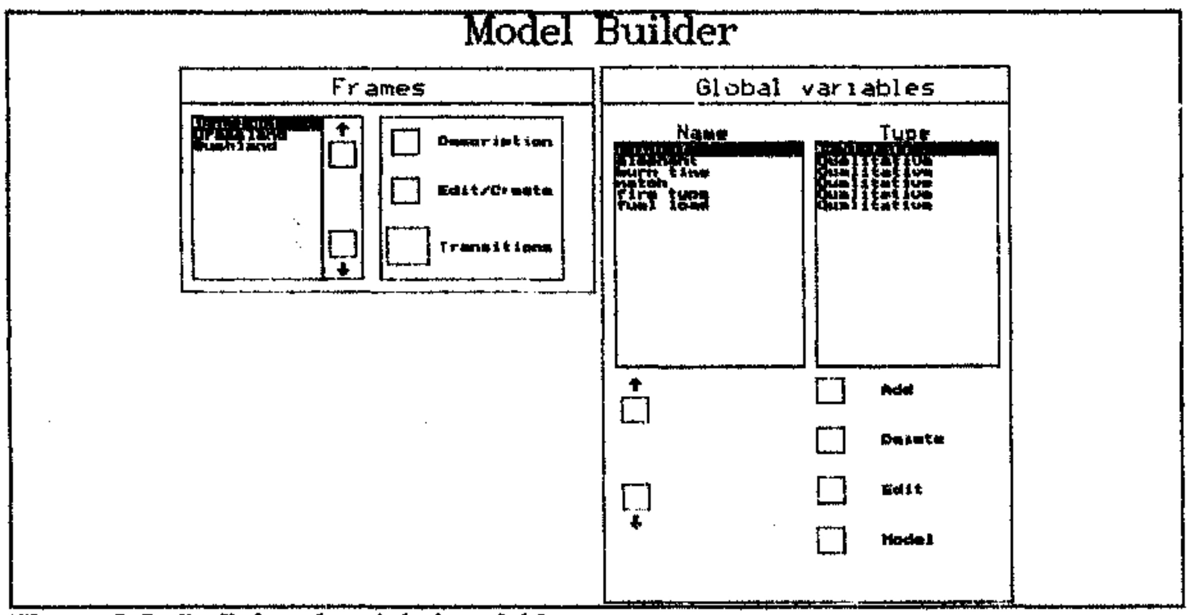


Figure 3.2: Defining the global variables

Figure 3.2 demonstrates the process of defining the frames of the model, and the global variables. For each variable, the user need only fill in the fields of a standard declaration template, as is illustrated in **Figure 3.3**:

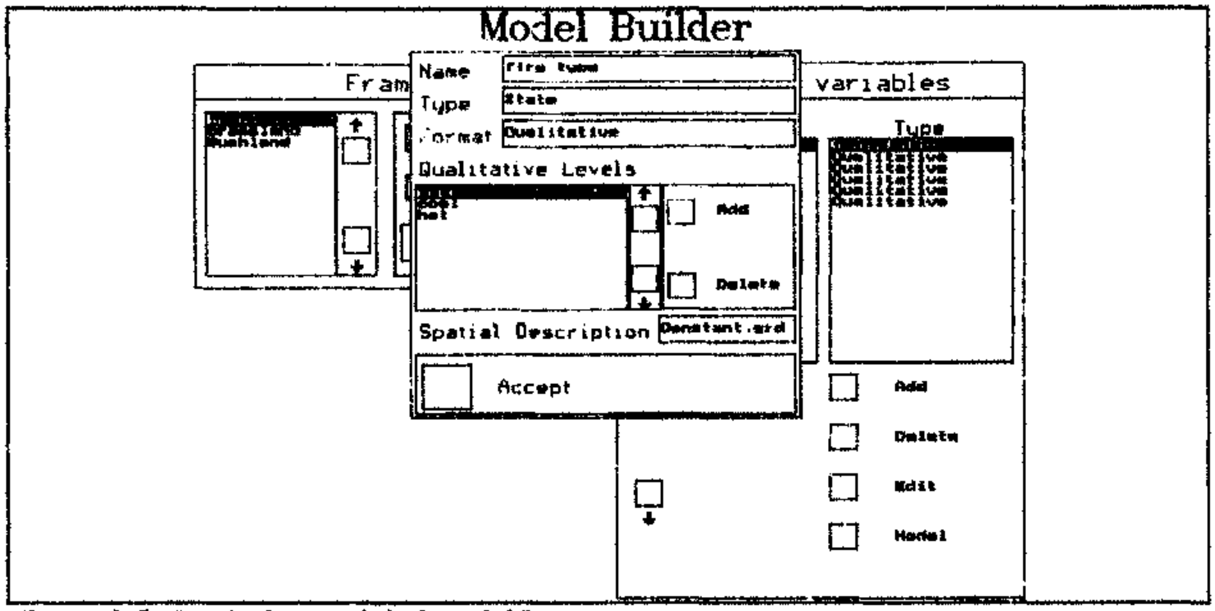


Figure 3.3: Declaring a global variable

There are several **types** of variables: A **driving** variable is an input variable whose value must be read in at each time step. A **state** variable is a general purpose variable which depends on all the other variables and whose behaviour must be modelled. A **random** variable has a non-uniform probability distribution that the user can specify. A predefined variable **RANDOM** is provided which returns a normally distributed random number between 0 and 1 as in most programming languages. A **time variable** has a predefined oscilliatory behaviour.

The format of a variable may be either **qualitative** (with a finite number of values) or **quantitative**. In the case of qualitative variables, a qualitative label may be provided for each state. A fractional increment may nevertheless be added to a qualitative variable, see **Chapter 5** for further discussion on this.

A limited two dimensional spatial capability is present in the first prototype: Any variable may have a grid file associated with it, and regions may be defined in terms of this grid file. The modelling environment allows each variable to have a different grid file if so desired; if two variables with different grid files are used together in the same context, a weighted mean is calculated for each based on their grid files. In the Brachystegia model, no spatial description is required, hence the specification of the grid file is CONSTANT.GRD (This grid file defines a one-by-one grid with a single spatial region, namely the entire universe).

Figure 3.4 demonstrates the definition of a frame, in this case the woodland frame:

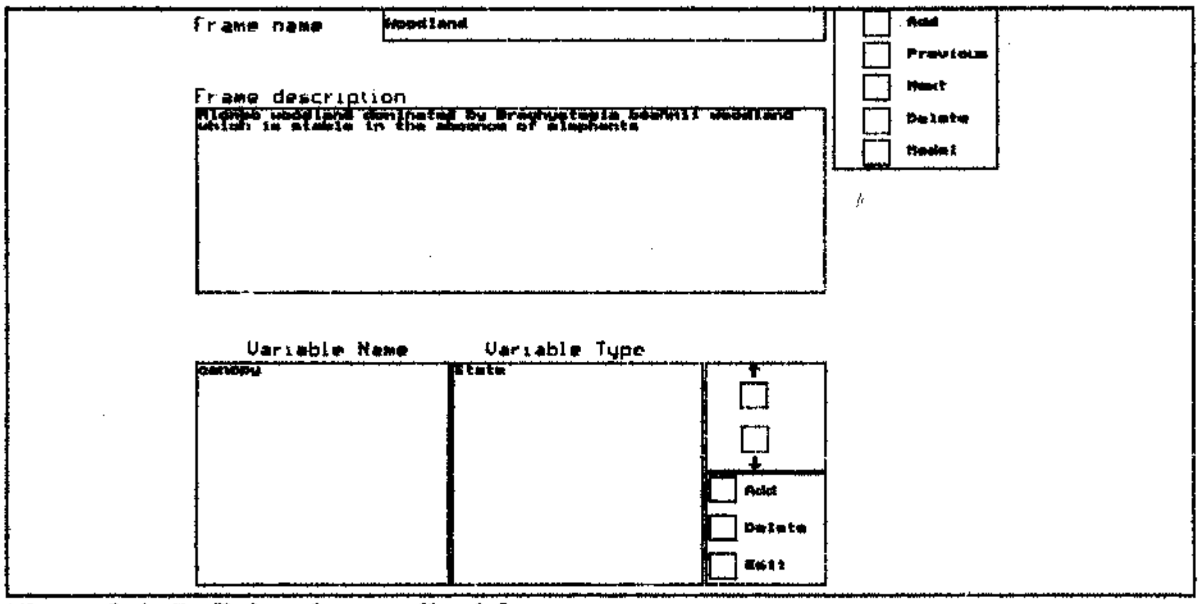


Figure 3.4: Defining the woodland frame

Figure 3.5 and **3.6** demonstrate how a matrix rule is entered into the modelling environment. The matrix is retrieved by clicking the mouse in the edit window of the rule in question.

Each rule may also have an informal description associated with it, where the user may make a note as to what assumptions have been made for that particular rule. Each rule is presented individually in the edit window, and at each stage the syntax of the rule may be checked using the **Check Syntax** button. In this way, common errors may be eliminated immediately instead of a list of error messages appearing at compile time.

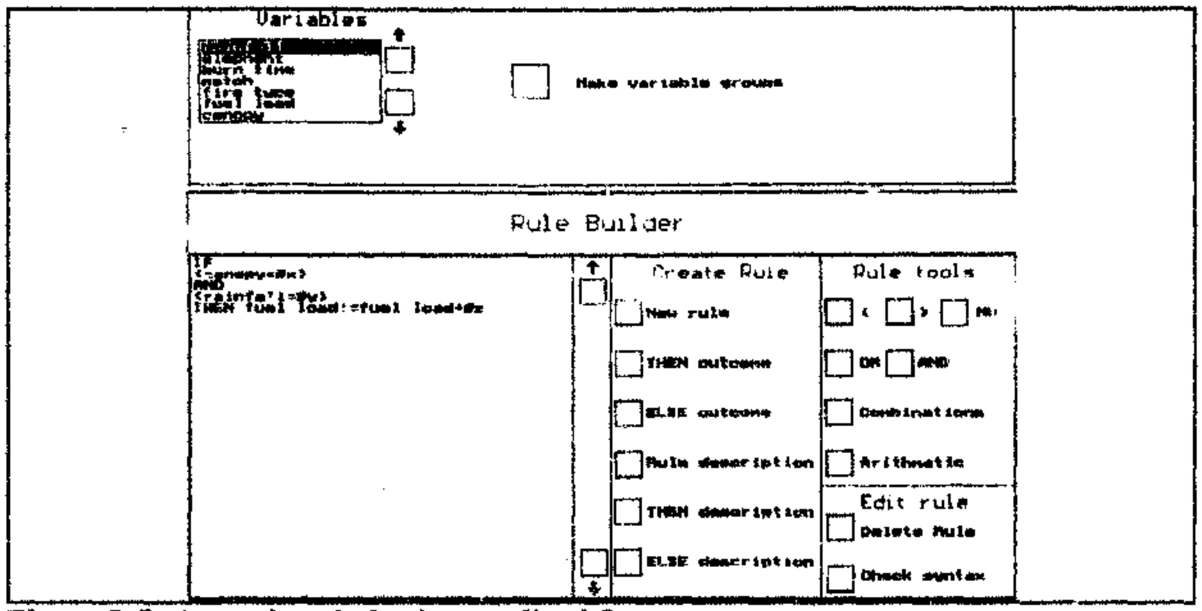


Figure 3.5: A matrix rule in the woodland frame

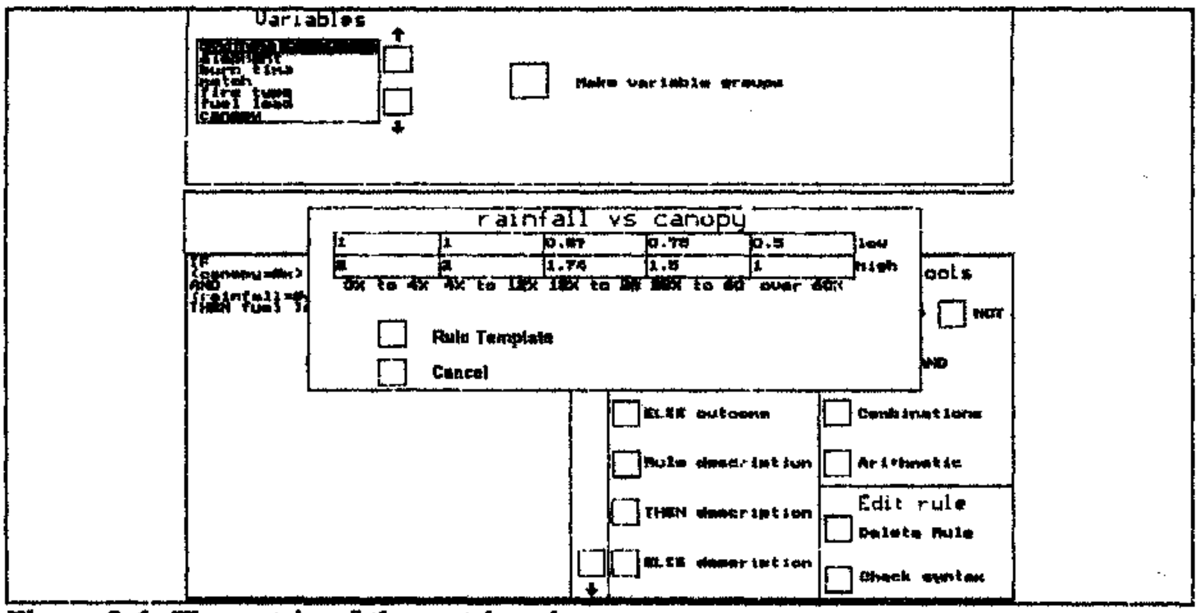


Figure 3.6: The matrix of the matrix rule

3.3.7 Active Rule Building

Since rules have a predictable format, it is not difficult to build into a modelling environment the ability to take the active role in the building of rules. An active role refers to the process by which the user is asked a number of questions, and based on these questions the modelling environment generates a rule. In the first prototype of a modelling environment, the user may either enter a rule manually (by clicking the mouse in the edit window), or initiate a question-asking sequence by clicking the mouse on the relevant variables in the variable list (see **Figure 3.5**). **Figure 3.7** demonstrates the first question asked after the user has clicked the mouse on the **burn time** variable in the variable list:

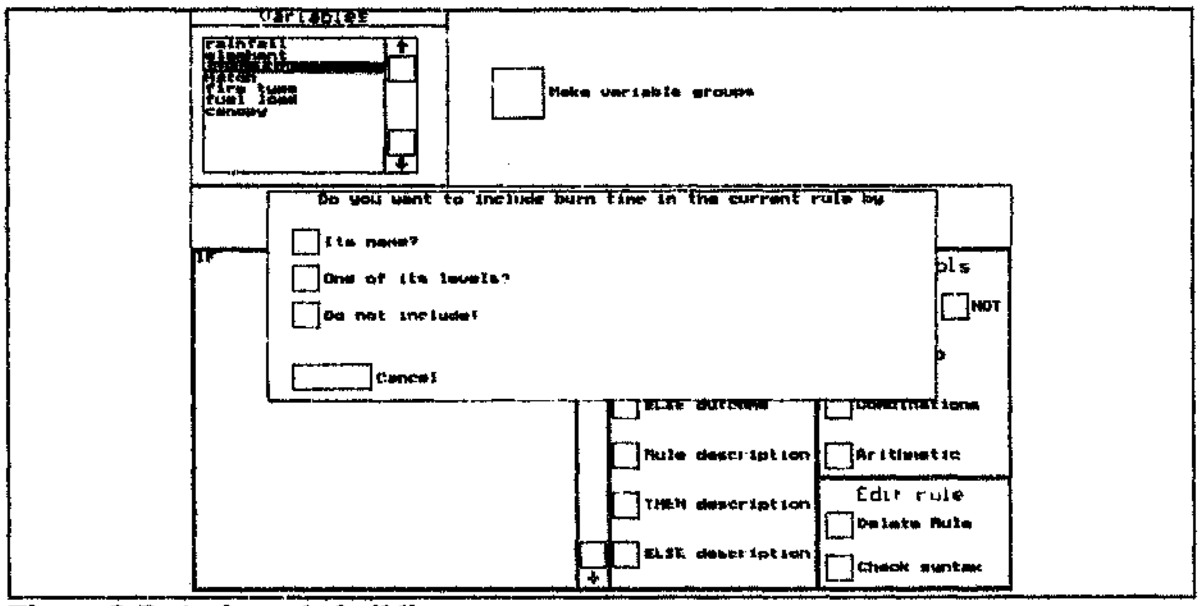


Figure 3.7: Active rule building

The exact sequence of questions depends on the answers at each stage, but a typical sequence of questions is:

Is the next rule nested in the current rule? **No**
Do you want to include the variable in the current rule by its name or one of its levels? **Name**
Do you want the current value of the variable? **No**
Refer to the value of the variable how many time units ago? **1**

Do you want the qualitative or quantitative value of the variable? **Quantitative**

Are you referring to the variable in the current spatial region? **Yes**

Figure 3.8 demonstrates how the transition rules are entered: In this case the condition for a switch from the Woodland frame to the Grassland frame has been given.

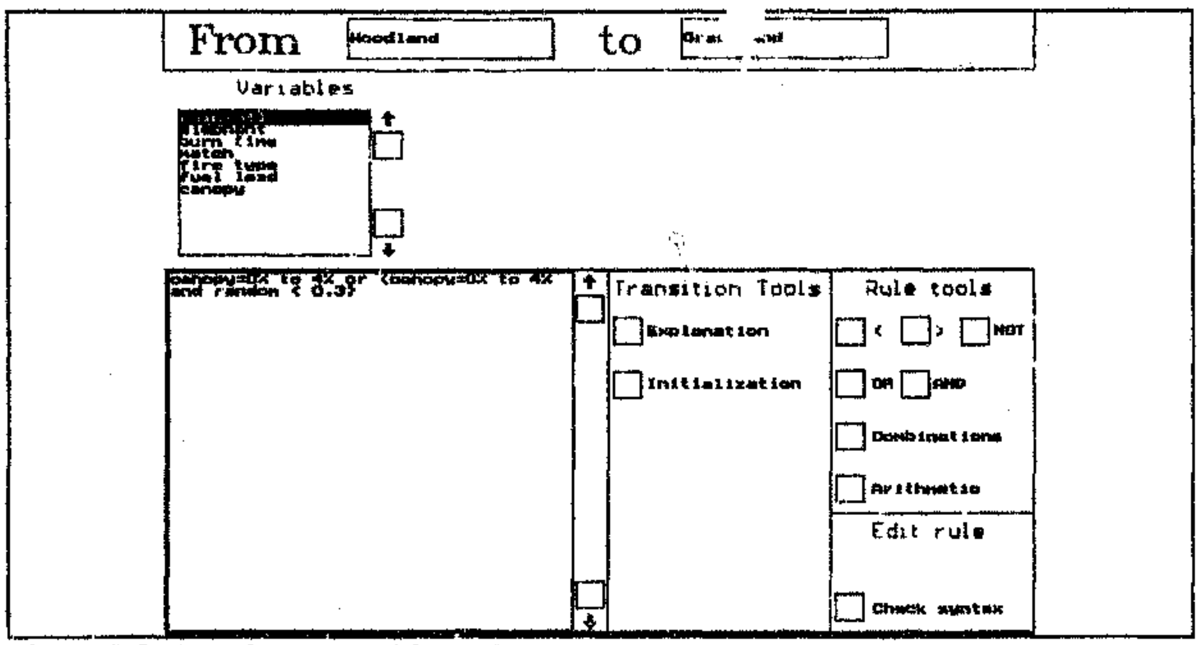


Figure 3.8: Entering a transition rule.

