Table of Contents

Go Back

A Possible Method for Assessing The Relative Values of Alternative System Dynamics Models

Geoff Coyle

Contact Person Professor Geoff Coyle,
Organisation Visiting Professor of System Dynamics, South Bank University,
London
Address (personal) 8, Cleycourt Road,
Shrivenham,
Swindon SN6 8BN,
United Kingdom

Telephone +44 1793 782817
Fax +44 1793 783188
Email geoff.coyle@btinternet.com

Abstract

Recent years have seen a debate within the system dynamics community about whether diagrammatic models can provide useful insights into problems involving feedback or whether one must always build a simulation model. Since building and testing a simulation requires much more effort than drawing a diagram the question arises of the measurement of the value added by that extra work or the value lost by forgoing it. However, there may be an analytical project also involves issues such as cost, replication of detail and so forth and different models might satisfy those requirements to a greater or lesser degree. The paper therefore uses another management science methodology, the Analytic Hierarchy Process, to propose a method of calculating an index of the relative value for money of competing system dynamics models. This may be a completely novel approach to some aspects of system dynamics practice.

Introduction

In two previous papers (Coyle, 1998, 2000) I pointed out that **all** system dynamics models start with a diagram of one form or other, whether it be a stock/flow diagram drawn by hand or on a PC screen or an influence diagram. I argued that such diagrams provide some degree of understanding about the problem and, in some cases, might be all that is needed to provide useful assistance to problem proprietors or in academic study. I then raised the issue of how one might assess the value added to an analysis by the extra effort required to build a fully tested simulation model.

Not surprisingly, those views have been criticised (Homer and Oliva, 2001) and it is fair to say that there is a widely held view that simulation is the *sine qua non* of system dynamics and that, unless one has simulated, one has not done system dynamics. Richardson (2001) has made a noteworthy contribution to resolving this issue of diagrams versus simulations and has sought to distinguish between ‘maps’ and ‘models’. His principal conclusion is that what he terms ‘maps’, or word-and-arrow diagrams which may identify stocks explicitly, can be valuable for structural insights, whereas ‘models’, formal quantified representations, confer dynamic insights. He states, and I do not differ, that dynamic insights can be derived from diagrams in only the very simplest cases.

Richardson’s implication that only a quantified simulation program is a model is, perhaps, too sharply drawn. Any model is a simplification of reality intended to be a tool for thought about a problem. In those terms, a cartographical map is a useful ‘model’ of geography. In system dynamics, diagrams of problems involving feedback have a substantial pedigree of refereed publication and textbook exegesis extending back to about 1977 (references to that work are given in Coyle, 2001) and such diagrams, being simplifications of reality intended to serve a useful purpose, are models in their own right. In this paper, therefore, the terms influence diagram, or ID, and simulation will be used to denote two types of feedback model, each with its own strengths, limitations and development costs. This paper will briefly review the issue of the value added by the extra cost of simulation but is mainly concerned with a proposed method for evaluating the extent to which different models might meet a defined need.

In order to explain the method, we must first define more clearly the problem it addresses. That requires a discussion of the uses of feedback diagrams and a more precise statement of the research issue. There then follows an explanation of the method and two worked examples. Having covered that ground, we will then comment on the practical use of the proposed method and, of course, its limitations.

The Uses of Influence Diagrams

Influence diagrams (or CLDs, though the two are not always the same. Coyle (1996) discusses IDs at length) have five uses:

1. Putting onto one piece of paper a view of a very complex problem. That can be remarkably valuable in its own right; the Angolan diagram (Coyle, Bate and Hamid, 1999) was described as “brilliant” by the client.

2. Serving as an agenda for discussing aspects of the problem.

3. Examination of the feedback structure, especially when the diagram has been redrawn at a higher level in the cone of diagrams (Coyle, 1996 and papers already cited), can give useful insights into a problem (Doman *et al*, 1995) though it does not, of course, enable dynamics to be predicted reliably.

