The qualitative and semiqualitative analysis of environmental problems

Simon Parsons
Advanced Computation Laboratory
Imperial Cancer Research Fund
P.O. Box 123
Lincoln's Inn Fields
London WC2A 3PX

Mirko Dohnal
Department of Chemical Engineering
University of Natal
King George V Avenue
Durban 4001
South Africa

ABSTRACT

The analysis of many biochemical engineering problems in environmental modelling is based upon the development and solution of sets of differential equations. A complete analytical solution of such a model requires that every numerical constant in this set of equations is precisely known. This paper describes the use of methods from artificial intelligence which permit the solution of such sets of equations when some constant values are unknown. The use of the methods are illustrated with the solution of a set of equations representing one model of an anaerobic fermentor, and a computer program that implements the methods is described.

SOFTWARE AVAILABILITY

Name of the software:	SEMI
Contact address:	Simon Parsons, Advanced Computation Laboratory Imperial Cancer Research Fund P.O. Box 123 Lincoln's Inn Fields London WC2A 3PX
Telephone:	0171 269 3636,
Fax:	0171 269 3186,
Email:	sp@acl.lif.icnet.uk.
Year first available:	1991
Hardware required:	IBM PC, or compatible machine.
Software required:	None.
Program language:	Pascal.

1. INTRODUCTION

The basis of the method by which much of modern engineering, especially in areas such as biochemical and environmental engineering, proceeds is by the identification and solution of models based upon sets of differential equations (Bailey & Ollis 1986). These equations describe the dynamic behaviour of various key variables, whose values, when the equations are solved, predict the state of the system at any instant. Often, knowledge of laws of nature such as the law of mass conservation, are sufficient to form the foundation of such a set of equations. As a result, given a particular system, it is often relatively easy to collect together a set of differential equations which can form the nucleus of a detailed mathematical model of that system. However, to take this nucleus and flesh it out with all the necessary information that will make it an accurate and realistic working model is far from simple.

This is because the laws of nature on their own are not sufficient to build a good model. What is needed are the precise values of the constants that relate the variables in the differential equations. Unfortunately, real environmental systems are horribly complicated and as a result, it is extremely difficult to measure the various constants with any accuracy. This is particularly true when the dynamic behaviour of such systems is considered. Indeed, the systems may be subject to complex relations with their surroundings (Serra *et al.* 1992) which may make it nearly impossible to isolate them without distorting any measurements made. Therefore it is often necessary to deal with sparse and inconsistent information about such systems, and it is likely to be extremely difficult, time-consuming and expensive to identify the value of every numerical constant, a point stressed by Steyer *et al.* (1992). Without the values of all the constants the set of equations have no practical value because without known values it is not possible to run a classical simulation, and for most real systems the set of equations cannot be solved by analytical means.

Thus incomplete and uncertain knowledge of the necessary numerical constants would seem to rule out the use of conventional methods in projects such as modelling the scaling up of laboratory fermentors in order to perform tasks like risk evaluation and cost estimation. It need not, however, prevent the use of artificial intelligence techniques. Several authors have applied rule-based methods, for instance (Baldwin, Martin, & Zhou 1993; Liong, Chou, & Law 1991), especially in the area of control (Serra *et al.* 1992; Steyer *et al.* 1991; Watts & Knight 1991). However, given the fact that many environmental models are based upon sets of equations, qualitative (Davis 1990; Weld & de Kleer 1990) and semiqualitative (Parsons &

Dohnal 1993) reasoning seem particularly appropriate. It is the intention of this paper to discuss how these techniques might be employed in environmental engineering, following the example of Steyer *et al.* (1992), but providing a general approach which can be applied to any model based on equations.

The structure of the paper is as follows. Section 2 introduces qualitative reasoning, giving a brief historical overview of the main developments, discussing applications based on them, and showing how it may be used to solve a model of an anaerobic fermentor in which the constants are not known. Section 3 then discusses some of the problems that have been noted with the qualitative approach, and mentions some solutions that have been proposed. Section 4 discusses semiqualitative reasoning, which is a generalisation of some of the ideas of qualitative reasoning to enable the use of any numerical information that is available. Section 5 describes a program that can be used for qualitative and semiqualitative reasoning using sets of equations, and Section 6 presents some conclusions about this method of analysis.

2. QUALITATIVE REASONING

2.1. The development of qualitative reasoning

The paper that is always cited as being the foundational work in qualitative reasoning is Hayes' Naive Physics Manifesto (1978) in which he urged practitioners of artificial intelligence to "put away childish things by building large scale formalizations" (Hayes 1985b). His suggestion was that real progress in the field would come about by attempting to model a large part of human commonsense knowledge about the real world, and his first attempt created an initial theory of liquid behaviour (Hayes 1985a). This work was built upon first order logic, the traditional tool of symbolic artificial intelligence. At the same time, and to some extent as a result of Hayes' proposal, work that modelled complex systems in a way that mirrored the kind of approach adopted by engineers was emerging.

There are, broadly speaking, three strands to this work, all having in common the fact that they deal with abstractions of real numbers into positive, negative and zero valued quantities rather than dealing with numbers themselves. The first approach is that of Kuipers (1984) who takes a set of differential equations, abstracts them to just consider their qualitative impact, and then uses them as a set of constraints on the possible values of the state variables. This approach has been implemented as the QSIM software system (Kuipers 1986). The second approach, taken by de Kleer and Brown (1984) and Williams (1984) is to build libraries of components, each of which has a well defined qualitative behaviour described by sets of qualitative differential equations, and connect components together to build a qualitative model. Some of this work is implemented as the ENVISION software system which takes its name from the process of "envisionment" by which behaviour is infered from the structure of the system. The final approach not only models components, but also the processes that they may undergo. Work on this approach is primarily due to Forbus (1984), and is closest in spirit to the work on naive physics. In addition, this approach goes further than the others in allowing sets of objects to have group behaviours over and above their individual ones, thus providing a far richer modelling language.