3.4 Informal versus Formal frame modelling

The great attraction of the informal frame modelling technique presented above is the ease with which models may be built, especially through the use of a specialized frame modelling environment. But that is not sufficient justification for pursuing this approach: It may be argued that rule based models with finite variables simply describe what is obvious and that the only benefit obtained from their use is during the modelling stage when the modeller is forced to formalize the qualitative knowledge he or she possesses about the system. Is there a limit to the possible resolution of models built using finite variables? Does the use of rules allow the building of models which can simulate ecosystems dynamics to an 'acceptable' degree of accuracy? If so, then under what conditions? These questions may appear to have little practical use since we simply do not have enough data to be able to model these ecosystems using conventional techniques. We are 'forced' into building qualitative rule-based models, whether they can be used for building highly accurate models or not. But by investigating them through a formal description of frame modelling, we can critically evaluate and determine the possible limitations of the technique. The formal structure may also suggest improvements in the algorithms used in these models.

The rest of this thesis is devoted to building a theoretical framework for the frame modelling technique and the use of this structure in looking at the questions posed above in detail.

3.5 Synopsis of Formal Frame Modelling

Before defining frames, we need to establish exactly what we mean by an ecosystem. A large portion of **Chapter 4** is devoted to the definition of a dynamical ecosystem system in as general a setting as possible. Referring to **Figure 3.9**, we construct a sample space Ω for the ecosystem where every point x in Ω refers to a possible change of the system in phase space. We assign a probability to certain subsets of Ω where each subset is interpreted as an experiment and the probability of that subset is the probability that the system is somewhere in that subset. This collection of subsets is denoted by $\mathcal{F}$. In this way, the dynamical behaviour of the system is completely non-deterministic. Mathematically, the set Ω together with the probability function on it is a probability or measure space with $\mathcal{F}$ the collection of measurable subsets.

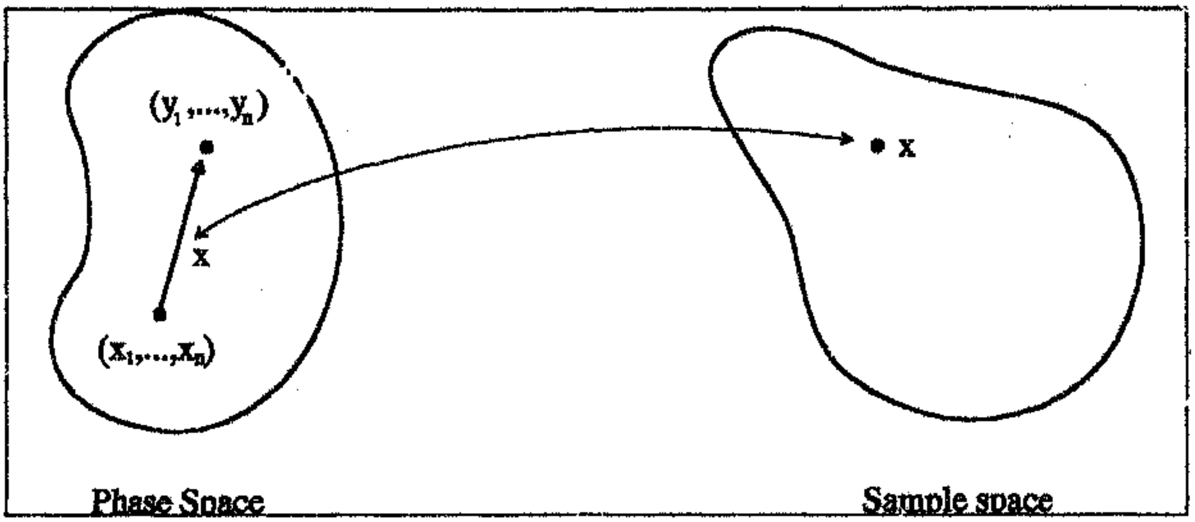


Figure 3.9: Definition of the sample space Ω

The dynamical behaviour of the system is determined by a function **T** which maps points of Ω into Ω. This function preserves the probability structure: If E is an experiment, then the probability of the time evolution of that experiment is equal to the probability of E. The transformation **T** is defined so that $\mathbf{T}^{-1}$ provides the forward moving time mechanism of the system: Looking backwards in time, we know where the system came from, hence **T** is a function. But since the system is non-deterministic, we may only assemble a set of possible future movements which are mapped by **T** to the current movement of the system. In other words, the most we can do is obtain the preimage $\mathbf{T}^{-1}\omega$ of the current system change ω.

Stability of the system is defined in as general a setting as possible. Referring to **Figure 3.10**, if E represents the set of possible configurations of the system then we assume that after a finite period of time, the system will be in one of the **attractors** Ω_i^*. The **attractors** of the system are simply defined to be those regions from which the system cannot escape without an external perturbation.

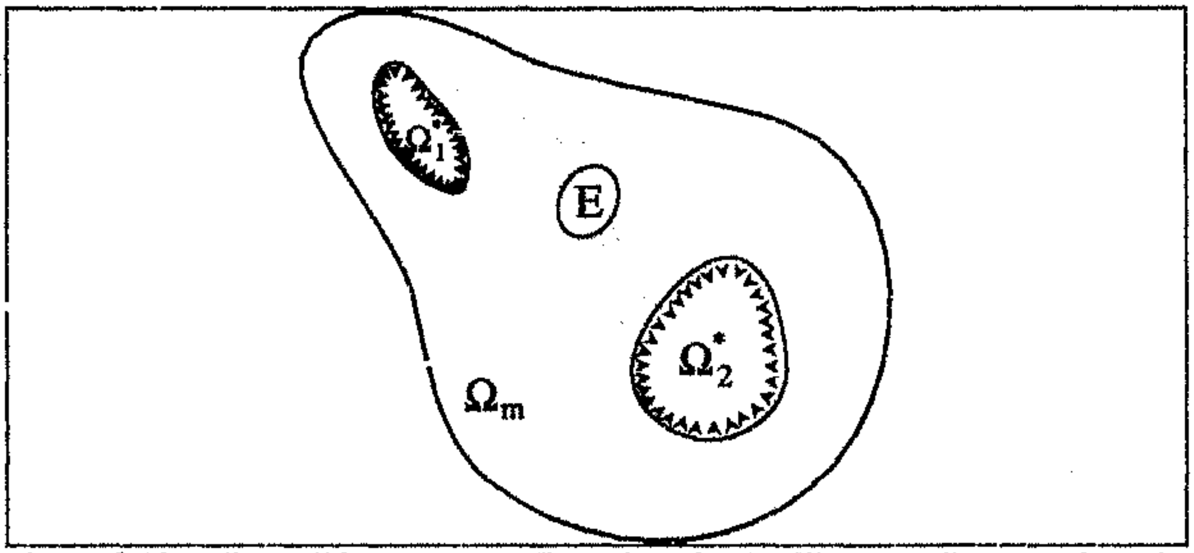


Figure 3.10: All possible system configurations in E will eventually move into the attractors

The **domain of attraction** Ω_i associated with the **attractor** Ω_i^* is then defined to be the set of all points from which the system has a non-zero probability of moving into the **attractor**. This definition allows the overlapping of the **domains of attraction**, as is illustrated in **Figure 3.11**:

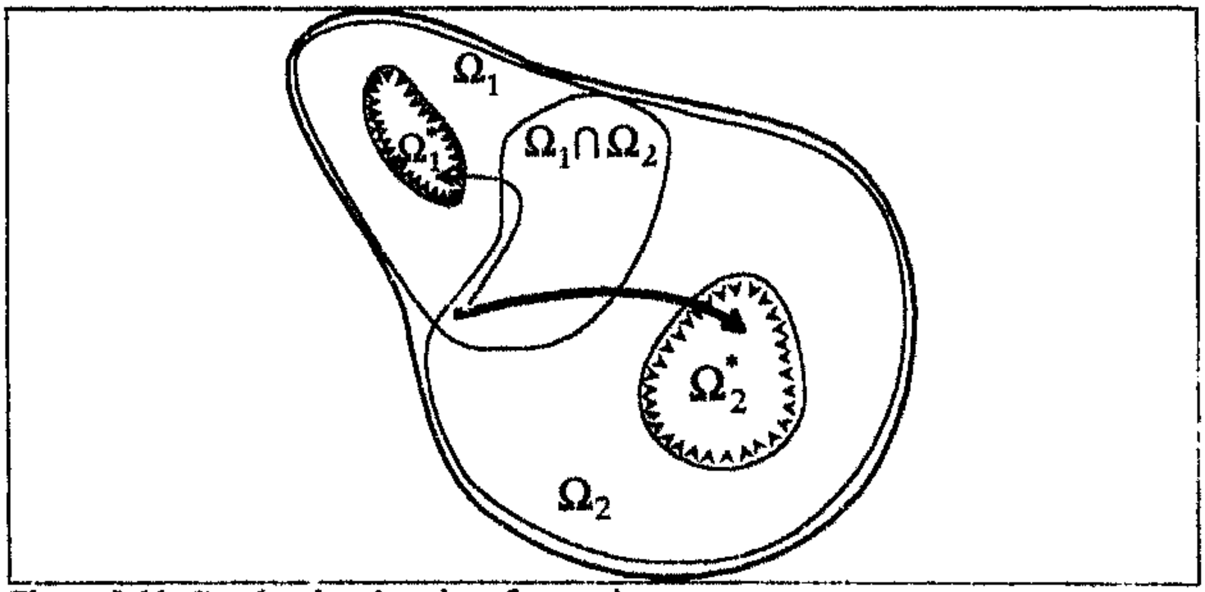
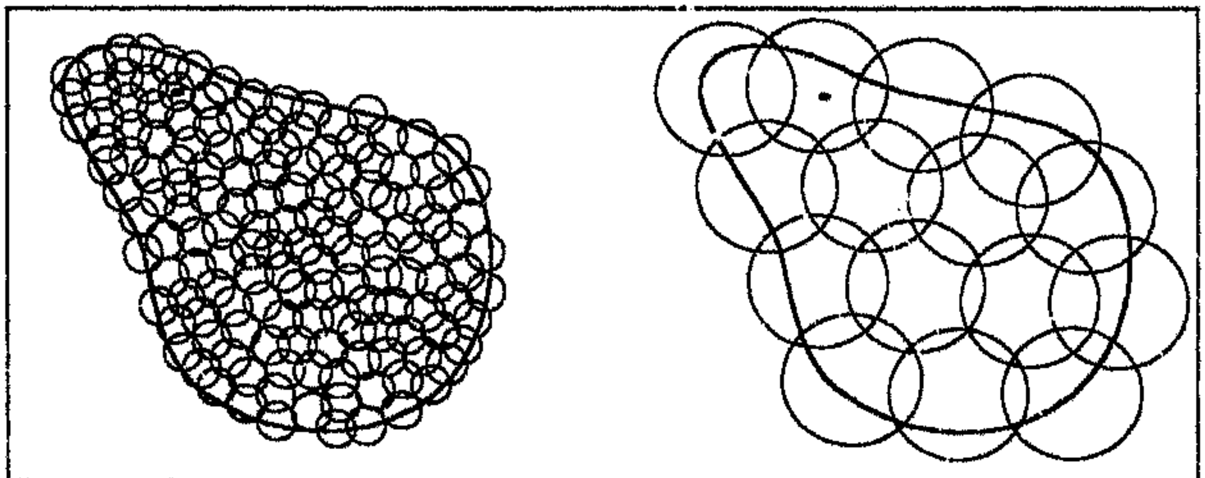


Figure 3.11: Overlapping domains of attraction

In the figure, the system may move from the same initial point towards two different domains of attraction. The probability of the movement is depicted graphically by the thickness of the arrow.

This definition of stability allows for all cases from the domain of attraction consisting almost entirely of the fuzzy region to the case where the domain of attraction is completely filled by the attractor. **Figure 3.11** shows the attractors and domains of attraction as being connected: In general, both may be disconnected. A **frame** is then defined in terms of these domains of attraction and the dynamical function **T** acting on the domain, as in the informal case.

Few would argue that the dynamical system defined above for the modelling of ecosystems is not sufficiently generalized. But as was explained in **Chapter 1**, the intention of frame modelling is **not** to determine **T**: We assume that such a **T** exists and represents the true dynamical behaviour of the ecosystem. A result to be proved in **Chapter 4** says that, in spite of the generality of the ecosystem dynamics, it is possible at each time step to obtain an approximate description (measured by the **information** or **entropy**) which is as close to the 'true' description as we like. Of great importance is the fact that this description may be given in terms of a **finite** set of symbols (a finite alphabet). **Figure 3.12** demonstrates how these symbols arise.



Covering of Ω by $\alpha=\{\alpha_i : i=1,...,100\}$ Covering of Ω by $\alpha=\{\alpha_i : i=1,...,15\}$

Figure 3.12: **The choice of alphabet depending on the information content of Ω**

On the left hand side, the sample space is covered with 100 open sets each of which is a symbol in the alphabet. The dot represents the instantaneous state of the system, and falls within the set α_{34} so that the symbol α_{34} may be used to describe the current state of the system. On the right hand side, we are far less demanding in knowing the state of the system so that a coarser covering set is being used. In this case, the alphabet has 15 symbols and the system can be described (although less accurately) by α_6.

In **Chapter 5**, the fact that a finite alphabet may be used to describe the ecosystem is used with great effect to demonstrate that qualitative rule based models may be used to model the dynamical ecosystem defined above. This is done by splitting the system variables into two sets: The input variables which are not modelled, and the output variables which must be modelled. The behaviour of the output variables is determined by **Probabilistic Finite Automata**, which are abstractions of non-deterministic finite state machines.

The structure given above achieves the task of demonstrating that frame based qualitative rule based models can be used to model ecosystems as accurately as we like, so that its theoretical value is high. But we are also interested in useful self-tuning algorithms for qualitative frame modelling, and from this practical modelling viewpoint, the structure built up from finite automata has a drawback: When developing models, the user may want to make references to past values held by variables (for example, the state of the vegetation may be most easily predicted in terms of the rainfall of the previous year). The structure detailed above does not allow this.

By adding memory to these automata (in the form of a stack), it is possible to allow the user to make references to past values if desired. **Figure 3.13** demonstrates the complete formal frame structure defined in **Chapters 4** and **5** using finite automata with stacks (**Probabilistic Pushdown Automata** or **PDAs**). Since this structure is static, it immediately suggests algorithms for tuning models using observed data. The variables are divided into the input group **I** and the output group **O**. The observed values of the output variables are not used by the completed model, but are used to tune the model (represented through the use of dotted lines). Every possible combination of the input and output variables including missing values, is enumerated by a combination alphabet

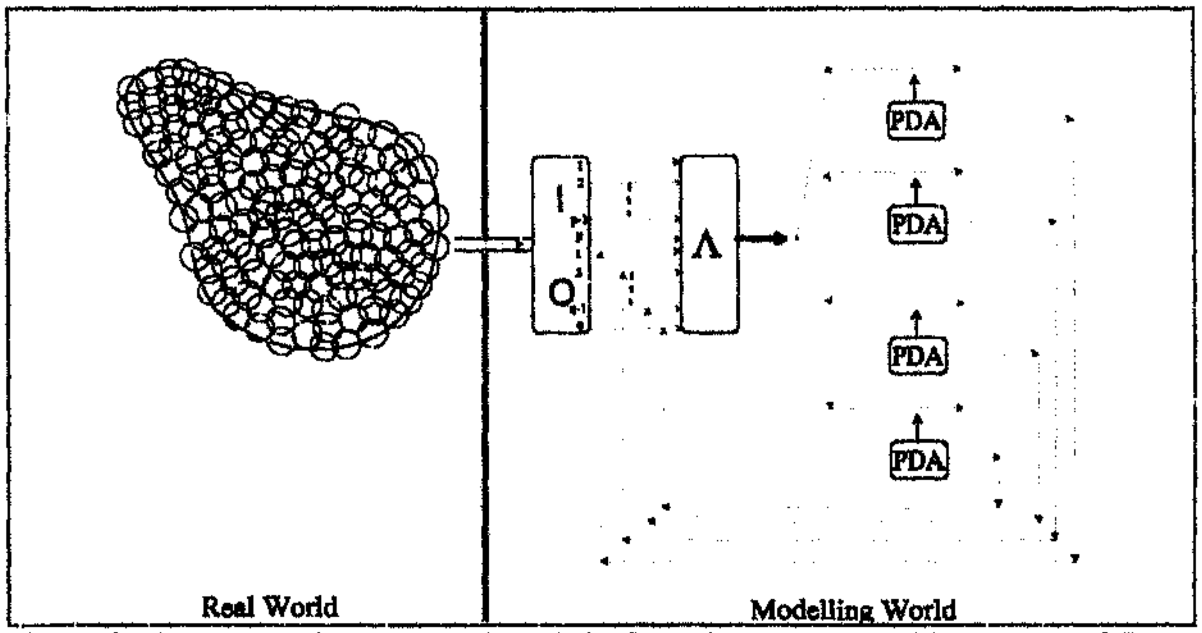


Figure 3.13: Graphical representation of the formal structures of Chapters 4 and 5

Λ. Each output variable is then modelled by a probabilistic pushdown automaton which depends on the current state of the system (as described by Λ), and the contents of the stack (history of the system movement). Tuning of the model then corresponds to the finite automata improving their transition functions through learning from the observed output data. At least two additional primitive symbols must be added to Λ to form Λ_p, and these symbols cannot be obtained through observation of the ecosystem dynamics through the communications channel: These symbols arise from a preliminary untuned model entered by the user in a **Context Free Language** (a PASCAL type language). (The class of languages accepted by pushdown automata is precisely the class of context free languages).

Figure 3.14 summarizes the important results to be proved in **Chapters 4** and **5**, based on three postulates. These results demonstrate that the informal frame modelling technique described earlier in this chapter may theoretically be refined to describe the ecosystem dynamics as accurately as we like. We are constrained only by the measuring difficulty. The formal theory goes further, however, and provides a theoretical basis or skeleton to be used for planning future implementation software.

