Interval Analysis in Automated Design for Bounded Solutions*

A. R. F. O'GRADY, I. D. L. BOGLE, E. S. FRAGA**

Department of Chemical Engineering, University College London, UK-WC1E 7JE London
e-mail: e.fraga@ucl.ac.uk

Received 11 April 2001

Jacaranda is an automated process synthesis package. It can be used to aid an engineer in making decisions in the early stages of process plant design. In order to search the solution space, variables such as pressures and component flow rates are mapped to discrete levels. This discretisation means that optimal structures may be missed. In addition, the user may have no idea of how finely to discretise a given variable. This paper describes the use of intervals to bound the possible values of discretised variables. The method was applied to a five-component separation problem using distillation. The pressure of streams and distillation units were set to be intervals. Interval arithmetic was used to bound the possible costs of unit designs. The list of best solution structures was determined for capital, operating and annualised cost criteria. It is demonstrated that the use of interval analysis within Jacaranda enables the user to ensure that the global optimum is present within the solutions obtained.

In the early stages of process plant design, an engineer may be faced with a large number of possible structures and alternative unit design parameters. The automated process synthesis package, Jacaranda [1], can be used to gain insight into the problem and identify the best flow sheets. It generates a list of the best solutions with respect to one or more criteria and is based on an implicit enumeration approach with dynamic programming. The system relies on discretisation for handling continuous variables such as pressure and component flow rates in streams.

The fact that these discretisations are necessary has several effects on the results. The solutions produced will consist of units and stream products each at one of the discrete levels set when the problem was formulated. The best solution for a given level of discretisation may not represent the globally optimum flow sheet structure in continuous space as potentially good values for the discretised variables may be missed between the discrete levels chosen.

Currently, the level of discretisation of streams and units is specified by the user. In some cases, user intervention may be appropriate. For example, some components may be more important than others for environmental or economic reasons. The base flow rate of these components would be set to lower values than the others. In addition, the engineer may know the pressure range of operation for distillation columns and base the range and level of discretisation on this knowledge. See *Laing* and *Fraga* [2] for a discussion on the iterative use of an automated procedure with particular emphasis on user interaction.

Typically, however, the engineer may have no insight on the level of discretisation required for a given problem. Furthermore, the solutions generated give no indication of the effect of the discretisation on the effectiveness of the search procedure. This paper describes a method of bounding the objective function values produced, and used for ranking the solutions, by Jacaranda.

THEORETICAL

Interval Analysis

The need for bounds on the solutions is as a result of discretisations made from continuous to discrete space. For example, the problem may involve separation by distillation. A solution may describe a flow sheet structure containing a series of distillation columns each at a discrete operating pressure. A range of pressures between 0.1 and 1 MPa with 10 discrete pressure alternatives could be discretised uniformly. The discrete values of pressure would be accurate to the nearest bar. An alternative at 0.2 MPa would actually represent a range of values with a lower bound of

*Presented at the 28th International Conference of the Slovak Society of Chemical Engineering, Tatranské Matliare, 21—25 May 2001.

**The author to whom the correspondence should be addressed.

1.5 MPa and an upper bound of 2.5 MPa. The same applies to all discretised variables, both in streams and units. It follows that the accuracy of the objective function value obtained will be limited by the discretisations.

Interval methods have been used previously within process engineering. They have been applied in order to find all roots to a problem with mathematical certainty [3]. This was tested on several chemical engineering simulation problems. Interval analysis has also been applied to the global optimisation of selected flow sheets [4].

The framework of Jacaranda allows any unit model to be written as long as it accepts a process stream and outputs another with information on pressure, phase, and composition. The generic nature of the framework allows us to consider the use of models that work with intervals instead of just real numbers. This section describes how interval methods have been used to bound the effects of discretisation on the pressures of streams and distillation units. The results in the following section will show how the method has been applied to a 5-component separation problem described by *Rathore et al.* [5].