This is, of course, just an introductory sketch of the field which is rapidly expanding, and therefore has, at the time of writing, eluded a definitive survey. Papers by Cohn (1989) and Coiera (1992) come closest to providing such a survey but fail to cover absolutely everything.

2.2. A brief demonstration of qualitative reasoning

The core of the first two approaches described above is the idea of qualitative differential equations. Rather than attempting to deal with a mass of numerical data, values are only distinguished as positive (+), zero (0), negative (-), or unknown (?). These values are sufficient to identify many of the interesting features of the behaviour of the important variables in a given system.

Briefly, this works as follows. Imagine that we have a very simple system which may be described by the equations:

$$\frac{dx}{dt} + k = x \tag{1}$$

$$\frac{d^2x}{dt^2} + x = 0 \tag{2}$$

where k is a positive constant and x is a substrate concentration. The qualitative abstraction of these equations, in which all numerical values are replaced by $+$, 0 or $-$ is:

$$\frac{dx}{dt} \oplus + = x \tag{3}$$

$$\frac{d^2x}{dt^2} \oplus x = 0 \tag{4}$$

where $\oplus$ is qualitative addition, as described by Table 1. To solve the pair of equations we look for sets of qualitative values that satisfy them, for instance:

$$x = + \tag{5}$$

$$\frac{dx}{dt} = + \tag{6}$$

$$\frac{d^2x}{dt^t} = - \tag{7}$$

since ? is an abbreviation for $+$ or 0 or $-$. In other words, x is positive, its first time derivative is positive, but its second time derivative is negative. This set of values tells us that the behaviour of the concentration over time will be to rise to some limiting value as in Figure 1. We may not know what the limit is, but we do know that the concentration will eventually level off, and this less precise information may be sufficient. Clearly if we are trying to establish that the substrate concentration has a maximum value then the information we are able to deduce is quite adequate, and

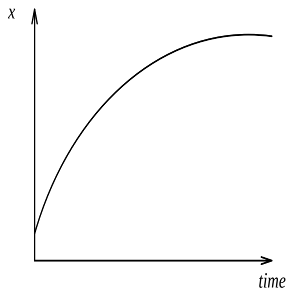

Figure. 1. The qualitative behaviour of x

$\oplus$	+	0	−
+	+	+	?
0	+	0	−
−	?	−	−

Table 1. Qualitative addition

in many cases the fact that we can learn something from qualitative reasoning far outweighs the fact that what we learn is not very detailed.

2.3. A more detailed example of qualitative reasoning

To illustrate the use of qualitative reasoning more completely, consider the following example. It is possible (Bailey & Ollis 1986) to write down a complex set of equations which fully describe the action of an anaerobic fermentor and which, when solved, provide a suitable model of its behaviour. Unfortunately, the results of this analysis hinge upon the values of a number of key constants whose values not only vary from fermentor to fermentor, but are also extremely difficult to measure. As a result it is difficult and expensive to provide accurate solutions from a conventional analysis. A qualitative analysis is, however, possible.

The following set of equations provide a simplified model of the behaviour of an anaerobic fermentor:

$$\frac{dx_1}{dt} + (k_{12} + k_{13})x_1 + k_{11}x_1x_5 = k_{21}x_2 \quad (8)$$

$$\frac{dx_2}{dt} + k_{21}x_2 = k_{12}x_1 \quad (9)$$

$$\frac{dx_3}{dt} - k_{63}x_4 = k_{13}x_1 \quad (10)$$

$$\frac{dx_4}{dt} + k_{43}x_4 = k_{11}x_1x_5 \quad (11)$$

$$\frac{dx_5}{dt} + k_{53}x_5 + k_{11}x_1x_5 = 0 \quad (12)$$

where x_1–x_5 are concentrations of various substrates, either those wastes being digested or the products of the digestion. The full model may also be solved using qualitative methods, but the additional detail adds nothing to the understanding of the technique. It is, of course, perfectly possible to apply the method to any equation based model. This particular model was chosen because it was easily available.

Solving this model using the Q SENECA software system

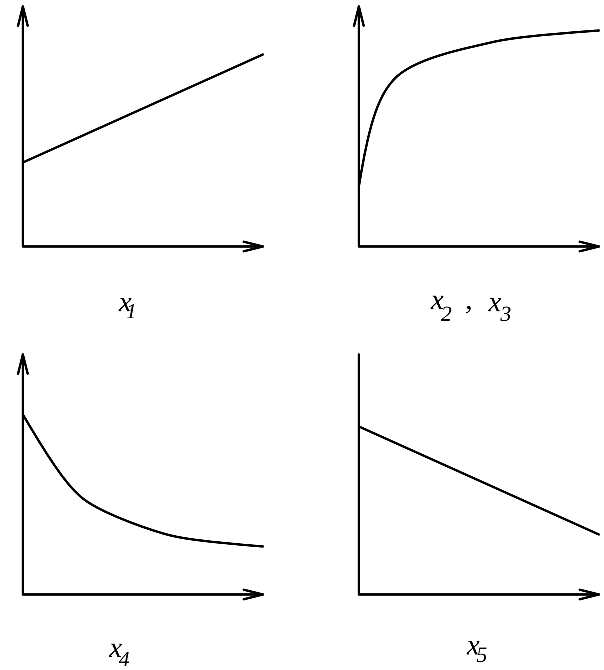

Figure. 2. The qualitative solution of the fermentor

(Dohnal 1991) gives, as is usually the case with qualitative models, a number of possible solutions. one of these is:

	x	$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$
x_1	+	+	0
x_2	+	+	−
x_3	+	+	−
x_4	+	−	+
x_5	+	−	0