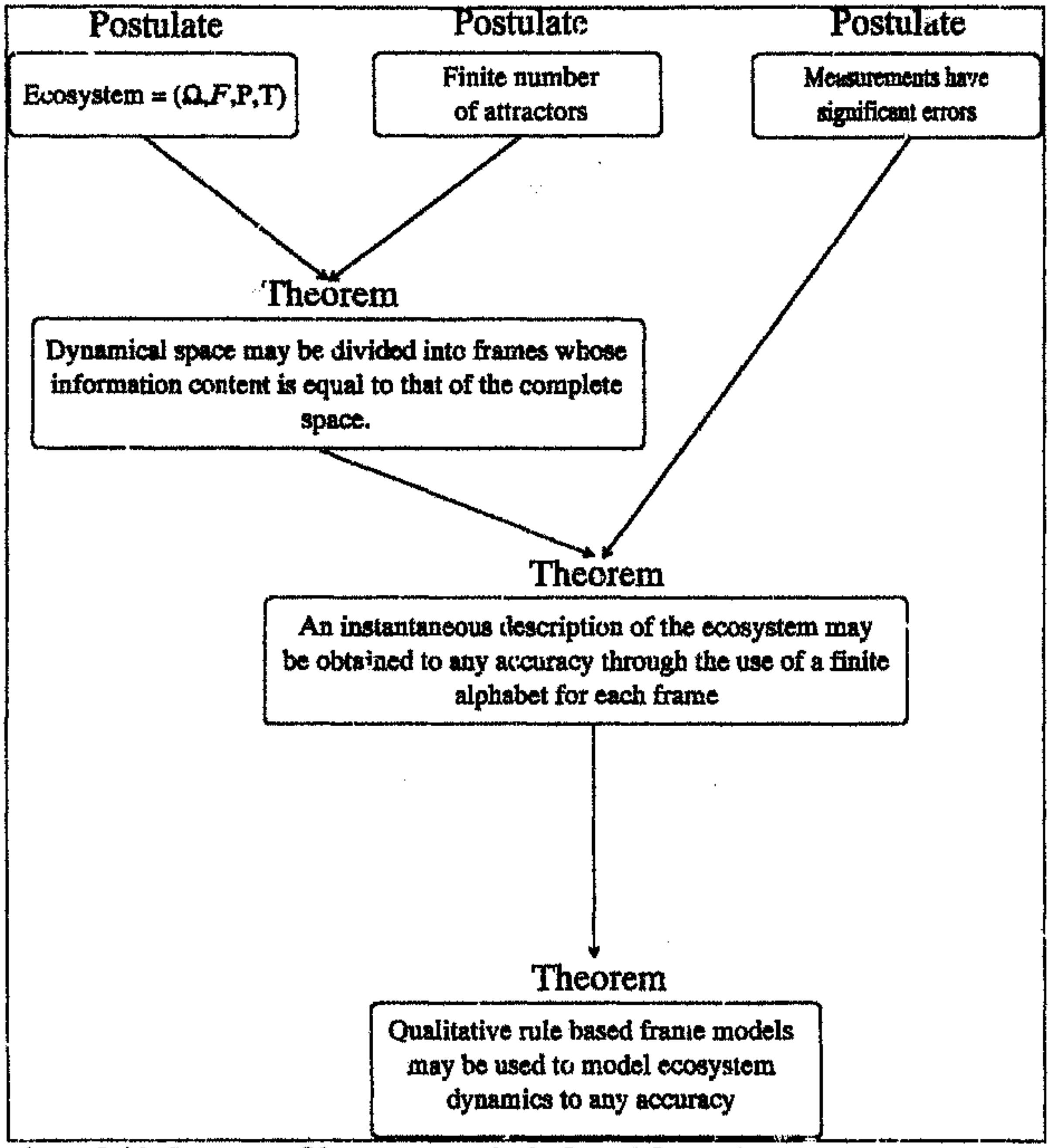


Figure 3.14: Summary of important results in formal frame theory

Chapter 4: Formal Frame Modelling

4.1 Introduction

In this chapter, the general dynamics view of ecosystems is discussed. As is usually the case, the less that is assumed about a system, the more sophisticated is the mathematics required to describe the dynamics within the system. A general description of ecosystem dynamics is no exception. Much of this chapter, therefore, is devoted to defining the components with which the postulates of the formal dynamics will be built. The formal ecosystem dynamics proposed in this chapter are preliminary in that they suggest a framework for investigating ecosystem dynamics and demonstrate the link with computer science theoretic frame-based modelling. It is expected that valuable understanding of theoretical ecosystem dynamics may be obtained by further exploring this route. Before looking at these postulates, the conceptual models in ecology which inspired them are discussed.

4.2 Succession in ecosystems

Ecological succession was one of the earliest conceptual models forwarded for describing ecosystem dynamics, and originated largely from observed changes in the spatial distribution of vegetation in disturbed areas, for example the pattern of revegetation in abandoned farmland[19]. Clements[20,21] applied the term to ecosystems as a whole, with succession within ecosystems being likened to the recovery of an organism from injury. The term was later applied to the adaptation of individual species rather than referring to a movement of the entire system[22], a view of succession which persists today. The classical ecological succession model is based on two key assumptions[23]:

> a. Species replacement occurs because of the less favourable conditions brought about by the species themselves, and the new conditions are more favourable for the replacement species.

b. A self-perpetuating climax state exists representing a natural end to the succession which occurs from the initial disturbance.

It is difficult to find examples of a system in a climax state, since major disturbances of ecosystems (such as fires, droughts and floods) appear to be more the norm than the exception.

This classical view was criticized by Egler[22] as not being applicable to all situations. He suggested that the succession depends very much on the initial conditions following a major disturbance (he referred to this as the **initial floristic composition** model in contrast to the classical **relay floristic** model), with the implication that if a species did not persist through a disturbance, then it could not be represented in the succession to follow. Connell and Slayter[23] later suggested that these models be combined by the recognition of three main pathways: The first is the **facilitation pathway** which is the **classical relay floristic pathway** where the earlier species facilitate the entry of the replacement species. The second is the **tolerance pathway** in which certain species become established whether or not they have been preceded by other species and a[illegible]rant of the presence of the other species. The third is the **inhibition pathway** in which [illegible] species cannot become established in the presence of earlier species. These early species may become dominant even if they are not regarded as late or climax succession species by inhibiting the establishment of later species. Based on these concepts and on the literature on succession, the following generalizations have been proposed by Noble and Slayter[24]:

a. The species composition immediately following a major disturbance depends directly on the **propagules** which either persisted through the disturbance or entered from elsewhere. (A propagule is defined by Noble *et al* to be a structure produced by an organism which becomes detached from the parent and gives rise to another individual).

b. Following a disturbance there is a period of growth characterised by little or no competition for space and resources.

c. Subsequent to the initial pulse, the presence of established plants slows down the recruitment.

d. Recruitment of additional species may be facilitated, restricted or unaffected by prior occupancy.

e. In the absence of further disturbances, the long lived species and those able to regenerate in the presence of their own adults become dominant.

Using these generalizations Noble and Slayter[24] developed a scheme for predicting major shifts in dominance within plant communities which has given satisfactory results when applied to terrestrial communities dominated by higher plants.

One of the ways in which these conceptual models have been implemented is through the use of Markov processes, especially the use of first order Markov processes. In a first order Markov process type model, the system is classified into a finite number N of states, and a transition probability assigned to each of the N^2 possible transitions between states. These probabilities are usually presented as a matrix, as is illustrated below for a Markovian succession process in a forest[19] containing Gray Birch (x_1), Blackgum (x_2), Red Maple (x_3) and Beech (x_4):

$$\vec{x}_{next}=A\vec{x} \tag{4.1}$$

where

$$\vec{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{4.2}$$

and

$$A = \begin{pmatrix} 0.05 & 0.01 & 0.00 & 0.00 \\ 0.36 & 0.57 & 0.14 & 0.01 \\ 0.50 & 0.25 & 0.56 & 0.03 \\ 0.09 & 0.17 & 0.31 & 0.96 \end{pmatrix} \tag{4.3}$$

If the process in (4.1) is performed until a stable distribution is obtained, the following is obtained.

$$\vec{x}_{final} = (0,5,9,86) \tag{4.4}$$

which compares favourably to the observed distribution in an old forest:

$$\vec{x}_{observed} = (0,3,4,93) \tag{4.5}$$

but certainly does not provide absolute verification.

The extension to higher order Markov processes is possible but this has not popular because of the difficulty (both mathematically and computationally) in dealing with the non-linear systems which result. While succession models may appear to provide reasonable simulation models at the statistical level, the flaws detailed below in the succession approach would seem to be too great to justify exploring successional models as a paradigm for ecosystem modelling.

a. Successional theories have been developed for vegetation dynamics and are inadequate for the modelling of more general ecosystems. By attempting to model ecosystems through a predominantly successional approach would be to make the implicit assumption that vegetation dynamics drive the system.

b. An implicit assumption of the succession model is that there is only one stable equilibrium state in the system, the climax state. This is perhaps the most serious flaw in successional models.

4.3 State-and-Transition models

Even in systems for which the succession model seems ideally suited, such as in dynamic rangelands, the succession model has been shown to be deficient[25]. For these systems, an alternative conceptual model has been proposed by Westoby, Walker and Noy-Meir[25] which on the surface appears to be very similar but has a number of fundamental differences. This conceptual model has been called a state-and-transition model and, like the successional model, categorizes the system into a finite number of states. But very different to the succession model is how each of these states is chosen: In the state-and-transition model each state is far broader than any state in the successional model and corresponds to a configuration of the system where the dynamic movement of the system is slow. The transitions refer to a shift between these states rather than species replacement of successional models. This conceptual model was demonstrated to be more effective in describing the dynamics of rangelands, and various examples were presented clearly indicating the usefulness of this approach.

This conceptual model has lead directly to the frame-based modelling technique and of particular importance to frame modelling is how each of these states is defined. In their paper, Westoby *et al* did not explicitly pin each state to a possible stable configuration of the system but divided the system dynamics into slowly moving states and relatively rapidly moving transitions. By doing this, they implicitly broadened the concept of stability in ecosystem dynamics to include those configurations of the ecosystem in which the dynamic movement of the system is slow. The formal frame theory to be presented develops this concept of stability is a rigorous manner.

4.4 General Dynamics in Ecosystems

Based on the experience gained from studying the conceptual models outlined above, the following requirements are proposed for developing a model of ecosystem dynamics:

a. The system must be built up from a non-deterministic base. It is far more difficult to build a deterministic framework for the ecosystem dynamics and then try to add non-determinism than the other way round.

b. The probability structure of the dynamics is independent of time. Put another way; if the ecosystem were observed for a fixed length of time, then reset to its initial conditions and observed again for the same length of time, the probability structure in both cases would be identical (a statistical analysis would indicate the indistinguishability). This is an extremely important assumption, as will become evident when the ecosystem dynamical space is studied in detail. What this means is that the basic processes (which may be wholly non-deterministic) and components of the ecosystem being studied are not changing with time. Over very long periods of time ecosystems **do** change through evolutionary processes, but this change is certainly insignificant over the time periods to be modelled. It is possible to relax this requirement by considering a family of dynamical spaces each of which is stationary (in the sense referred to above) and indexed by time, but that would result in dramatically increased mathematical complexity and consequently will not be considered in this thesis. On the other hand, it may be argued that evolutionary change in ecosystems does not constitute a fundamental change, but instead represents in a large part the relabelling of processes and components (i.e. the family of spaces mentioned above would be isomorphic). In this sense, the dynamics would be stationary or very close to stationary.

c. The general dynamics must be sufficiently general to include all other modelling approaches as special cases.

There is a branch of mathematics called Ergodic theory (the theory of measure preserving transformations) which is concerned with precisely the type of dynamical system characterized by the requirements listed above. Requirements **a** and **c** are satisfied by working with a general probability space (simply a measure space where the measure of the universe set is unity). Requirement **b** is satisfied by working with probability preserving transformations.

4.5 General dynamical spaces

The following postulates are proposed for the ecosystem dynamics. The motivation for each and notation used will be discussed in detail in this chapter.

Postulate 1: A dynamical ecosystem is completely specified by the dynamical quadruple $(\Omega,\mathcal{F},P,T)$.

Postulate 2: In the dynamical ecosystem $(\Omega,\mathcal{F},P,T)$, there exist a finite number of **attractors** Ω_i^* with their associated domains of attraction Ω_i where $i=1,\ldots,n$. Each pair $[T_i,\Omega_i]$ where T_i is the restriction of T to Ω_i is defined to be a **frame**.

Postulate 3: For each Ω_i there is a discrete channel $[\Sigma_i(A_i),P_i(\omega,\ ),\Sigma_i(B_i)]$ of limited capacity C_i through which all information about T_i must pass. The alphabets A_i and B_i may be chosen as desired, the only fixed parameter of the channel is C_i.

The rest of this chapter is concerned with the definition of the terms used in these postulates, and the results leading up to the **Fundamental Theorem of Frame Modelling**. The requirement that the channel associated with each formal frame $[T_i,\Omega_i]$ be discrete may appear to be too great a hindrance, but it will be shown in this chapter that such a channel can be used to transmit the information to an arbitrary degree of accuracy:

Fundamental theorem of frame modelling: Given a dynamic ecosystem $(\Omega,\mathcal{F},P,T)$ with N **attractors**, the following are true:

a. The system may be decomposed into formal frames $[T_i,\Omega_i]$ where $i=1,\ldots,N$ such that the total information content of the formal frames is equal to the information content of the entire space.

b. An instantaneous description of the ecosystem may be obtained using N channels (associated with each formal frame) of limited capacity, to an arbitrary accuracy and using a finite alphabet. The accuracy is bounded only by the measuring error.

As will be demonstrated in **Chapter 5**, the use of a finite alphabet to transmit the information is of immense value when it comes to developing algorithms for modelling ecosystem dynamics.

4.5.1 Basic definitions

The underlying space to be used is a general probability space:

Definition: A general probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is an arbitrary set, $\mathcal{F}$ is a collection of subsets on Ω and P is nonnegative real valued function on Ω and the following are true:

a. $\mathcal{F}$ is closed under countable unions, complements and contains Ω (i.e. $\mathcal{F}$ is a σ-algebra).

b. If $\{E_n\}$ is a pairwise disjoint collection of members of $\mathcal{F}$, then

$$P(\bigcup E_n) = \sum P(E_n) \tag{4.6}$$

(i.e. P is a measure).

c. $P(\Omega) = 1$

The set Ω may be interpreted as the sample space, and the collection $\mathcal{F}$ as the set of possible experiments or outcomes, with a probability assigned to each member of $\mathcal{F}$. Partitions of Ω form a very important tool in dealing with probability spaces, and in the case of dynamical system are related to stochastic processes within the space (see page 58).

Definition: A partition ξ of a probability space $(\Omega,\mathcal{F},P)$ is a collection of subsets of Ω which are disjoint and whose union is Ω. The elements of ξ are often referred to as the **atoms** of the partition. The factor space $(\Omega_\xi,\mathcal{F}_\xi,P_\xi)$ is the probability space formed on the set Ω_ξ which contains all the atoms of ξ, $\mathcal{F}_\xi$ which contains all the $\mathcal{F}$-measurable atoms of ξ and P_ξ which is the restriction of P to Ω_ξ.

It is possible to define a partial ordering on the set of all partitions of Ω: If two partitions ξ and ζ are chosen so that every atom of ξ can be constructed from a union of atoms of ζ, then ζ is said to be a refinement of ξ and this is denoted by $\xi\leq\zeta$ (intuitively ζ does have more elements than ξ). Given any set of partitions $\{\xi_\alpha:\alpha\in A\}$, it is possible to form the supremum[26] or common refinement denoted by $\vee_\alpha\xi_\alpha$ with $\xi_\alpha\leq\vee_\alpha\xi_\alpha$ for all $\alpha\in A$, and similarly the infimum denoted by $\wedge_\alpha\xi_\alpha$ may be defined.

It is not necessary for Ω to have any particular structure (such as a topology) to obtain the general results, but often in the literature probability spaces are assumed to be Lebesgue spaces since these have a far richer mathematical structure, are well understood and most random phenomena can be modelled as a factor space of them[26]. For reference, a Lebesgue space is defined[26]:

Definition: A probability space $(\Omega,\mathcal{F},P)$ is a Lebesgue space if it is isomorphic to a probability space which is the disjoint union of a countable set x_i, $i\in\mathbf{N}$ each of positive measure $p(x_i)$ and the probability space $([0,s],\mathcal{F}_1,M)$ where

$$s=1-\sum_{i=1}^{\infty}p(x_i) \tag{4.7}$$

$\mathcal{F}_1$ is the collection of Lebesgue measurable sets on $[0,s]$ (M is the Lebesgue measure).

Intuitively, we would expect that most random phenomena could be constructed from the unit interval [0,1] with the most common real measure, the Lebesgue measure. There is no real loss in generality by assuming all probability spaces to be Lebesgue spaces, and in what follows no distinction will be made between the two: We assume all probability spaces to be Lebesgue spaces.

Another common way of looking at probability spaces is in terms of the Hilbert space $L^2(\Omega,P)$ which is the Hilbert space of all functions over Ω square integrable in terms of the measure P, that is the Lebesgue integral

$$\int f^2 dP \tag{4.8}$$

exists and is bounded. This viewpoint is especially useful when looking at invertible dynamical systems (see **Section 4.5.2**) as it allows a linear space to effectively replace a non-linear space. In frame modelling, one of the fundamental results is the demonstration that the entropy of the dynamical system may be decomposed into a direct sum, which may in turn be interpreted as indicating that the subspaces over which the entropy is decomposed are independent is some sense. For this reason, we must consider conditional entropies which in turn requires knowledge of how conditional probabilities are dealt with in general probability spaces. In elementary probability theory, the conditional probability of an event E_1 given event E_2 is defined to be

$$P(E_1 | E_2) = \frac{P(E_1 \bigcap E_2)}{P(E_2)} \tag{4.9}$$

This concept of conditional probability may be cast in a more useful form by looking at the decomposition of an experiment using a measurable partition of the space. The basic idea is simple: Given a countable measurable partition ξ, each member $E\in\xi$ defines a space $(E,\mathscr{F}(E),P(\ \ E))$ where $P(\ \ E)$ is a measure acting on $\mathscr{F}(E)$. For any $G\in\xi$

$$P(G) = \sum P(E)P(G \bigcap E | E) \tag{4.10}$$

In terms of the factor space, this may be written

$$P(G) = \int_{\Omega_\xi} dP_\xi P(G|E) \tag{4.11}$$

since each atom of ξ is identified as a point in Ω_ξ. It can be shown that this decomposition can be done for any measurable partition.

Theorem : Suppose ξ is a measurable partition of the space $(\Omega,\mathcal{F},P)$ and N_ξ the projection associated with ξ. For almost all $c\in\Omega_\xi$ (in terms of P_ξ) there exists a probability measure $P^\xi(\ \ c)$ on $\mathcal{F}$ such that

$$P(G)=\int_{\Omega_\xi} P^\xi(F|c)dP_\xi=\int_{\Omega} P^\xi(F|N_\xi(\omega))dP \tag{4.12}$$

The projection N_ξ maps a point ω in Ω to the atom in ξ containing ω. The function $P^\xi(\ \ |N_\xi(\omega))$ is also denoted by $P^\xi(\omega,\ )$, so that a similar form to **(4.9)** may be obtained:

$$P(E_1\bigcap E_2)=\int_{E_2} P^\xi(\omega,E_2)dP \tag{4.13}$$

We may now define what is meant by a dynamical system:

Definition: A dynamical system is a quadruple $(\Omega,\mathcal{F},P,T)$ where $(\Omega,\mathcal{F},P)$ is a probability space, and $T:\Omega\to\Omega$ is a measurable transformation such that for any member $E\in\mathcal{F}$, $T^{-1}E\in\mathcal{F}$ and $P(T^{-1}E)=P(E)$. (**T** is probability preserving).

This is the most general mathematical model of a dynamical system, and includes as special cases all other models of dynamical systems. **T** is an example of a **metric endomorphism** and specifies the time development of the dynamics. If T^{-1} exists and is also a metric endomorphism, then the dynamical system is said to be **invertible** and **T** is referred to as a **metric automorphism**. The conditions under which **T** is invertible may be relaxed in the usual way through the use of the notion of **equivalence almost everywhere** (ae) (where non-equivalence over subsets of measure zero is ignored):

Definition: A dynamical system $(\Omega,\mathcal{F},P,T)$ is invertible if the range of **T** is almost all of Ω (i.e. the set difference between the two has measure zero) and the restriction of **T** to a subset of Ω of measure 1 is invertible and this inverse map is a metric endomorphism.

