Structural Dominance Analysis and Theory Building in System Dynamics

Christian Erik Kampmann\textsuperscript{1} and Rogelio Oliva\textsuperscript{2*}

\textsuperscript{1}Department of Innovation and Organizational Economics, Copenhagen Business School, Denmark
\textsuperscript{2}Mays Business School, Texas A\&M University, College Station, Texas, USA

We provide a review of different approaches to linking model structure to observed behaviour with a particular view towards using models for theory building. We argue that theory building cannot be based upon pure simulation and model building alone: the inference from system dynamics models invariably uses concepts and analogies from simple feedback systems and models. Strengthening the analytical foundation for this inference will therefore have a direct impact on the strength of system dynamics as a theory-building tool. We identify four approaches to establish this link (traditional, pathway participation, eigenvalue and eigenvector), assess the strengths and weaknesses of each approach, and point to challenges and tasks ahead. We find that the eigenvalue and eigenvector approaches carry significant potential but that a more solid theoretical foundation of the method is required. However, since a ‘grand unified theory’ will never be possible, all tools will be based on approximations and it is only in their practical use that we can discover their real value. Copyright © 2008 John Wiley & Sons, Ltd.

Keywords system dynamics; linear model analysis; loop eigenvalue elasticity analysis; structural dominance analysis

INTRODUCTION

In recent years, there has been a growing interest among system dynamics researchers in developing methods for formal quantitative tools to help modellers understand the relationship between observed model behaviour and the elements of the model structure influencing this behaviour in large models. This process is obviously of great importance to both practitioners and theorists. For the former group, an understanding of the link between system structure and observed behaviour is the key to finding leverage points for policy initiatives. For the theorist, the system dynamics paradigm builds on the notion that structure causes behaviour, that is, a system dynamics theory of a particular phenomenon is an account of how certain feedback loops cause certain dynamic patterns of behaviour to appear. This qualitative understanding is often at least as important as the particular numerical predictions obtained, even

\*Correspondence to: Rogelio Oliva, Mays Business School, Texas A&M University, College Station, TX 77843-4217, USA. E-mail: roliva@tamu.edu

Copyright © 2008 John Wiley & Sons, Ltd.
in applied studies. Yet the rigor of such an account depends directly on the rigor with which the structure–behaviour link can be made in a given model.

The purpose of this paper is to give an overview of the different approaches taken to the structure–behaviour problem, to outline the apparent strengths and weaknesses of each approach with a view towards using system dynamics for theory building, and to point to some main challenges and tasks ahead in order to make the methods useful for a wider audience of system dynamics practitioners and researchers.

Understanding model behaviour is closely related to the process of model testing and validation, for which there is a well established tradition and an extensive literature in the field (e.g. Forrester and Senge, 1980; Barlas, 1989). Indeed there is no sharp line between model building, testing, validation and analysis – in system dynamics practice, these phases overlap (Forrester and Senge, 1980). The present paper is focused exclusively on the tools used in linking structure to behaviour, that is, how behaviour patterns may be attributed to feedback loops (or external driving forces) and how the relative importance of different feedback loops may shift over time (shifting ‘loop dominance’).

Traditionally, system dynamicists have relied on trial-and-error simulation, changing parameter values or switching individual links and feedback loops on and off, to discover important system elements. The tradition is well developed and includes a set of principles for partial model formulation (Homer, 1983; Oliva, 2003). The intuition guiding this effort relies on simple feedback systems with one or a few state variables, where the behaviour is fully understood.

In large-scale models with perhaps hundreds of state variables, however, the traditional approach shows significant limitations. In practice, model building and analysis is often done using a ‘nested’ partial model testing approach where one goes from the level of small pieces of structure to entire subsystems of the model, with frequent re-use of known formulations and partial models. Although this approach does carry a long way, it can be very difficult to discover feedback mechanisms that transcend model substructures in ways not anticipated by the modeller in the original dynamic hypothesis. Thus, there is a danger that observed behaviour is falsely attributed to certain feedback mechanisms when in fact another set of feedbacks is driving the outcome.

Clearly, a more rigorous theory for the link between feedback structure and behaviour in general large-scale systems would be of great value. By way of introduction, the following section of the paper discusses what constitutes an ‘explanation’ of the link between structure and behaviour and the fundamental analytical limits to this ideal, given that the systems are usually nonlinear. It is important to recognize that these limits exist, that is, that we cannot hope to completely ‘automate’ the model analysis. Therefore, the system dynamics modelling process will always involve critical inquiry and interaction with the model. The goal of analytical tools must be to support the modeller’s intuition and insight with rigorous mathematical methods. We then proceed to discuss four main approaches used in system dynamics to explain the link between structure and behaviour, which we have termed the traditional approach, the pathway participation approach, eigenvalue elasticity analysis, and eigenvector analysis, respectively. The paper is intended to give an overview of the strategies adopted by various researchers. For a detailed description of how to perform and interpret the various analyses we point the reader to the citations we provide in each section. In a subsequent section, we provide a brief example of how a formal analytical tool can reveal errors of feedback loop inference. The paper concludes with a discussion of the most important tasks ahead and speculates on the prospects for more widespread use of analytical tools in the future.