4. It can show the context of other problems such as the Y2K problem for UTILITY Corporation (Coyle, 1998, 2000)

5. Last, but far from least, it can be used as the basis for building a simulation model where it is desirable and feasible to do so.

Uses 1 to 4 apply regardless of the format of the diagram (stock/flow or ID) though IDs are sometimes easier to explain to the client or to other academic disciplines. The value of uses 1 to 4 cannot reasonably be denied so I have posed the question:

What is the value added by the extra effort of step 5 (and the effort might be substantial) or the value lost by omitting that step in a given analysis?

It has, of course, been asserted, without evidence or quantification, that simulation virtually always adds so much value that the effort of building a simulation is justified because, without it, dynamics cannot be predicted. That begs the question by assuming that the prediction of dynamics is the object of the exercise. It can be argued that it is really to shed light on problems and, in any given case, the analyst has the responsibility of using the most appropriate tool of the many which are available to us. Even within its own terms, though, the assertion that a simulation is nearly always essential (some argue that it is *always* essential) does not define the type of simulation. Should it be aggregate and broad-brush with, say, 60 variables, or very detailed with, perhaps, thousands, or even tens of thousands, of variables, or somewhere between?

This paper will suggest a possible approach to a metric for discriminating between alternative models and hence for providing a possible answer to the research question.

Problem Statement

Before proposing that solution it will be as well to state the problem more precisely but, in what follows, numbers and the factors considered *are simply examples to demonstrate a method*. They are not intended to represent all, or any, SD practice. Please read them as illustrative data, and no more than that.

Consider three possible models of a particular problem.

1. Model X is an ID with, say, 50 variables and some parameters. It can be drawn and reported to the client or the academic community within, perhaps, one week. (That time scale is perfectly realistic and is based on my own experience.)

2. Model Y is a simulation with about the same number of variables and parameters and, given usual SD practice, probably some nonlinearities. It will take someone a month or more to develop the equations, *rigorously* test them, conduct policy design and report back.

3. Model Z is a very detailed simulation involving hundreds if not thousands of variables (perhaps as arrays) and much numerical data. *Rigorous and complete* testing of all the equations and their behaviour will be a nightmare and it will require a few people for some months to complete all the work.

All SD models start with a diagram and Y and Z are no exceptions. In those cases, though, the diagram is not an end in itself, as it is with model X, but with modern iconic software is a necessary portal without which equations cannot be written.

Such variants are not at all fanciful. The World Model (Forrester, 1971) is clearly Type Y. It requires about four pages to list its equations and definitions. Jay is, of course, a virtuoso and it did not take him long to build it, but World2 (Meadows *et al* 1972) lists 17 people as the research team and required many months of effort. It does not have thousands of variables, but it is clearly close to Type Z (bearing in mind that these model types are only illustrations). However, a Type X 'world model' has also been drawn (Coyle, 1984).

In this case, and no doubt in very many others, it is reasonable to conceive of different models being possible and the issue is to calculate some sort of index of their relative value in meeting the needs of a client or of the academic community, while taking into account their differing characteristics. The method will still be useful even to those who assert that we must always simulate and that not to do so is not to practice system dynamics (I am not sure who has the authority to say what is, and is not, system dynamics other than that it involves a point of view rooted in feedback). Such a person would, though, have to deny the value of the first four uses of diagrams, of whatever type, which does not seem to be a reasonably tenable stance.

The Suggested Method

The suggested method is based on the Analytic Hierarchy Process, AHP (Saaty, 1980), often referred to, eponymously, as the Saaty method. Before showing a worked example in the next section, it is necessary to explain the principles.