As an example of how other dynamic models can be generated using this approach, we consider Markov processes which are of particular importance as discussed in **Section 4.2** because of their use in ecological modelling. Before doing so, it is necessary to define exactly what is meant by a stochastic process:

> **Definition:** A stochastic sequence in a measure space is a collection of random variables defined on a probability space and indexed by a subset of the integers. (A random variable is a measurable function on the sample space). If the stochastic sequence is measure preserving, then it is a stationary stochastic sequence.

As an example, if x is a random variable on a dynamical system $(\Omega,\mathscr{F},P,T)$, and if we define

$$x_n(\omega)=x(T^n(\omega)) \qquad \textbf{(4.14)}$$

then $\{x_n\}$ is a stationary stochastic process. In a probability space, members of $\mathscr{F}$ correspond to different possible experiments and so intuitively it would be expected that a stochastic sequence relates to a countable subset of this collection (a measurable partition). This is a very useful way of looking at stochastic processes:

> **Theorem**[26]: Stationary stochastic sequences are equivalent to measurable partitions of a dynamical system.

A Markov process in particular may be constructed using a Markov partition ξ which satisfies the condition that the conditional probability of ξ given $\bigvee_{1,n} T^{-i}\xi$ is equal to the conditional probability of ξ given $\mathbf{T}^{-1}\xi$. At first glance, this definition may appear to be very far from the conventional view of Markov processes. But by looking at $\mathbf{T}^{-1}\xi$ as the experiment ξ being conducted one time step later, the condition above is consistent with the usual view of a first order Markov process only depending on the current state of the system, and not on the history of how the system got there. The interpretation of $T^{-1}\xi$ as being the experiment ξ one time step later may be confusing when first seen, since the operator **T** would at first appear, in a sense, to be the 'wrong way around'. There is a mathematical justification for doing this[26,27,28], but on further reflection the reason why it must be so should become apparent: In a non-deterministic system, we cannot exactly predict where the system will be going but we can see why it came from where it did. In other words, **T** is only definitely a function when looking backwards in time. When looking forward in time, the most we can expect is to be able to assemble the possible outcomes, that is determine the preimage of the current state of the system. Consequently, by only demanding that **T** be a function when looking into the past, we are ensuring that the dynamical system is as general as possible and that **T** accurately represents development of the dynamic system with time. Of course, in the case of an invertible dynamical system, both **T** and $\mathbf{T}^{-1}$ exist and are functions.

A further assumption is often made about **T** when dealing with dynamical systems, and that is to assume that **T** is ergodic. For **T** to be ergodic, for any member E of $\mathcal{F}$ with $\mathbf{T}^{-1}E=E$ either P(E)=0 or P(E)=1. In other words, the only processes not affected by **T** are trivial. The reason why assuming ergodicity is so useful becomes evident when looking at the time and space means of an arbitrary integrable function f:

The time mean of f at E is given by

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}f(T^i(E)) \tag{4.15}$$

since each $f(\mathbf{T}^i(E))$ represents a time step at that region. The space mean of f is given by

$$\int_E f\,dP \tag{4.16}$$

A very important theorem may be proved for ergodic systems which states that these means are equal if and only if **T** is ergodic.

Birkoff Ergodic Theorem[27]: If T is a measure preserving operator on a measure space and f is an integrable function over the space (i.e. $f \in L^1(m)$), then

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \qquad (4.17)$$

converges almost everywhere to a function $f^* \in L^1(m)$, $f^* \bullet T = f^*$ almost everywhere and if the measure of the entire space is finite, then

$$\int f^* dm = \int f dm \qquad (4.18)$$

4.5.2 Linearizing the ecosystem dynamics

In **Chapter 2**, the linearization of a system of differential equations was discussed. The prime purpose in linearizing the system dynamics is to obtain an eigenvalue equation such as in **(2.10)**:

$$\frac{d\vec{x}}{dt} = A\vec{x} \qquad (4.19)$$

By calculating the eigenvalues of A, much may be deduced about the stability of the system. But as was pointed out, the approximation used to obtain **(4.19)** depended on the assumption that the system is close to equilibrium, a condition which is totally unacceptable in the modelling of dynamic ecosystems. In the general dynamics we have been developing, there is way of interpreting the dynamics as a linear space in a manner which does not require unjustified approximations. in fact the correspondence is exact. The way this is done is very simple: For each transformation T acting on a probability space $(\Omega, \mathcal{F}, P)$ we define the operator U_T acting on the Hilbert space $L^2(\Omega, P)$ by

$$U_T f = f \cdot T \qquad (4.20)$$

This operator is linear, which means that every measure preserving transformation on Ω may be mapped to a linear operator. Just as in the finite dimensional vector case, an eigenvalue analysis may be performed: In the case of Hilbert spaces, the set of complex numbers λ such that $U_T - \lambda 1$ does not have a bounded inverse defines the spectrum of U_T. In Hilbert space theory, an eigenvalue analysis is greatly aided if the operator in question is unitary (for example, in the case of a unitary transformation each element of the spectrum is on the unit circle centred at the origin). If the ecosystem dynamical system is invertible (i.e. **T** is an automorphism), then U_T is a unitary operator[26]. In other words, if the ecosystem is invertible (reversible in the probabilistic sense) it may be modelled by a linear system about which much is known, and the mapping is true globally, not just in the neighbourhood of an equilibrium point.

4.5.3 Entropy and Information in a dynamical system

When dealing with random phenomena, a quantitative measurement of uncertainty is often required. In thermodynamics and statistical mechanics in general, this quantity is given by the **entropy**, so it is natural to try to define the entropy of a general dynamical system. In thermodynamics, entropy is not explicitly defined in a rigorous way, but is rather assumed to exist along with a number of relations involving it and the other thermodynamics variables. For a general dynamical system, there have been two main attempts at defining an entropy and those are the **topological entropy** and the **measure theoretic entropy**. The connection between these two entropies was eventually demonstrated[27] so that it is really just a matter of convenience when choosing which to work with. Before looking in detail at the definition of information and entropy, it is illuminating to look at the commonsense meaning of information.

4.5.4 Form of an information function

The form of the entropy and information functions may be determined (up to a constant) by a simple intuitive example: Suppose a computer is required to process and store data from the outside world. This information would ultimately be made up of numbers, and the most obvious quantitative measure of the information in that data is provided by the question 'How many Kilobytes does this information occupy?'. In other words, the most obvious measure of the information is the number of bits required to store the information in binary format. The number of bits required to store a number N is simply the integer value of the number $1+\log_2 N$ By changing the base of the logarithm and combining all constants into a single constant k, we would therefore expect that a reasonable measure of the **information** would be given by:

$$I=k \log N \tag{4.21}$$

This is exactly the form obtained for estimates of the entropy in statistical mechanics. From an abstract point of view, it is possible to determine this form of the **information** (or **uncertainty**) by defining the **information** to have a number of properties we would expect:

Definition: The **uncertainty** is a real valued function satisfying the following:

a. An event whose probability of occurring equals 1 has zero uncertainty.
b. If for two events E_1 and E_2, $P(E_1)<P(E_2)$ then E_1 has a greater uncertainty than E_2.
c. The uncertainty of two simultaneous but independent events is equal to the sum of their uncertainties.

Suppose Φ is a function on the interval [0,1] which satisfies the three conditions given above. Then it is clear that Φ must be a monotonic decreasing function on [0,1] and must satisfy the condition

$$\Phi(xy)=\Phi(x)+\Phi(y) \tag{4.22}$$

A well-known result from analysis is that the only measurable function obeying **(4.22)** is the log function (up to a constant):

$$\Phi(t)=-b \log t \tag{4.23}$$

for $t\in[0,1]$ and where we define $\Phi(0)=\infty$. Consequently, a function satisfying the conditions above is given by:

$$I(E)=\Phi(P(E)) \tag{4.24}$$

Based on the forms suggested above, we may provide a definition of uncertainty which is consistent to the traditional view.

> **Definition**: Suppose ξ is a countable measurable partition of the space $(\Omega,\mathscr{F},P)$. The uncertainty function of this partition is given by
>
> $$I(\xi)(\omega)=-b\sum_{A\in\xi} 1_A(\omega)\log P(A) \tag{4.25}$$
>
> where $\omega\in\Omega$. If ζ is another measurable partition then the conditional uncertainty of ξ given ζ is given by
>
> $$I(\xi/\zeta)(\omega)=-\sum_{A\in\xi} 1_A(\omega)\log P^{\zeta}(\omega,A) \tag{4.26}$$

The function $1_A(\omega)$ is simply 1 if $\omega\in A$ and 0 otherwise. In other words, the uncertainty of a partition at a particular point in the sample space is just the uncertainty of the atom A of the partition containing ω.

The **entropy** of a partition may now be defined as the expected value of the **uncertainty** of the partition.

4.5.5 Entropy of a random process

Definition: The entropy $H(\xi)$ of a measurable partition ξ of the space $(\Omega,\mathscr{F},P)$ is given by

$$H(\xi)=\int_{\Omega} I(\xi)dP \tag{4.27}$$

which may be expanded to

$$H(\xi)=-b\sum_{A\in\xi} P(A)\log P(A) \tag{4.28}$$

The conditional entropy $H(\xi/\zeta)$ of the partition ξ given the partition ζ is the expected value of the conditional uncertainty and is given by[26]:

$$H(\xi/\zeta)=\int_{\Omega_\zeta} H(\xi\bigcap E)dP_r \tag{4.29}$$

This view that the **entropy** of a partition is the average amount of uncertainty in the process represented by the partition is consistent with the usual view of **entropy**. It may be shown that by measuring the **entropy** in bits, a lower bound is obtained to the expected number of yes-no questions required to determine the outcome of the process. This fact is not so much useful as illuminating since it gives a clear physical interpretation to **entropy**. The demonstration of this result is also very similar to the proof of the **Coding Theorem** which is so important in demonstrating the link between the formal dynamics and computer science theoretic frame-based modelling. For this reason, this interpretation of entropy is demonstrated in detail below:

Suppose the process whose outcome is being sought has n possible outcomes $\{x_1,x_2,..,x_n\}$ each of which has probability p_i. For each of these outcomes, we let N_i be the number of questions required to determine whether x_i is the outcome. We must firstly decide on a questioning

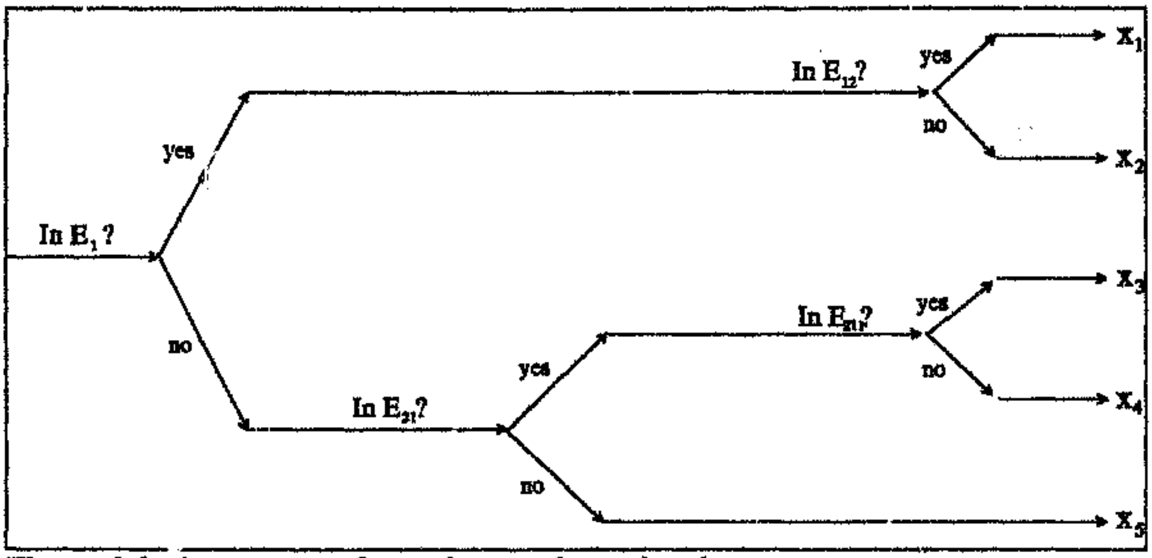


Figure 4.1: A sequence of questions to determine the outcome

scheme, and for this example we choose the most obvious one. To determine which outcome will be chosen using as few questions as possible, each question should remove as much uncertainty as possible in the remaining outcome space at each step. To do this, we divide the outcomes into two sets E_1 and E_2 such that their probabilities are as close to 1/2 as possible (since we do not know before asking the question in which set the outcome to be chosen lies). We may now ask the first question 'Is the outcome in E_1?' Depending on the result of this question, we repeat the process with either the set E_1 or E_2. Eventually the sets being considered will each contain one outcome so that the question asking process terminates and the outcome has been found. **Figure 4.1** demonstrates a possible sequence of questions for an experiment with five outcomes. The partitions used in this scheme are:

$$\begin{gathered}\xi_1=\{E_1=\{x_1,x_2\},E_2=\{x_3,x_4,x_5\}\}\\ \xi_2=\{E_1,E_{21}=\{x_3,x_4\},E_{22}=\{x_5\}\}\\ \xi_3=\{E_1,E_{211}=\{x_3\},E_{212}=\{x_4\},E_{22}\}\\ \xi_4=\{E_{12}=\{x_3\},E_{212}=\{x_4\},E_{22}\}\end{gathered} \tag{4.30}$$

The expected number of yes-no questions required to determine the outcome is given by

$$E=\sum_{i=1}^{n} p_i N_i \tag{4.31}$$

By defining

$$B=\sum_{i=1}^{n} 2^{-N_i} \tag{4.32}$$

and

$$q_i=\frac{2^{-N_i}}{B} \tag{4.33}$$

it can be demonstrated that (**lemma 2.8** in Martin[26])

$$H=-\sum_{i=1}^{n} p_i \log_2 p_i \leq -\sum_{i=1}^{n} p_i \log_2 q_i \tag{4.34}$$

By a straightforward computation, it may be shown that this inequality is equivalent to

$$H \leq E+\log_2 B \tag{4.35}$$

Since $B \leq 1$ we must have $H \leq E$. This demonstrates that the entropy is an indication of the minimum number of yes-no questions required to determine the outcome of a process.

The entropy given above can be shown to have all the usual properties of entropy as used in thermodynamics. A result of particular importance to us is the entropy of compound experiments:

Theorem[26]: If ξ and ζ are measurable partitions of a probability space $(\Omega, \mathcal{F}, P)$ then

$$H(\xi \vee \zeta)=H(\xi)+H(\zeta/\xi) \tag{4.36}$$

$H(\zeta/\xi)=H(\zeta)$ if and only if ζ and ξ are independent. In other words, the entropy of a compound experiment is equal to the sum of the entropies of the individual experiments if and only if these experiments are independent.

Associated with the conditional entropy is the concept of the mutual uncertainty between two partitions ξ and ζ which is a measure of the average gain in uncertainty about ξ given the result of ζ.

Definition: The mutual uncertainty between two measurable partitions ξ and ζ is defined to be

$$I(\xi;\zeta)=H(\xi)-H(\xi/\zeta) \qquad (4.37)$$

4.5.6 Entropy of a dynamical system

The **entropy** of a dynamical system $(\Omega,\mathscr{F},P,T)$ is given by the **entropy** of time mechanism **T** of the system which is in turn defined to be the greatest rate of generating information using **T** and using all possible finite processes in the dynamical universe $(\Omega,\mathscr{F},P,T)$. This is the **entropy** (also referred to as the **Kolmogoroff-Sinai Invariant**) of greatest value since it is an isomorphism invariant: if two dynamical systems are isomorphic then they have the same **entropy**.

Definition: The rate of uncertainty generated in the measurable partition ξ by **T** is defined to be

$$h(T,\xi)=\lim_{n\to\infty}\frac{1}{n}H\left(\bigvee_{j=0}^{n-1}T^{-j}\xi\right) \qquad (4.38)$$

and is also referred to as the **entropy** of **T** given ξ.

This is the form we would expect h(T,ξ) to have since the above is simply the average uncertainty in ξ for any trial using **T**. All trials of finite length are included in calculating this average, with $\mathbf{T}^{-1}$ providing the time mechanism as usual. The **entropy of T** is then given by:

Definition: The entropy h(T) of the dynamical space $(\Omega, \mathcal{F}, P, T)$ is given by

$$h(T) = \sup \{ h(T,\xi) : \xi \text{ is a finite measurable partition} \} \qquad (4.39)$$

4.6 Communication theory

The definition of **entropy** given above actually came after the concept of **entropy** as a quantitative measure of uncertainty was used with great effect in the theory of communication. The formal theory of communication we will be referring to was introduced by C.E. Shannon[29] in his classic paper "A mathematical theory of communication" in 1948. The elements of a communication system are given in **figure 4.2**.

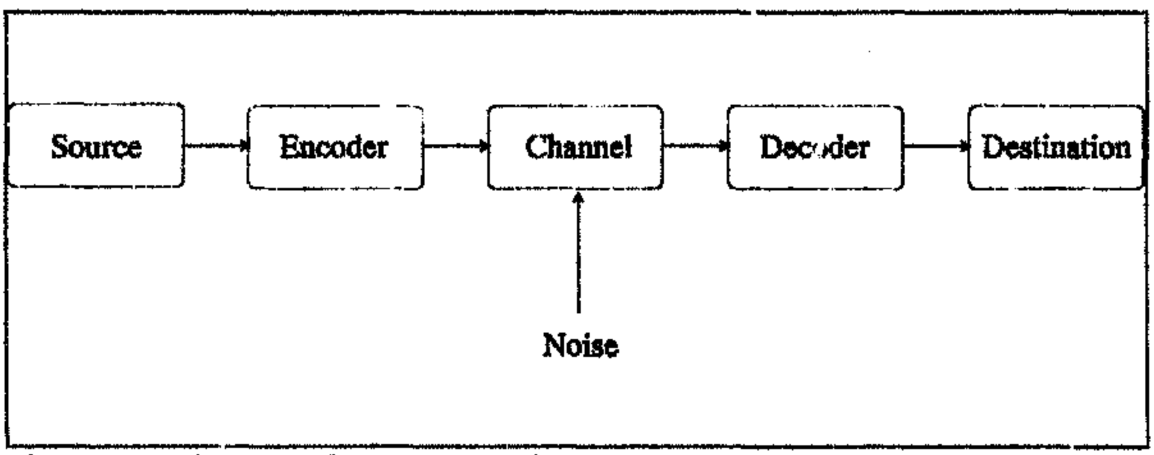


Figure 4.2: Elements of a communication system

The source is a stochastic sequence in a finite set S, which may be expressed[24] as the quadruple $(\Sigma(S), \mathcal{F}_s, \mu, T_s)$ where $\Sigma(S)$ is the set of all double infinite sequences in S (i.e. the set of all functions from the integers to S), $\mathcal{F}_s$ is the product σ-field, μ the probability measure associated with the process (μ determines the source) and T_s the so-called shift transformation. In general, T_s need **not** be μ preserving. Using a source alphabet S, the only quantity which changes depending on the source chosen is the probability measure μ which explains the following notation:

Definition: A discrete source in a finite set S is given by $[\Sigma(S),\mu]$ where $\Sigma(S)$ is the set of all doubly infinite sequences of elements of S and μ is a probability measure obtained from the joint distributions of the process.

Since the underlying space is stationary (probability preserving), we may assume that $\mathbf{T}_s$ is probability preserving as well so that all the sources dealt with are stationary. If $\mathbf{T}_s$ is also ergodic then the source is described as an ergodic source, and similarly for other properties (such as the Bernoulli property[26,27,28]).