which tells us, for instance, that x_1 increases linearly while x_3 is rising to a limit and x_4 is falling to a limit. This behaviour is summarised in Figure 2. The use of the set of solutions that is provided by the system is that they define all the possible states of the system*and from these it is possible to predict all the reasonable transitions from one state to another, where transitions are restricted to those changes that do not involve a jump, reducing the number of possible transitions. For example, in the state given above, x_4 is positive, $\frac{dx_4}{dt}$ is negative, and $\frac{d^2x_4}{dt^2}$ is positive. Thus x_4 is decreasing, and so it is possible for this state to make a transition into a state in which x_4 is zero, and then into one in which x_4 is negative. Alternatively, it is possible to make a transition into a state in which $\frac{dx_4}{dt}$ is zero because $\frac{d^2x_4}{dt^2}$ is positive. However, because both x_1 and $\frac{dx_1}{dt}$ are both positive in the current state, it is not possible to make a transition into a state in which x_1 is negative without first going through a number of intermediate states.

Given that the initial state of the system is known, it

*Qualitative systems are generally complete but not sound so that they generate all possible physical eventualities as well as some situations that cannot occur.

is possible to use the set of allowable state transitions to predict how the system will behave over time. This allows the design of the system to be validated, by demonstrating that undesired states, can never be reached or that oscillations are bound to occur (de Kleer & Brown 1984). The same information can be used for diagnosis by determining how a given abnormal state was reached. It is also possible to use the state transitions for control. If some of the variables may be controlled, it is possible to identify how they should be altered in order to reach a particularly desirable state.

2.4. Applications of qualitative reasoning

Qualitative methods have been widely used, a fact that is not surprising when one considers that much of the original work was driven by the desire to model real systems. Two of the early papers on qualitative reasoning (de Kleer 1984; Williams 1984) were concerned with the analysis of digital circuits, and similar attempts with a distinct flavour of naive physics are provided by Davis (1984), Genesereth (1984) and Barrow (1984), while Mohammed and Simmons have considered the use of qualitative simulation in modelling the fabrication of semiconductor devices (Mohammed & Simmons 1986). More recently, Ormsby *et al.* (1991) have applied similar techniques to the diagnosis of faults in the electrical systems of automobiles.

However, electronics is not the only area in which qualitative methods may be applied. Falkenheimer and Forbus (1987) have tackled the simulation of a more complex system, namely the steam plant of a naval vessel, albeit with a number of simplifying assumptions, and Kuipers (1987) has used QSIM to model process in the human body, a system which is arguably more complex than any man-made artifact, and is certainly less well understood. Ardizzone *et al.* (1988) have modelled cell growth with a qualitative system, and Farley and Lin (1991) have modelled economic systems.[†]

There are also several applications of qualitative reasoning in areas closely related to environmental engineering. Hangos *et al.* (1992) discuss the control of a distillation column using a qualitative model to both determine what actions will give the correct response and predict the result of a given action. Hurme *et al.* (1991) make a similar analysis of a chemical recycling process, but with the aim of using the qualitative simulation to identify the conditions under which profit can be maximised. Finally Koivisto *et al.* (1989) discuss the qualitative diagnosis of a chemical reactor.

3. PROBLEMS WITH QUALITATIVE REASONING

Despite the undoubted success of qualitative methods, there are some problems with qualitative reasoning that

[†]Qualitative reasoning has a long and distinguished tradition in economics, having first been considered many years before the inception of artificial intelligence.

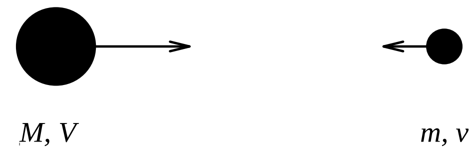

Figure. 3. Two colliding masses

$\ominus$	+	0	−
+	?	+	+
0	−	0	+
−	−	−	?

Table 2. Qualitative subtraction

make it unsuitable for modelling certain systems. These problems stem from the limited number of values that any constant or variable can adopt. Raiman (1986) illustrates this with a simple example from mechanics. Consider two masses which collide while travelling towards one another along the same line (Figure 3). One has a large mass M and velocity V, the other has a small mass and velocity m and v. The net momentum from right to left above is given by the law of the conservation of momentum as:

$$MV_{net} = MV - mv \tag{13}$$

Since M, V, m and v are all positive values, they all have qualitative value +, and the net leftwards momentum is established by the calculation:

$$MV_{net} = + \otimes + \ominus + \otimes + \tag{14}$$

where $\ominus$ is the operator representing the difference of two qualitative values (see Table 2), and $\otimes$ is the operator representing the product of two such values (see Table 3). It is clear from Table 3 that the product of two positive values will itself be positive so that the calculation reduces to:

$$MV_{net} = + \ominus + \tag{15}$$

Now, Table 2 summarises the fact that the difference of two values which are only known to be positive can be either positive, negative or zero, depending on the relative sizes of the values. Thus qualitative reasoning can only deduce that the overall leftwards momentum will be ?, while intuitively we can see that it will be + because MV is much larger than mv.

The problem of coping with situations such as this, that cause difficulties for qualitative reasoning techniques but which can be easily resolved by humans, has been addressed by the artificial intelligence community. The first solution was proposed by Raiman (1986) in the paper in which he pointed out the problem. He introduced a system called FOG which allowed the representation of "order of magnitude" concepts. Thus it allows the statement that, for instance, A is negligible with respect to B, $A\ Ne\ B$, or that A has the same sign and order of magnitude as B, $A\ Co\ B$. These relations are then used to define a set of inference rules such as:

⊗	+	0	−
+	+	0	−
0	0	0	0
−	−	0	+

Table 3. Qualitative multiplication

$$\begin{array}{ccc} A & Ne & B \\ B & Co & C \\ \hline A & Ne & C \end{array}$$

So that if A is much smaller than B, which is about the same size as C, then A is much smaller than C. In all Raiman provides 30 such rules of inference, giving a semantics for the approach which is based on non-standard analysis, and FOG has been used in the modelling of analog circuits (Dague, Raiman, & Devès 1987).