Interval Arithmetic

Interval mathematics was first introduced by *Moore* [6]. An interval is a closed bounded set of real numbers, $X = [a, b]$, where $a \leq x \leq b$. An interval of zero width (*i.e.* with the same values for both lower and upper bounds) is called a *degenerate* interval. In the discussion that follows, interval variables will be denoted by upper case letters and real variables by lower case letters. The bounds of the interval are shown by square brackets enclosing the real lower bound followed by a comma and then the real upper bound *e.g* $[a, b]$.

A set of arithmetic operations can be defined for intervals that correspond to the operations on real numbers. If X and $Y$are both intervals, X or Y will yield an interval containing every possible number that can be calculated resulting from the operation of each $x \in X$ on each $y \in Y$. A set of rules for interval arithmetic have been developed from this definition [7].

Interval Functions

An interval function will yield an interval when applied to one or more interval arguments. An interval function, F, is said to be an interval extension of a real function, f, if

$$F(x) = f(x) \quad \forall x \ni \Re \tag{1}$$

If the arguments of F are degenerate intervals, then F is an interval extension of f if they are equal when evaluated.

The *natural* interval extension of a function f is to replace the variables of the real function with interval variables but there are, in fact, an infinite number of interval extensions of a function.

An interval function is said to be *inclusion monotonic* if $X_i \subset Y_i$, $i = 1, \ldots, n$ implies that

$$F(X_1, \ldots, X_n) \subset F(Y_1, \ldots, Y_n) \tag{2}$$

Interval functions, containing a sequence of interval addition, subtraction, multiplication, and division operators, are inclusion monotonic [7] if the interval extension retains the same form when evaluating X and Y. An example of this is presented in [8].

If an interval function, $F(X_1, \ldots, X_n)$, is an inclusion monotonic interval extension of a real function $f(x_1, \ldots, x_n)$, then $F(X_1, \ldots, X_n)$ contains all the possible values of $f(x_1, \ldots, x_n)$, for all $x_i \in X_i$ $(i = 1, \ldots, n)$ [7]. This result will prove useful in bounding the value of the global optimum in an optimisation procedure.

Dependency

The interval returned by an interval function depends on the form that the function takes. For example,

$$F_1(X) = X^2 - X - 3$$
$$F_2(X) = (X - 1/2)^2 - 3\frac{1}{4}$$

are both interval extensions of

$$f(x) = x^2 - x - 3$$

yet they do not yield the same result when evaluated

$$F_1([1,\ 2]) = [-4,\ 0]$$
$$F_2([1,\ 2]) = [-3,\ -1]$$

F_2 produces sharper bounds for the range of f over the interval [1, 2] than F_1. This is due to the *dependency* phenomenon associated with interval arithmetic. Generally, the more often a given variable occurs within a function, the wider the bounds become. In fact, F_2 yields the exact range of f for $X = [1, 2]$ as X only occurs once in the function. When evaluating interval functions, dependency should be kept to a minimum so as to keep the bounds as sharp as possible.

Thick and Thin Functions

The term "parameter" will be used to refer to the constant values, either real or interval, within a function. The argument is the value of the function variable at which the function is evaluated. A *thick* function has interval valued parameters whereas a *thin*

function has only real valued (or degenerate interval valued) parameters. A thin interval function evaluated on a degenerate interval argument will return a degenerate interval; a thick function would return an interval value. The procedures described below will all deal with thick functions.

Interval Analysis within Jacaranda

The motivation for using interval analysis within Jacaranda is to generate upper and lower bounds for the optimisation criteria values obtained. These bounds are required because the optimisation procedure within Jacaranda is based on discretisation. Streams are mapped to discrete space through the discretisation of the pressure of the stream and by mapping the component flow rates to an integral multiple of some base flow rate. The discrete mappings of streams is used by the dynamic programming aspects of the Jacaranda system which allows for the efficient re-use of computation. Unit design parameters are also discretised. The combinations of the different discrete values for the parameters define the design alternatives available to the search procedure.

In the first instance, we have decided to consider interval analysis to analyse the effect of discretisation on the pressure, both of streams and as unit operating conditions. Pressure is chosen because there is no *a priori* discretisation available to the user and, hence, the choice of discretisation parameters is arbitrary.

