IMPROVING CAUSAL MAPPING PRACTICE USING THE SYSTEM DYNAMICS 'FRONT-END' TOOL

Alan Charles McLUCAS

University College, University of New South Wales

School of Civil Engineering,

University College UNSW @ ADFA,

Australian Defence Force Academy,

Northcott Drive,

CAMPBELL ACT 2600

AUSTRALIA

+ 61 2 6268 8332 / +61 2 6268 8337

a.mclucas@adfa.edu.au

Abstract

Causal loop diagrams are graphical aids for depicting hypotheses about causality and feedback mechanisms. Though widely used, causal loop diagramming has significant limitations. Causal loop diagrams can contain numerous untested assumptions about causality.

This paper demonstrates how causal loop diagramming practice can be made more robust. The proffered technique described in this paper suggests how we might improve the ways we contemplate cause and effect. Applying the technique, and using the tool, offers new opportunities for testing assumptions about multi-factorial causal influences. The paper suggests that our first attempts at building remedial strategies for complex systemic problems might be completed without building quantitative stock and flow models. The paper demonstrates how causal loop diagrams can be kept free of errors in logic. It also raises a number of important issues regarding the ways we view combinations of potentially confounding causal influences, and identifies the need for further research in this area.

Introduction

The need for the tool described in this paper was identified during the conduct of research described in McLucas (2001). Further to the analysis of feedback mechanisms, the tool mentioned and techniques described are intended to help us investigate the combining causal influences to produce an effect at a point in a causal loop diagram, whether they are expressed as 'soft' variables or a mix of 'hard' and 'soft' variables. They are also used to test the logic of the causal loop diagram. The tool's graphical user interface is not described in this paper.

Prior to the development of the System Dynamics 'Front End' Tool, the depictions of cause and effect in cause maps, causal loop diagrams and influence diagrams were limited to simple linear relationships. The recent work of Kim (2000), and Kwahk and Kim (1999), which only accommodated linear causal relationships, has been taken a most significant step forward with the development of this tool. Systemic problems containing a wide range of continuous causal relationships can be subjected to 'first pass' analysis to give us insight into where we might apply our management effort. Relationships approximated by the SD 'Front End' Tool may be linear, sigmoid, exponential, hyperbolic, or polynomial with multiple extrema. The significant constraints of depicting causal linkages as '+' or '-', 's' or 'o'[1], or simply as linear increasing or decreasing in their causal influence, have been overcome.

This paper describes a tool [2] that is directed to enhancing SD practice and the understanding decision-makers have of about system dynamics, through:

a. exploiting knowledge of domain experts who may be available for a limited time, such as during meetings or workshops conducted for the purpose of addressing a particular problem;

b. providing a vehicle for fostering dialogue, discourse and critical analysis;

c. surfacing and testing assumptions that various stakeholders might have;

d. rapidly building an understanding of dynamic behaviour underlying current difficulties, or which may impact on the development of business strategies;

e. enabling first pass analysis of the dynamics when time may not be available to build quantitative system dynamics models;

f. determining whether more detailed analysis through system dynamics modelling and simulation, is needed;

g. undertaking preliminary analysis as the basis for creating a business case for the commitment of significant resources; and

h. rapid prototyping and first pass testing, in a limited time frame, of a set of

---

[1] These conventions for indicating polarity are commonly used in causal loop diagramming.

[2] The SD 'Front End' Tool mentioned in this paper is the subject of an Australian Innovation Patent.

strategies for urgent implementation.

Reliance on Causal Loop Diagrams in System Dynamics Practice

Despite known limitations of causal loop diagramming, described by Richardson (1976; 1985), it is very widely used in systems thinking and system dynamics modelling. As an indication of the extent of this reliance on causal loop diagramming, Sterman (2000) uses over a hundred diagrams that are either causal loop diagrams or have the attributes of causal loop or influence diagrams, to explain causality. The system thinking publications of Senge (1990), Senge *et al*. (1994), Kim (various) actively promote analysis based on causal loop diagrams.

Often, causal loop diagram construction relies on a great deal already being known about the problem at hand. If knowledge is limited or understanding is superficial, there can be significant risks to building meaningful diagrams.

Whilst extreme, one well-respected SD practitioner in a recent personal communication wrote: "I would maintain that I can draw any causal loop diagram I want and justify any conclusion I want with a perfectly straight face that would be bullet proof to anyone with less gile than me."

Particularly in the early stages of finding out about a problem situation, or when system dynamics experience is limited, hypotheses need to be formulated and tested in ways that quickly get to the core of the problem.

The tool and technique described in this paper go beyond simply achieving consensus about causality. They facilitate surfacing and testing of hypotheses embodied in the diagram.

To date, quantification of cognitive maps or causal loop diagrams has been limited to assigning simple polarity and strength of influence to the causal links. Whilst reasons for generally limiting relationships to linear ones are not clearly enunciated in the literature, the following reasons are suggested:

a. difficulty associated with setting initial values, particularly when causal linkages are non-linear;

b. calculating a polynomial fit of a curve to available data can involve significant computational effort;

c. confidence in curve fitting depends on how comprehensive the collected data sets are;

d. curves fitted to sparse data can be imprecise; and

e. conventional algebra provides limited opportunity to influence the fit of a particular curve through application of additional user defined rules such as … "the curve is asymptotic to a line $y = ax + b$ for values of y above y_1 ".

Getting the Most from Causal Loop Analysis

Proceeding prematurely to building system dynamics models without first investigating the nature and scope of the problem in detail, can be counterproductive (McLucas and Linard, 2000; McLucas, 2001). Equally, spending inordinate time and effort in problem conceptualisation can be wasteful.

The tool and techniques described here enable problem conceptualisation, extensive qualitative analysis and limited quantitative analysis. This preliminary analysis also helps reveal just which data sets need to be collected and to what degree of accuracy. Collection of data can commence even before the preliminary analysis is completed. This is important because data collection can be the rate determining activity. Indeed, lack of data or lack of data in an appropriate form can preclude quantitative system dynamics modelling (McLucas, 2001: 304, 312). Given that the required data can be collected concurrent with the early stages of model building, products from quantitative modelling can flow more quickly. This tool offers the opportunity to undertake the following:

a. rapid preliminary analysis of dynamic behaviour, noting that often the opportunity to undertake such analysis might be restricted to a workshop or meeting attended by those with specialist domain knowledge;

b. exploration of changes over time, previously possible only through the building of system dynamics models;

c. identification of pressure points to which management effort might be applied;

d. sensitivity analysis, that is, identification of significant sources of influence and the extent to which, singly or in combination, they produce change;

e. identification of the causes of shifts in feedback loop dominance; and

f. analysing the confounding correlation that can exist between a number of cause and effect relationships.

Limitations of 'SD' Front End Tool

It is not intended that the tool described here be a replacement for conventional system dynamics modelling and simulation. It fills an identified gap between qualitative and quantitative analysis and serves as a precursor, and informant to the design and conduct of, quantitative modelling and simulation activities.

In its current form, the tool does not have facility to handle delayed causality on a single link, or total delay in a feedback loop. Whilst planned enhancements, McLucas (2001), will overcome particular limitations up to a point, enhancements to provide for time-series inputs, such as sinusoidal waves, are not planned on the basis that models requiring this level of sophistication are best-built using conventional system dynamics modelling.

Continuous Functions Only

It is assumed that each function linking cause and effect is continuous. Where functions might be discontinuous, quantitative SDM would be required.