FOG has also been discussed by Dubois and Prade (1991) who have considered the problem that is caused by the use of non-standard analysis as a basis for a semantics—namely that the results are only valid in the limit. In order to cope with situations in which $A\ Co\ B$ does not mean that A and B are infintely close together, they propose a new interpretation in terms of an interval on the ratio of A to B. This allows them to validate the inference rules, and allows a sensible limit on the chaining of inferences such as:

$$\begin{array}{ccc} 30 & Co & 31 \\ 31 & Co & 32 \\ \hline 30 & Co & 32 \end{array}$$

to be established that prevents the derivation of 30 Co 1000 without the need for an arbitrary cut-off. More recent work on FOG has provided a means of obtaining smooth changes between the orders of magnitude that are recognised (Dague 1993b), and to incorporate numerical information (Dague 1993a).

Another scheme for order of magnitude reasoning is due to Mavrovouniotis and Stephanopoulos (1987; 1989) who have formalised the representation of relations such as $A > B$ to give a system called O[M] that they claim is expressive enough for all engineering problems. The semantics of the relations is provided in terms of the bounds on the ratio between A and B, and two possible interpretations are given. The first is mathematically correct, but conservative, and the second is heuristic but more humanly aggressive in the inferences it sanctions. O[M] has been applied to problems in process engineering (Mavrovouniotis & Stephanopoulos 1988). It is possible to show that the O[M] approach can be handled using a formal system for reasoning using intervals (Parsons 1993), and Travé-Massuyès and Piera have shown that interval systems can be used as the basis of a general approach to order of magnitude reasoning (Travé-Massuyés & Piera 1989).

There have been several other attempts to extend qualitative reasoning using limited amounts of numerical information, including (Dormoy 1988; Féray Beaumont 1991; Kuipers & Berleant 1988; Steyer, Queinnec, & Simoes 1992; Sticklen, Kamel, & Bond 1991). In the remainder of this paper we present another such approach —a generalisation of qualitative reasoning, known as semiqualitative reasoning, which allows qualitative and quantitative information to be used in the solution of sets of differential equations.

4. SEMIQUALITATIVE REASONING

4.1. A brief demonstration of semiqualitative reasoning

In semiqualitative reasoning the values of variables and constants are restricted to a set of $2k+1$ intervals (Parsons & Dohnal 1993). This set of intervals covers all numbers from ∞ to $-\infty$, and the intervals are continuous and non-overlapping, so that any real number falls into one, and only one, interval. The intervals are symmetric about zero, which is a distinguished value, and there are k positive and k negative intervals. The boundaries of the intervals may be set by an arithmetic or geometric progression, or may be chosen to reflect what are considered to be interesting values. Since the set of values used in qualitative reasoning corresponds to the set of semiqualitative intervals obtained for $k = 1$, it is clear that semiqualitative reasoning is a generalisation of qualitative reasoning.

A basic understanding of how semiqualitative reasoning may be used to solve sets of differential equations may be obtained from a simple example. Consider that the following set of equations is a model of a physical system:

$$x_1 + x_2 = x_3 \quad (16)$$

$$x_1.x_4 = x_3 \quad (17)$$

$$\frac{dx_4}{dt} = x_5 \quad (18)$$

The model is solved just as the qualitative one was by finding a set of values for the variables and their derivatives from the set of all possible values[‡] so that the equations are satisfied. If we have the set of intervals depicted in Figure 4 then the following five triplets describe one set of assignments of values to the five variables, and thus one conceivable state of the system:

	x	$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$
x_1	$[0, 20]$	$[0, 10]$	0
x_2	$[20, 100]$	$[0, 10]$	$[-20, -100]$
x_3	$[10, 20]$	$[20, 100]$	$[20, 100]$
x_4	$[0, 10]$	$[0, 10]$	0
x_5	$[10, 20]$	$[0, -10]$	0

This state is not however a physically possible state of the system since it is not a solution of equations that describe the system. This is because x_3 is determined from $x_1 =$

[‡]The set of all possible values is not restricted to the set of $2k+1$ intervals. All compositions of contiguous intervals are also permitted values.

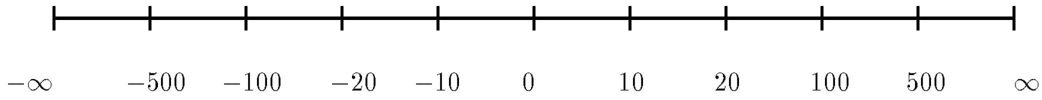

Figure. 4. An initial set of semiqualitative intervals

$[0, 20]$ and $x_2 = [20, 100]$ from the first equation which gives a value of $[0, 20] \oplus [20, 100] = [20, 500]$, where $\oplus$ gives the result of adding two intervals using interval arithmetic (Moore 1966) (in this case $[20, 120]$) and then finding the smallest interval or composition of intervals that holds the result. This value of x_3 contrasts with that of the proposed solution, in which $x_3 = [10, 20]$, and this contradiction rules out the solution. By similar means it is possible to identify all the sets of five triplets which are solutions of the set of equations, and these correspond to all the semiqualitative states of the model.

By allowing variables to take on a wider range of values, semiqualitative reasoning permits the use of those numerical values that are known, and this means that it generates more precise solutions than are possible using qualitative reasoning. However, the fact that it is not necessary to have any more information than whether a quantity is positive, negative or zero means that semiqualitative reasoning is very robust, and may be used in situations where conventional methods cannot be used.