New stream and distillation models have been implemented that can handle and manipulate pressure intervals rather than single real values. Each distillation column design produces a lower, nominal and upper bound on the objective function which may be, for example, capital cost. The nominal value is based upon the discrete real values of the feed stream pressure and the distillation column operating pressure. The lower and upper bounds are determined by solving the model using interval arithmetic with the pressure as an interval, in both streams and columns.

The Distillation Unit Model

The distillation model is based upon the *Fenske* [9],

$$N_{\min} = \frac{\log \dfrac{y_{\mathrm{lk}} x_{\mathrm{hk}}}{x_{\mathrm{lk}} y_{\mathrm{hk}}}}{\log \dfrac{\alpha_{\mathrm{lk}}}{\alpha_{\mathrm{hk}}}} \tag{3}$$

Underwood [10],

$$\sum_{i=1}^{n} \frac{\alpha_i x_{\mathrm{f}i}}{\alpha_i - \theta} + q - 1 = 0 \tag{4}$$

and *Gilliland* [11]

$$R_{\min} + 1 = \sum_{i=1}^{n} \frac{\alpha_i y_i}{\alpha_i - \theta} \tag{5}$$

correlations. Capital and operating cost models are provided by *Rathore et al.* [5].

The column model assumes semi-sharp separation. Non-key components pass completely into the top and bottom products. The key components are split according to the fractional recovery specified. This was set to be 98 % in all cases. The component flow rate discretisation is set to 10 % of the component flow rates in the feed stream. As a result, the semi-sharp column acts as a sharp separator when the discretisation procedures are applied to the outputs of the units. Product tanks accept streams that are over 90 % pure in any of the components. This is consistent with the level of component discretisation.

Heat exchangers are costed based upon the heat transfer area required. Continuous utilities are available [5]. A constant temperature difference of 8.5 K between utilities and the process streams is assumed in order to calculate the exchanger area.

The unit model is presented with a feed at a pressure within a certain interval. Each particular distillation unit alternative also has a pressure range associated with it. The design generated yields interval values for the height, diameter, and heat exchanger areas for all possible stream and distillation pressures by the application of interval arithmetic in the design calculations.

The Root of a Thick Interval Function

It is necessary to solve eqn (4) to determine θ, a value between the relative volatilities of the keys. If interval analysis is used, this value is itself an interval, Θ. This is not only because the relative volatility of component i, α_i, varies with pressure but also because the column is to be designed for a pressure interval. The parameter q is a measure of the fraction of the feed that is vapour. This is calculated by comparing the enthalpy of the feed at its current pressure to its enthalpy at the column pressure. Since both the feed and the column are within certain pressure intervals, the enthalpy will also be an interval. Q will contain the range of possible real values of q. Note, the range of α is not as sharp as theoretically possible because of dependency due to the interaction of pressure intervals. This will be discussed further in the next section.

The evaluation of Q is a thick function. The solution is obtained by a bisection method. This is the most convenient option as it is known that Θ must be between the relative volatilities of the keys and it is assumed that there is only one root between these bounds. Real values of θ within the root interval, Θ must meet the criterion that $f(\Theta) \cap [-\varepsilon, \varepsilon] \neq 0$. This

means that for any value of θ, there is at least one possible design for some combination of real values within the Q and α intervals.

RESULTS

A Separation Case Study

The problem attempted [5] is the separation of a 5-component hydrocarbon mixture. This problem has previously been attempted using Jacaranda [5] but the results obtained gave no indication of the effect of discretisation or certainty that the global optimum lay within the bounds of the discrete structure obtained. The operating pressure for the distillation units was in the range of 0.1 and 3.2 MPa. The number of discrete levels for the operating pressure was varied between runs and the effect on the criteria values noted. The criteria for optimisation were capital, operating and annualised (capital amortised over two years) costs. For each criterion, solutions were obtained by ranking according to the value of the lower bound and also the nominal bound.

Fig. 1 shows the capital cost of the three best solutions ranked according to the lower bound on the capital cost. The positions of bars on each line correspond to lower, nominal, and upper values of cost. The costs of the solutions are shown for various degrees of discretisation. As the number of discrete points increases the bounds on the solution become tighter.