The mathematics of the AHP is briefly explained in the Annex to this paper but its essence is to construct a matrix expressing the relative values of a set of attributes. For example, what is the relative importance to a client of the cost of a study as opposed to the detail with which his problem is represented? He (or, of course, she) is asked to choose whether cost is very much more important, rather more important, as important, and so on down to very much less important, than representation of detail. Each of these judgements is assigned a number on a scale. One common scale (adapted from Saaty) is:

Intensity of importance	Definition	Explanation
1	Equal importance	Two factors contribute equally to the objective
3	Somewhat more important	Experience and judgement slightly favour one over the other.
5	Much more important	Experience and judgement strongly favour one over the other.
7	Very much more important	Experience and judgement very strongly favour one over the other. Its importance is demonstrated in practice.
9	Absolutely more important.	The evidence favouring one over the other is of the highest possible validity.
2,4,6,8	Intermediate values	When compromise is needed

A basic, but very reasonable, assumption is that if A is absolutely more important than B and is rated at 9, then B must be absolutely less important than A and is valued at 1/9.[i]

These pairwise comparisons are carried out for all factors to be considered, usually not more than 7, and the matrix is completed. The matrix is of a very particular form which neatly supports the calculations which then ensue (Saaty was a *very* distinguished mathematician).

The next step is the calculation of a list of the relative weights, importance, or value, of the factors, such as cost and detail, which are relevant to the problem in question (technically, this list is called an eigenvector). If, perhaps, cost is very much more important than detail, then, on a simple interpretation, a detailed model is not required. The second is the calculation of a Consistency Ratio (CR) to measure how consistent the judgements have been relative to large samples of purely random judgements. If the CR is much in excess of 0.1 the judgements are untrustworthy because they are too close for comfort to randomness and the exercise is valueless or must be repeated. It is easy to make a minimum number of judgements after which the rest can be calculated to enforce a perhaps unrealistically perfect consistency.

The AHP is sometimes sadly misused and the analysis stops with the pairwise comparison of relative importance (sometimes without even computing the CR!) but the AHP's true subtlety lies in the fact that it is, as its name says, a *Hierarchy* process. The first eigenvector has given the relative importance attached to requirements, such as cost and detail, but, in our case, different models might contribute to differing extents to the satisfaction of those requirements. Thus, subsequent matrices can be developed to show how Models X, Y and Z respectively contribute to the needs of the problem owners. The matrices from this lower level in the hierarchy will each have their own eigenvectors and standard matrix calculations will produce an overall vector of the relative merits of X, Y and Z vis-à-vis the client's needs (or the requirements of academic publication and career progress). The significance of that will be discussed at the close of this paper.

A Worked Example

It is now appropriate to illustrate the concept, reiterating that the factors and the data used are chosen simply to demonstrate a set of calculations. The illustration is in terms of a consultant developing a proposal which will best meet a client's needs.

We start with the requirements for the analysis and suppose that the client has four factors in mind: the expense of the study; the ability of its results to be understood and explained to other people in the firm; the ability of a simulation to reflect detail; and the desirability of having reasonably reliable predictions of dynamics, or at least some insight into dynamics. These requirements are denoted E, U, R and P respectively. Corresponding factors could be adduced for an academic planning how best to embark on a new research topic.

The factors chosen should be independent, as required by Saaty's mathematics. At first sight, E and R are not independent but, in fact, what is really shown is that the client would prefer not to spend too much money but is willing to do so if the results justify it.

We first provide an initial matrix for the client's pairwise comparisons in which the principal diagonal contains entries of 1, as each factor is as important as itself.

	E	U	R	P
E	1			
U		1		
R			1	
P				1

There is no standard way to make the pairwise comparison but let us suppose that the client (perhaps with the help of the consultant) decides that U, understandability, is slightly more important than cost. In the next matrix that is rated as 3 in the cell U,E and 1/3 in E,U. He also assesses that cost is far more important than the ability to reproduce detail, giving 5 in E,R and 1/5 in R,E, as shown below. (It matters not whether or not we think that these are the right relative weights to attach, and this client is not necessarily being irrational. His aversion to expense means that he wishes to try before he buys too much, which is exactly what a recent client of mine, who was unfamiliar with SD, wished to do.)