4.2. A more detailed example of semiqualitative reasoning

As a more realistic example of the use of semiqualitative reasoning, this section describes the semiqualitative simulation of an anaerobic digester using the model discussed in Section 2.3. As with the qualitative model, the fact that a specific example is used should not distract attention from the fact that the method can be used to solve any equation based model. The model used for the semiqualitative analysis consists of the same set of differential equations as before, plus those numerical values that are known. These are the values of the following constants:

$$k_{11} = 100 \quad (19)$$

$$k_{12} = 1.5 \quad (20)$$

$$k_{13} = 5.0 \quad (21)$$

$$k_{21} = 3.0 \quad (22)$$

$$k_{43} = 1.0 \quad (23)$$

$$k_{53} = 0.3 \quad (24)$$

$$k_{63} = -1.0 \quad (25)$$

This model is first solved with the boundaries of the semiqualitative intervals set as in Figure 5. These are a default set of boundaries suitable for a first attempt at an analysis. In order to limit the number of solutions, we can specify additional constraints by stating the values of some of the variables and their derivatives:

	x	$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$
x_1	$[0, 20]$	$[0, 10]$	0
x_2	$[10, 1000]$	?	?
x_3	$[0, 20]$	?	?
x_4	$[10, 1000]$	?	?
x_5	$[10, 20]$	$[0, -10]$	0

Note that ? is shorthand for the interval $[-\infty, \infty]$. This set of constraints may be considered as a query, in this case asking the question:

> When x_1 is present in a concentration of less than 20, what are the ways in which it is possible to achieve a linear $\left(\frac{d^2x_1}{dt^2} = 0\right)$ increase of concentration of x_1 of less than 10 units per unit time while x_5 is present with a concentration of between 10 and 20, and changes linearly $\left(\frac{d^2x_5}{dt^2} = 0\right)$ at a rate of less than 10 units per unit time? Meanwhile x_3 is known to have a positive concentration of less than 20, while that of x_2 and x_4 is between 10 and 1000. The way that these last three variables change with time is not known.

Solving the model with this set of values gives the solution:

	x	$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$
x_2	$[20, 100]$	$[20, 100]$	$[0, -10]$
x_3	$[10, 20]$	$[500, 1000]$	$[-20, -100]$
x_4	$[500, 1000]$	$[-500, -1000]$	$[500, 1000]$

which gives us a reasonably detailed idea of what values the substrate concentrations should have, and less detailed but still useful information on how they will change over time. The solution is also a definite improvement on the information we obtained from the qualitative solution of the same model, though, as one would hope, both solutions are the same when only the signs of the values are considered. In fact, in general, there will be a number of semiqualitative solutions for every qualitative one each with slightly different interval values for variables and their derivatives.

The analysis may be refined. For instance, if we want to further investigate the value and first derivative of x_2, say, we could choose a new set of intervals, choosing the upper limit of the third positive interval to be 50 instead of 100 as in Figure 6. With the same set of constraints as before the following solution is generated:

	x	$\frac{dx}{dt}$	$\frac{d^2x}{dt^2}$
x_2	$[20, 50]$	$[20, 50]$	$[0, -10]$
x_3	$[10, 20]$	$[500, 1000]$	$[-20, -50]$
x_4	$[500, 1000]$	$[-500, -1000]$	$[500, 1000]$

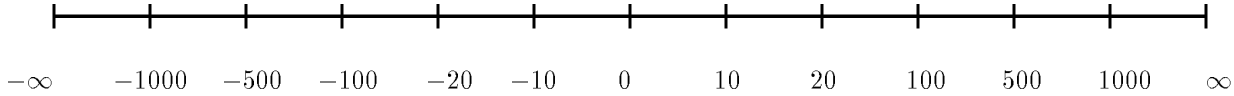

Figure. 5. A second set of semiqualitative intervals

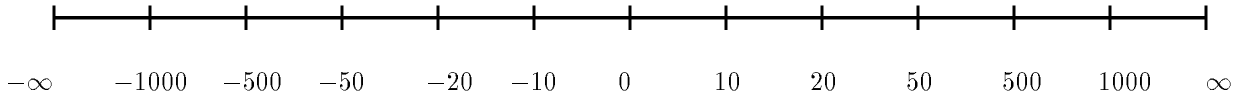

Figure. 6. A third set of semiqualitative intervals

Identification	Decription
$M1$	Addition
$M2$	Multiplication
$M3$	Derivative
$M4$	$X > Y$
$M5$	$DX > DY$
$M6$	$DDX > DDY$

Table 4. Functional blocks: DX represents $\frac{dx}{dt}$*, and DDX* $\frac{d^2x}{dt^2}$

which shows that by making certain intervals narrower, it is possible to make the solution more accurate, in that the intervals in the solution become narrower also. This process could, of course, be repeated. We could split the interval $[20, 50]$ into $[20, 35]$ and $[35, 50]$ in order to further narrow down the value of x_2, or split $[500, 1000]$ into $[500, 750]$ and $[750, 1000]$ to get a better idea of the value of x_4.

5. A PROGRAM FOR SEMIQUALITATIVE ANALYSIS

5.1. An introduction to the program

Having seen the kind of results that semiqualitative analysis can generate, we consider a software system which can perform a semiqualitative analysis on a set of differential equations. By restricting the set of intervals to three, it can also be used for qualitative analysis—indeed this was how the results presented in Section 2.3 were produced. In order to explain what the program does, and how it is operated, we discuss in detail the process of solving the example discussed in Sections 2.3 and 4.2.