It is important to stress that these bounds are not as tight as possible due to dependency. Nevertheless, the bounds strictly contain all the possible values of the objective function for a particular solution as a function of the operating pressure intervals chosen by the distillation units. This bounding information can be used to determine whether a particular flow sheet

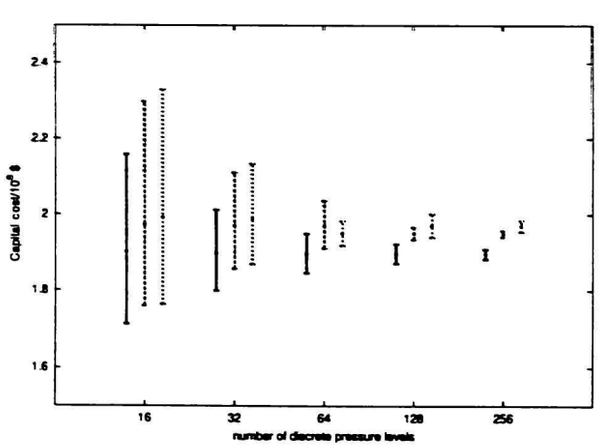

Fig. 1. Top three solutions ranked according to the lower bound for capital cost. 1st best – solid lines, 2nd best – long-dash lines, 3rd best – short-dash lines.

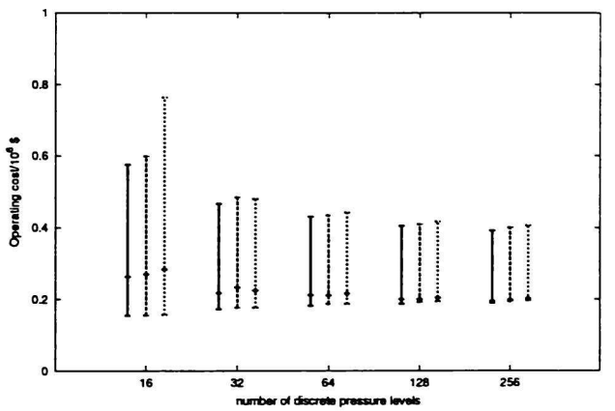

Fig. 2. Top three solutions ranked according to the lower bound of the annual operating cost. The meaning of lines as in Fig. 1.

structure could potentially include the global optimum solution. Note, this is only true if we assume that the other discretisations performed by Jacaranda have a negligible effect on the objective function value. Although not true in general, for this problem the effect of pressure discretisation is much larger than the effect from other discretisation variables and, furthermore, the discretisation parameters for the other variables can be chosen with more confidence through the use of engineering insight.

Fig. 2 shows the bounds on the annual operating cost for solutions ranked according to the lower bound on the operating cost. As with capital cost, increasing the level of discretisation sharpens the bounds. As a percentage of the nominal value, the bounds on operating cost are much wider than those of capital cost. In particular, the difference between the upper and nominal values is large compared to the difference between the lower and nominal values due to the equations used to determine heating and cooling requirements. Specifically, the large bounds are due to dependencies in the calculation of the heat balance around the column. The enthalpies of the streams are intervals as they are functions of the operating pressure. However, the feed enthalpy is a function of the stream pressure. The combination of these leads to large bounds.

Jacaranda currently discretises variables and determines a solution in terms of discrete values. This gives no assurance that optimal solutions are not missed by overly coarse discretisation. Bounded results can provide this assurance: if the upper bound of the objective function value of the best solution value is smaller than the lower bound of the second ranked solution, the global minimum must be the structure represented by the best solution, with operating conditions within the interval used by that solution. This is a useful result as it allows us to identify the discretisation level to use to generate the best solution. However, even more information can be gleaned from

the nominal value (calculated from the stream and unit pressure mapped to real values).

Throughout these experiments, the discrete pressures allowed in streams and columns were kept consistent. The pressure of a stream leaving a column would not change due to the mapping to discrete space as it would already be at one of the stream pressure levels allowed. The nominal value of the unit's operating pressure is never mapped to another value, so the nominal value of an optimisation criterion is a feasible value. This is, of course, only true if there were no other discretisations but, as mentioned above, we have assumed that these other discretisations are negligible in comparison with the pressure discretisations.