Even Experienced Modellers Easily Misled by Causal Loop Diagrams

An important driver for the development of the SD 'Front End' tool is Sterman's statement (Richardson, 1985: 158) that even experienced modellers are easily misled by causal loop diagrams. If experienced modellers are easily misled, then inexperienced ones are at significant risk of making errors of interpretation during problem conceptualisation and preliminary analysis.

Despite their limitations, it seems most unlikely, that causal loop diagrams will ever be replaced in any widespread sense by the more logically robust influence diagrams, such as are advocated by Coyle (1996).

Improved Causal Loop Analysis

If causal loop diagrams can be given greater utility to support the important activities of surfacing and testing assumptions about what underlies the observed reference modes of behaviour, system dynamics practice could be significantly improved. The need for the SD ‘Front End’ tool derives from concerns raised by Coyle (2000), Ford and Sterman (1998), and Nuthmann (1994).

A Tutorial in Causal Analysis – Availability and Use of Illicit Drugs

The supply and use of illicit drugs creates innumerable, undesirable, side effects for society. These include a drain on medical resources for treatment and rehabilitation, burglaries and a black market in stolen goods. The product of a workshop to consider the ways of managing drug-related issues might look like Figure 1:

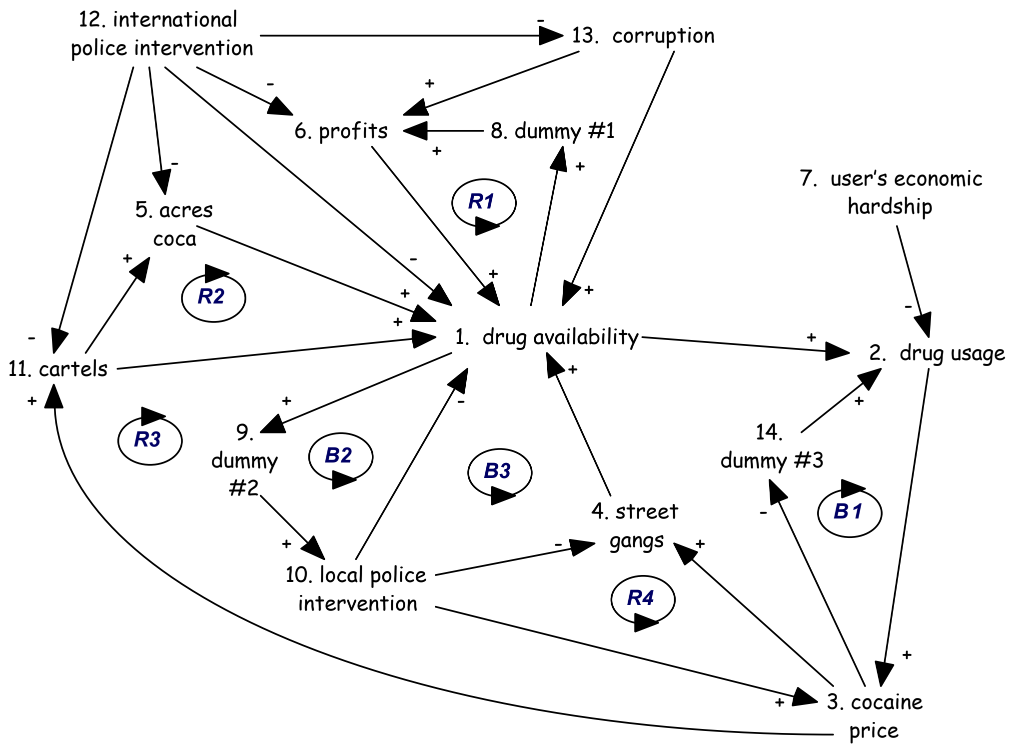

Figure 1: Typical Problem: Illicit Drug Usage [3] [4] [5]

---

[3] This causal loop diagram is the product of research by Taber (1991: 83-87).

[4] Dummy nodes 8, 9, and 14 have been added to remove bi-directional arrows, which Taber used in his original version of the diagram. This has no effect on the logic of the diagram and serves to make it easier to identify individual feedback loops. This also avoids double arrows, which might have positive polarity in one direction and negative polarity in the other, found confusing by some. For the software development team, it also made the sequence of

Making Estimates of Causality

Kwahk and Kim (1999) use 'causal impact' questionnaires to establish whether the relationship between cause and effect is increasing or decreasing, very strong, strong, or weak. They apply simple weightings as a result. Kosko (1993) uses a similar technique. These approaches are limiting, if for no other reason than they necessarily suggest linear causal relationships. Non-linear relationships are much more likely to occur: this is a persistent observation over nearly 40 years of system dynamics practice.

Consider the following example of the relationship between 1. drug availability and 2. drug usage (usage of illicit drugs by existing users). It is unlikely that hard data about either drug availability or drug usage or the relationship between the two, would exist.

However, educated guesses may be made. The following are likely to be known, or could be established.

a. The number of drug busts affected by police and drug agencies.

b. The quantities of drugs seized during raids.

c. The number of drug overdoses attended by paramedics or handled by hospital emergency medical staff.

d. The number of deaths attributed to drug overdoses.

e. Size of the general population.

f. Estimated size of the drug-using population. For example, 1 in 1,000 of the general population may be estimated to be users of illicit drugs such as cocaine.

The following might be reasonably deduced, or deduced from available information sources:

a. Zero supply equates to zero usage: the 1. drug availabilty / 2. drug usage curve passes through the origin.

b. Even when availability reaches glut proportions, only so many people will use drugs. That is, the curve is asymptotic to a vertical line depicting the estimated maximum size of the drug-using population. This number might be taken as the totality of the estimated population, plus an arbitrary percentage, say 10%, for growth during the period over which this study is to be conducted. In a city of 300,000, it might be estimated that there is a maximum of 330 cocaine users (one in 1,000 plus 10%).

c. It is estimated that the maximum 330 users consume y_1 grams of cocaine per year. Note that we are not interested in absolute numbers, *per se*, but the estimates are needed to enable normalisation of scales on the axes of each graph. Maximum values on each axis are normalised to unity. Rationale

---

calculations easier to establish. The familiar causal loop diagramming convention of using single arrows is, thus, retained.

[5] Kim (2000), Kosko (1993; 1997) and Kwahk and Kim (1999) all assume cause and effect relationships are linear. In contrast, non-linear cause and effect relationships are assigned to the illicit drug problem described in this paper.

behind estimations made and normalisation calculations must be recorded for future reference.

Such reasoning enables the construction of basic relationship diagrams, such as Figure 2, below. A scattergram of estimates, combined with known data, might appear as shown.

From such a graph, we are seeking to determine for each input, what the appropriate output value might be, noting that this may change as a result of each iterative calculation, as occurs when simulating changes over time. Ford and Sterman (1998) describe, then demonstrate the application of a methodology, in a group setting, for aiding experts in the process of explicating their tacit knowledge. This is directly applicable to the process, described above, of estimating and formalising causal relationships for further analysis using the SD 'Front End' Tool.