The program works by considering the relations between the semiqualitative variables of the set of equations as a series of constraints upon their value. The analysis then consists of taking the known values, and propagating these through the network of constraints, seeing how they affect those values that are initially undefined. In order to do this the program needs some way of specifying the relationships between variable values, and this is done by means of a series of functional blocks. (Table 4). Each block in the semiqualitative system has one or two inputs, and a single output, and specifies that a particular relationship holds between the inputs and the output. For instance the $M1$ block has two inputs and a single output, and specifies that the output must be equal to the semiqualitative sum of the inputs. In some ways to talk of input and output is a little misleading, since the propagation need not take place from input to output. Indeed what happens is that a graph is constructed whose arcs are semiqualitative variables and whose nodes are functional blocks, and the known constraints propagated around until the values are as refined as possible. In this use of functional block and the propagation of constraints on the values of variables, the program can be seen to be a generalisation of the systems described by Kuipers (1986) and de Kleer and Brown (1984) to a full set of semiqualitative values.

5.2. Rewriting the model

The first stage in the analysis is to write the equations that describe the model in form in which they may easily be specified using the functional blocks. Initially they are written as a series of variables related only by addition and equality. There is no subtraction block since subtraction causes problems in interval arithmetic, and any equation written using subtraction may be rewritten using addition. This generates a new set of equations:

$$x_6 + x_{11} + x_{19} = x_{12} \quad (26)$$
$$x_7 + x_{12} = x_{14} \quad (27)$$
$$x_{15} + x_{16} = x_8 \quad (28)$$
$$x_9 + x_{16} = x_{19} \quad (29)$$
$$x_{10} + x_{17} + x_{19} = x_{18} \quad (30)$$

which may be directly written down in terms of functional blocks. There are further equations which relate the variables in the above equations to each other and the variables whose values are specified in the query:

$$x_6 = \frac{dx_1}{dt} \quad (31)$$
$$x_7 = \frac{dx_2}{dt} \quad (32)$$
$$x_8 = \frac{dx_3}{dt} \quad (33)$$
$$x_9 = \frac{dx_4}{dt} \quad (34)$$
$$x_{10} = \frac{dx_5}{dt} \quad (35)$$
$$x_{11} = 6.5x_1 \quad (36)$$

$$x_{12} = 3.0x_2 \quad (37)$$
$$x_{13} = 100x_1 \quad (38)$$
$$x_{14} = 1.5x_1 \quad (39)$$
$$x_{15} = 5.0x_1 \quad (40)$$
$$x_{16} = 1.0x_4 \quad (41)$$
$$x_{17} = 0.3x_5 \quad (42)$$
$$x_{18} = 0 \quad (43)$$
$$x_{19} = x_{13}x_5 \quad (44)$$
$$x_{20} = x_{19} + x_{11} \quad (45)$$
$$x_{21} = x_{19} + x_{17} \quad (46)$$

Having done this, it is simple to transform the set of equations to a network of functional blocks. At this stage it is easy to verify that none of the variables are redundant—if every variable connects into the network, as it does in this case, every variable plays an important part in the model. The network may also be used to determine if the model consists of a number of autonomous sub-systems. If it does, the network will be made up of a number of separate sub-networks.

The full set of equations, written in terms of functional blocks, form one part of the input to the program. The second part of the input is the semiqualitative query mentioned above, which sets the limit on the values of the variables in the original set of equations, in this case variables x_2–x_4. These limits may be any pair of boundaries of the semiqualitative intervals allowing composite intervals to be used as values. These boundaries themselves form the third part of the input. The fourth and final part of the input is a list of variables whose value are required in the output. In the example, since we are interested in the values of x_2–x_4 this part of the input will contain the names of these variables and their first and second derivatives, because we want to know the value of all three.

5.3. What the program does

The first thing that the program does is to compile a set of combinator tables from the set of semiqualitative intervals, a measure designed to make it more efficient when it comes to applying the mathematical constraints. These tables are compiled by considering the pairwise combination of every possible set of intervals. Thus for the set of intervals in the example it would first consider adding $[1000, \infty]$ to $[1000, \infty]$, which gives $[1000, \infty]$ since the addition of two numbers in the interval cannot lie outside the interval. Next it would try $[1000, \infty]$ and $[500, 1000]$ which would again give $[1000, \infty]$, and continue until it had added every interval to every other, including adding compound intervals such as $[20, 500]$ and $[-10, 10]$ to get $[10, 1000]$. Clearly this is a lengthy process, and provides a powerful argument for reducing the number of semiqualitative intervals, or at least keeping it as small as possible, but when it is complete the program can construct a look-up table which will provide very swift arithmetic operations later in execution. The program then assembles similar lookup tables for the other functional blocks.

Next the program decides on an order in which to test the values of the variables. This is done in such a way that the most constrained variable has its value propagated first, so that once its value is established the conceivable values of all the related variables may be evaluated as swiftly as possible. Again, establishing this order takes a little time, but it pays dividends in the long run.

After these two steps the program begins the process of propagating the constraints, essentially following the algorithm of Figure 7, for each and every variable, where N is the number of levels of derivative of the variable in question (so $N = 3$ in our example since we have value, first derivative and second derivative), i is the current level of derivative under consideration, M is the number of semiqualitative intervals that the value of the ith derivative can take on, and k is the index of the current value that is being considered for the ith derivative. Finally, after applying all the constraints to all the variables and establishing their possible value, the system outputs a list of all the possible interval values of all the derivatives of all the variables listed in the final part of the input.

6. CONCLUSIONS

This paper has described the use of qualitative and semiqualitative analysis, two methods from artificial intelligence that may be usefully applied to a wide range of environmental problems. Indeed they may be applied to any problem for which an equation based model may be written. The methods combine the strengths of the human ability to reason about the qualitative behaviour of systems, exemplified by statements such as "if volume decreases then pressure must increase", with the use of whatever numerical data is available. As a result the methods can be used to model complex physical systems for which sets of differential equations may be written, but for which the exact values of numerical constants are not known. Clearly the methods are not capable of providing exact answers when working with inexact data, that is impossible, but they do permit the most exact possible answers to be established from the available information. The use of the methods was illustrated on a particular model of the behaviour of an anaerobic fermentor, and the use and operation of a program that can perform both qualitative and semiqualitative analysis was discussed.