	E	U	R	P
E	1	1/3	5	
U	3	1		
R	1/5		1	
P				1

The client similarly judges that understandability, U, is much more important than replication of detail, R (rating = 5), and the same judgement is made as to the relative importance of P vis-à-vis R. The completed matrix, which we will term the Client Preference Matrix (CPM), is:

	E	U	R	P
E	1	1/3	5	1
U	3	1	5	1
R	1/5	1/5	1	1/5
P	1	1	5	1

The eigenvector (a column vector but written as a row to save space), which we will call the Client Value Vector (CVV), is calculated by standard methods as (0.232, 0.402, 0.061, 0.305); this client values understandability most of all, likes the idea of dynamic prediction, is just not interested in detail and is not all that worried about cost. The CR is 0.055, well below the critical limit of 0.1, so he is consistent in his choices.

We now turn to the three potential models, X, Y and Z. We now need four sets of pairwise comparisons but this time in terms of how well X, Y and Z perform in terms of the four criteria, E, U, R and P.

The first table is with respect to E, expense, and ranks the three models as :

	X	Y	Z
X	1	5	9
Y	1/5	1	3
Z	1/9	1/3	1

This means that X is considerably better than Y in terms of cost as it will be much cheaper to produce, and similarly for Z. Actual cost figures could be used but that would distort this matrix relative to others in which qualitative factors are assessed. The eigenvector for this matrix is (0.751, 0.178, 0.071), very much as expected, and the CR is 0.072, so the judgements are acceptably consistent.

The next three matrices are respectively judgments of the relative merits of X, Y and Z with respect to understandability, reproduction of detail and reliability of prediction of dynamics:

Understandability:

	X	Y	Z
X	1	1	5
Y	1	1	3
Z	1/5	1/3	1

Eigenvector (0.480, 0.406, 0.114), CR=0.026

Replication of detail

	X	Y	Z
X	1	1/3	1/9
Y	3	1	1/3
Z	9	3	1

Eigenvector (0.077, 0.231, 0.692), CR=0 (perfect consistency)

Prediction of Dynamics

	X	Y	Z
X	1	1/9	1/5
Y	9	1	2
Z	5	1/2	1

Eigenvector (0.066, 0.615, 0.319), CR= 0.

The reason why Y scores better than Z on this criterion is that Z, being a large model, will be much more difficult to debug so its predictions might contain numerical inaccuracies.

The final stage is to construct a matrix of the eigenvectors for X, Y and Z

	E	U	R	P
X	0.751	0.480	0.077	0.066
Y	0.178	0.406	0.231	0.615
Z	0.071	0.114	0.692	0.319

This matrix, which we call the Option Performance Matrix (OPM), summarises the respective capability of the three models in terms of what the client wants. Reading down each column, it somewhat states the obvious: X is far better than Y and Z in terms of cost; it is a little better than Y in terms of understandability, as the client will have to take somewhat on trust that Y's calculations are correct; however, X is of limited value for representing fine detail and predicting dynamics. These are not, however, absolutes; *they relate only to the set of criteria chosen by this hypothetical client*. For someone else, who perhaps valued only the ability to describe the system and wanted to avoid cost, the three models would score quite differently.

Those results are only part of the story and the final step is to take into account the client's judgements as to the relative importance of E, U, R and P. For a client whose only requirement was for dynamic prediction, model Y would be ideal. Someone who valued only replication of detail would need model Z. This client is, however, more sophisticated, as, I suspect, are most clients, and has already expressed his assessment of the relative weights attached to E, U, R and P in the Client Value Vector (0.232, 0.402, 0.061, 0.305). Finally, then, we need to weight the value of achieving something, R, say, by the respective abilities of X, Y and Z to achieve R, that is to combine the Client Value Vector (CVV) with the Option Performance Matrix (OPM). Technically, the calculation is to post-multiply the OPM by the CVV to obtain the priority, or client satisfaction, vector (0.392, 0.406, 0.204). This vector might be called the Value For Money vector (VFM). In matrix algebra, OPM*CVV=VFM or, in words,

performance*requirement= value for money.

In those terms, this suggested method might have wide applicability.

The three numbers in the VFM are the final result of the calculation, but what do they mean?