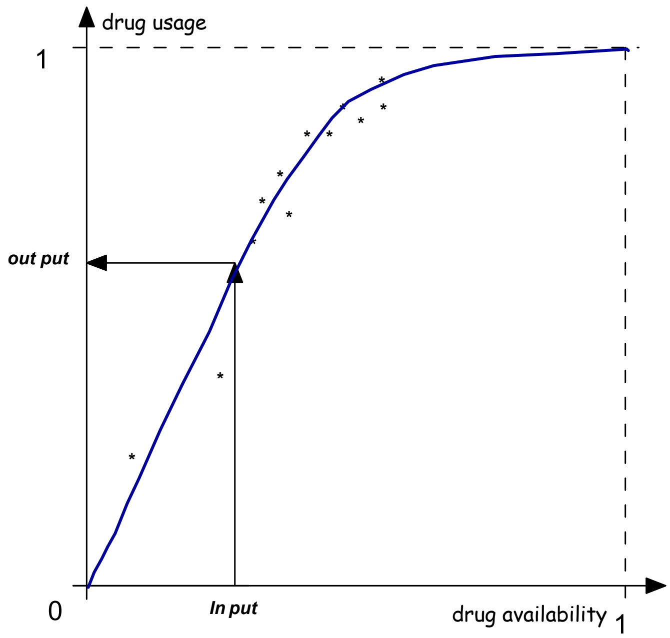

Figure 2: Basic Relationship Diagram

Intialisation of the Causal Loop Diagram – A Critical Step

Initialisation of the diagram is exceedingly important:

a. Estimates made of nodal values are critical determinants of the viability of the diagram. If these estimates are erroneous then initialisation will be difficult to achieve, or it will be meaningless.

b. Careful consideration is needed in the creation of the causal relationship graphs. Making such estimates demands considerable rigour and discipline. Insight into how causal relationships produce observed modes of behaviour comes with experience in system dynamics. This is an unfortunate reality for the novice who must take extra care in formulating hypotheses about the shape of the graphs and critically examining how those relationships contribute to the observed behaviour.

c. Concurrent adjustments of estimates of nodal values and shape of causal relationship graphs may be needed. It is essential that adjustments be made through an iterative process. The aim is to create revisions resulting in nodal values and causal relationship curves that are consistent, that is, influences combining to produce the respective nodal state values.

Initialisation of the Model – Iterative Re-alignment of Estimates Made and Mental Models

The initialisation sequence is not intuitively obvious, and it is stressed that when using the software tool, the user is guided through the necessary steps and prompted for response as needed. As initialisation of the model proceeds, there will be cycles of estimating, validating and adjusting. This will produce change to both the estimates made to initialise the model, and the mental models of those making the estimates. It may be necessary to make changes to the diagram to correct illogical structures. Illogical structures, such as logical omissions or incorrect polarity are frequently detected during initialisation.

The process of initialisation can lead to confusion over which adjustments should be made first, adjustments to the estimates of nodal values or the form of causal relationships. Estimates of nodal values should be among the first products of group modelling activities, nominal group technique or Delphi. Achieving agreement about these must be the first priority. Estimating the form of the causal relationships should be a secondary activity. Estimates of the shape of the causal relationships may change numerous times during initialisation, and during the subsequent analysis.

Setting Initial Values – Making Estimates of Nodal Values

The first step is to establish estimates of the state values at each node. Values are recorded along with explanations of why particular values were chosen. See Table 1:

Node No. and Node	Value [6]	Basis of Determining Estimated Value
1. drug availability	0.5	Median value between best and worst observed over the past three years.
2. drug usage	0.6	Based on records of recent treatment of overdoses by hospitals and paramedics.
3. cocaine price	0.7	Based on estimate provided by Police drug squad.
4. street gangs	0.5	Based on estimate provided by Police drug squad.
5. acres coca	0.4	Drug Enforcement Agency (DEA) estimate provided through local Police.
6. profits	0.6	Nominal group estimate.
7. user economic hardship	0.3	Based on estimates provided by Health and Social Security.
8. dummy	0.0	Not applicable.
9. dummy	0.0	Not applicable.
10. local police interdiction	0.4	Local Police activity reports.
11. cartels	0.6	DEA estimate provided through local Police.
12. international police intervention	0.45	Nominal group estimate, supported by DEA reports.
13. corruption	0.3	Nominal group estimate, supported by DEA reports.
14. dummy	0.00	Not applicable.

Table 1: Estimated Values at Each Node

Fundamentally, initialisation is a guided trial and error activity, following the scheme described below under the heading 'Sequence of Calculation', the results of which are at Tables 2 and 3, below. The influences from each of the causal relationship curves are summed to produce calculated nodal values. These are compared with the estimates at Table 1. The causal relationships for every causal link must be estimated, following the process described above. These causal relationships appear as at Figure 3, below.

[6] Values are normalised on the basis of maximum expected values: see 'Making Estimates of Causality'. Maximum expected, or worst-case value is taken as 1.0.

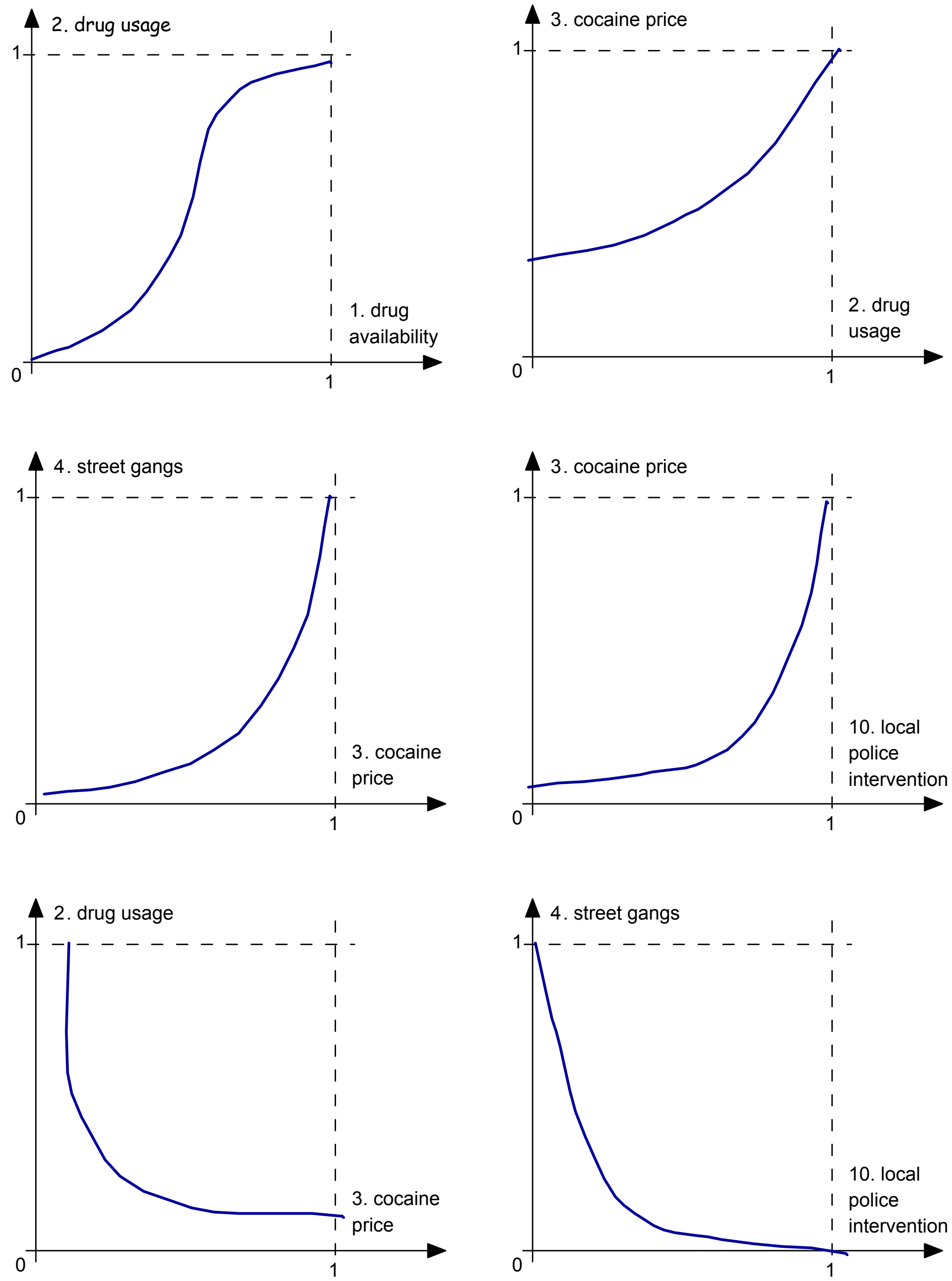

Figure 3: Sample Causal Relationship Curves for the Illicit Drug Use Problem [7]