CHARACTERIZING LINEAR AND NONLINEAR SYSTEMS

A system dynamics model can be represented mathematically as a set of ordinary differential equations

\[ \frac{dx(t)}{dt} = \dot{x}(t) = f(x(t), u(t)), \quad (1) \]
where \( x(t) \) is a column vector of \( n \) state variables (levels) \( (x_1(t), \ldots, x_n(t)) \), \( u(t) \) is a column vector of \( p \) exogenous variables or control variables \( (u_1(t), \ldots, u_p(t)) \), \( f() \) is a corresponding vector function, and \( t \) is simulated time. In this paper, we restrict our attention to the state variables (levels) of the model for notational convenience, ignoring the auxiliary variables. Mathematically, a model can always be brought to the reduced form (1), but in practice, the auxiliary variables give a more intuitive account of the analysis. In general, \( f \) is a nonlinear function of its arguments, and we speak of a nonlinear system. Conversely, if \( f \) is a linear function, we speak of a linear system.

Given the model structure (1), knowledge of the initial conditions \( x(0) \), and the path of the input variables \( u(t) \), the behaviour of the model is completely determined. It is in this sense that the model structure (1) constitutes a ‘theory’ of the time behaviour \( x(t) \) (Forrester, 1961; Sterman, 2000). But we are interested in methods that yield a more compact explanation, short of having to simulate the entire model structure.

In its ultimate general form, however, this dream is beyond reach: Since the days of 19th century mathematician Henri Poincaré, we have known that it is impossible to find general analytical solutions to nonlinear systems. Furthermore, the development of nonlinear dynamics and chaos theory has proven that such systems, even when they have very few state variables, can produce highly complex and intricate behaviour that would be impossible to anticipate, let alone analyse, directly from their structure (e.g. Richardson, 1988; Ott, 1993). Thus, in the absence of a ‘grand unified theory’ of dynamical systems, we shall always have to rely on simulation to discover the dynamics implied by the structure.

What we can aim for is a set of tools that will guide intuition and help identify dominant structure in the model. By dominant structure we mean particular feedback loops, or possibly external drivers, that are in some sense ‘important’ in shaping the behaviour of interest. To the extent that we can both rigorously define and identify such dominant structures, we choose to say that we have found a ‘theory’ of the observed behaviour.

Although the term ‘behaviour’ may appear rather loose, experience and reflection tells us that there is a limited number, perhaps a dozen or so, of relevant behaviour patterns that dynamical systems can exhibit. Some of these behaviours, like exponential growth, exponential adjustment, and damped or expanding oscillations, are typical of linear systems. Others, like limit cycles, quasi-periodic motion, mode-locking, and chaos, can only be exhibited by nonlinear systems.

Common to the approaches considered in this paper is that they are based on tools from linear systems theory, that is they approximate the nonlinear model (1) with a linearized version, using as a first-order Taylor expansion around some operating point \( x(t_0) = x_0, u(t_0) = u_0 \), that is

\[
\dot{x}(t) \approx f(x_0, u_0) + \frac{\partial f}{\partial x}(x(t) - x_0) + \frac{\partial f}{\partial u}(u(t) - u_0), \tag{2}
\]

or, by redefinition of the variables

\[
x \rightarrow x-x_0-f(x_0,u_0) \ (t-t_0) \quad \text{and} \quad u \rightarrow u-u_0,
\]

\[
\dot{x}(t) \approx Ax(t) + Bu(t), \tag{3}
\]

where \( A \) is an \( n \times n \) matrix of partial derivatives \( \frac{\partial f_i}{\partial x_j} \) and \( B \) is an \( n \times p \) matrix of partial derivatives \( \frac{\partial f_i}{\partial u_j} \), taken at the operating point \( f(x_0,y_0) \).

For the linear system (3), there is a well-developed and extensive theory of the system behaviour as a function of its structure, expressed in the matrices \( A \) and \( B \). One may broadly distinguish two parts of the theory, named classical control theory (e.g. Ogata, 1990) and modern linear systems theory (e.g. Chen, 1970; Luenberger, 1979). We return to the classical control theory in the next section.