First, in simple terms, they mean that Model X, which scores 0.392, seems to come out slightly worse in terms of its ability to meet the client's needs than does Model Y at 0.406. The large simulation, Z, is well behind at 0.204 and would do rather badly at satisfying the client's requirements in this illustrative case.

Secondly, the three decimal places are, in practical terms, illusory, and X and Y are equal at 0.4. In the real world, the consultant could confidently recommend Y, the moderate size simulation, knowing that it was as good as X and would be a larger contract. The client might well prefer X, as it would give some results quickly.

Thirdly, and perhaps most importantly, the vector of the relative merits of X, Y and Z follows ineluctably from judgements made by the client as to his requirements and by the analyst as to the capabilities of differing models. There is a strong audit trail from output back to inputs. Of course, anyone who understands the AHP mathematics might be able to fiddle the judgements so as to guarantee a preferred outcome, but that is unavoidable expect by vigilance and the Delphi approach discussed below.

A Second Example

Another client has a different set of objectives. In his view, E is more important than U, but R and P are respectively much more important and absolutely more important than expense. He also judges that U is more important than R, that predicting dynamics is more important than understanding and that replication of detail is more important than prediction of dynamics. That leads to the Client Preference Matrix:

	E	U	R	P
E	1	3	1/5	1/9
U	1/3	1	3	1/3
R	5	1/3	1	3
P	9	3	1/3	1

The discussion of his objectives does indicate a certain amount of confusion and that is confirmed by the calculations. The eigenvector, or Client Value Vector, turns out to be (0.113, 0.169, 0.332, 0.395) but the Consistency Index is 0.94, vastly in excess of the cut-off of 0.1 and indicating that the client's valuations have, for all practical purposes, been made at random. (This example illustrates the immense importance of the calculation of the CR, a step which is sometimes omitted in careless use of the AHP.)

A set of objectives such as these is a recipe for a disastrous consultancy assignment or an unproductive research project. The explanation above foreshadows that but the calculation has confirmed just how incoherent the objectives are, and has done so in a way which might not have been so clear by verbal discussion. All too often, this sort of confusion remains hidden in the mind of the client or, and even worse, in the separate minds of different interest groups within the client organisation. That is a virtual guarantee of trouble, no matter how skilful the consultant at client management during the project. Using this technique to persuade the client or researcher to be explicit might help to avoid problems.

What is to be done in such a case?

For very good reasons, a consultant is usually reluctant to walk away from a proffered assignment or a researcher to abandon an apparently promising line of enquiry, though in both cases that is *sometimes* the right course of action. The sensible course is to try to work out a consistent set of objectives. That could be supported by the software mentioned below enabling one to say "we're not quite consistent yet".

Let us suppose that rethinking the objectives leads to the new matrix:

	E	U	R	P
E	1	1/3	1/9	1/5
U	3	1	1	1
R	9	1	1	3
P	5	1	1/3	1

As before, it is meaningless to debate whether or not these are 'good' judgements; they reflect the client's mental model of the significance of the problem to the wider objectives of the firm.

The new matrix produces the Client Value Vector, or eigenvector, (0.262, 0.454, 0.226) and has a CR of 0.068, well on the safe side of the cut-off of 0.1; the judgements are now strongly consistent as opposed to the first set which were virtually random.

Weighting this CVV by the OPM previously calculated gives a Value for Money vector of the relative merits of models X, Y and Z of (0.220, 0.360, 0.420). Model X, the diagram as an end in itself, is now well out of the running and model Z, the large simulation, is rather better than model Y, the small simulation, but not dominantly so. A wise conclusion might be to do model Y as a first phase but carefully designed so that progression to Z will be as easy as possible, should the client wish to do so after having seen the results of Y. This suggested AHP method has thus helped to resolve confusion about objectives and, furthermore, indicated a sensible way ahead.