These estimates become benchmarks for initialisation of the causal loop diagram wherein it is assumed that influences at a node are all taken into account by simple addition. This is an important, and possibly erroneous assumption that requires further investigation (Coyle, 1999: 2000, McLucas, 2001; Nuthmann, 1994). Depending upon results of future research, it is expected that a range of alternate combinatorial algorithms will be required.

---

[7] Values on each axis are normalised. 1.0 corresponds to worst case or maximum observed historical value

Discipline Needed in Defining Causal Relationships

Estimating causal relationships may be done in real time, such as during a workshop, defining causal relationships requires considerable discipline. In the absence of hard data, employing a nominal group, questionnaires as used by Kwahk and Kim (1999) or workbooks used by Vennix (1996: 114), or the methodology described by Ford and Sterman (1998) in group model building, are suggested as alternatives most likely to produce reliable estimates. Rather than encouraging reliance on personal judgement and intuition in making estimates, the SD 'Front End' Tool forces assumptions about the form or each and every causal relationship, be made explicit. This is important as the Illicit Drugs example evinces. In this case, there are 23 causal relationships, and the number can increase rapidly as diagrams become more complex.

The influences produced by these causal relationships must combine at the node to produce the values estimated at Table 1. If they do not, each estimated nodal value and each causal relationship must be critically analysed. Only when the inputs to each node sum to the estimated values, can the calculations be made for each loop in the diagram. Whilst not included in the SD 'Front End' Tool at this time, a scheme, which employs genetic algorithms to automate the calculation process, is proposed. The sequence of calculations is described below.

Creating Correct Sequence of Calculations

The original causal loop diagram, Figure 1, is prepared showing only the nodal identifier numbers and the various feedback loops, and redrawn. See Figure 4, below.

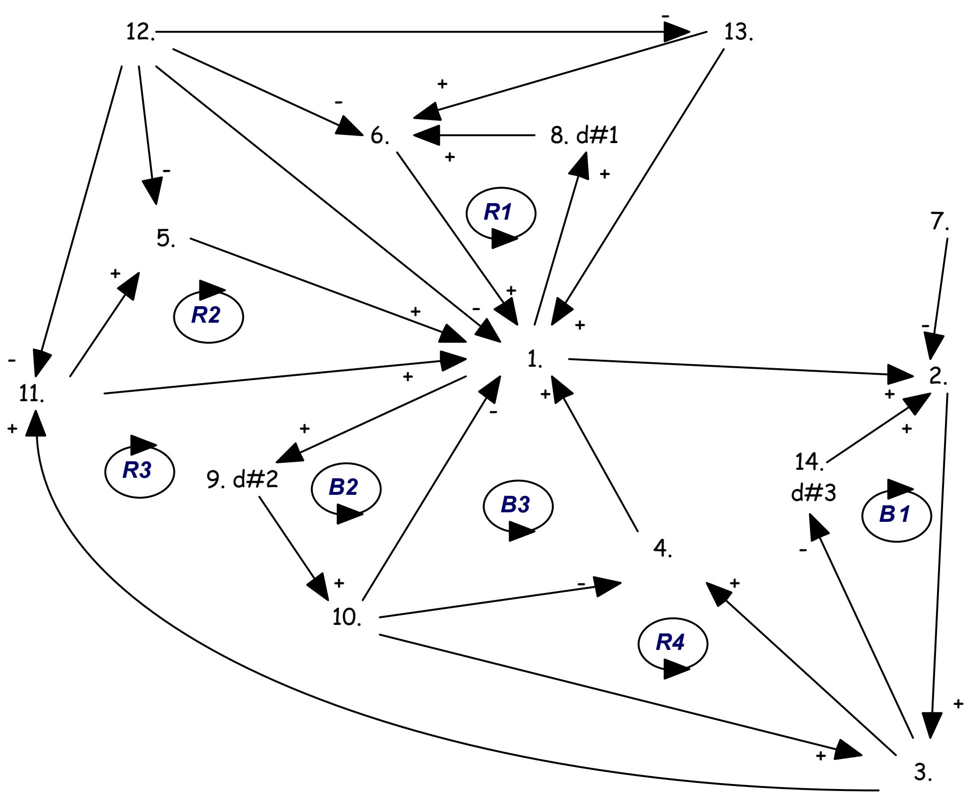

Figure 4: Drugs Causal Loop Diagram Showing IDs and Feedback Loops

Polarity is shown as '+' or '-' as appropriate, and only used as a reminder of the general shape of the curve. All inputs to a node are added: polarity does <u>not</u> suggest addition or subtraction of influence at the node concerned. In the absence of a sign, polarity is assumed to be positive. The *assumption that influences at a node are additive* will be discussed at the end of the paper.

Where feedback loops are involved, the order in which calculations are made is critical:

a. A focal node is selected. This node becomes the start and finish for each cycle of calculations. We might also consider selecting this node on the basis of its importance as an indicator of the effectiveness of any strategy we might develop. In this case, the focal node chosen is 1. drug availability. Any node can be the focal node. Whilst points of interest may change during analysis, the focal node remains as initially defined. However, a specific node is nominated to ensure the logic of setting up the order of calculation is correctly established and maintained, noting that the products of one cycle of calculations will need to be stored and used as the starting values for the next iteration.

b. Identify the longest feedback path, that is, identify the longest (or equal longest) path, which leads back to the start. It must be possible to track from beginning to end of this path without traversing any link more than once. If there is an equally long path, from this point on, it is treated as a subordinate feedback loop.

A clockwise convention is followed for the main (longest) loop. This may require the diagram be re-drawn. Redrawing of subordinate loops is a matter of personal choice. If those loops are short, redrawing should be unnecessary. The longest feedback loop is as depicted at Figure 5, with only the direct inputs shown.

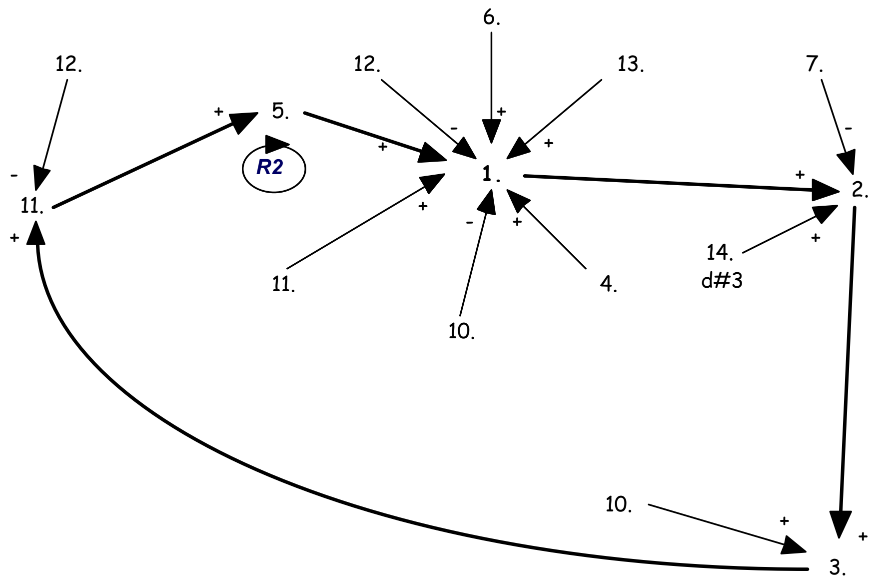

Figure 5: Longest Feedback Loop Identified

Subordinate loops are depicted at Figure 6, below.