Modern control theory or linear systems theory (LST) is concerned with the dynamical properties of the system as a direct function of the system matrices \( A \) and \( B \). A key element in this theory is the notion of the system eigenvalues, that is the eigenvalues of the matrix \( A \). If, for simplicity, we restrict ourselves to the endogenous dynamics of the system (set \( u = 0 \)), we can write the solution
to (3) as

\[ x_i(t) = c_{i1}\exp(\lambda_1 t) + c_{i2}\exp(\lambda_2 t) + \cdots + c_{in}\exp(\lambda_n t), \quad i = 1, \ldots, n. \]  

(4)

where \( \lambda_1, \ldots, \lambda_n \) are the \( n \) eigenvalues of the matrix \( A \) and \( c_{ij} \) are constants that depend upon the eigenvectors and the initial condition of the system. In other words, the resulting behaviour is a weighted sum of distinct behaviour modes, \( \exp(\lambda t) \). If an eigenvalue is real, the corresponding behaviour mode is exponential growth (if \( \lambda > 0 \)) or exponential adjustment (if \( \lambda < 0 \)). Complex-valued eigenvalues come in complex conjugate pairs \( \lambda = \tau \pm i\omega \) which give rise to oscillations of frequency \( \omega \) that are either expanding (if \( \tau > 0 \)) or damped (if \( \tau < 0 \)). In this manner, the eigenvalues serve as a compact and rigorous characterization of the behaviour (of linear systems).

At any point in time, any system, linear or nonlinear, may be approximated by the expression (4). Whether it remains a good approximation depends upon how much and how quickly the eigenvalues change due to the nonlinearities in the function \( f \). If they are more or less constant for significant periods of time, we may speak of quasi-linear systems that are well approximated by the linear system. In some cases, however, the eigenvalues change so rapidly that it makes little sense to characterize the behaviour by Equation (4). (See Kampmann and Oliva, 2006 for further discussion.)

The concept of dominant structure, on the other hand, is less clear. To even claim that certain parts of the model are more important than others is perhaps to go too far. Richardson (1986) suggested a taxonomy of approaches to the notion of dominant structure, where he classified the approaches along three dimensions, model reduction versus structure contribution, time graphs versus frequency response versus eigenvalues, and linear versus nonlinear. The focus here will be on Richardson’s loop contribution or, more generally, structure contribution approaches. The structure contribution approach reflects the intuitive idea that if one removes the structural element under consideration, for example by weakening a link or switching off a feedback loop, and the behaviour then ‘disappears’, one would say that the element in some sense ‘causes’ the observed behaviour. This notion underlies the traditional trial-and-error simulation approach. This approach can be refined by instead considering marginal (infinitesimal) changes in structure, for example in the strength of a particular link. It is then possible to derive rigorous analytical results for the resulting change in behaviour expressed as the eigenvalues of the linearized model. One would then say that if a change in a system element has a relatively large effect upon the behaviour pattern of interest, this element is ‘significant’ in ‘causing’ the behaviour. This refinement is what underlies the last two approaches we describe below.

TRADITIONAL CONTROL THEORY APPROACHES

The first set of methods, which we call the traditional approach, has been used for decades and is part of the standard curriculum in system dynamics teaching at the graduate level. It involves using the concepts from classical control theory (Ogata, 1990) to very simple systems with only a few state variables.

The starting point is the simple first- and second-order positive and negative feedback loops found in any introductory treatment of system dynamics. The advantage of the approach is its simplicity. Although it serves at a guide to intuition, however, the obvious shortcoming is that it applies rigorously only to simple systems. There have been some attempts to treat higher-order systems by adding a few feedback loops (Graham, 1977), but the step to large-scale models is beyond this method given its inherent limitations.

In a way, Graham’s objective was the opposite of what seems to be the ideal pursued by the other methods: he sought to derive a set of ‘principles’ based on simple feedback structures, as an intuitive guide or metaphor for understanding behaviour. Thus, the goal is to create an intuitive understanding of the mathematical results of classical control theory rather than to
develop a new formal theory that can explain behaviour.

Graham distils a number of principles that are based on the metaphor of a ‘disturbance’ travelling along the chain of causal links in a feedback loop and getting amplified, damped, and possibly delayed in the process. For major negative feedback loops, which are known to tend to produce oscillation, adding minor negative loops and cross-links, or shortening the delay times, increases the damping. Conversely, adding positive loops in to the oscillatory system tends to lengthen the period of oscillation whereas the effect on the damping depends upon the delays in the positive loop. Using the metaphor of pushing a child on a swing, where the timing of the push affects whether the swings are amplified or attenuated, it becomes clear that the timing of the propagation of a disturbance has as much importance for its effect on the damping as its strength. For analyzing the behaviour of positive feedback loops, Graham suggested calculating the open-loop steady-state gain, a measure of the amplification around the loop. A gain greater than unity will result in exponential growth while gains less than 1 will give exponential adjustment (levelling off or decay).

In the context of oscillating systems, system dynamics has also employed concepts from classical control theory, in the form of the frequency response methods. The frequency response is determined from the transfer function of the system, \( G(i\omega) \), which is a complex-valued function that specifies how an input signal \( u(t) \) with frequency \( \omega \) results in an output signal \( x(t) \) that may be phase shifted (delayed), and either amplified or attenuated. For linear systems, \( G \) can be calculated directly from the system matrices in (3) – the transfer function (matrix) is \( G(i\omega) = B(i\omega I - A)^{-1} \), where \( I \) is the identity matrix (see, e.g. Chen, 1970).