The argument used above when comparing the upper bound of the best solution with the lower bound of the second best solution can also be applied using the nominal value of the best solution. As this nominal value is a guaranteed attainable value, it is an upper bound on the global optimum. Therefore, if the nominal value of the best solution is smaller than the lower bound of second ranked solution, the global optimum value must be between the lower and nominal values of the best solution. We can ignore the range of values above the nominal value for all solutions and Figs. 3 and 4 present the results for the three finest levels of discretisation in this light. Fig. 3 shows that, using 64 discrete pressure levels, the global minimum can be identified. With 256 discrete pressure levels, we can also distinguish between the 2nd and 3rd best solutions.

Fig. 5 shows the results using an annualised cost criterion, the sum of the operating cost with the capital cost amortised over two years. With 128 discrete pressure levels, the upper bound for the top ranked structure is smaller than the lower bound on the minimum of the second ranked structure. The best structure is therefore guaranteed to be the global optimum. Using 256 discrete levels, the second and third best solutions can be identified with certainty. Solutions were also ranked according to the nominal value of the three criteria used. For the finer, and generally for all, levels of discretisation, the structures obtained were ranked in the same order as was obtained by ranking according to the lower bound. The solutions were different in the choice of operating conditions.

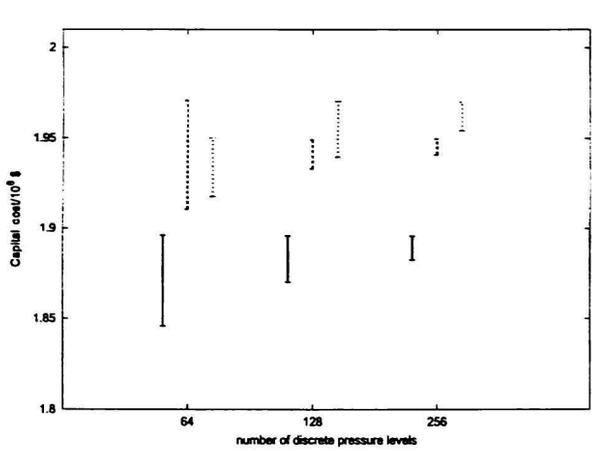

Fig. 3. Top three solutions ranked according to the lower bound of the capital cost, using the nominal value as an upper bound. The meaning of lines as in Fig. 1.

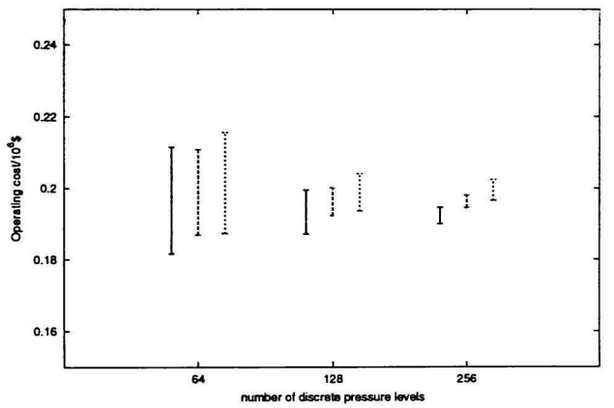

Fig. 4. Top three solutions ranked according to the lower bound of the operating cost using the nominal value as an upper bound. The meaning of lines as in Fig. 1.

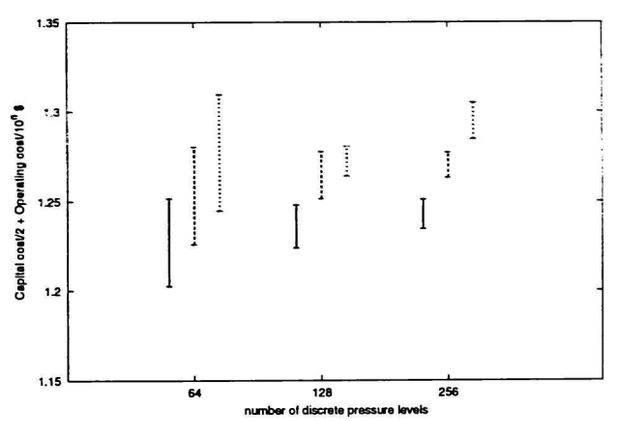

Fig. 5. Top three solutions ranked according to the lower bound of the annualized cost, using the nominal value as an upper bound. The meaning of lines as in Fig. 1.