Comment on the Method

Reiterating yet again that the factors chosen and data used were no more than a demonstration of a method, the ability to compute the relative merits of alternative models seems to be close to a solution to the earlier value-added question. Someone who does not think that it is a valid question might discard Model X, the system description diagram (though that stance involves also rejecting the first four uses of diagrams) and use the suggested method simply to compare two or more possible simulation models to see which was the better, or best, option.

It does, though, seem reasonable to suppose that there must be more than one criterion against which possible models of a given problem might be assessed. A client who values the ability to reproduce detail but is heedless of cost is hard to imagine, as would be an academic who started a research project without any consideration of the demands on increasingly scarce time.

The four factors used here, E, U, R and P were, of course, purely to demonstrate a calculation, but how might factors be determined in a real case? They could be an *ex cathedra* statement from someone in authority, but a more rational approach might be discussion with a small group, first in Focus Group mode to identify factors and then as a simple Delphi to obtain the Client Preference Matrix. Recall, though, that Delphi is a controlled debate and the reasons for extreme values are debated, not to force consensus, but to improve understanding.

Uses Of The Method

A consultancy firm might be able to use this technique, or a refinement of it, as a sales development tool. It requires discussion of what the client wants, and how much he values the achievement of those objectives, leading to a proposal, not for *a* model, but for the model best suited to meeting those needs. It would be trivial to set up software to record the discussion and do the calculations, perhaps displaying any inconsistent judgements very prominently.

Clearly, if a potential contract is only worth £1000 the client values expense above all else and, in any case, even the slight the effort of the AHP is hardly worthwhile. For a contract worth £100000, the effort could be very valuable by showing the client that his problem is being taken very seriously and that the consultant is thinking very carefully about how the client's needs can best be satisfied. One feels, therefore, that this approach could be a valuable marketing tool for consultants. It might confer an advantage over competing consultants who do not make such assessments.

An academic planning a new research venture might also use this approach, perhaps to justify a research grant application (research grant committees might use it to evaluate applications!). The scholar might value aspects such as demands on scarce research time, the personal benefit of getting another publication in time for the next meeting of the Tenure Committee, and whatever else that person judges to be relevant factors in terms of career development (cynical) or contribution to a body of knowledge (altruistic).

Limitations of the AHP Method

The AHP is, of course, in the same vein as Multi-attribute Utility analysis (MAU) and it is well –known that there are difficulties in obtaining estimates of utility preferences in MAU. The AHP overcomes those, to some extent, by the calculation of the consistency index so, although the preferences are subjective, they cannot be any other, at least it is known how consistent they are. Indeed, the second worked example displayed that aspect very clearly.

The AHP is also criticised for producing extreme values in the eigenvector. For instance, an eigenvector of {0.45,0.45,0.1} seems to imply that two factors are 4.5 times as important as the third but to do so is to misinterpret the eigenvector. The best interpretation is in terms of the verbal meanings in the table of weightings discussed earlier; the first two items are of equal importance and each is 'much more important' than the third. Such an interpretation is easily applied to any other eigenvector.

Summary

The introduction to this paper referred to the question of the value-added by building a simulation or the value foregone by not building one and relying simply on the understanding which can be generated by a diagram. Some might argue that no worthwhile understanding can be gained from a diagram and that simulation is a *sine qua non* for system dynamics work. For such a view, Model X could not be considered, though that involves the judgement that the first four uses of a diagram have no value, so we extended the question to consideration of the relative merits, *in terms of meeting a client's needs,* of alternative forms of simulation.

The proposed method is based on the Analytic Hierarchy Process and has been shown to be capable of computing the relative merits of competing models. This seems to be applicable both as a sales tool in consultancy and in the planning of academic research.

The AHP is, perhaps, not widely known in the system dynamics community, so this paper describes a novel approach, drawing on another analytical technique to address an apparent problem in the practice of SD. That may vindicate an earlier comment (Coyle, 1998) that some knowledge of other operational research disciplines is useful in SD.

The ideas presented here are by no means the last word on the issue. The calculations are trivial and soundly rooted in matrix algebra but the subtleties of the method may call for more exploration. What, for instance, are 'good' things for a client to require from a model? How can independence between those factors best be assured? How can objective assessments be made as to the capabilities of alternative models? How is the whole process best conducted?