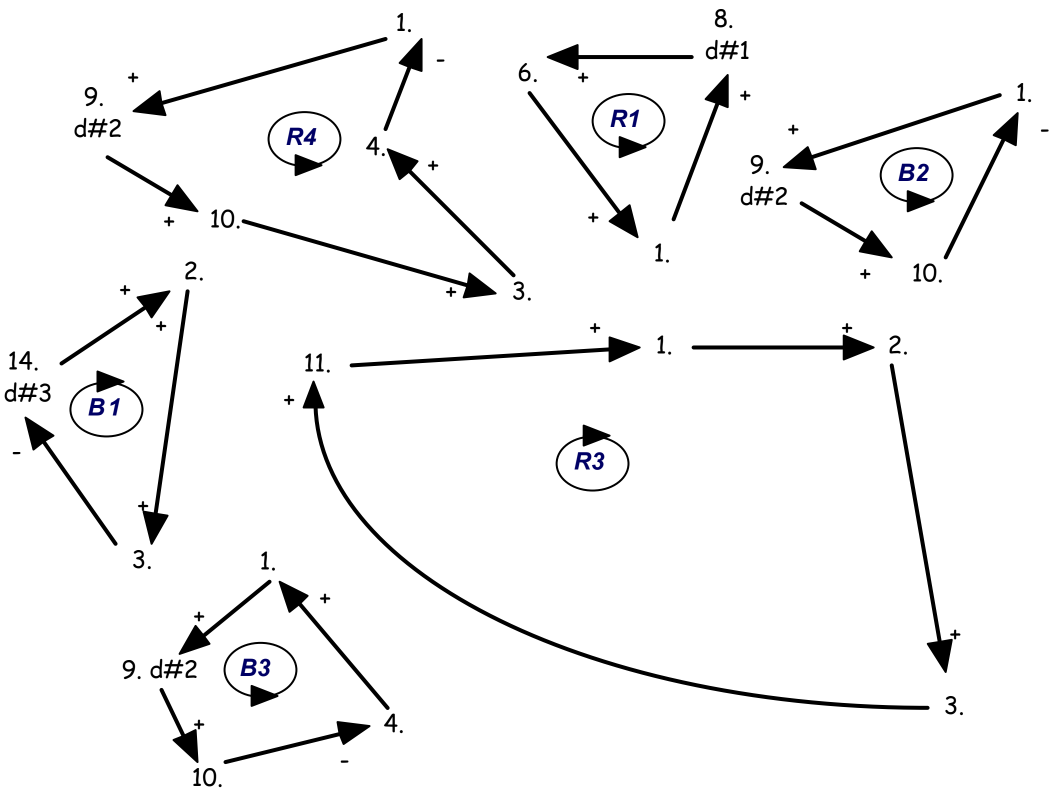

Figure 6: Ancillary Feedback Loops

Sequence of Calculations

The same sequence of calculations must be followed both for initialisation and simulation. Calculations are made in the specified sequence to ensure all influences are included and there is no double counting. When calculations for side loops are carried out, the value to be input to the main loop is held until the next iteration. This sequence is shown Table 2, where *node (iteration)* depicts the nodal value calculated during a particular iteration, and δ *(iteration)* $n_{1,}n_{2}$ depicts the influence (1) of 1 on 2. In the special case of initialisation, adjustments are made until the value calculated for node 1 at the end of the first iteration 1(1) is equal to the initial estimate, 1(0).

As the main loop is navigated, only influences originating from 'upstream' nodes are included in the calculations. For example, when at node 3, upstream nodes having an influence are nodes 1, 2 and 7. Node 7 is not connected, except via node 2 and its influence has already been counted. However, node 1 is upstream from node 10, which influences node 3. So 1(0) is used to calculate 10(1) which, in turn, is used to calculate the influence (1) of node 10 on node 3. Influences originating downstream from the node where current calculations terminate must be held for the next iteration.

Inputs	Combining of Influences	Outputs [8]
1(0) ⇒ 7(0) ⇒ 3(0) ⇒	$+\delta(0)_{1,2}$ – influence (0) of 1 on 2 $+\delta(0)_{7,2}$ – influence (0) of 7 on 2. $+\delta(0)_{3,2}$ – influence (0) of 3 on 2	⇒ 2(1)
2(1) ⇒	$+\delta(1)_{2,3}$ – influence (1) of 2 on 3 $+\delta(1)_{10,3}$ – influence (1) of 10 on 3	⇒ 3(1)
10(1) ⇒	$+\delta(1)_{10,4}$ – influence (1) of 10 on 4 $+\delta(1)_{3,4}$ – influence (1) of 3 on 4	⇒ 4(1)
3(1) ⇒ 12(0) ⇒	$+\delta(1)_{3,11}$ – influence (1) of 3 on 11 $+\delta(1)_{12,11}$ – influence (1) of 12 on 11	⇒ 11(1)
12(0) ⇒	$+\delta(0)_{12,5}$ – influence (0) of 12 on 5 $+\delta(0)_{11,5}$ – influence (0) of 11 on 5	⇒ 5(1)
12(0) ⇒	$+\delta(0)_{12,13}$ – influence (0) of 12 on 13	⇒ 13(1)
12(0) ⇒ 13(1) ⇒	$+\delta(0)_{12,6}$ – influence (0) of 12 on 6 $+\delta(1)_{13,6}$ – influence (1) of 13 on 6 $+\delta(1)_{1,6}$ – influence (1) of 1 on 6	⇒ 6(1)
12(0) ⇒ 11(1) ⇒ 6(1) ⇒ 10(1) ⇒ 4(1) ⇒	$+\delta(0)_{12,1}$ – influence (0) of 12 on 1 $+\delta(1)_{13,1}$ – influence (1) of 13 on 1 $+\delta(1)_{6,1}$ – influence (1) of 6 on 1 $+\delta(1)_{10,1}$ – influence (1) of 10 on 1 $+\delta(1)_{4,1}$ – influence (1) of 4 on 1	⇒ 1(1)

Table 2: Sequence of Calculations

[8] The output from one stage becomes in the input to the next, and subsequent, stages of the calculations. The last output in this column becomes the starting value for the next iteration. In the software tool, the sequence of calculations is established by a series of questions posed to the user via dialog boxes. There is no need for the user to establish the sequence of calculations *ab initio*.

Initialisation calculations from Table 2, in the form of an Excel ™ spreadsheet are shown at Tables 3 and 4, below.