REFERENCES

Ardizzone, E.; Bonadonna, F.; Gaglio, S.and Nicolini, C.; Ruggiero, C.; and Sorbello, F. 1988. Qualitative modelling of cell growth processes. *Applied Artificial Intelligence* 2:251–263.

Bailey, J. E., and Ollis, D. F. 1986. *Biochemical Engineering Fundamentals*. New York: McGraw Hill.

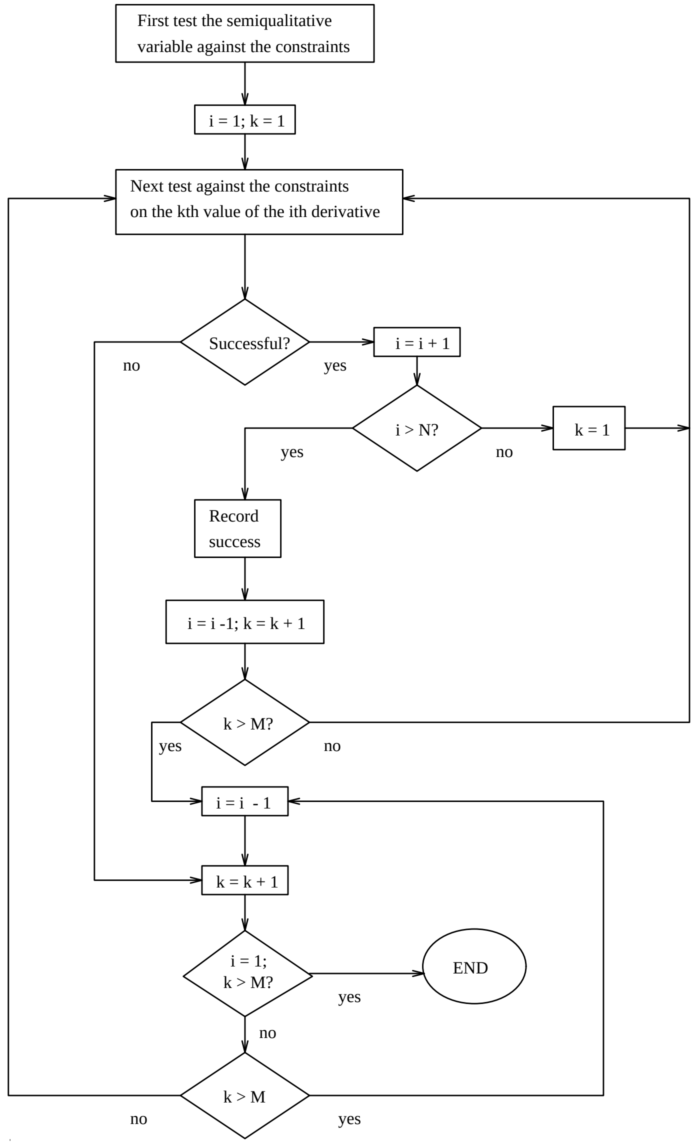

Figure. 7. The basic algorithm used by the semiqualitative program

Baldwin, J. F.; Martin, T. P.; and Zhou, Y. 1993. A Fril knowledge base for the management of uncertainty in performance assessment of hazardous waste repositories. ITRC 188, Advanced Computing Research Centre, University of Bristol.

Barrow, H. 1984. VERIFY: a program for proving correctness of digital hardware. *Artificial Intelligence* 24:437–491.

Cohn, A. G. 1989. Approaches to qualitative reasoning. *Artificial Intelligence Review* 3:177–232.

Coiera, E. 1992. The qualitative representation of physical systems. *The Knowledge Engineering Review* 7:55–77.

Dague, P.; Raiman, O.; and Devès, P. 1987. Troubleshooting: when modelling is the trouble. In *Proceedings of the 6th National Conference on Artificial Intelligence.*

Dague, P. 1993a. Numeric reasoning with relative orders of magnitude. In *Proceedings of the 11th National Conference on Artificial Intelligence.*

Dague, P. 1993b. Symbolic reasoning with relative orders of magnitude. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence.*

Davis, R. 1984. Diagnostic reasoning based on structure and behaviour. *Artificial Intelligence* 24:347–410.

Davis, E. 1990. *Representations of commonsense knowledge.* San Mateo, CA.: Morgan Kaufmann.

de Kleer, J., and Brown, J. S. 1984. A qualitative physics based on confluences. *Artificial Intelligence* 24:7–83.

de Kleer, J. 1984. How circuits work. *Artificial Intelligence* 24:205–280.

Dohnal, M. 1991. A methodology for common-sense model development. *Computers in Industry: Forum on Fuzziness* 24:1–18.

Dormoy, J.-L. 1988. Assembling a device. *Artificial Intelligence in Engineering* 3:216–226.

Dubois, D., and Prade, H. 1991. Semantic considerations on order of magnitude reasoning. In Singh, M. G., and Travé-Massuyés, L., eds., *Decision Support Systems and Qualitative Reasoning*. Amsterdam: Elsevier Science Publishers.

Falkenheimer, B., and Forbus, K. D. 1987. Setting up large-scale qualitative models. In *Proceedings of the 6th National Conference on Artificial Intelligence*.

Farley, A. M., and Lin, K.-P. 1991. Qualitative reasoning in microeconomics: an example. In Singh, M. G., and Travé-Massuyés, L., eds., *Decision Support Systems and Qualitative Reasoning*. Amsterdam: Elsevier Science Publishers.