The approach nicely demonstrates the ‘endogenous viewpoint’ that behaviour (oscillations) is generated internally by the system. As an analytic tool for large-scale systems, however, the method does not seem to produce any additional insights. Thus, we may conclude that the classical approaches serve mostly as intuitive metaphors to guide the analyst rather than as full analytical tools. Indeed, it is the authors’ impression that they are rarely used outside the classroom.

### PATHWAY PARTICIPATION METRICS

The pathway participation method (Mojtahedzadeh, 1996; Mojtahedzadeh et al., 2004) represents a further development of an original suggestion by Richardson (1984/1995) to provide a rigorous definition of loop polarity and loop dominance. Richardson motivated this with the common confusion associated with positive feedback loops, which may exhibit a wide range of behaviours (Graham, 1977), as Barry Richmond noted with his characteristic humour:

“Positive loops are ... er, well, they give rise to exponential growth ... or collapse ... but only under certain conditions ... Under other conditions they behave like negative feedback loops ...” (Richmond, 1980).

Richardson proposed that the polarity of a loop be defined as the sign of the expression

\[
\frac{\partial x_i}{\partial x_i} = \frac{\partial f_i(x, u)}{\partial x_i},
\]

in the model (1), with a positive sign indicating a positive loop and vice versa. When several loops operate simultaneously, the sign of the expression indicates whether the positive or negative loops dominate. Note, however, that the definition only applies to minor loops (i.e. loops involving a single level). Put differently, it only considers the diagonal elements of the matrix \( A \) in the linearized system (3). Richardson (1984/1995) demonstrates how even with this limitation, analyzing the system with this metric can (sometimes) yield insights into behaviour of higher-order systems.

The expression (5) hints that it is relevant to consider the curvature, that is the second time derivative, \( \ddot{x} \), of a variable when looking for dominant structure. Although he does not define it as such, this is effectively the focus of Mojtahedzadeh’s pathway method. The sign of
the expression $\dot{x}/\dot{x}$, which Mojtahedzadeh denotes the total pathway participation metric or PPM, indicates whether the behaviour appears dominated by positive or negative loops, much in line with Richardson’s definition of dominant polarity. A zero curvature indicates a shift in loop dominance.

Mojtahedzadeh’s method proceeds by decomposing the PPM into its constituent terms separating the influence of each of the system’s state variables on the behaviour of $x_i$. By explicitly considering auxiliary variables $y$ in the model, one may further decompose each term $\partial f_i/\partial x_j$ into a sum of terms corresponding to a causal chain or pathway. Mojtahedzadeh then considers each possible pathway and defines the dominant pathway as the one with the largest numerical value and the same sign as PPM. Having selected this dominant pathway, which originates in the state variable $x_j$, the procedure is repeated for that state variable $x_j$ and so forth, until one either reaches one of the already ‘visited’ state variables (in which case a loop has been found) or an exogenous variable (in which case an external driving force has been found). Thus, the procedure may result in three alternative forms of dominant structure: a ‘pure’ minor or major feedback loop, a pathway from a feedback loop elsewhere in the system, or a pathway from an exogenous variable. By dividing the observed model behaviour into different phases according to the sign of the first and second derivatives and then applying the method just described at different points in during these phases, one identifies how the dominant structure changes over time.

The PPM method is still mostly used at an early explorative stage on rather simple models, where it does appear to aid insight into the dynamics (e.g. Oliva and Mojtahedzadeh, 2004). The method has been implemented in a software package, Digest (Mojtahedzadeh et al., 2004), yet its use by practitioners still seems limited.

From the studies performed so far, it is clear that the main strength of this method is its relative computational simplicity (it does not require computing eigenvalues, which is a numerically demanding task), and the intuitive and direct connection it makes between the observed behaviour and the influencing structural elements. Unlike the other approaches that operate in the ‘frequency domain’, the method considers the time path of a specific variable directly.

There are, however, some important outstanding issues that remain to be clarified. First, the method is not suitable for oscillatory systems. The problem is easy to recognize when one considers how the PPM measure will vary over the course of a sinusoidal outcome from a linear system: the sign of the PPM will shift twice during each cycle, indicating that the behaviour is alternately dominated by positive and negative loops, even though the relative strength of the system loops, and hence the loop dominance, remain constant all the time. Richardson (1984/1995) already alluded to this problem by noting that the measure only considers the diagonal elements in the system matrix in (3). This is a significant limitation, given the prevalence and importance of oscillation in system dynamics analysis.

A second limitation of the current implementation of PPM is that it uses a depth-first search for the single most influential pathway for a variable. This strategy does not capture the situation where more than one structure may contribute significantly to the model behaviour and, through the depth-first algorithm, may miss alternative paths that could prove to yield a larger total value of the metric. This problem could likely be addressed by modifying the search algorithm and is most likely of minor importance.