Input		Influences		Output		Comment
1(0)	0.5	(0) 1 on 2	0.4			Drug Availability at start of iteration. Input permitted.
7(0)	0.3	(0) 7 on 2	0.1			Input permitted
3(0)	0.7	(0) 3 on 2	0.1	2(1)	0.6	
1(0)	0.5	(0)1 on #2 on 10	0.4	10(1)	0.4	
2(1)	0.6	(1) 2 on 3	0.6			
	0.4	(1) 10 on 3	0.1	3(1)	0.7	
10(1)	0.4	(1) 10 on 4	0.2			
3(1)	0.7	(1) 3 on 4	0.3	4(1)	0.5	
3(1)	0.7	(1) 3 on 11	0.4			
12(0)	0.45	(0) 12 on 11	0.2	11(1)	0.6	Input permitted
12(0)	0.45	(0) 12 on 5	0.2			
		(0) 11 on 5	0.2	5(1)	0.4	
12(0)	0.45	(0) 12 on 13	0.3	13(1)	0.3	
12(0)	0.45	(0) 12 on 6	0.25			
13(1)	0.3	(1) 13 on 6	0.15			
		(1) 1 on 6	0.15	6(1)	0.55	
12(0)	0.45	(0) 12 on 1	0.05			
11(1)	0.6	(1) 13 on 1	0.1			
(6)1	0.55	(1) 6 on 1	0.15			
10(1)	0.4	(1) 10 on 1	0.1			
4(1)	0.5	(1) 4 on 1	0.1	1(1)	0.5	Drug Availability at end of iteration.

Table 3: Initialisation Calculations Summary

Link 1-2 (input 1. Drug Availability)	0.5	1	2	3	4	5	6	7	8	9
Lower Limit		0	0.24	0.34	0.42	0.46	0.52	0.58	0.62	0.7
Upper Limit		0.239	0.339	0.419	0.459	0.519	0.579	0.619	0.699	0.999
Value Returned		0.06	0.16	0.25	0.31	0.4	0.5	0.6	0.7	0.92
Link 1-2 Influence Calculated	0.4	0	0	0	0	0.4	0	0	0	0
Link 7-2 (input 7. User Ec Hardship)	0.3	1	2	3	4	5	6	7	8	9
Lower Limit		0	0.24	0.34	0.42	0.48	0.52	0.58	0.63	0.7
Upper Limit		0.239	0.339	0.419	0.479	0.519	0.579	0.629	0.699	0.999
Value Returned		0.05	0.1	0.14	0.18	0.22	0.27	0.31	0.38	0.6
Link 7-2 Influence Calculated	0.1	0	0.1	0	0	0	0	0	0	0
Links 3-14-2 (input 3. Cocaine Price)	0.7	1	2	3	4	5	6	7	8	9
Lower Limit		0	0.12	0.14	0.23	0.32	0.38	0.48	0.52	0.62
Upper Limit		0.119	0.139	0.229	0.319	0.379	0.479	0.519	0.619	0.999
Value Returned		0.7	0.54	0.42	0.34	0.26	0.2	0.16	0.12	0.1
Links 3-14-2 Influence Calculated	0.1	0	0	0	0	0	0	0	0	0.1
Link 2-3 (input 2. Drug Usage)	0.6	1	2	3	4	5	6	7	8	9
Lower Limit		0	0.24	0.34	0.42	0.46	0.52	0.58	0.62	0.7
Upper Limit		0.239	0.339	0.419	0.459	0.519	0.579	0.619	0.699	0.999
Value Returned		0.34	0.4	0.44	0.5	0.54	0.57	0.6	0.66	0.78
Link 2-3 Influence Calculated	0.6	0	0	0	0	0	0	0.6	0	0

Table 4: Initialisation Calculations - Selected Causal Relationships [9]

The Initialised Illicit Drugs Diagram

The initialised illicit drugs diagram is at Figure 7, below.

[9] Columns marked ‘1’ through ‘9’ contain the input values, within selected input bands – between ‘Lower Limit’ and ‘Upper Limit’ to give the ‘Value Returned’, that is the corresponding output values on each of the selected curves. This is shown as the ‘Link *X-X* Influence Calculated’ for the selected curve.

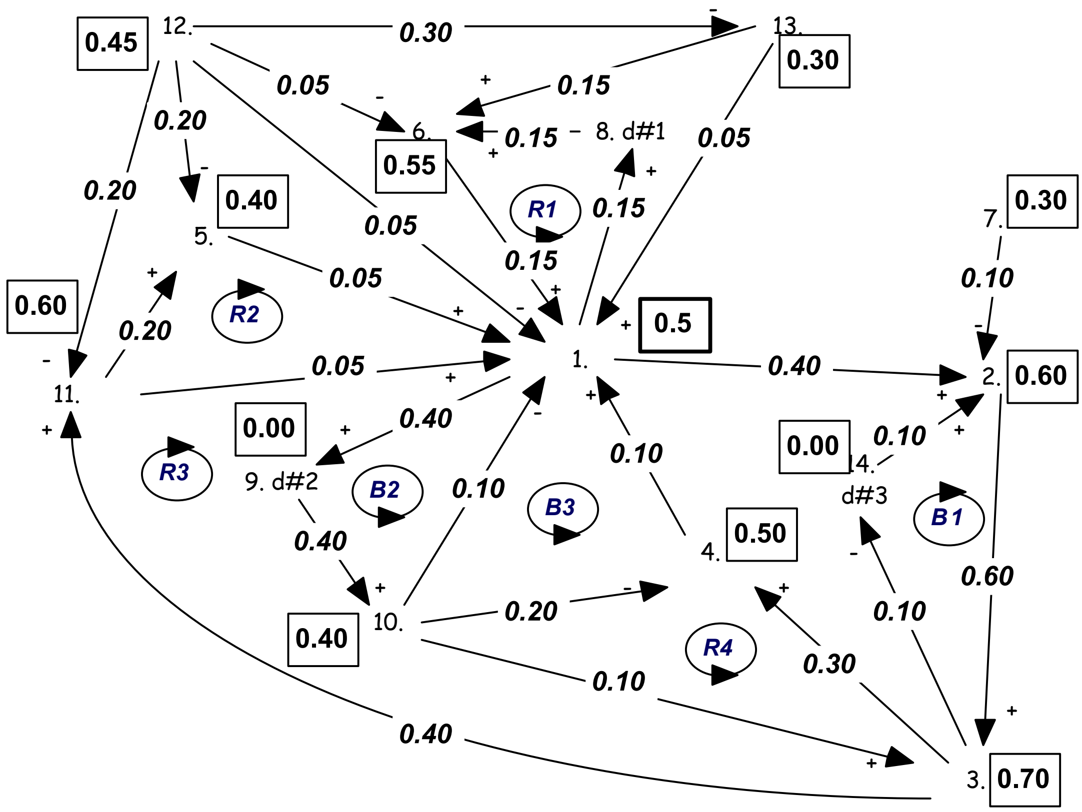

Figure 7: Initialised Illicit Drugs Causal Loop Diagram [10] [11]

Stepping Through the Simulation

The simulation to combat availability of drugs is run step-by-step. Once the diagram has been initialised, a scenario is developed to reflect desired remedial strategy, for example:

a. Increase international police intervention progressively, in a series of steps, from 0.45 to 0.60.

b. Making a one-off change to the shape of the relationship 1⇒10 which reflects increasing local police surveillance and intelligence gathering capabilities.

An example of the calculations is at Table 5, below, and graph of change of 1. drug availability as a result is at Figure 8.

[10] Initial estimates of nodal values, taken from Table 1, are shown in bold with rectangular borders.

[11] Initial estimates of causal relationships are shown in bold italics.

