Integrating Qualitative and Numerical models in binary distillation design

Nikitas M. Sgouros
Qualitative Reasoning Group
The Institute for the Learning Sciences
Northwestern University
1890 Maple Ave., Evanston IL 60201
E-mail: nikitas@ils.nwu.edu

Abstract

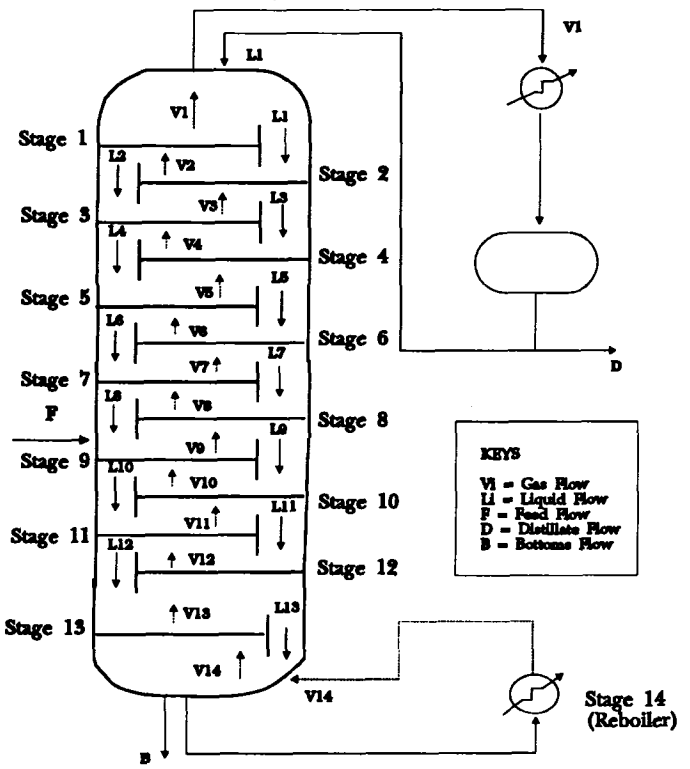
Recently qualitative reasoning about physical systems has been widely used in modelling and diagnosis tasks. In spite of this popularity there have been relatively few applications of it in design. This is unfortunate since QR techniques are capable of capturing the first principles on which design in a given domain is based and of guiding and monitoring decisions made during the design process. This paper describes how appropriate qualitative and numerical models are automatically selected and used in a system that focuses on chemical process design. The system applies a general propose-critique-modify method in order to solve problems relating to the design of binary distillation columns.

Introduction

Although QR techniques have been very popular in modelling and diagnosis tasks this has not been true for design. This is unfortunate since QR formalisms and methods have much to offer in design tasks, mainly because:

1. Formalisms for QR [4], [8] provide standardized languages that are capable of representing the first principles used in modelling physical systems. In addition, domain theories expressed in qualitative terms are flexible enough for use in more than one task (e.g. design, diagnosis, control).
2. QR provides ways to organize and access multiple, alternative models of physical systems based on the simplifying or operating assumptions that were introduced in each one of them [5], [1].
3. QR formalisms are able to model the effects that parameter value modifications introduce in a system and therefore are able to help design decisions. In addition, qualitative models can be used for monitoring design decisions and for detecting contradictions at the right time (e.g. values for parameters which are supposed to be always positive in the qualitative model but they aren't as a consequence of an erroneous design decision).

Figure 1: A typical distillation column



This paper describes how appropriate qualitative and numerical models are integrated in a system that focuses on chemical process design. The qualitative models that we use are based on Qualitative Process Theory (QPT) [8]. QPT is especially suitable for modelling chemical processes because its process-centered ontology is able to capture the first principles on which unit operations in chemical engineering are based [2].

The distillation process

Distillation is one of the most widely used separation processes in chemical engineering. It involves the separation of the components of a mixture based upon

differences in their tendencies to evaporate at a given temperature. In a binary (two-component) mixture the component with the highest tendency to evaporate is called the *volatile* component of the mixture. The other component is called the *non-volatile* component of the mixture. We will say that the vapor and liquid phases of a binary mixture are in *equilibrium* with each other if the temperature, pressure and chemical potentials for each component in the mixture are assumed to be uniform in both phases, i.e. these variables have no spatial or temporal gradients. This does not mean, however, that the liquid and vapor compositions for each component in the mixture are the same in equilibrium. In many cases the volatile component tends to concentrate in the vapor phase leaving the liquid phase richer in the non-volatile component. This provides the basis for separation in distillation.