A third limitation is the emphasis on identifying a single ‘dominant’ structure. In reality, of course, many loops and pathways influence a variable’s behaviour simultaneously. Reducing the consideration to a single one of these may miss important features of the structure–behaviour relationships. It is more appropriate to consider the relative importance of alternative pathways, yet the method does not address how one would partition the behaviour among pathways – only among individual links. Thus, while the notion of pathways seems an interesting and
useful idea, ultimately it may be more effective to use a list, ranked in order of magnitude, of the pathways that influence a variable.

Finally, the method shares a weakness with the traditional method in that it considers primarily partial system structures rather than global system properties. In contrast, the two eigenvalue methods to which we now turn are based on a rigorous characterization of the entire system (at a given point in time).

**EIGENVALUE ELASTICITY ANALYSIS**

The third method may be termed eigenvalue elasticity analysis (or EEA for short) and builds upon the tools from ‘modern’ linear systems theory (LST), applied to the linearized model (3). The method is concerned with the structural elements that significantly affect the system eigenvalues or behaviour modes – the values $\lambda_i$ in (4). Specifically, it measures influence by the elasticity of an eigenvalue $\lambda$ with respect to some parameter $g$ in the model, defined as $\varepsilon = (\partial \lambda / \partial g)(g/\lambda)$, that is the fractional change in the eigenvalue relative to the fractional change in the parameter. The advantage of this fractional measure is that it is dimensionless, that is independent upon the choice of units, including the time scale unit. Sometimes, the influence measure is used instead, defined as $\mu = (\partial \lambda / \partial g)g$, which has dimension [1/time] and so depends upon the choice of time unit, but it is generally easier to interpret for complex-valued eigenvalues and avoids numerical problems with very small or zero eigenvalues (see Kampmann, 1996; Saleh et al., 2008).

The idea behind EEA was first introduced in system dynamics by Forrester (1982) in the context of economic stabilization policy. Though the criteria for behaviour stabilization are not new, the EEA method is unique in its attempt to use them to gain qualitative intuitive understanding of the system. A significant step in this direction was first suggested by Forrester (1983) with the notion that the elasticities of any link in the model (corresponding to elements of the matrix $A$ in the linearized system (3)), can be interpreted as the sum of elasticities of all feedback loops containing that link. We have chosen to name this approach loop eigenvalue elasticity analysis (LEEA).

Kampmann (1996) provided a rigorous definition of LEEA and also pointed to the fact that feedback loops are not independent. In other words, given the possibly very large number of loops in a given model (Kampmann demonstrated how the theoretical maximum number of loops grows combinatorically with the number of variables), it only makes sense to speak of individual contributions of a limited set of loops, which Kampmann termed the independent loop set (ILS). He proved that a fully connected system (where there is a feedback loop between any pair of variables – the typical case in system dynamics models) with $N$ links and $n$ variables has a total of $N-n+1$ independent loops and provided a procedure for constructing this set and calculating the loop elasticities.

Kampmann’s analysis points to a fundamental issue relating to the notion of feedback loops as a way to explain behaviour: the significance assigned to a particular loop depends upon the context (the chosen ILS). In other words, feedback loops are derived and relative concepts rather than fundamental independent building blocks of systems. Oliva (2004) further refined the definition of independent loop sets by introducing the shortest independent loop set (SILS) along with a procedure for constructing the set. Although a SILS is not generally unique, experience seems to suggest that it is easier to interpret (Oliva and Mojtahedzadeh, 2004). Yet, the issue remains that feedback loops are relative concepts.

The EEA/LEEA method has been applied in a number of contexts (e.g. Gonçalves et al., 2000; Saleh and Davidsen, 2001a,b; Gonçalves, 2003; Abdel-Gawad et al., 2005; Güneralp, 2006; Kampmann and Oliva, 2006; Saleh et al., 2008), but remains a tool employed only by specialists and in fundamental research, not least because it has not been incorporated into standard software packages. Its potential for widespread practice remains unexplored.