Input		Influences		Output		Comment
1(0)	0.47	(0) 1 on 2	0.4			Drug Availability at start of iteration. Input permitted.
7(0)	0.3	(0) 7 on 2	0.1			Input permitted
3(0)	0.7	(0) 3 on 2	0.1	2(1)	0.6	
1(0)	0.47	(0)1 on #2 on 10	0.4	10(1)	0.4	
		(0)1 on #2 on 10	0.49	10(1)	0.49	
1=yes; 0=no	1	(0)1 on #2 on 10	0.49	10(1)	0.49	Increased surveillance capability. Input permitted
2(1)	0.6	(1) 2 on 3	0.6			
	0.49	(1) 10 on 3	0.13	3(1)	0.73	
10(1)	0.49	(1) 10 on 4	0.1			
3(1)	0.73	(1) 3 on 4	0.3	4(1)	0.4	
3(1)	0.73	(1) 3 on 11	0.57			
12(0)	0.6	(0) 12 on 11	0.16	11(1)	0.73	Input permitted
12(0)	0.6	(0) 12 on 5	0.18			
		(0) 11 on 5	0.44	5(1)	0.62	
12(0)	0.6	(0) 12 on 13	0.25	13(1)	0.25	
12(0)	0.6	(0) 12 on 6	0.21			
13(1)	0.25	(1) 13 on 6	0.15			
		(1) 1 on 6	0.1	6(1)	0.46	
12(0)	0.6	(0) 12 on 1	0.04			
11(1)	0.73	(1) 13 on 1	0.1			
(6)1	0.46	(1) 6 on 1	0.11			
10(1)	0.49	(1) 10 on 1	0.1			
4(1)	0.4	(1) 4 on 1	0.08	1(1)	0.43	Drug Availability at end of iteration.

Table 5: Simulation Calculations

At line 6 in Table 5, there is a user selectable input for the increased level of local police surveillance and intelligence gathering capability. This is selected at the start of the third iteration, or simulation step. See Figure 8.

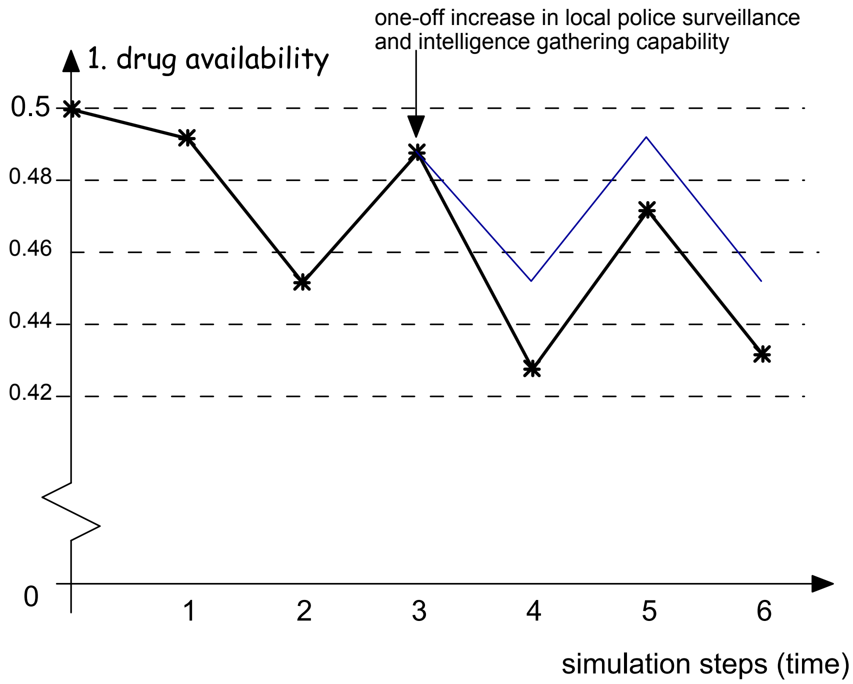

Figure 8: Simulated Impact of Strategies to Combat Illicit Drug Availability

Analysis of Simulation Results

International Police intervention produced an immediate reduction of drug availability from 0.5 to a level, which oscillated between 0.45 and 0.49. When local Police interdiction increased as a result of greater awareness of drugs on the street, achieved through increased surveillance and intelligence gathering capability, 1. drug availability dropped further. The light line in Figure 8 shows the level of 1. drug availability created by international police intervention alone. Combined impact of 12. international police intervention and 10. local police interdiction is shown at simulation steps 4, 5, and 6.

Sensitivity analysis revealed:

a. No contribution to the reduction of 1. drug availability for values of 7. user economic hardship less than 0.30.

b. Levels of 12. international police intervention, alone, above 0.60 had no effect.

c. Increasing 10. local police interdiction through increased surveillance and intelligence gathering capabilities, which changed the shape of the cause and effect curve as shown at Figure 9, had the effect of creating greater levels of police activity for given levels of 1. drug availability.

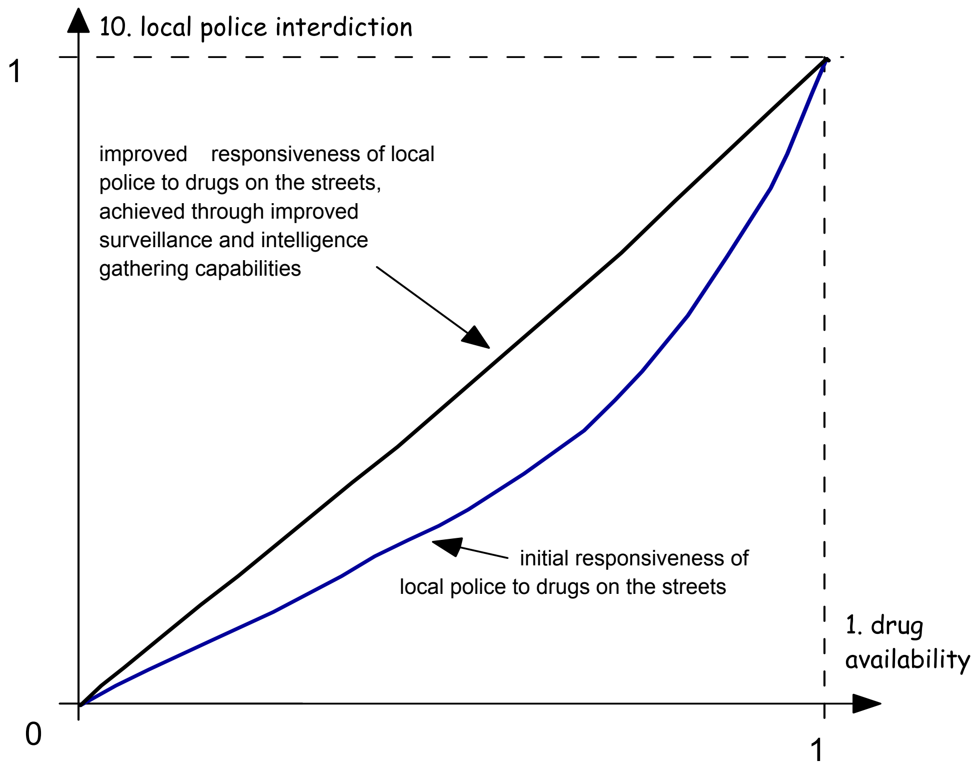

Figure 9: Improved Local Police Interdiction Capability

SD 'Front End' Tool in Brief - a Critique

The SD 'Front End' tool was designed to be used in conjunction with Iterative and Interactive Strategy Development (IISD) (McLucas, 1998), to provide support to rapid analysis of complex systemic problems, when the opportunity presents to bring together a number of people with domain expertise for a limited period. It will work in any situation where causality is being investigated. The main threat to achieving results described here is initialisation of the causal loop diagram. This is considerably more difficult than might appear at first, for the following reasons:

a. views about causality, such as strength of cause and effect relationships, can vary significantly between individuals;

a. this process is computationally demanding, certainly if done manually – although, as suggested earlier, these calculations be accelerated by employing a scheme which, or example, uses genetic algorithms for calculation of a 'workable' set of initial values;

b. it assumes that influences combine at a node by simple addition – situations can arise where such addition gives a value at a node, which is greater than the state value estimated – to correct this requires:

(1) artificially constraining the sum of influences at a node [12];

(2) re-defining and re-drawing of causal relationships - which happens in routinely in quantitative SD, and is part of the iterative SD process, or

(3) use of alternative combinatorial algorithms – the need for this suggests that the assumption that influences are combined by simple addition is questionable.