If V is the alphabet with which we wish to encode the source $[\Sigma(S),\mu]$, then the encoding process refers to the process of mapping messages in $[\Sigma(S),\mu]$ into the new space $(\Sigma(V),\mathscr{F}_V)$ with the code specifying how this is done.

Definition: A code for encoding the source $[\Sigma(S),\mu]$ using a finite alphabet V is a measurable function from $\Sigma(S)$ to $\Sigma(V)$.

Usually the codes considered are restricted to those which are time independent. Another property often assumed for a code ϕ is that ϕ is invertible with respect to μ (i.e. ϕ is invertible on a subset of $\Sigma(S)$ which has measure 1).

The channel provides a means for a message from the input device to be produced on the output device which may be of another type.

Definition: A channel using input alphabet A and output alphabet B is denoted by $[\Sigma(A),P(\omega,\ ),\Sigma(B)]$ where $\omega\in\Sigma(A)$ is an input message and the function

$$P(\ ,C):\Sigma(A)\to\mathbb{R} \tag{4.40}$$

is measurable. The channel is stationary if

$$P(T_A\omega,T_BC)=P(\omega,C)\quad \forall\omega\in\Sigma(A),C\in\mathscr{F}_B \tag{4.41}$$

and memoryless if at any point in time, the choice of the output symbol depends only on the current input symbol.

For our purposes, we need only work with stationary memoryless channels. In the case of a memoryless channel, an expression for P may be easily obtained by simply considering the mapping using the source alphabet A which is isomorphic to $\{1,2,\ldots,n\}$ and the output alphabet B which is isomorphic to $\{1,2,\ldots,m\}$. We may form a matrix C_{mn} where C_{ij} is the conditional probability of the channel producing i as an output given that the input symbol was j. Then for each message $\omega=(\ldots j_{-2} j_{-1} j_0 j_1 j_2 \ldots)$ the probability measure $P(\omega,\ )$ may be written explicitly as[27]:

$$P(\omega,\{\overline{\omega}\in\Sigma(B):\overline{\omega}_k=i_k,\ p\le k\le q\})=\prod_{k=p}^{q} C_{i_k j_k} \tag{4.42}$$

A very important characteristic of a channel is its capacity in transmitting information, where the rate of transmission in a channel is defined to be

Definition: The rate of information about the source $[\Sigma(A),\mu]$ processed by the channel $[\Sigma(A),P(\omega,\ ),\Sigma(B)]$ is defined to be

$$R(\mu,P)=\lim_{n\to\infty}\frac{1}{n}I(\xi_A^{(n)};\xi_B^{(n)}) \tag{4.43}$$

where for each n

$$\xi_A^{(n)}=\bigvee_{j=0}^{n-1} T_A^{-j}\xi_A,\quad \xi_B^{(n)}=\bigvee_{j=0}^{n-1} T_B^{-j}\xi_B \tag{4.44}$$

The capacity of the channel is defined to be the supremum of all these rates:

$$C(P)=\sup_{\mu} R(\mu,P) \tag{4.45}$$

The fundamental theorem in communication theory is the following:

Noisy Channel Coding Theorem[26,28]: Let $[\Sigma(A),P(\omega,\ ),\Sigma(B)]$ be a discrete memoryless channel with capacity C and let $[\Sigma(S),\mu]$ be an ergodic source with entropy h. If $h<C$ then for $\varepsilon>0$ there exists a code $\phi:\Sigma(S)\to\Sigma(A)$ such that the rate of transmission of the source over the channel $R(\mu\phi^{-1},P)$ is greater than $h-\varepsilon$. If $h>C$ then there exists no code ϕ for which the rate of transmission reaches h.

4.7 Formal frame modelling

4.7.1 Stable behaviour in ecosystem dynamics

In **Chapter 2**, the shortcomings of assigning the property of stability solely to the points of equilibrium of a system were discussed. As was mentioned there, the next level of generalization is provided by considering attractors as studied in chaotic dynamics. There is a strong link between chaotic dynamics and formal frame-based modelling as will be demonstrated. What exactly do we mean by chaotic dynamics? In a classical chaotic system, if two identical systems are started with initial conditions which are not equal but very similar, the subsequent behaviour diverges sharply (in 'perfectly' chaotic systems this divergence is exponential with time). These systems are usually given by differential equations and so are usually deterministic. This may prompt one to ask why deterministic systems need be labelled chaotic at all since theoretically the deterministic equations can be solved to any desired degree of accuracy. But, as with frame modelling, the designation of these systems as chaotic makes sense when real problems such as measurement accuracy are taken into account. If the system is such that the smallest measurement error δ is large enough so that two identical systems whose initial conditions differ by δ diverge very rapidly, then knowing the initial conditions in the system is not sufficient for determining the subsequent behaviour. The system would then appear to behave chaotically (in the literal sense) in spite of the fact that the underlying dynamics are theoretically deducible. Chaotic systems are also characterized by very complex patterns similar to the stable limit cycle referred to in **Chapter 2**, these patterns are obviously examples of stable phenomena. The following is a partial list of the types of stable behaviour[30].

Point attractors: These are the equilibrium points considered in **Chapter 2** and are the most stable in the sense that if the system is not perturbed, it will remain exactly at these point attractors forever.

Limit Cycles: These attractors were also considered in **Chapter 2** and are characterized by simple closed curves.

Toroidal Flow: These are attractors which are characterized by an orbit on the surface of a torus. The orbit may be periodic so that after an integral number of complete revolutions, the orbit repeats itself. If the orbit never repeats itself, it is described as being quasi-periodic.

[illegible]ractors: Strange attractors are among the most intensively studied [illegible] chaotic dynamics and are characterized by orbits which are neither periodic nor quasi-periodic. Their Hausdorff dimension is low and non-integral (Hausdorff dimension to be defined later).

Turbulence: Here the behaviour of the system is highly erratic. The system may spend an unusually long time in a turbulent region so that the turbulent region appears to exhibit some degree of stability.

If there were a quantitative measure of stability, then in terms of this measure we would expect that the attractors listed above be ranked in decreasing order of stability. Such a quantity may be provided locally by the fractal dimension or by the entropy, interestingly these are related.

4.7.2 The connection between entropy and fractal dimension

In trying to develop the idea of dimension and extend its usefulness from well behaved sets (such as vector spaces), to more irregular sets, Hausdorff introduced the notion of the **fractal dimension** of a metric space. The **fractal dimension** of a metric space M is defined in terms of the outer measures:

> **Definition**: Let α be a positive real number and M a metric space, and let a ρ-covering of M be a countable covering of M by closed spheres S_i each of diameter less than ρ. We define the α-dimensional outer measure of M to be
>
> $$O_\alpha(M,\rho) = \inf_{\rho\text{-coverings}} \sum_i (\text{diam } S_i)^\alpha \qquad (4.46)$$
>
> The limit as [illegible]ds to zero exists and is denoted by
>
> $$O_\alpha(M) = \lim_{\rho \to 0} O_\alpha(M,\rho) \qquad (4.47)$$

The last limit exists because of the way $O_\alpha(M,\rho)$ is defined: By decreasing ρ, the infimum is over a smaller collection of coverings so that it cannot possibly decrease (and may increase). Since the infimum is bounded below, the limit must exist. The Hausdorff dimension is defined by considering the behaviour of $O_\alpha(M)$ as a function of α. The following theorem may be proved:

> **Theorem**[28]: The quantity
>
> $$\dim M = \sup\{\alpha : O_\alpha(M) = \infty\} = \inf\{\alpha : O_\alpha(M) = 0\} \qquad (4.48)$$
>
> exists, is unique and is defined to be the **Hausdorff dimension** of M.

This definition of dimension is consistent with the conventional one when applied to conventional 'well-behaved' spaces. As an example, it can be shown[28] that the **Hausdorff**

dimension of a smooth (not necessarily flat) surface in a 3-D space is two, exactly what we would expect. But when calculated for irregular sets (the most obvious example being the Mandelbrot set), the **Hausdorff dimension** does not necessarily give integer values, which immediately indicates why it is so often referred to as the **fractal dimension**. There is a very interesting connection between **topological entropy** and the **fractal dimension**: it can be shown that these two quantities are just two different interpretations of the same concept[27]. The topological entropy is related in turn to the measure theoretic entropy, this connection is exploited in **Section 4.7.4**.

Definition: Let X be a compact topological space. For any cover α of X, let $N(\alpha)$ denote the smallest finite subcover of α (i.e. the subcover of α with the fewest members). The **topological entropy** of α is defined to be

$$H(\alpha)=\log N(\alpha) \tag{4.49}$$

This definition of the topological entropy makes sense since it has the log form we expect and also depends on the smallest finite subcover (as a function of any cover) which intuitively gives an indication of the information or uncertainty in the space. Instead of dealing with measure preserving transformations, in the topological context we consider continuous mappings (which are similar in that the inverse image of a open or closed set is open or closed respectively). In almost exactly the same way as the measure theoretic case, the entropy of the continuous map **T** is defined as:

Definition: If α is an open cover of the topological space X and **T** is a continuous map on X, then the entropy of **T** relative to α is defined to be

$$h(T,\alpha)=\lim_{n\to\infty}\frac{1}{n}H\left(\bigvee_{i=0}^{n-1}T^{-i}\alpha\right) \tag{4.50}$$

and the topological entropy of **T** is defined to be

$$h(T)=\sup_{\alpha} h(T,\alpha) \qquad (4.51)$$

where the supremum is over all open covers of X.

(The common refinement of a family of covers is defined in the same way as for partitions and provides a partial ordering $\leq$ on a sequence of refining sets).

The connection between measure theoretic entropy and topological entropy is demonstrated by the following theorem:

Theorem[27]: Let $T:X\rightarrow X$ be a continuous map on a compact metric space X, and M(X,T) denote the space of all probability measures on X which are preserved by T then

$$h(T)=\sup\{h_{\mu}(T)\,|\,\mu\in M(X,T)\} \qquad (4.52)$$

where h(T) is the topological entropy and $h_{\mu}(T)$ the measure theoretic entropy with respect to the measure μ.

The quantity, therefore, which may be seen providing a quantitative measure of stability in the list of attractors given earlier is the information, uncertainty, entropy or the fractal dimension of the attractors. (Another widely used quantity is provided by the Lyapunov exponents[31]). Strange attractors are characterized by low fractal dimensions[30,32] (many strange attractors in 3D systems have orbits which are 'almost' on a two dimensional surface so that their fractal dimension is close to 2), and turbulence by even lower fractal dimensions. Various methods have been proposed for estimating the fractal dimensions and entropies of strange attractors[32], they are all based on the definition of entropy: The phase space in the neighbourhood of the attractor is partitioned into cubes of size ε, and used to estimate the entropy and fractal dimension.

4.7.3 Formal basis of frame based modelling

In the formal frame modelling of ecosystems, the sample space Ω may be constructed as follows: We assume that the ecosystem may be described by a finite collection of state variables represented by the sets X_i which are subsets of $\mathbf{R}$. These state variables are observables of a physical ecosystem, and not 'badly behaved' artificial variables (such as delta functions).

We define

$$\Omega=\prod_{i=1}^{n} X_i \times \prod_{i=1}^{n} X_i \tag{4.53}$$

(It is entirely possible to form such an Ω from the product of arbitrary spaces, but that level of abstraction is certainly not necessary since we are concerned with the modelling of physical ecosystems). For a point $t\in\Omega$, we interpret

$$t=(x_1,\ldots,x_n,y_1,\ldots,y_n)\in\Omega \tag{4.54}$$

as being a movement from $(x_1,\ldots,x_n)$ to $(y_1,\ldots,y_n)$, with the probability of any such movement being preserved through the time evolution of the system. To obtain a dynamical system for the ecosystem, we need to construct a collection $\mathcal{F}$ of subsets of Ω and a probability measure P acting on $\mathcal{F}$ such that the time mechanism **T** preserves the probability measure of each member of $\mathcal{F}$. We interpret each element E of $\mathcal{F}$ as an experiment being performed in order to determine the current state of the system.

> **Postulate 1**: A dynamical ecosystem is completely specified by the dynamical quadruple $(\Omega,\mathcal{F},P,T)$.

The sample space Ω also has a topology associated with it which is inherited from the variables used to define Ω. For this reason, we may assume with no loss in generality that the σ-algebra $\mathcal{F}$ contains all open sets of this topology (if it does not, we simply expand $\mathcal{F}$ so that it does). We may construct such a σ-algebra using the following theorem:

Theorem[33]: If M is any collection of subsets of X, there exists a smallest σ-algebra M^* in X such that $M \subseteq M^*$.

The σ-algebra generated in this manner from the open sets of a topological space X is called the collection of Borel sets of X: We are assuming that $\mathscr{F}$ contains the Borel sets of Ω.

In the preceding section, we saw that the entropy of an attractor does give some quantitative measure of the stability of an attractor. In order to investigate the entropy of this dynamical system, we need the following:

Definition: For each i=1,..,n, we define T_i^* to be the restriction of **T** to Ω_i^* and $\mathbf{T}_i$ to be the restriction of **T** to Ω_i.

A definition for stability within ecosystems will now be given which has an interesting relationship with the entropy:

Definition: An ecosystem $(\Omega, \mathscr{F}, P, T)$ is said to have **stable domains of attraction** if it has the following properties:

a. We assume that Ω may be partitioned into the mutually disjoint sets Ω_i^* (where i=1,..,n) and Ω_m with the property:

$$\forall x \in \Omega_i^*,\ \forall j \in N,\ T^{-j}x \subset \Omega_i^* \qquad (4.55)$$

Each set Ω_i^* is referred to as an **attractor** of the system. By definition

$$\Omega = \bigcup_{i=1}^{n} \Omega_i^* \cup \Omega_m \qquad (4.56)$$

b. If $E \in \mathscr{F}$ is any experiment with $E \subset \Omega_m$, we assume that there exists a $p \in \mathbf{N}$ such that

$$T^{-p}E \cap \Omega_m = \emptyset \tag{4.57}$$

and define $\Omega_i \supseteq \Omega_i^*$ to be

$$\Omega_i = \{x \in \Omega : \exists q \text{ with } T_i^{-q}x \subset \Omega_i^*\} \tag{4.58}$$

Each Ω_i is referred to as a **domain of attraction** of the system.

The condition in **(4.55)** says that once the system is in attractor Ω_i^*, it cannot escape (unless, of course, there is an external perturbation which disturbs the system). Condition **(4.57)** says that if the system is not in one of the attractors, it will eventually move into an attractor region. Condition **(4.58)** then defines the domains of attraction according to the possible attractors the system can move into from any particular point. This definition of stability is highly generalized, since the domains Ω_i may overlap and we do not specify how the system must behave within the attractor or even the spatial description of the attractor. Each set Ω_i^* may be anything from a point (a point attractor), to being the whole set Ω_i. Note that by condition **(4.57)**, the union of Ω_i is equal to the whole of Ω.

How are the entropies of **T** and $\mathbf{T}_i$ related? From the definition of entropy, we know that

$$h(T) \leq \sum_{i=1}^{n} h(T_i) \tag{4.59}$$

We also know that

$$h(T) = \sum_{i=1}^{n} h(T_i^*) + h(T_m) \geq \sum_{i=1}^{n} h(T_i^*) \tag{4.60}$$

by the disjointness of Ω_i^* and Ω_m. The definition of stability given above allows us to go further and to prove the following:

Theorem: The entropy of T is equal to the sum of the entropies of T_i

Proof: To show this, let α be a cover of Ω_i (and consequently of Ω_i^*). Let p be a number such that

$$T_i^{-p}(a)\subset\Omega_i^* \ \forall a\in\alpha \tag{4.61}$$

(p exists by definition since $\bigcup_{a\in\alpha} a\in\mathcal{F}$). Then

$$\bigvee_{j=0}^{n-1} T_i^{-j}(\alpha)\leq\bigvee_{j=0}^{p}\left[(T_i^{-j}(\alpha)\cap\Omega_i^*)\bigcup(T_i^{-j}(\alpha)\cap(\Omega_i-\Omega_i^*))\right]\vee\bigvee_{j=p+1}^{n-1} T_i^{*-j}(\alpha) \tag{4.62}$$

In the last term, $T_i=T_i^*$ over Ω_i^*. The first term has been split into two disjoint sets resulting in those sets which are neither proper subsets of Ω_i^* nor of $\Omega_i-\Omega_i^*$ being split into subsets of Ω_i^* and $\Omega_i-\Omega_i^*$ respectively. The right hand side is consequently a refinement of the left hand side, hence the inequality. The right hand term is then equal to or refined by

$$\left[\bigvee_{j=0}^{p} T_i^{*-j}(\alpha)\bigcup\bigvee_{j=0}^{p} T_i^{-j}(\alpha)\cap(\Omega_i-\Omega_i^*)\right]\vee\bigvee_{j=p+1}^{n-1} T_i^{*-j}(\alpha) \tag{4.63}$$

again using the fact that $T_i=T_i^*$ over Ω_i^*, and the behaviour of T_i^* within Ω_i^*. This shows that

$$\bigvee_{j=0}^{n-1} T_i^{-j}(\alpha)\leq\bigvee_{j=0}^{n-1} T_i^{*-j}(\alpha)\bigcup\bigvee_{j=0}^{p} T_i^{-j}(\alpha)\cap(\Omega_i-\Omega_i^*) \tag{4.64}$$

using disjointness. By the definition of the topological entropy of T_i with respect to the cover α, we obtain the following:

$$h(T_i,\alpha)\leq\lim_{n\to\infty}\frac{1}{n}\left[H\left(\bigvee_{j=0}^{p} T_i^{-j}(\alpha)\cap(\Omega_i-\Omega_i^*)\right)+H\left(\bigvee_{j=0}^{n-1} T_i^{*-j}(\alpha)\right)\right] \tag{4.65}$$

where we have used the fact that $H(A\cup B)=H(A)+H(B)$ if A and B are disjoint. The first term vanishes in the limit, so that we have:

$$h(T'_i,\alpha) \leq h(T_i^*,\alpha) \tag{4.66}$$

which implies that the entropy of T_i never exceeds the entropy of T_i^*. We then have the following:

$$h(T) \leq \sum_{i=1}^{n} h(T_i) \leq \sum_{i=1}^{n} h(T_i^*) \leq h(T) \tag{4.67}$$

which proves the theorem.

This is a very interesting result since it shows the close connection between the entropy and the domains of attraction of the system as we have defined them. It was shown in **Section 4.5.5** that the measure theoretic entropy of the union of two partitions is equal to the sum of the entropies of each partition if and only if those partitions are independent. The above theorem shows that the frame domains are independent in this sense even though they may overlap.

In informal frame modelling, each frame refers to an area of attraction of the system with the associated dynamical processes, which is exactly the definition we give for the formal case:

> **Definition**: Each pair $[T_i,\Omega_i]$ is defined to be a **formal frame** of the ecosystem. The subsets Ω_i are called the **frame domains** of the ecosystem.

One of the premises of frame based modelling is that these generalized frame domains do indeed exist (a premise which is motivated by the observation of multiple stable configurations in ecosystems, see **Chapter 2**):

> **Postulate 2**: In the dynamical ecosystem $(\Omega,\mathcal{F},P,T)$, there exist a finite number of **attractors** Ω_i^* with their associated frame domains Ω_i where $i=1,\ldots,n$.

Each Ω_i is a domain of attraction of the ecosystem, but it should be noted that these sets are

not necessarily mutually disjoint, or even connected: These domains are highly generalized.