One might be sceptical that a method derived from linear systems theory may have any use for the nonlinear models found in system dynamics.
Kampmann and Oliva (2006) considered the question of what types of models the method would be particularly suited for. They defined three categories of models, based upon the behaviour they are designed to exhibit: (1) linear and quasi-linear models, (2) nonlinear single-transient models and (3) nonlinear periodic models. The first category encompasses models of oscillations, possibly combined with growth trends, with relatively stable equilibrium points (e.g. the classical industrial dynamics models in Forrester, 1961). Nonlinearities may modify behaviour (particularly responses to extreme shocks) but the instabilities and growth trends can be analysed in terms of linear relationships. Kampmann and Oliva concluded that LEEA showed the most promise and potential for this class of models because the analytical foundations are solid and valid, and because the method has the ability to find high-elasticity loops even in large models very quickly without much intervention on the part of the analyst. In single transient behaviour pattern models (e.g. Forrester, 1969, 1971; Sterman, 1981), nonlinearities usually play an essential role in the dynamics, but it is possible to divide the behaviour into distinct phases where certain loops tend to dominate the behaviour. In this class of models, LEEA also shows promise by measuring shifts in structural dominance by the change in elasticities; yet it requires more input from the analyst (e.g. in defining the different phases of the transition) and it has no obvious advantage over other methods, such as PPM. The third class, nonlinear periodic models, are those that exhibit fluctuating behaviour in which nonlinearities play an essential role, such as limit cycles, quasi-periodic behaviour, mode-locking and interactions among cyclical modes, or chaos (see, e.g., Richardson, 1988). Here the utility of the method is much less clear and depends upon the specifics of the model in question.

Compared to the former two methods, the EEA/LEEA is mathematically more general and rigorous, though many of the mathematical issues in the method remain to be addressed, as we summarize below. This rigor is also the main strength of the method, since it provides an unambiguous and complete measure of the influence of the entire feedback structure on all behaviour modes.

An emerging challenge is the computational intensity in calculating eigenvalues and elasticities. This is not so much an issue of computer time and memory space as of the stability of numerical methods. Kampmann and Oliva (2006) found that the numerical method used sometimes proved unstable, yielding meaningless results. Clearly, there is a need to explore this issue further, possibly building upon the developments in control engineering.

A more fundamental weakness is the difficulty in interpreting the results: eigenvalues do not directly relate to the observed behaviour of a particular variable. The concepts of eigenvalues and elasticities are rather abstract and unintuitive (Ford, 1999). There is a need for tools and methods that can translate them into salient, intuitive and parsimonious measures. A possible route may be to use (linear) filtering in the frequency domain to define a behaviour of interest. For example, an analyst may be concerned with structures causing a typical business cycle (3- to 4-year oscillation) and, by specifying a filter that ‘picks out’ that range of fluctuation, could obtain measures for structures that have elasticities in that range. Because filters are linear mathematical operators, all the analytical machinery of the LEEA method will also apply in this case – a significant advantage.

Using filters will also solve an issue that appears in large-scale models, namely the presence of several identical or nearly identical behaviour modes. Saleh et al. (2008) do consider the analytical problems associated with repeated eigenvalues, where it becomes necessary to use generalized eigenvectors, and where new behaviour modes appear that involve power functions of time. A filter essentially constitutes a weighted average of behaviour modes and in this fashion avoids the ‘identity problem’ of non-distinct eigenvalues.

The most serious issue, in our view, is how the results are interpreted using the feedback loop concept. As mentioned, the concept is relative (to a choice of an independent loop set). Moreover, practice reveals that the number of loops to consider is rather large and that the loops
elasticties often do not have an easy or intuitive explanation. A lot of care must be taken when interpreting the results. For instance, Kampmann and Oliva (2006) found that ‘phantom loops’ – loops that cancel each other by logical necessity and are essentially artefacts of the equation formulations used in the model – could nonetheless have large elasticities and thus seriously distort the interpretation of the results. These kinds of problems may not be intractable, but their resolution will require careful mathematical analysis.

Finally, a problem with EEA and LEEA is that they only consider changes to behaviour modes, not the degree to which these modes are expressed in a system variable of interest. This issue is addressed by also considering the eigenvectors of the system, which is the foundation for the analysis in the next section.

EIGENVECTOR (EVA) AND DYNAMIC DECOMPOSITION WEIGHTS (DDW) ANALYSIS

The last set of methods, which are still in early development, we have termed the eigenvector-based approach (EVA). EVA attempts to improve the EEA/LEEA method by considering how much an eigenvalue or behaviour mode is expressed in a particular system variable. The logic of the method and how LEEA and EVA complement each other is presented in Figure 1. As shown by Kampmann (1996), in a sense there is a one-to-one correspondence between eigenvalues and loop gains, whereas the eigenvectors arise from the remaining ‘degrees of freedom’ in the system. The observed behaviour of the state variables in the model is then the combined outcome of the behaviour modes (from the loop gains) and the weights for each mode (from the eigenvectors) in the respective state variable.