Success is critically dependent upon on the selection of initial nodal values.

Reliably estimating the causal relationships involves considerably greater risk for two reasons:

a. potentially, there can be large numbers of causal relationships involved; and

b. we are not very good at attributing observe effects to underlying causes.

Poor choice in defining cause and effect relationships can mean that it becomes exceedingly difficult, or impossible, to initialise the diagram. Whilst this might be interpreted as a potential weakness in the design of the tool, in reality it reveals a major strength: the diagram cannot be initialised unless two key criteria are met:

a. the diagram is logical, and

b. realistic estimates of cause and effect are made.

This need for estimates of cause and effect compatible with estimates of nodal state values, forces repeated creation and testing of hypotheses about the cause and effect relationships. It also forces mental models regarding cause and effect to be surfaced and tested. In itself, this is a highly important outcome.

The SD 'Front End' tool is <u>not</u> intended to be a replacement for quantitative system dynamics modelling and it should not be used to simulate more than a few iterations. Practical limits to its usefulness, as a simulation tool, have not been established, as yet. At this stage, it is more valuable as a research tool than a practical tool for use in a workshop or consulting situation.

Overcoming Current Limitations of the SD 'Front End' Tool – Calculating Delayed Feedback

To create a delayed feedback loop where accumulation occurs, the influence value is simply held for the appropriate number of iterations. This involves the addition of a dummy loop, which includes a node where state values are held temporarily. Graphical depictions of each of the relationships for each link are created as described in this paper. Initial calculations are conducted as normal, but when simulations are conducted the influence value is held until required, a nominated number of iterations. The calculated output values are held in a temporary register until needed for the appropriate iteration.

[12] From a computational viewpoint, this is readily overcome by applying weightings to influences at a node; the sum of weightings being 1.0.

Incorporation of a scheme, which used genetic algorithms to help in determining initial values, is proposed. Incorporation of alternative combinatorial schemes, which themselves require detailed testing, is proposed.

Summary

In this paper, the rationale behind a tool for analysing causal loop diagrams was demonstrated. It was shown that it is possible to analyse causal loop diagrams and produce our 'first pass' remedial strategies for complex, dynamic problems, as a result. The key limitation in the use of this tool lies with our own limited ability to make reliable, meaningful estimates of non-linear cause and effect. Using this tool:

a. forces the enunciation of hypotheses regarding the nature of causal relationships,

b. does not artificially constrain causality to linear relationships,

c. supports the testing of hypotheses regarding causality,

d. enables analysis of the veracity of the assumption (Coyle, 2000; Nuthman 1994) that causal influences at a node can be added, and

e. will support future analysis of the confounding effect produced by the cross correlation between causes and effects.

Whilst further development is required, this tool promises to help bridge the gap between qualitative and quantitative analysis of dynamic, systemic, complex problems. An unprecedented level of rigour has been added to the creation, analysis and interpretation of causal loop diagrams. This has the potential to change the way problem conceptualisation and first pass problem analysis, are conducted. It enables application of a level of rigour absent in such forms of influence diagramming, until now, with the exception of the form of influence diagramming espoused by Coyle (1996).

This paper demonstrated that direct simulation of causal loop diagrams involving non-linear relationships is possible. Just how valid and practical this is requires further investigation. Whilst simulation capability is strictly limited, it is seen to be useful in helping to identify where to apply resources and management effort.

The value of using this tool and techniques lies in the removal of primary reliance on judgment and intuition in the analysis of causal loop diagrams. The result is capability to identify where to direct remedial strategies, the form of strategies that might be employed, and the first pass testing of those strategies through rudimentary simulation.

This tool also provides a vehicle for testing assumptions regarding the additive combining of causal influences in a causal loop diagram.

<u>References</u>:

Coyle, R.G. 1996, 'System Dynamics Modelling: A Practical Approach', Chapman and Hall, London.

Coyle, R. G. 2000, 'Qualitative and quantitative modelling in system dynamics: some research questions', in: *System Dynamics Review,* vol. 16, no. 3, (Fall) 2000, System Dynamics Society, Wiley.

Ford D. N. and Sterman J. D., 1998, 'Expert knowledge elicitation to improve formal and mental models', in: System Dynamics Review, Vol 14, No 4, Winter 1998: 309-340.

Kim, D.H. 1990, 'Reinforcing and balancing loops: Building blocks of dynamic systems', in: *The Systems Thinker*, vol. 1, no. 1: 3.

Kim, D.-H. 2000, 'A simulation method for cognitive maps', in: *Proceedings of International Conference of Systems Thinking in Management Conference,* Deakin University, Australia, Nov 2000.

Kosko, B. 1993, 'Fuzzy Thinking: The New Science of Fuzzy Logic', Harper Collins, London.

Kwahk, K.-Y. and Kim, Y.-G. 1998, 'Supporting business process redesign using cognitive maps', in: *Decision Support Systems* 25, 1999: 155-178, Elsevier.

McLucas, A.C. 1998, 'Integrating soft and hard systems analysis: Seeking a practical framework for addressing strategic problems', in: *Proceedings of SE'98: Systems engineering pragmatic solutions to today's real world problems,* Systems Engineering Society of Australia, Oct 1998.

McLucas, A.C. 2001, 'An investigation into the integration of qualitative and quantitative techniques for addressing systemic complexity in the context of organisational strategic decision-making,' PhD Dissertation, University of New South Wales, July 2001.

McLucas, A.C. and Linard K.T. 2000, 'System dynamics practice in a non-ideal world: modelling Defence preparedness', in: *Proceedings of System Dynamics 2000, International System Dynamics Conference,* System Dynamics Society, Bergen, Norway, August 2000.

Nuthman, C. 1994, 'Using human judgement in system dynamics models of social systems', in: System Dynamics Review, vol. 10, no. 1 (Spring 1994): 1-27.

Richardson, G.P. 1985, 'Problems with causal-loop diagrams (1976)', in: *System Dynamics Review*, vol. 2, no. 2: 158-170.

Senge, P. 1990, 'The fifth discipline: The art and practice of the learning organisation', Doubleday, New York.

Senge, P., Roberts, C., Ross, R.B., Smith B.J. and Kleiner, A., 1994, 'The fifth discipline field book: Strategies and tools for building a learning organisation', Nicholas Brealey Publishing, London.

Sterman, J. D. 2000, 'Business dynamics: Systems thinking and modelling for a complex world', Irwin McGraw-Hill.

Taber, R. 1991, 'Knowledge Processing with Fuzzy Cognitive Maps', in: *Expert Systems with Applications*, vol. 2, no.1, 83-87.

Vennix, J.A.M. 1996, 'Group model building: Facilitating team learning using system dynamics', John Wiley and Sons, Chichester, UK.

Back to the Top