> **Postulate 3**: For each Ω_i there is a channel $[\Sigma_i(A_i), P_i(\omega,), \Sigma_i(B_i)]$ of limited capacity C_i through which all information about T_i must pass. The alphabets A_i and B_i may be chosen as desired, the only fixed parameter of the channel is C_i.

In this postulate, we explicitly recognize the limitations imposed by the measuring difficulty by demanding that the information pass through noisy channels of limited capacity. These channels also provide the connection between the general ecosystem dynamics, and the qualitative frame models.

The challenge in frame modelling is to find a way to transmit the information content of each **formal frame** $[T_i, \Omega_i]$ (the entropy of each T_i acting on Ω_i) as accurately as possible. The following is proposed as a method for doing just that.

4.7.4 The fundamental theorem of frame modelling

We have assumed the presence of a discrete communications channel (with good reason as will become apparent later), which requires a method of transmitting the information as accurately as possible while only using a finite alphabet. It will be shown, that this is possible to an **arbitrary** degree of accuracy:

> **Theorem**: For each T_i, there exists a discrete source whose entropy is arbitrarily close to the entropy of T_i if this entropy is less than the capacity C_i of the channel associated with $[T_i, \Omega_i]$.

> **Proof**: Because of the topological structure already present in Ω_i it is easier to work with the topological entropy (which we have seen is equivalent to the measure theoretic entropy). To do this, we must first extract a compact metric subspace Ψ of

Ω. But this is easily done since Ω has been constructed from the phase space of the ecosystem variables (each of which is real valued) and consequently inherits a metric from the individual variable spaces X_i. To obtain a compact subspace of Ω, we need only bound the variables on both sides. This can be justified since we are then bounding the system variables (which are the observables of a real physical ecosystem and consequently would be bounded) and these bounds may be as large in absolute value as desired. Given $\varepsilon>0$, we need to show that there exists a measure μ on Ψ so that

$$|h_{\mu}(T)-h(T)|<\epsilon \tag{4.68}$$

and a finite cover α of Ψ so that

$$|h_{\mu}(T)-h(T,\alpha)|<2\epsilon \tag{4.69}$$

Both (4.68) and (4.69) follow by definition since

$$h(T)=\sup_{\beta} h(T,\beta) \tag{4.70}$$

implies that there is a finite cover α of Ψ with

$$h(T)-\epsilon<h(T,\alpha)\leq h(T) \tag{4.71}$$

Also since

$$h(T)=\sup_{v} h_{v}(T) \tag{4.72}$$

there is probability measure μ with

$$h(T)-\epsilon<h_{\mu}(T)\leq h(T) \tag{4.73}$$

Using these inequalities, we obtain (4.68) and (4.69).

We have thus obtained an entropy $h_{\mu}(T_i)$ which is arbitrarily close to $h(T_i)$. Also, we have obtained a discrete source $[\Sigma(\alpha_i),\mu_i]$ since α_i is finite and can make use of the

Noisy Coding Theorem where the channel is given by $[\Sigma_i(\alpha_i), P_i(\omega_i), \Sigma_i(\alpha_i)]$. This demonstrates that the only restriction to the flow of information from the general dynamical space is the capacity of the channels (the constraint imposed by the difficulty in performing measurements on the ecosystem), and that this information may be transmitted using a finite alphabet to an arbitrary degree of precision.

The **fundamental theorem of frame modelling** may now be proved:

Fundamental theorem of frame modelling: Given a dynamic ecosystem $(\Omega, \mathscr{F}, P, T)$ with N **attractors**, the following are true:

a. The system may be decomposed into formal frames $[T_i, \Omega_i]$ where $i=1,...,N$ such that the sum of the entropies of the formal frames is equal to the entropy of the entire space.

b. An instantaneous description of the ecosystem may be obtained using N channels (associated with each formal frame) of limited capacity, to an arbitrary accuracy and using a finite alphabet. The accuracy is bounded only by the measuring error.

Proof: **a**. This part has already been proved above.

b. The quantitative measure of the information content is given by the entropy: By part **a**, we know that the entropy of the entire space may be decomposed into a sum of the entropies of the formal frames. It has been shown above that a discrete source using a finite alphabet may be found for each frame which approximates the real entropy to an arbitrary accuracy. This proves the theorem.

What if the channel associated with the formal frame has a capacity which is less than the entropy of the frame? In this case, the **Noisy Coding Theorem** says that there is **no** code capable of transmitting the information. But through the use of **source coding**[26,28], the source may be mapped into an approximation of itself whose entropy does not exceed the capacity of the channel. In this case, the minimum error is given by the difference between this approximated entropy and the channel capacity.

Chapter 5: Computational Aspects of Frame Modelling

5.1 Introduction

In this chapter, we are concerned with the computational aspect of frame modelling, as depicted on the left hand side of **Figure 1.1**. The primary aims of this chapter, therefore, are to show the connection with the formal frame dynamics discussed in **Chapter 4**, to develop algorithms and to design a modelling environment for the building of frame models. The first two are of a theoretical nature and follow naturally from the last chapter, and so will be dealt with first in **Section 5.2**. Based on the theoretical requirements, a specification for the frame modelling environment is proposed in detail in **Section 5.3**.

The following important theorem will be proved in this chapter:

> **Theorem: Stochastic Qualitative Rule Based Frame Models** may be used to model a dynamical ecosystem (as defined in **Chapter 4**) arbitrarily closely. The accuracy is bounded only by the measuring difficulty.

5.2 Theoretical structures

5.2.1 Building the input and output streams

In **Chapter 4**, it was shown that for each formal frame $[T_i,\Omega_i]$, an approximate source can be found whose information content is as close to that of the dynamics in that frame as desired. The source is discrete and uses a finite alphabet α_i, which is a finite cover of the space, so that the discrete channel $[\Sigma(\alpha_i),P_i(\omega,\ ),\Sigma(\alpha_i)]$ may be used to transmit the information. The symbols (elements of the cover α_i) transmitted to the destinations of each channel then give an instantaneous description of the ecosystem at any particular point in time. By building a discrete description of the ecosystem, we have also implicitly used a discrete time variable. The basic unit of this time variable depends on the rate of information flow in the channels, and over a fixed period of time may be defined to be the minimum time between two symbols appearing at the destination. In what follows, we do not refer to the time duration between symbols since we need only view the symbols as being indexed by time.

In frame modelling, the primary objective is to build a model which simulates the behaviour of the output of each channel as closely as possible. What exactly is meant by speaking of the simulation of the channel outputs? By simulation, we mean that the variables of the ecosystem may be divided into two groups, the input group of variables whose values are known (such as the rainfall in the Brachystegia model in **Chapter 3**), and the output group of variables whose values are to be modelled and which all depend in some way on the input group of variables, and each other.

Definition: The set of variables $\{X_i:i=1,..,n\}$ used to form Ω may be divided into two disjoint groups **I** and **O**: The **input group** of variables $\mathbf{I}=\{I_i:i=1,..,p\}$ and the **output group** of variables $\mathbf{O}=\{O_i:i=1,..,q\}$. The **input group** is the collection of those variables whose values are known, and which need not be modelled. The **output group** is formed with the remaining variables. Let P_i acting on Ω be the ith projection:

$$P_i(t_1,\dots,t_i,\dots,t_{2n})=t_i \;\; \forall t=(t_1,\dots,t_{2n})\in\Omega \tag{5.1}$$

and let $N(I_i)$ be the index of the first appearance of $I_i=X_j$ in the cartesian product of Ω. Let $N(O_i)$ be defined similarly for the output variables.

Then for each $i=1,..,p$, the **state set** of I_i is defined to be the finite set

$$\eta_i=\{P_{N(I_i)}(a)\cap I_i : a\in\alpha\} \tag{5.2}$$

and similarly for each $i=1,..,q$, the state set of O_i is defined to be the finite set

$$\theta_i = \{P_{N(O_i)}(a) \cap O_i; a \in \alpha\} \tag{5.3}$$

Each element of η_i and θ_i is also referred to as a **qualitative level** or **qualitative state** of I_i and O_i respectively.

Note that we have effectively taken the starting point of each system change in Ω. This could be modified to the end point or to the average if necessary. From the definition of $\{X_i, i=1,...,n\}$, each member of **I** and **O** is a compact subset of **R**, as is each qualitative level. It should be noted that by this definition of a qualitative level, the qualitative levels of a particular variable need not be ordered since they are not necessarily connected (this indicates why they are referred to as being qualitative). For consistency, we will consider the state sets to be ordered even if only by the indices of their elements.

Referring to **Figure 5.1**, the definition given above corresponds to the output of each channel being used to generate a stream of symbols (the qualitative levels η_i and θ_i) for each variable in **I** and **O**. This decomposition is essential for the development of algorithms for the implementation of the practical frame modelling technique.

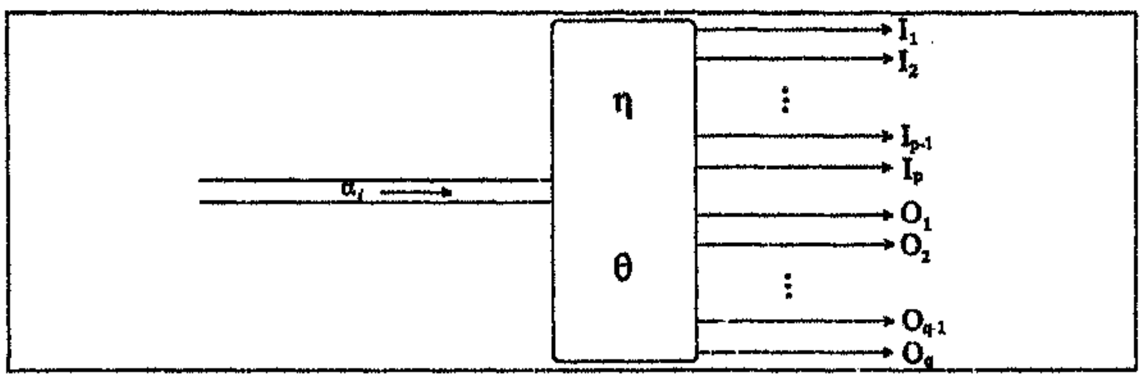


Figure 5.1: Definition of the qualitative levels of the Input and Output variables.

5.2.2 The use of finite automata

Both from a theoretical as well as a computational point of view, it is very useful to reduce the controlling components of a model into an assembly of units each as small as possible. Each unit may be viewed as a black box controlling some process. For this reason, we need the following basic concepts from the theories of **Automata** and **Formal Languages**.

5.2.2.1 Basic Definitions and Results

The most elementary object in the theory of formal languages is the **symbol**, which may be organized into **strings** of symbols of finite length. The length of a string **s** is denoted by $|\mathbf{s}|$, and the empty string by ϕ. By definition, the empty string has zero length, $|\phi|=0$.

> **Definition:** An **alphabet** is a finite set of symbols, and a **formal language** the set of strings of symbols selected from an alphabet. Given an alphabet Σ, the language Σ^* is the set of all strings of finite length formed using Σ.

The basic unit in computer science for describing finite systems is the **finite automaton** which is a mathematical construction with discrete inputs and outputs and which can be in any one of a discrete set of possible configurations or states. Through the use of discrete channels, we are interpreting the dynamical ecosystem as a finite system so that we may use the powerful mathematical machinery of theoretical computer science in the investigation of the frame based modelling technique. To demonstrate what is meant by a finite state system, it is useful to consider an example. A well known problem often used to demonstrate the use of deductive AI computer languages such as Prolog is the problem of getting a man, a wolf, a goat and a cabbage from one side of the river to the other using a boat. The constraints to the solution arise from the fact that, if left alone, the wolf will eat the goat and likewise the goat will eat the cabbage if left unattended. By assigning the symbols M,W, G and C to the man, wolf, goat and cabbage respectively, the state of this system at any given time may be summarized by two strings: The first string indicates who is on the left side of the river and the second string indicates who is on the right side. **Figure 5.2** is a directed graph giving the two shortest possible solutions to the problem. This is an example of a finite automaton, and **Figure 5.2** an example of a transition diagram for a finite automaton. As pictured in

Figure 5.3, another way of viewing a finite automaton is as a device monitoring the symbols in a channel. The symbols in the channel move from left to right, and the device changes state depending on the symbol being read. This view is closer to the mathematical definition of a finite automaton than a directed graph representing a finite automaton.

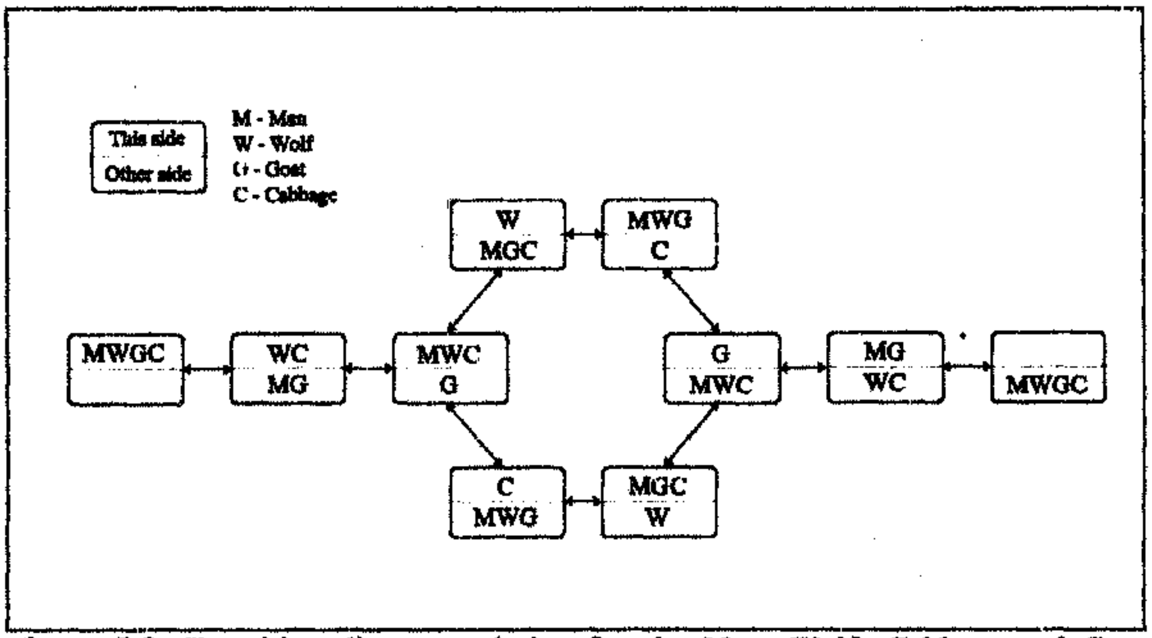


Figure 5.2: Transition diagram solution for the Man, Wolf, Cabbage and Goat problem.

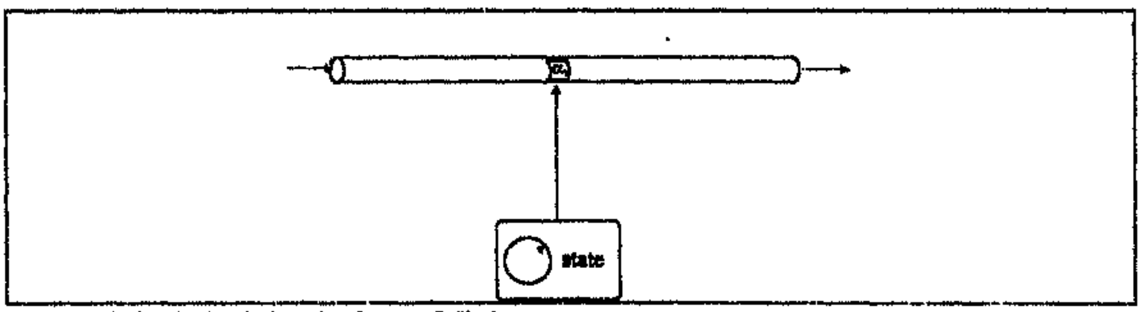


Figure 5.3: 'Black box' view of finite automaton.

Definition: A **Deterministic Finite Automaton** (DFA) is a quintuple (Q,Σ,δ,q_0,F) where Q is a finite set of states, Σ a finite alphabet, δ a function $\delta:Q\times\Sigma\rightarrow Q$, $q_0\in Q$ the initial state and $F\subseteq Q$ the set of final states.

The function δ is the transition function which specifies which state the DFA will change to depending on its current state and the current symbol being read. If the current state is **q**, and the current symbol **a** then the new state is δ(q,a). The transition function δ is extended to accept strings as a parameter as follows:

Definition: The transition function δ is extended to

$$\hat{\delta}:Q\times\Sigma^*\rightarrow Q \tag{5.4}$$

by defining

$$\hat{\delta}(q,\phi)=q \tag{5.5}$$

and by the recursive relation

$$\hat{\delta}(q,wa)=\delta(\hat{\delta}(q,w),a) \tag{5.6}$$

where **w** is any string.

The extension to δ is also denoted by δ since they are equal when the string is a single symbol and can be distinguished by their parameters when the string is longer. The definition given is for a **deterministic finite automaton** because there is one and only one possible transition per pair (q,a) where $q\in Q$ and $a\in\Sigma$. The **non-deterministic automaton** is defined in much the same away:

Definition: A **Non-deterministic Finite Automaton** (NFA) is a quintuple (Q,Σ,δ,q_0,F) where Q is a finite set of states, Σ a finite alphabet, δ a function $\delta:Q\times\Sigma\rightarrow\wp(Q)$, $q_0\in Q$ the initial state and $F\subseteq Q$ the set of final states. A **Non-Deterministic Finite Automaton with ϕ-moves** is defined as above, but the transition function is given by $\delta:Q\times(\Sigma\cup\{\phi\})\rightarrow\wp(Q)$

As for the deterministic case, the transition function may be extended:

Definition: The transition function δ is extended to

$$\hat{\delta}: Q\times\Sigma^{*}\rightarrow \wp(Q) \tag{5.7}$$

by defining

$$\hat{\delta}(q,\phi)=\{q\} \tag{5.8}$$

and by the recursive relation

$$\hat{\delta}(q,wa)=\{p\,|\,\exists r\in\hat{\delta}(q,w)\ \mathit{with}\ p\in\delta(r,a)\} \tag{5.9}$$

where **w** is any string.

The only difference between a DFA and a NFA is the transition function δ, which in the case of a NFA is mapped into $\wp(Q)$ (the power set of Q). The power set of Q is the set of all subsets of Q, and has $2^{o(Q)}$ elements because Q is finite. An NFA with ϕ-moves is simply an NFA where a transition is possible when there is no input symbol. The transition function appears confusing at first, since for any particular transition, the δ function may specify that the automaton assume a multiple number of states. From a theoretical point of view, we do not know which state the NFA will move to and do not need to. One way of interpreting the behaviour of an NFA is to view the set of possible destination states as being multiple copies of the device reading the contents of the channel. In contrast to **Figure 5.3**, a **non-deterministic automaton** may contain several devices (instead of just one) reading the current symbol in the channel. With each symbol read, new devices may be created and others destroyed. In frame modelling, we extend the transition function in order to obtain a probability distribution:

Definition: An **Probabilistic Non-Deterministic Automaton** is defined to be an NFA with distribution function $\delta_p: Q\times\Sigma\times Q\rightarrow\mathbf{R}$. For each $s\in\delta(q,\sigma)$ where $q\in Q$ and $\sigma\in\Sigma$, $\delta_p(q,\sigma,s)$ is the probability that the **Probabilistic Non-Deterministic Automaton** will move from state q to state s on reading input symbol σ.