A number of researchers have attempted to develop EVA methods. Some emphasize the curvature (second time derivative) of the behaviour, similar to the starting point of the PPM method (Saleh and Davidsen, 2001a,b; Saleh, 2002; Güneralp, 2006). The slope or rate of change \( \dot{x}(t) \) of a given variable \( x \) in the linearized system may be written by

\[
\dot{x}(t-t_0) = w_1 \exp(\lambda_1(t-t_0)) + \cdots + w_n \exp(\lambda_n(t-t_0)),
\]

where the weights \( w_i \) are related to the eigenvectors. Then, differentiating with respect to time, one finds that the curvature at time \( t_0 \) is

\[
\ddot{x}(t_0) = w_1 \lambda_1 + \cdots + w_n \lambda_n.
\]

One may therefore interpret (7) as the sum of contribution from individual behaviour modes. Güneralp (2006) suggested using the terms on the right-hand side of (7) as weights to combine elasticities of individual behaviour modes \( e_i \) with respect to some system element (like a link gain or a loop gain) into a weighted sum as a measure of the overall significance of that system element. He further normalized the elasticity measure by the sum of elasticity measures for other system elements creating a measure that varies between \(+1\) and \(-1\). His results shed an alternative light on the behaviour of these models, though in our opinion, there is no dramatic improvement in intuition. In particular, the mathematical meaning, consistency and significance of the doubly normalized measure needs to be clarified. It is still too early to tell what the most useful approach will be, but one may note that the emphasis on the curvature shares the basic weakness of the PPM approach in dealing with oscillations.

Other researchers have looked directly at the dynamic decomposition weights (DDW), the \( w_i \)'s in

---

**Figure 1. Schematic view of eigenvalue and eigenvector analysis approach**

Copyright © 2008 John Wiley & Sons, Ltd.
(6 and 7). Thus, the focus is on the relative weight of the modes for a particular variable, from a policy criterion perspective, similar to Forrester’s original focus and the starting point for the EEA analysis (Gonçalves, 2006; Saleh et al., 2006, 2008).

Saleh et al. (2008) explore using the method for stabilization policy in two simple business cycle models. They explore the policy design space by assessing the influence of model parameters on \( \omega_i \) and identify leverage points by focusing on parameters that most affect the weights of the behaviour modes for the variable of interest. Their analysis highlights the fact that changes in parameters affect both the eigenvalues themselves and their relative presence in the behaviour of a given system variable. Hence, while it is still ‘early days’ as far as eigenvector-based methods are concerned, it is clear that tools for policy analysis will need to consider both eigenvectors and eigenvalues.

**AN EXAMPLE\(^1\)**

To illustrate briefly the significance of using formal analytical tools for theory building, we provide an example from a well-known model in the field: the simple long wave model, as originally published by Sterman (1985). The model has only three state variables (capital, capital supply line and order backlog), yet it is highly interconnected and its ILS contains 16 feedback loops (see Figure 2 for a stock and flow diagram and the SILS used in this analysis). The model shows surprisingly intricate dynamics and quickly settles into a limit cycle with a period of approximately 50 years (Figure 3 shows the behaviour of few key variables).

In the original paper, Sterman emphasized the role of capital self-ordering (Loop 16) in generating this economic long wave. Self-ordering occurs because orders for capital bloat the backlog, raising desired production and thus desired capital stock, leading to further orders from the capital-producing sector to itself. Furthermore, the model includes a ‘hoarding’ loop: capital orders are adjusted according to the variable delivery delay for capital. If the current delivery delay is longer, the capital sector orders more capital, which creates a positive feedback loop: as orders for capital go up, this bloats the backlog, lengthening the delivery delay, leading to still further increases in capital ordering (Loop 15). If one were to follow the intuitive guidelines suggested by Graham (1977), the presence of this feedback loop should exacerbate the cycle: increasing the amplitude and lengthening the period.

Performing loop eigenvalue elasticity analysis (LEEA), however, reveals a more subtle explanation for the observed dynamics. Figures 4 and 5 show, respectively, the behaviour of the three eigenvalues of the linearized system and the gains of selected loops from the SILS in Figure 2.

The strong nonlinear effects in the model are evident in large changes in both the eigenvalues and the loop gains of the system. During the initial few years of the upswing of the cycle (Phase I in Figures 3–5) and during the first few years of the downturn (Phase III), the behaviour is dominated by two large real positive eigenvalues of roughly the same magnitude (see Figure 4). During the later part of the upturn (Phase II), the two large real eigenvalues change into a complex conjugate pair, with a positive real part of roughly the same magnitude as before. Finally, during the long decay of capital (Phase IV), the dominant eigenvalue is the one corresponding to the decay time of the capital stock (labelled z3 in Figure 4).

The self-ordering loop is indeed involved in generating large fluctuations, but only during the relatively brief Phases I and III. At all other times, the loop is shut off by nonlinearities in the system, as seen in Figure 5. During Phase II capital orders continue to grow, as mirrored in the large real part of the complex eigenvalues, but driven by other positive loops than self-ordering. Phase IV constitutes a long decay of the capital stock once it exceeds its desired capacity, where behaviour must be dominated by the negative loop controlling capital decay.