Féray Beaumont, S. 1991. Qualitative model-based reasoning: an application to distillation process. In Singh, M. G., and Travé-Massuyés, L., eds., *Decision Support Systems and Qualitative Reasoning*. Amsterdam: Elsevier Science Publishers.

Forbus, K. D. 1984. Qualitative process theory. *Artificial Intelligence* 24:85–168.

Genesereth, M. R. 1984. The use of design desciptions in automated diagnosis. *Artificial Intelligence* 24:411–436.

Hangos, K. M.; Csáki, Z.; and Bay Jørgensen, S. 1992. Qualitative model-based intelligent control of a distillation column. *Engineering Applications of Artificial Intelligence* 5:431–440.

Hayes, P. 1978. The naive physics manifesto. In Michie, D., ed., *Expert Systems in the Microelectronic Age*. Edinburgh: Edinburgh University Press.

Hayes, P. 1985a. Naive physics I: ontology for liquids. In Hobbs, J. R., and Moore, R. C., eds., *Formal Theories of the Commonsense World*. Norwood, NJ.: Ablex Publishing Corporation.

Hayes, P. 1985b. The second naive physics manifesto. In Hobbs, J. R., and Moore, R. C., eds., *Formal Theories of the Commonsense World*. Norwood, NJ.: Ablex Publishing Corporation.

Hurme, M.; Dohnal, M.; and Järveläinen, M. 1991. A method of qualitative economic optimisation and its application to a practical engineering problem. In Singh, M. G., and Travé-Massuyés, L., eds., *Decision Support Systems and Qualitative Reasoning*. Amsterdam: Elsevier Science Publishers.

Koivisto, R.; Dohnal, M.; and Likitalo, A. 1989. Deep and shallow knowledge integration: a case study of an ai diagnosis of a chemical reactor. In *Proceedings of the 2nd Scandinavian Conference on Artificial Intelligence*.

Kuipers, B., and Berleant, D. 1988. Using incomplete quantitative knowledge in qualitative reasoning. In *Proceedings of the 7th National Conference on Artificial Intelligence*.

Kuipers, B. 1984. Commonsense reasoning about causality: deriving behaviour from structure. *Artificial Intelligence* 24:169–203.

Kuipers, B. 1986. Qualitative simulation. *Artificial Intelligence* 29:289–338.

Kuipers, B. 1987. New reasoning methods for artificial intelligence in medicine. *International Journal of Man-Machine Studies* 26:707–718.

Liong, S. Y.; Chou, W. T.; and Law, C. C. 1991. An expert system for storm water management modelling and its application. *Engineering Applications of Artificial Intelligence* 3:367–375.

Mavrovouniotis, M. L., and Stephanopoulos, G. 1987. Reasoning with orders of magnitude and approximate relations. In *Proceedings of the 6th National Conference on Artificial Intelligence*.

Mavrovouniotis, M. L., and Stephanopoulos, G. 1988. Formal order-of-magnitude reasoning in process engineering. *Computers in Chemical Engineering* 12:867–880.

Mavrovouniotis, M. L., and Stephanopoulos, G. 1989. Order-of-magnitude reasoning with O[M]. *Artificial Intelligence in Engineering* 4:106–114.

Mohammed, J., and Simmons, R. 1986. Qualitative simulation of semiconductor fabrication. In *Proceedings of the 5th National Conference on Artificial Intelligence*.

Moore, R. E. 1966. *Interval Analysis*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Ormsby, A. R. T.; Hunt, J. E.; and Lee, M. H. 1991. Towards an automated fmea assistant. In Rzevski, G., and Adey, R. A., eds., *Applications of Artificial Intelligence in Engineering VI*. Computational Mechanics Publications.

Parsons, S., and Dohnal, M. 1993. Qualitative, semi-qualitative and interval algebras and their application to engineering problems. *Engineering Applications of Artificial Intelligence* 5:553–560.

Parsons, S. 1993. Using interval algebras to model order of magnitude reasoning. *Artificial Intelligence in Engineering* 8:87–98.

Raiman, O. 1986. Order of magnitude reasoning. In *Proceedings of the 5th National Conference on Artificial Intelligence*.

Serra, P.; Lafuenta, J.; Mareno, R.; de Prada, C.; and Poch, M. 1992. Development of a realtime expert system for wastewater treatment plants. In *Proceedings of the IFAC/IFIP/IMACS International Symposium on AI in Real Time Control*.

Steyer, J. P.; Pourciel, J. B.; Simoes, D.; and Uribelarrea, J. L. 1991. Qualitative knowledge modelling used in a real time expert system for biotechnological process control. In Singh, M. G., and Travé-Massuyés, L., eds., *Decision Support Systems and Qualitative Reasoning*. Amsterdam: Elsevier Science Publishers.

Steyer, J. P.; Queinnec, I.; and Simoes, D. 1992. Biotech: a real time application of artificial intelligence for fermentation processes. In *Proceedings of the IFAC/IFIP/IMACS International Symposium on AI in Real Time Control*.

Sticklen, J.; Kamel, A.; and Bond, W. E. 1991. Integrating quantitative and qualitative computations in a functional framework. *Engineering Applications of Artificial Intelligence* 4:1–10.

Travé-Massuyés, L., and Piera, N. 1989. The orders of magnitude models as qualitative algebras. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*.

Watts, P., and Knight, B. 1991. Fault diagnosis in aspex: an expert system for the control of the activated sludge process. *Engineering Applications of Artificial Intelligence* 4:151–155.

Weld, D., and de Kleer, J. 1990. *Readings in qualitative reasoning about physical systems*. San Mateo, CA.: Morgan Kaufmann.

Williams, B. C. 1984. Qualitative analysis of MOS circuits. *Artificial Intelligence* 24:281–346.