We assume that all NFAs to be probabilistic NFAs in what follows, since the inclusion of the distribution function does not affect the theory of NFAs from the computer science theoretic viewpoint. We are interested in the type of language accepted by the finite automata, where we define:

Definition: A string **w** is **accepted** by a finite automaton (Q,Σ,δ,q_0,F) if $\delta(q_0,w)=f$ where $f\in F$ is one of the final states. The **language accepted** by the finite automaton M is the set

$$L(M)=\{w|\delta(q_0,w)\in F\} \quad (5.10)$$

A language is described as being **regular** if it is accepted by some finite automaton.

The reason why the language accepted by a finite automaton is called **regular** becomes apparent when the relationship between finite automata and the set of **regular expressions** is demonstrated:

Definition: Given two subsets X and Y of Σ^* where Σ is a finite alphabet, the concatenation of X and Y is defined to be the set

$$XY=\{xy|x\in X,\ y\in Y\} \quad (5.11)$$

The regular expressions over a finite alphabet Σ are defined using the following recursive relations:

a. The empty set $\varnothing$ is a regular expression.
b. The set $\{\phi\}$ denoted by ϕ is a regular expression.
c. For each $a\in\Sigma$, $\{a\}$ denoted by **a** is a regular expression.
d. If x and y are regular expressions denoting languages X and Y, then so are $(x+y)$, (xy) and (x^*) which denote the sets $X\cup Y$, XY and X^*.

By definition, * has a higher precedence than + or concatenation, and concatenation has a higher priority than +.

The following theorem may be proved relating all the concepts given above. This theorem is summarized by **figure 5.4**.

Theorem[34]: The equivalence of DFAs and NFAs

a. If L is accepted by an NFA, then there exists a DFA which accepts L.

b. If L is accepted by an NFA with ϕ-moves, then there exists an NFA without ϕ-moves which accepts L.

c. If r is a regular expression, then there exists an NFA with ϕ-moves which accepts L(r).

d. If L is accepted by a DFA, then L is a regular expression.

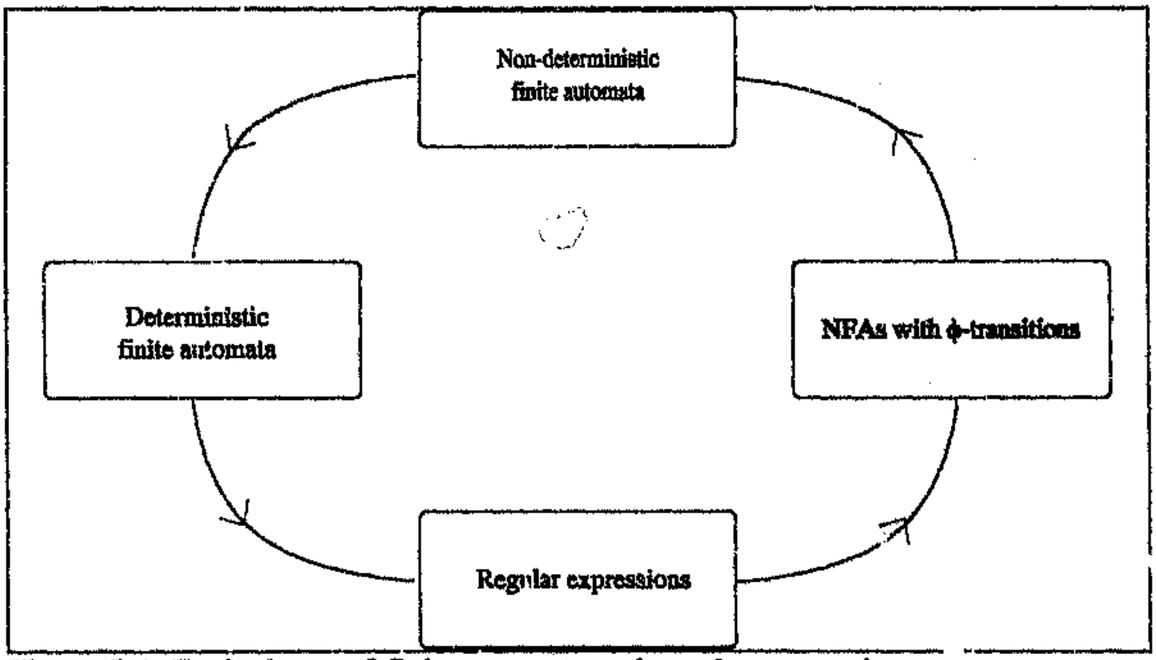


Figure 5.4: Equivalence of finite automata and regular expressions.

This theorem indicates that finite automata are equivalent to regular expressions and that they may be viewed (from a computer science viewpoint) as deterministic or non-deterministic as required, since they are equivalent in that they accept the same class of languages. Part **a** of the proof is proved by simply constructing a DFA whose states enumerate the subsets which the NFA moves to after reading each symbol. From the frame modelling point of view, of course, we require the probability distribution function we have associated with each NFA.

5.2.2.2 Modelling using NFAs

Ve seek a method for reproducing the behaviour of the variables in the output group **O**, and the following is proposed as being such a method:

The finite alphabet Σ to be used is simply an enumeration of all the possible combinations of the values of the variables in the input and output groups **I** and **O**. We add the symbol ϕ to each state set to represent the absence of that variable in the combination: $\eta_i^* = \eta_i \cup \{\phi\}$ for $i=1,..p$ and $\theta_j^* = \theta_j \cup \{\phi\}$ for $j=1,...,q$.

> **Definition**: The **combination alphabet** of the frame is defined in terms of the associated state sets η_i^* and θ_j^* by
>
> $$\Lambda = \{\lambda_{i_1 \ldots i_p j_1 \ldots j_q} = \alpha_{i_1} \ldots \alpha_{i_p} \alpha_{j_1} \ldots \alpha_{j_q} \mid \alpha_{i_k} \in \eta_i^* \text{ where } i=1,\ldots,p \text{ and } \alpha_{j_k} \in \theta_j^* \text{ where } j=1,\ldots,q\} \tag{5.12}$$
>
> and has order
>
> $$o(\Lambda) = \prod_{i=1}^{p} o(\eta_i^*) \prod_{j=1}^{q} o(\theta_j^*) \tag{5.13}$$

By definition, the stream of symbols thus generated by the dynamical system make up a regular set which means that there is an NFA which accepts this stream. For each output O_j in the output group **O**, we then have an NFA $(\theta_j^*, \Lambda, \delta_j, q_j, F_j)$ which depends on the current values of the input and output variables, and whose state corresponds to the new value of O_j. **Figure 5.5** gives a graphical representation of these structures.

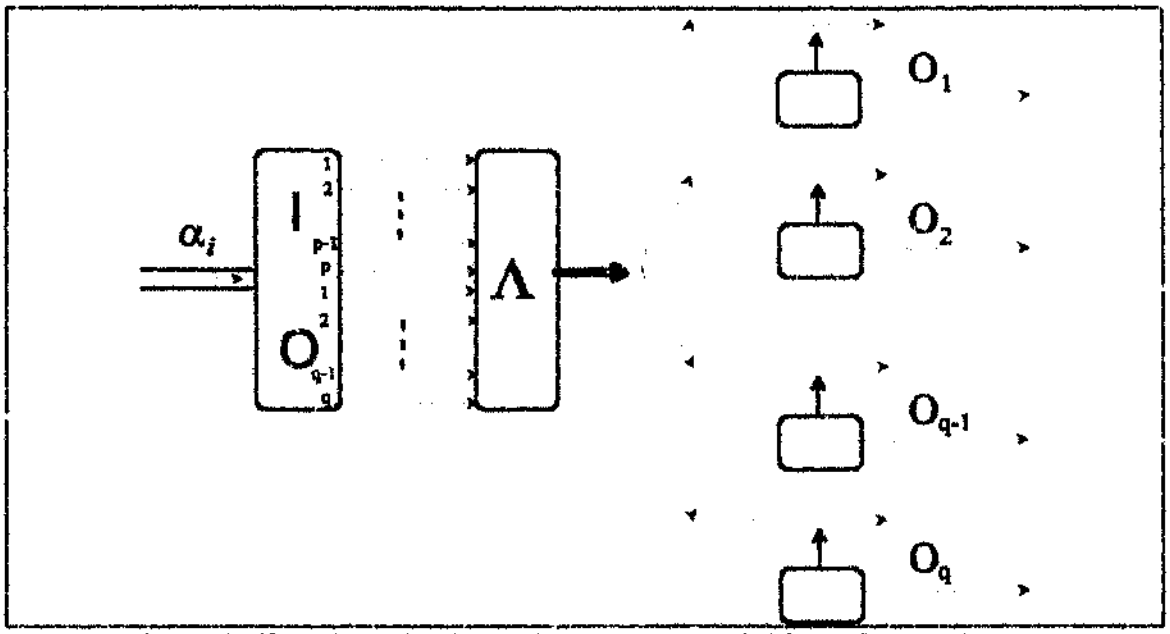


Figure 5.5: Modelling the behaviour of the output variables using NFAs.

The idea is very simple: For each combination of the input variables, there is a corresponding response in the output variable O_i.

A subtle but very important property of this system arises from the way the channels were defined for each frame in **Chapter 4**. These channels are memoryless, which means that the information flow corresponds to an instantaneous description of the system. The result is that in the model presented above, the NFAs depend only on the current state of the system, and the transition functions are not affected by the history of how the system got there at all. In other words, the instantaneous description of the ecosystem is sufficiently detailed so that the history of how the system got there is not required. From the theoretical dynamics viewpoint, if the ecosystem moves to the same point in phase space using two different routes, the next movement is independent of the route used to get there. From a modelling viewpoint, this is too restrictive since what we know about a system may be naturally expressed by making references to the past values held by the variables (an example is given by the Brachystegia model in **Chapter 3**). It is possible to allow the reference to past variables by making a modification of the finite automata:

5.2.3 The use of Pushdown Automata

5.2.3.1 Basic Definitions and Results

A pushdown automaton is a finite automaton with a stack, which is simply a memory structure operating on the principle **Last-In-First-Out** (LIFO). A stack may be viewed as a stack of cards in a box exposing only the top card. The only way a card can be added or removed is through the top of the stack so that the last card added is always the first card removed. The importance of this fundamental structure is evident from the fact that almost all CPU's have machine level instructions for manipulating a stack both directly and indirectly. We must add at least two more symbols to the alphabet Σ, the first must place the current symbol onto the stack (referred to as **pushing** the symbol on the stack) and the second must remove the symbol at the top of the stack (referred to as **popping** the symbol from the stack). In general, the transition function depends on the symbol at the top of the stack as well as the current symbol.

Definition: A **pushdown automaton** (PDA) is denoted by the 7-tuple $(Q,\Sigma,\Gamma,\delta,q_0,Z_0,F)$ where

a. Q is a finite set of states.

b. Σ is the input alphabet.

c. Γ is the stack alphabet.

d. $q_0 \in Q$ is the initial state.

e. $Z_0 \in \Gamma$ is the initial stack symbol.

f. $F \subseteq Q$ is the set of final states

g. δ is a function

$$\delta: Q\times(\Sigma \cup \{\epsilon\})\times\Gamma \rightarrow \wp(Q\times\Gamma^{*}) \qquad (5.14)$$

mapped into **finite** subsets of $Q\times\Gamma^{*}$.

If **q** is the current state of the PDA, **a** the current input symbol and **Z** the symbol at the top of the stack, then $\delta(\mathbf{q,a,Z})$ gives a set of possibilities $\{(\mathbf{q}_i,\mathbf{Z}_i)\}$ where each pair gives a new state for the PDA and a new string of stack symbols. The leftmost symbol of the string Z_i becomes the stack top, and the other symbols placed sequentially below it. Note that this is a non-deterministic automaton since the transition function maps the current configuration to a set of possible new states. In exactly the same way as for NFAs, we require a probability distribution for the transition set in frame modelling:

> **Definition**: An **Probabilistic Pushdown Automaton** is defined to be a PDA with distribution function $\delta_p: Q\times(\Sigma\cup\{\epsilon\})\times\Gamma\times Q\rightarrow\mathbf{R}$ where δ_p is interpreted as for NFAs.

We assume all PDAs to be probabilistic PDAs in what follows. It is possible to define a deterministic PDA, but the language accepted by such an automaton is not equivalent to that accepted by the non-deterministic PDA. The language accepted by pushdown automata is a very important one:

5.2.3.2 Context Free Languages

Context Free Grammars (CFGs) originate from an attempt to describe the structure of natural languages, for example the following rules can be used to build a certain class of english sentences:

<Sentence>	→	<Noun Phrase> <Verb Phrase>
<Noun Phrase>	→	<Adjective><Noun Phrase>
<Noun Phrase>	→	<Noun>
<Verb Phrase>	→	<Verb Phrase><Noun Phrase>
<Verb Phrase>	→	<Verb><Adverb>

The rules given above are called **productions** and define the language recursively in terms of primitive symbols called **terminals**. To generate a set of sentences, we need only provide rules for the **terminals**, such as

<Noun>	→	book
<Adjective>	→	heavy
<Verb>	→	look

and so on. Although we could provide a complex set of productions for a natural language which generates most sentences, this formal structure is not sufficient for providing a full specification of natural languages. In particular, a CFG does not refer to the semantic value of a construction. The result is that meaningless sentences are possible. But CFGs can be used to describe the structure of many programming languages quite satisfactorily, since every construction is also valid in the semantic sense. An example of the success of CFGs in describing structures in programming languages is given by the **productions**

<expression>	→	(<expression>)
<expression>	→	<expression>+<expression>
<expression>	→	<expression>*<expression>
<expression>	→	**identifier**

which may be used to generate all arithmetic expressions. Programming languages such as Pascal and C are easily described using CFGs.

> **Definition**: A **Context Free Grammar** G is a quadruple (V,T,P,S) where V is a finite set of **variables** and T a finite set of **terminals** and $V \cap T$ is empty. P is a finite set of productions of the form
>
> $$v \rightarrow \alpha \text{ where } \alpha \in (V \cup T)^* \tag{5.15}$$
>
> and the variable $S \in V$ is the **Start Symbol**. If β and γ are two strings such that γ is obtained from β with zero or more productions, then

$$\beta \rightarrow_G \gamma \tag{5.16}$$

The **language generated** by G is defined to be

$$L(G) = \{w \mid w \in T^* \text{ and } S \rightarrow_G w\} \tag{5.17}$$

The equivalence between **Context Free Languages** and **Pushdown Automata** is given by the following theorem:

> **Theorem**[34]: The class of languages ac[illegible] by PDAs is precisely the class of **Context Free Languages**.

5.2.3.3 Modelling using PDAs

The structure is very similar to that proposed using NFAs, with pushdown automata forming the basic components. The **combination alphabet** Λ must be extended to include symbols for manipulating the stack, and the simplest possible addition is to add a **push** symbol π^+ for pushing the current symbol onto the stack, and a **pull** symbol π^- for removing the symbol at the top of the stack.

$$\Lambda^p = \Lambda \cup \{\pi^+, \pi^-\} \tag{5.18}$$

For each O_i in **O** we define a PDA $(\theta_i, \Lambda^p, \Lambda^p, \delta, q_{i0}, Z_{i0}, F)$. The transition function for a PDA depends on the current symbol being read as well as the symbol at the top of the stack, so it is possible to build a system referring to the history of the system dynamics. It is clear that a stack may be used for ensuring that the PDAs depend on past values, to see this we need only consider the worst case scenario where the stack must be filled with every single combination since the beginning of the simulation. In general, the past combinations must be pushed onto the stack in the reverse order they are referred to. The system is presented diagrammatically in **Figure 5.6**.

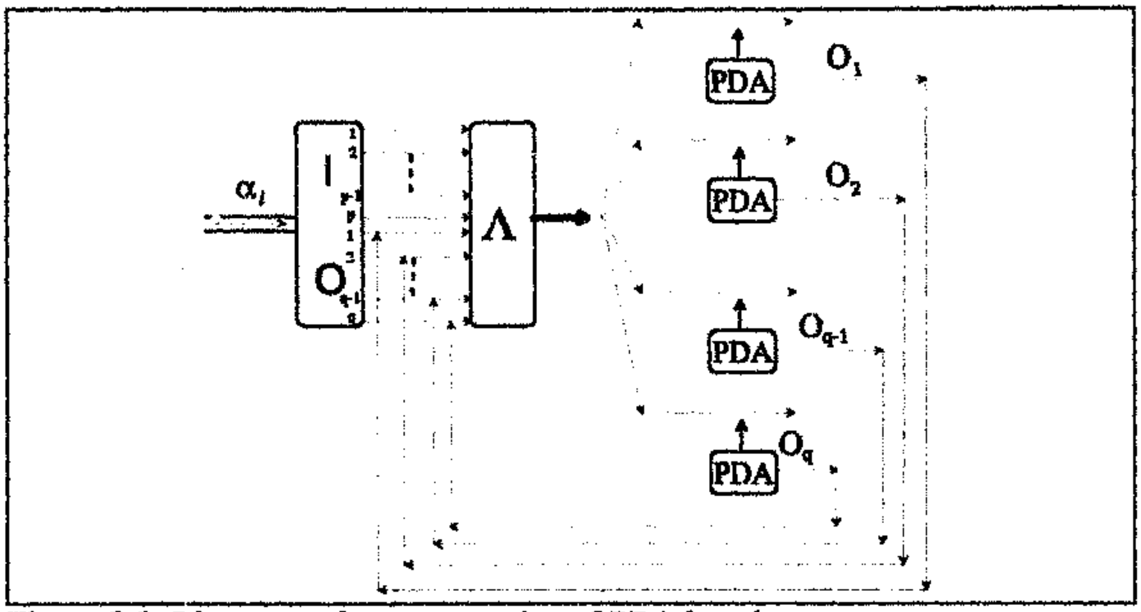


Figure 5.6: Diagrammatic representation of PDA based structure.

Unlike the situation using NFAs, the input stream to the PDA structure given above is not built up from the values of the system values. Where and how do the symbols π^{+} and π^{-} required to manipulate the stack arise? They arise from an approximate model the user must provide, written in a Context Free Language and making references to the necessary past values. In **Chapter 6**, a specification for an improved modelling environment will be provided which is based on this structure.

5.3 Specification of a frame modelling environment

5.3.1 Introduction

It was demonstrated in **Chapter 4** that a finite alphabet may be used to give an instantaneous description of the ecosystem to an arbitrary degree of precision. In **Section 5.2.2** the use of **Non-deterministic Finite Automata** for the modelling of the variables in the output group **O** was discussed. The second model presented in **Section 5.2.3** does not add any artificial variables to **I** and **O**, but nevertheless allows explicit references to past values through the use of **Pushdown Automata** as the controlling structure. From a practical viewpoint, the latter approach is more useful than the former since it allows the use of a Context Free Language and includes the former approach as a special case. In what follows, we consider the consequences of both models in the design of a frame modelling environment.

5.3.2 Rule based systems

The structure considered in **Section 5.2.2** naturally leads to rule-based models, where a **rule** is given by:

Definition: A **qualitative rule** is a statement of the form

IF **boolean expression** THEN **outcome**

where **boolean expression** may be any combination of the **boolean variables** using the boolean operators **AND** ($\wedge$), **OR** ($\vee$) and **NOT** ($\neg$). An **outcome** is an assignment to one of the variables.

A **stochastic rule** is a rule with a set of possible outcomes and a probability distribution on that set.

A **Qualitative rule based model** is a set of qualitative rules.