---

\(^1\)The example is in large part a reproduction of Kampmann’s original analysis of the same model (Kampmann, 1996), but using the SILS method to construct the independent loop set. The model, as well as the computations to identify the SILS and perform the LEEA, is available for independent analysis and inspection at: http://iops.tamu.edu/faculty/roliva/research/sd/.
The analysis of the loop influence on the relevant behaviour modes reveals further insights. Figure 6 shows, for a point in time during each of the four phases of the cycle, a scatter plot of the absolute value and real part of the influence measures on the largest eigenvalue for a subset of the independent loops (loops with very small influence measures have been removed to simplify the figure). We find this representation to be useful in focusing attention on the loops most influential in the mode of interest (the largest absolute value on the x-axis) while informing about the direction of the loops' influence in the y-axis (a negative influence measure implies a stabilizing influence). The figure clearly reveals the significance of the self-order loop (16) during Phases I and III, and shows how it drops from view during the other two phases. Further, it reveals that during Phase II, which lasts approximately 8 years, the growth...
of capital orders is mainly determined by three positive loops: the ‘Capital Expansion’ loop (3) arises because, when orders are limited by the nonlinear function $g()$, the order rate is anchored in depreciation $d$ and, consequently, on the capital stock itself; the ‘Economic Growth’ loop (9) reflects the standard physical capital accumulation in a growing economy; and the ‘Delivery Delay’ loop (11) reflects the effect of the depletion of the backlog, thus accelerating the delivery time for capital orders and further increasing the capital available to fill orders. During Phase IV, the plot reveals that the Capital Decay loop (1) dominates the gradual fall in the capital stock.

In this manner, the analysis qualifies the role of self-ordering in the model: though it is significant for powering the swings of the cycle, it is surrounded by other feedback mechanisms and, one is tempted to say, paradoxically the self-ordering loop is absent from the model most of the time, a significant fact that was not revealed in Sterman’s original analysis.

As mentioned in the description of the model above, it contains a ‘hoarding’ loop that one might expect will further exacerbate the cycle. However, in his thorough analysis in the original paper, Sterman (1985) showed that the hoarding mechanism makes virtually no difference to the behaviour of the system. A less careful analyst might easily have made the false inference that the loop must be important but not thinking of testing the hypothesis explicitly. In the LEEA, by contrast, the result follows automatically, as evidenced in Figure 6, where the loop never has a significant influence measure. Furthermore, while Sterman simply notes the fact that hoarding does not affect behaviour in the model, the LEEA tool points to an intuitive explanation: much of the time, it is shut off by the nonlinearities that also deactivate the self-ordering loop, as seen in Figure 5. And during the brief periods when both loops are free to operate, the power of self-ordering is so strong that it overshadows any effect of the hoarding loop.

This brief example hopefully demonstrates how the formal tools can both provide a more rigorous analysis, catching errors of inference and deepening the understanding of exactly what parts of the structure lead to the observed behaviour.
As mentioned above, we cannot hope for a ‘grand unified theory’ that will automatically provide modellers with ‘the’ dominant structure. Given the analytical intractability of nonlinear high-order systems found in our field, the most we can hope for is a set of tools that will guide the analysis and aid the development of the modeller’s intuition.

This is not to say that formal methods should not be pursued. On the contrary, we have the view that any analysis tool must be based upon a solid mathematical foundation. Without a clear understanding of the mathematical meaning of the measures used, one is vulnerable to arbitrary interpretations. All of the tools described in this article would benefit from further mathematical understanding and development before one invests the effort of turning them into polished software packages. Understanding how and why the tools work the way they do is crucial. This will require further work in at least the following areas:

1. fundamental questions, such as the significance of the independent loop set and how loop elasticity measures depend upon this

Figure 6. Loop influence on eigenvalue 2 – long wave model loop influence of a loop gain g on eigenvalue λ is equal to $\frac{\delta \lambda}{\delta g} |g|$
choice, how the different methods are related, and the further theoretical development of parsimonious measures, such as linear filters,

(2) specific puzzles relating the ‘pathological cases’, such as ‘phantom loops’, ‘figure-8 loops’ (two negative loops that interact to effectively yield a positive loop), and non-distinct eigenvalues,

(3) technical challenges, relating to numerical computational efficiency, accuracy and stability in eigenvalue methods and

(4) a classification of the types of problems and models appropriate for each method, through extensive testing and comparison of the methods to an array of different models.

Only after these areas have been further explored will the time come to submit the methods for wider application for the ultimate test of their real-world utility by relatively user-friendly and polished software. Furthermore, to speed this process, we advocate the open-source philosophy that is currently dominant among those of us working on these methods, where code, models and documentation are made freely available on-line.

On the more creative side, it would be interesting to explore alternative forms of visualizing the various influence measures developed. For instance, one could imagine that links between variables in a model diagram ‘glow’ in different colours and intensities depending upon their effect on a behaviour pattern in question. This is not just a question of fancy user interfaces: the function of these tools will be as intuitive consistent aids to understanding, not analytical ‘answering machines’. In this light, the visualization is as important as the analytical principles behind it. Given the power of the human eye in finding patterns in visual data, this could be a significant next step.

REFERENCES