CONCLUSION

The use of interval analysis with Jacaranda allows the user to gain some insight into the effect of discretisations on the solutions obtained. If a clear gap exists between the nominal value of one solution and the lower bound of a subsequent solution, there is a guarantee that the better solution is indeed better, at least with respect to the discretisation parameter investigated. The bounds generated for the optimisation

criteria are not tight in some cases. Nevertheless, they are valid bounds: the criteria values cannot be outside this range for a given structure and discretisation parameter values within the interval chosen. Furthermore, as the number of discrete levels increases, the bounds on the criteria values become sharper.

If the objective function is minimised for the lower bound, then the global minimum is bounded. If the stream and distillation model pressure discretisations are kept consistent, the presence of the global optimum can be assured. Its presence can be identified if the lower bound of the second best solution is greater than the nominal value calculated for the best solution. In effect, the nominal value is an upper bound on the minimum for that solution. Any uncertainty that discretisation of the parameter analysed may have caused another optimal solution structure to be missed is removed. If minimisation is carried out on the nominal value, the presence of a solution for this value of the objective function is assured. A lower bound, however, is not necessarily attainable due to dependency in the interval analysis. This issue will be addressed in future work to further sharpen the bounds on the solutions.

In terms of the optimum solution, some structures can be discounted on the basis of their bounds. In the future, this property may be used within an adaptive procedure to automatically set the level of discretisations used. Structures that did not contain the global optimum could be discarded before increasing the level of discretisation until one structure or a ranked list of best structures remained. In this work only one variable, pressure, was included in the interval analysis procedures. In the future, the effect of the discretisation of component flow rates and other unit operating conditions will be examined. When all discrete variables are incorporated into the interval analysis procedures, we will be able to guarantee the identification of the globally optimum flow sheet.

Acknowledgements. The authors gratefully acknowledge the funding provided by the EPSRC, Grant No. B/95/AF/2011 and GR/M83643, which made this work possible.

SYMBOLS

$N_{\min}$	minimum number of stages
x	mole fraction of liquid
y	mole fraction of vapour
α	relative volatility
θ	the root of eqn (*4*) between α_{lk} and α_{hk}
q	the ratio of the heat required to vapourise 1 mole of the feed to the molar latent heat of the feed
$R_{\min}$	the minimum reflux ratio
n	number of components

Subscripts

F	feed
lk	light key
hk	heavy key

REFERENCES

1. Fraga, E. S., Steffens, M. A., Bogle, I. D. L., and Hind, A. K., An object oriented framework for process synthesis and optimization, *Foundations of Computer-Aided Process Design.* (Malone, M. F., Trainham, J. A., and Carnahan, B., Editors.) AIChE Symposium Series *323* (96), 446 (2000).
2. Laing, D. M. and Fraga, E. S., *Comput. Chem. Eng. 21*, S53 (1997).
3. Schnepper, C. A. and Stadtherr, M. A., *Comput. Chem. Eng. 20*, 187 (1996).
4. Byrne, R. P. and Bogle, I. D. L., *Ind. Eng. Chem. Res. 39*, 4296 (2000).
5. Rathore, R. N. S., van Wormer, K. A., and Powers, G. J., *AIChE J. 20*, 491 (1974).
6. Moore, R. E., *Interval Analysis.* Prentice Hall, Englewood Cliffs, New Jersey, 1966.
7. Hansen, E., *Global Optimization Using Interval Analysis.* Dekker, New York, 1992.
8. Caprani, O. and Madsen, K., *Computing 25*,147 (1980).
9. Fenske, M. R., *Ind. Eng. Chem. 24*, 482 (1932).
10. Underwood, A. J. V., *Chem. Eng. Prog. 44*, 603 (1948).
11. Gilliland, E. R., *Ind. Eng. Chem. 32*, 1220 (1940).