A rule is cast into a more useful form using the following theorem:

Theorem: A finite set of **rules** involving finite variables may be expanded to an equivalent set of rules which contains **AND** operators only, and where the boolean variables involve only expressions of the form $\mathbf{x}=\mathbf{x}_i$ where $\mathbf{x}_i$ is one of the possible values of $\mathbf{x}$.

Proof: To show this, we consider an arbitrary **rule** in the original set:

IF **boolean expression** THEN **outcome**

By repeated use of the equivalences

$$\neg(a \wedge b) \equiv \neg a \vee \neg b, \quad \neg(a \vee b) \equiv \neg a \wedge \neg b \qquad (5.19)$$

we rewrite the condition part so that the **NOT** operators operate only on the boolean variables: Each boolean variable is of the form $\mathbf{x}<\mathbf{x}_i$, $\mathbf{x}\leq\mathbf{x}_i$, $\mathbf{x}=\mathbf{x}_i$, $\mathbf{x}\geq\mathbf{x}_i$, $\mathbf{x}>\mathbf{x}_i$ where $\mathbf{x}$ is a finite variable and $\mathbf{x}_i$ one of its states. We do not allow the comparison of two different variables, as the source α of their values is not necessarily ordered. Since the variables are finite, any boolean variable involving <, >, ≤ or ≥ may be expanded to a boolean expression containing = only. For example, the boolean variable $\mathbf{x}<\mathbf{x}_i$ can be rewritten as

$$(x<x_i) \equiv \bigvee_{j<i} (x=x_j) \qquad (5.20)$$

The process is similar for removing the **NOT** operators:

$$\neg(x=x_i) \equiv \bigvee_{j\neq i} (x=x_j) \qquad (5.21)$$

By using the equivalence

$$a \wedge (b \vee c) \equiv (a \wedge b) \vee (a \wedge c) \qquad (5.22)$$

the **boolean expression** may be rewritten as

$$\bigvee_{i=1}^{n} e_i^{\wedge,=} \qquad (5.23)$$

where n is the number of **OR** operators in the **boolean expression**, and each e_i is a boolean expressions containing only boolean variables of the form $\mathbf{x}=\mathbf{x}_i$ and **AND** operators. The **rule** may then be expanded into n rules "IF e_i THEN **outcome**" where i goes from 1 to n. and each e_i is in the form required by the theorem.

We are now in the position to proof a very important theorem which justifies the use of qualitative rule based frame models:

Theorem: **a**. Qualitative rule based models are equivalent to the NFA structure defined in **Section 5.2.2.**
b. Stochastic Qualitative Rule based models may be used to model a dynamical ecosystem (as defined in **Chapter** 4) arbitrarily closely.

Proof: **a**. Given a qualitative rule based model, the above theorem may be used to expand the rule set into an equivelent rule based model whose rules are in the form described above. But the condition part of each rule is then precisely an element of the combination alphabet Λ, and any element of Λ may be expressed as the condition part of a rule.

By further subdividing the expanded rule set into subsets of rules affecting the same variable, it is clear that each of these subsets is a detailed description of the transition function δ of the NFA $(\theta_j^*,\Lambda,\delta_j,q_j,F_j)$ associated with the output variable $O_j \in \mathbf{O}$. The converse is also true.

b. Follows by the **Fundamental Theorem of Frame Modelling** proved in **Chapter 4**, and part **a**.

The first prototype of the frame modelling environment is based on this theorem.

5.3.3 First prototype of a frame modelling environment

A specialized frame modelling environment has already been implemented based on the theory contained in **Section 5.3.2**: An example of a model implemented using this software has been given in **Chapter 3**. The basic structure of this prototype frame modelling environment is discussed below: Refer to the software documentation for a more detailed description.

The basic unit of the frame modelling environment (referred to as the shell in what follows) is the **rule**, but we extend the definition of a rule to allow multiple THEN and ELSE outcomes, and the nesting of rules within rules. Rules are used to determine the dynamics within the frames, and the transitions between frames. By definition, each frame has its own rule set associated with it which is only used to determir the system dynamics once the system is in the particular frame. A separate set of rules is used to determine the switching between frames, and operates at a higher le l than the internal frame rules. In the shell, a global rule set is also allowed which is considered no matter which frame the system is in at any given moment. This has been in done in recognition of the fact that important processes may be common to every frame, and to repeat these processes for each frame would clearly be a needless repetition of the same submodel.

The main purpose of **Chapter 4** and the preceding contents of **Chapter 5** was to demonstrate that by using only finite variables, it is possible to model a dynamic ecosystem as defined in **Chapter 4** as accurately as we like through the use of the qualitative rule based frame modeiling technique. When it comes to the practical implementation of a shell, it would be pedantic and unnecessarily inflexible to insist that all variables be finite. In **Chapter 4**, it was proved that a finite alphabet may be used to transmit the information content of a frame $[\mathbf{T}_i,\Omega_i]$ with an error which is only bounded below by the capacity of the measurement channel. But that says nothing about the accuracy of the observed individual components of the ecosystem. When we look at specific systems, there will be components which may be measured to a high degree of accuracy, and this possibility is catered for in the shell by allowing real valued (quantitative) variables. The finite variables may have their levels enumerated by qualitative labels (such as 'Very low', 'Low', 'Medium', 'High' and 'Very high'). Other variable types have been mentioned in **Chapter 3**.

The channels used for in the formal theory each frame are defined to be one way, since we are interested in an instantaneous description of the ecosystem. But when building a model, we simulate the behaviour of some of these variables (the output group of variables **O**) and so are forced to consider the accuracy both in measuring a variable and the accuracy of the modelled effect of a particular process on the dynamic behaviour of the variable. An example would be a variable **tree** having only four meaningful levels **seedling** (height less than a metre), **small** (height between 1 and 3 metres), **medium** (height between 3 and 5 metres) a **tall** (height over 5 metres). However, the process of growing may involve an increment of at most 50 cm per time unit, so that several time units may be required for the variable to switch from one level to the next. This is allowed for in the shell by permitting a fractional increment to be added to a variable even if that variable is qualitative.

As in the case with the rules, the shell further distinguishes between global and local variables. The global variables are those variables which are required by the frame models in two or more of the frames (an example would be **rainfall** which we would want to refer to in all the frames). The local variables of a particular frame exist in that frame only. The shell further allows global quantitative variables to be interpreted locally as a qualitative variable. For example, if **rainfall** is a quantitative variable with a value between 0 and 1000 (in millimetres), then the shell would allow **rainfall** to be interpreted as a qualitative variable having three levels (**low** 0-250, **medium** 250-650, **high** 650-1000) in the first frame, and as a five level qualitative variable in the second frame (**very low** 0-150, **low** 150-250, **medium** 250-650, **high** 650-850, **very high** 850-1000).

There is a limited spatial component to the current version of the shell: Two dimensional models are possible using the shell, with the spatial distribution described in terms of a grid or in terms of regions which are also defined on grids. The shell allows different variables to have different spatial descriptions, a realistic condition in the measurement of important variables in real ecosystems. When two variables of different spatial descriptions are used together in the same context, an approximation is calculated for each by using a weighted average based on the region definitions and underlying grids of both variables.

Since quantitative variables may be used in the rule sets, the shell allows the use of any arithmetic expressions (including the **exp**, **log**, **trig** and **sqrt** functions) as is common in most programming language.

Past values of any variable may be referred to easily in any rule without the user having to create artificial variables to hold these values. An example is given by **rainfall 2 time periods ago** which would return the value of the rainfall two time periods previous to the current one. The values held by a variable in different spatial regions are also referred to in the same manner.

5.3.4 Systems involving a Context Free Language

The theory presented in **Section 5.3.3** suggests another technique for building frame models. The structure was built using **Pushdown Automata** as the controls which, equivalently, allows the use of a **Context Free Language** in building the model. The use of a CFL is superior since it does not require the use of additional artificial system variables for the storage of past information, so that the modeller may make references to past values if desired. The consequence of this approach is that a PASCAL type language may be used for the model: Instead of the basic unit being a rule as is the case for DFAs, the model would then be built up using the more complex structures possible in CFLs.

Both the NFA and PDA structures given earlier presented have a bonus because of their static nature: This property allows us to investigate algorithms for optimizing or fine-tuning the model (to be elaborated on in **Chapter 6**). We have in effect reduced the problem of tuning an ecological model to the problem of finding the best estimate of the probability distribution δ_p of a set of PDAs. For each combination symbol in Λ and value of the output variables (which are the states of the PDA), we need only find the probability distributions of the set of possible values the output variables will next assume. This may be done, for example, by using real data for all the variables or perhaps even an estimate of the way we would expect the output variables to change. The ability of a frame modelling environment to fine-tune itself by learning from observed or desired data is an exciting concept which will be investigated further.

Chapter 6: Conclusion

6.1 Review

In this thesis, we have investigated the qualitative frame based modelling technique which has been proposed by Starfield. Software was developed so that qualitative rule based frame models could be investigated in detail. Through the practical building of models, it has become clear that this technique is easily applied and closely follows the format of what is known about many ecosystems which cannot be modelled succesfully through the use of conventional techniques. Perhaps more so than any other modelling techique, qualitative frame models are easily refined by subdividing frames, or increasing the number of qualitative levels in the variables.

The aim of this thesis was to determine the validity of using qualitative rule based frame models, and to investigate any possibly generalizations of the technique. In ecological modelling, a common way to validate a modelling paradigm is to compare the theoretical predictions of a model built using the technique with the observed behaviour of an ecosystem. Such a verification approach is not ideal with qualitative frame models: A preliminary frame model (an example being the Brachystegia model presented in **Chapter 3**), has coarse qualitative variables with few qualitative states so that although it may predict the overall behaviour of the system in a qualitative manner, we have no idea how much further the model could be refined. We know that the technique works for the coarse preliminary models but would like to know if these models could be refined theoretically to make the error between the model and the real system arbitrarily small.

The approach we have followed is to provide a set of postulates which detail what we are assuming about the underlying ecosystem. For reference they are repeated here:

> **Postulate 1**: A dynamical ecosystem is completely specified by the dynamical quadruple $(\Omega,\mathcal{F},P,T)$.

Postulate 2: In the dynamical ecosystem $(\Omega,\mathcal{F},P,T)$, there exist a finite number of **attractors** Ω_i^* with their associated domains of attraction Ω_i where $i=1,\ldots,n$. Each pair $[T_i,\Omega_i]$ where T_i is the restriction of **T** to Ω_i is defined to be a **frame**.

Postulate 3: For each Ω_i there is a discrete channel $[\Sigma_i(A_i),P_i(\omega,\ ),\Sigma_i(B_i)]$ of limited capacity C_i through which all information about T_i must pass. The alphabets A_i and B_i may be chosen as desired, the only fixed parameter of the channel is C_i.

In the first postulate, we made the assumption that the ecosystem is a non-deterministic dynamical system and assumed little about the sample space Ω other than that it is a compact topological space and has a measure (with a σ-algebra of measurable sets) associated with it. Such a system is sufficiently general to include all other models for dynamic ecosystems.

An attactor was defined to be a region within the sample space from which the system cannot escape without being disturbed externally. The domain of attraction associated with an attractor was defined to be the set of points from which the system **may** move into the attractor. In the second postulate, we assumed that there are a finite number of attractors. (The theory would be equally valid for a countably infinite number of attractors, but any practical software implementation could deal only with a finite number of attractors)

In the third postulate, we made the difficulty in measuring ecosystems explicit in the technique, by requiring that any observations must be made through a noisy limited capacity discrete communication channel.

Using these postulates, the **Fundamental Theorem of Frame Modelling** was proved

Fundamental theorem of frame modelling: Given a dynamic ecosystem $(\Omega,\mathcal{F},P,T)$ with N **attractors**, the following are true:

a. The system may be decomposed into formal frames $[T_i,\Omega_i]$ where $i=1,\ldots,N$ such that the sum of the entropies of the formal frames is equal to the entropy of the entire space.

b. An instantaneous description of the ecosystem may be obtained using N channels (associated with each formal frame) of limited capacity, to an arbitrary accuracy and using a finite alphabet.

This theorem proved that the ecosystem may be divided into a finite number of **frames** such that the total sum of the information content of each frame is exactly equal to information content of the entire system, even though the **frames** may overlap. It also proved that it is possible to describe each frame as accurately as we like using a finite alphabet, with the error being bounded below by the measuring error or capacity of the communication channel.

In **Chapter 5**, this theorem was used to prove that qualitative rule based frame models can be refined:

> **Theorem:** Stochastic Qualitative Rule based frame models may be used to model a dynamical ecosystem (as defined in **Chapter 4**) arbitrarily closely.

This justifies the use of qualitative rule based frame models for the modelling of ecosystems which cannot be modelled using conventional techniques. The structures developed in **Chapter 5** go further and suggest improvements to the modelling technique:

6.2 Future Work

On the practical front, a second modelling environment is currently being developed based on the theoretical structure built up from PDAs in **Chapter 5** and is being written for the Microsoft Windows and IBM OS/2 platforms. The proposed modelling environment would be capable of accepting the models of the first prototype, but would also allow the use of a specialized PASCAL type language for specifying the core of the model. From the user perspective, this would be a far more flexible environment to model in. What exactly do we mean by a core of a model? In **Section 5.2.3**, a structure was detailed based on PDAs (which we have extended by including a probability distribution function δ_p) in a manner not requiring the addition of artificial system variables. This static structure allows the development of learning algorithms (outlined informally in Section 5.3.3). When developing a model, the modeller would break up the problem into small units of approximately uniform uncertainty. Each unit would be a rule, or a more complex structure built using the specialized frame language and would have a confidence associated with it (a confidence of 100% would indicate that the unit in question must not be changed as it is believed to be exact). Based on actual or desired data, the model would then tune itself while taking these confidences into account. The 'core' model (with 100% confidence) would not be altered, while the complete model would be 'tuned' automatically. In summary, the second modelling environment is to have the following characteristics:

a Spatial heterogeneity would be specifically catered for in the second environment, and would allow vector based spatial regions (instead of the current pixel based spatial regions).

b The self-tuning of models using observed data would be an important component of the second environment.

c. A specialized PASCAL type language would be developed for the user to specify the core model (which may be further optimized by self-tuning).

On the theoretical front, the following will be investigated in greater detail:

a. The AI concept of frames would be further explored, leading to improved flexible user interfaces, and enhancement of the learning algorithms for tuning models.

b. The algorithms would be studied in greater detail, especially in reference to the computational time required in modelling the ecosystem to a given accuracy, and the time taken for the tuning algorithms to converge (and the conditions under which they do converge).

In conclusion, the frame based modelling technique has proved to be fascinating both on the theoretical and practical fronts, and the formal theory has suggested a number of areas to explore in the future.

References

1. Starfield, A.M., *Private Communication*.

2. Starfield, A.M., 1990, Qualitative, rule based modelling, *Bioscience*, 40:601-604.

3. Starfield, A.M., Farm, B.P., Taylor, R.H., 1989, A rule-based ecological model for the management of an estuarine lake, *Ecol. Model*, 46:107-119.

4. Minsky, M., 1975, A framework for representing knowledge, In P.H. Winston, editor, *The Psychology of Computer Vision*.

5. Holly, T.D., Sinclair, A.R.E., McGlade, J., 1990, Elephants and fire as causes of multiple stable states in the Serengeti-Mara woodlands, *J. of An. Eco.*, 59:1147-1164.

6. May, R.M., 1976, Patterns in multi-species communities, In R.M. May, editor, *Theoretical Biology Principles and Applications*, Blackwell, Oxford.

7. Starfield, A.M., Bleloch, A.L., 1991, *Building Models for Conservation and Wildlife Management*, 2nd edition, Burgess International Group, Minneapolis.

8. Grinell, J., 1924, Geography and Evolution, *Ecology*, 5:225-229.

9. Grinell, J., 1928, The presence and absence of animals, *Univ. Calif. Chronicle*, 30:429-450.

10. Elton, C., 1942, *Animal Ecology*, Sidgwick and Jackson, London.

11. Pianka, E.R., 1976, Competition and Niche Theory, In R.M. May editor, *Theoretical Biology Principles and Applications*, Blackwell, Oxford.

12. May, R.M., 1974, *Stability and Complexity in Model Ecosystems*, 2nd edition, Princeton University Press.

13. Holling, C.S., 1966, The strategy of building models of complex ecological systems, In K.E.F. Watt, editor, *Systems analysis in Ecology*, Academic Press, New York.

14. Holling, C.S., 1968, The tactics of a predator, In T.R.E. Southwood, editor, *Insect Abundance*, Blackwell, Oxford.

15. Holling, C.S., 1978, editor, *Adaptive Environmental Assessment and Management*, Wiley, Chichester.

16. Levins, R., 1966, The strategy of model building in population biology, *Amer. Sci*, 54:421-431.

17. Levins, R., 1968, Ecological Engineering: Theory and Practice, *Quart. Rev. Biol.*, 43:301-

305.

18. Kolomogorov, A.N., 1936, Sulla Teria di Volterra della lotta per l'Esistienza, *Giorn. Instituto Ital. Attuari*, 7:74-80.

19. Slatyer, R.O., editor, 1976, Dynamic Changes in Terrestrial Ecosystems, *Proceedings SCOPE MAB workshop*, Santa Barbara, California.

20. Clements, F.E., 1916, Plant Succession, *Carnegie Inst. Washington, Publ. 242*, 512 pp.

21. Horn, P.S., 1976, Succession. In R.M. May, editor, *Theoretical Biology Principles and Applications*, Blackwell, Oxford.

22. Egler, F.E., 1954. Vegetation Science concepts, *Vegetatio*, 4:412-417.

23. Connell, J.H., Slatyer, R.O., 1977. Mechanisms of succession in natural communities and their role in community stability and organization, *Amer. Nat.*, 111:1119-1144.

24. Noble, I.R., Slatyer, R.O., 1980, The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbances, *Vegetatio*, 43:5-21.

25. Westoby, M., Walker, B., Noy-Meir, I., 1989, Opportunistic management for rangelands not at equilibrium, *J. of Range Management*, 42(4):266-273.

26. Martin, N.F.G., England, J.W., 1981, *Encyclopaedia of mathematics and its applications Vol 12: Mathematical theory of Entropy*, Addison-Wesley, Reading.

27. Walters, P., 1982, *An introduction to Ergodic theory*, Springer-Verlag, New York.

28. Billingsley, P., 1965, *Ergodic theory and information*, John Wiley & Sons, New York.

29. Shannon, C., 1948, A mathematical theory of communication, *Bell System Tech. J.*, 27:379-423, 623-656.

30. Schaffer, W.M., Kot, M., 1986, Differential systems in ecology and epidemiology, In A.V. Holden, editor, *Chaos*, Princeton University Press.

31. Wolf, A., 1986, Quantifying chaos with Lyapunov exponents, In A.V. Holden, editor, *Chaos*, Princeton University Press.

32. Grassberger, P., 1986, Estimating the fractal dimensions and entropies of strange attractors, In A.V. Holden, editor, *Chaos*, Princeton University Press.

33. Rudin, W., 1987, *Real and Complex Analysis*, 3rd edition, McGraw Hill.

34. Hopcroft, J.E., Ullman, J.D., 1979, *Introduction to automata theory, languages and computation*, Addison-Wesley.

35. O' Neill, R.V., DeAngelis, D.L., Waide, J.B., Allen, T.F.H., 1986, *A Hierarchical Concept*

of Ecosystems, Princeton University Press.

25×

Author: Quadling Mark Sherwood.

Name of thesis: Frame modelling of dynammic ecosystems.

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2015

LEGALNOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page)for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.