Finally, I reiterate again that the factors used and the judgements made are *solely* to show how calculations can be carried out.

Notes and References

Coyle R G (1984), ‘East and West and North and South’, *FUTURES*, Vol.16, No. 6, pp594-609.

Coyle R G (1998), ‘The practice of system dynamics: milestones, lessons and ideas from 30 years experience’, *System Dynamics Review*, Vol. 14, No. 4, pp 343-365.

Coyle R G (2000), ‘Qualitative and quantitative modelling in system dynamics: some research questions’, *System Dynamics Review* 16:3, 225-244.

Coyle R G (2001), ‘Rejoinder to Homer and Oliva’, *System Dynamics Review*, 17:4, .

Coyle R G, M J Bate and S Hamid (1999), ‘Strategic assessment of present and future security environments: A systems approach’, *Defense Analysis*, 16:3, 299-314.

Doman A, M Glucksman, N Mass, M Sasportes (1995), ‘The dynamics of managing a life insurance company’, *System Dynamics Review*, Vol. 11, No. 3, pp 219-232.

Forrester J W (1971) *The Limits To Growth*, Cambridge, Mass, MIT Press (since republished by Productivity Press)

Homer J and R Oliva (2001), ‘Maps and models; a response to Coyle’, *System Dynamics Review*, 17:4, .

Meadows D H *et al* (19772), *World Dynamics*, London, Earth Island Ltd, and elsewhere.

Richardson G P (2001), ‘Mapping and Modeling: THE answer to the controversies’, *Proceedings of the 2001 System Dynamics Conference*.

Saaty T L (1980), *The Analytic Hierarchy Process*, McGraw Hill International.

Wolstenholme E F and R G Coyle (1983), ‘The development of system dynamics as a rigorous procedure for system description’, *Journal of the Operational Research Society*, Vol. 34, No. 9, pp 569-581.

Annex – the Mathematics of the Analytic Hierarchy Process

The mathematical basis of the AHP can be explained in fairly simple outline for the purposes of this paper. For a full treatment, the mathematically undaunted should refer to Saaty's book.

Consider n elements to be compared, C_1 … C_n and denote the relative 'weight' (or priority or significance) of C_i with respect to C_j by a_{ij} and form a square matrix A=(a_{ij}) of order n with the constraints that a_{ij} = 1/a_{ji}, for i   j, and a $_{ii}$ = 1, all i. Such a matrix is said to be a reciprocal matrix.

The weights are consistent if they are transitive, that is a_{ik} = $a_{ij}a_{jk}$ for all i, j, and k. Such a matrix might exist if the a_{ij} are calculated from exactly measured data. Then find a vector ù of order n such that $Aw = lw$. For such a matrix, ù is said to be an eigenvector (of order n) and ë is an eigenvalue. For a consistent matrix, $l = n$.

For matrices involving human judgement, the condition a_{ik} = $a_{ij}a_{jk}$ does not hold as human judgements are inconsistent to a greater or lesser degree. In such a case the ù vector satisfies the equation Aù= ë$_{max}$ù and ë$_{max}$ n. The difference, if any, between ë$_{max}$ and n is an indication of the inconsistency of the judgements. Finally, a Consistency Index can be calculated from (ë$_{max}$-n)/(n-1). That needs to be assessed against judgments made completely at random and Saaty has calculated large samples of random matrices of increasing order and the Consistency Indices of those matrices. A true Consistency Ratio is calculated by dividing the Consistency Index for the set of judgments by the Index for the corresponding random matrix. Saaty suggests that if that ratio exceeds 0.1 the set of judgments may be too inconsistent to be reliable. In practice, CRs of more than 0.1 sometimes have to be accepted. If ë$_{max}$ = 1 then the judgements have turned out to be consistent.

---

[i] In other uses of Saaty, I have experimented with scales going up to 25.

Back to the Top