Figure 1 depicts a typical 14-stage column for a continuous distillation process. In this example we have a binary mixture (*feed*) entering the column between stages 8 and 9. We also have a total condenser at the top of the column and a partial reboiler at the bottom. In each one of the stages in the column the liquid and vapor phases of a mixture come into contact, causing some of the liquid to evaporate and some of the vapor to condense. In the column the liquid flows down due to the force of gravity while the vapor flows upward under the force generated by a slight pressure drop from stage to stage. The vapor that reaches the top of the column is condensed in the total condenser and part of the resulting liquid (*reflux*) is returned back to the tower while the rest of it is retrieved as the distillate product. An analogous situation occurs in the partial reboiler where part of the liquid at the bottom of the column is retrieved as the bottom product of the process and the rest of it is vaporized in the reboiler and returned back to the tower. The net effect of the process is an increase in the concentration of the volatile component in the vapor and of the non-volatile component in the liquid.

In this paper we will describe how our design system is able to solve a common situation in binary distillation columns referred to as the *design* problem [9]. What is usually known in a design problem are the desired separation and the flow at some point in the column (usually the reflux). What is required is the number of stages and (probably) the location of the feed point in the column, i.e. the column to accomplish the given separation at a given reflux flow rate.

The Design system

Inputs

There are six inputs to the design system:

1. A domain theory describing the distillation process.
2. The numerical model corresponding to the domain theory given in (1).

```
(qpe::defEquation
 ;;; The name for the equation
 Stripping-and-Rectifying-Section-Liquid-Flows
 ;;; The model fragment on which it depends.
 ;;; This is actually a view in the domain theory.
 (Steady-State-Conditions-for-Feed-Stage
  ?column ?stage ?in-l-f ?out-l-f
  ?distillate-flow ?bproduct-flow ?feed-flow
  ?substance1 ?substance2 ?phase ?c-phase
  ?bproduct-phase ?c-stage ?r-stage)
 ;;; The quantities that take part in the equation
 ((Flow-Rate ?out-l-f) (Feed-Rate ?feed-flow)
  (Flow-Rate ?in-l-f))
 ;;; The qualitative relations that correspond to this equation
 ;;; (Optional).
 ((Qprop (Flow-Rate ?out-l-f)
         (Feed-Rate ?feed-flow))
  (Qprop (Flow-Rate ?out-l-f)
         (Flow-Rate ?in-l-f)))
 ;;; The actual numerical form for the equation.
 (= (A (Flow-Rate ?out-l-f))
    (+ (A (Feed-Rate ?feed-flow))
       (A (Flow-Rate ?in-l-f)))))
```

Figure 2: An example of an equation used in the numerical model for a binary distillation column. In this and the following figures: *?column* is the current distillation column, *?stage* is the current stage, *?in-l-f* is the liquid flow entering into the stage from above, *?out-l-f* is the liquid flow leaving the stage, *?distillate-flow*, *bproduct-flow*, *?feed-flow*, *?phase*, *?c-phase* and *?bproduct-phase* are the flows and the phases of the feed, the distillate and the bottom product respectively, *?c-stage* and *?r-stage* are the condenser and the reboiler of the column and finally *?substance1* and *?substance2* correspond to the components of the binary mixture processed in the column.

3. A set of equations that provide rough estimates for some of the process parameters.
4. A set of initial values for some of the parameters of the system along with an initial scenario description for the process.
5. A set of rules that describe procedures for modifying the scenario description according to the values of the parameters that describe the structure of the column.
6. A set of design heuristics that guide the design decisions of the system.

The following sections describe each input in more detail.

The domain theory The qualitative model that we use for design describes the steady-state operation of an ideal staged distillation column for binary mixtures. Steady-state operating assumptions are common in the early stages of chemical process design since they greatly simplify the numerical models for describing the process. This is true for the qualitative models as well. The domain theory that is used by the system is simpler than the one presented in [10] that describes

the dynamic behavior of binary columns.

The numerical model The numerical model for the behavior of the column at steady-state is derived from the model that describes the dynamic behavior of the process. Each of the equations in the model is attached to a model fragment (entity, view, process or perspective) in the domain theory and it becomes active whenever the later is active. This organization allows the system to select numerical models that are consistent with the qualitative theory used in describing the process. The parts of the qualitative theory that are active at any point in the design process are determined, in turn, by the simplifying or operating assumptions employed in the design task. This organization allows the system to access and organize its models in terms of the assumptions it uses. Figure 2 provides an example of the form used for representing every equation in the numerical model.

Whenever there is only one unknown in an equation the system tries to solve it by using a simple set of rules that describe general algebraic transformations. For the particular application we are describing this component does not have to be very sophisticated since most of the equations describing binary distillations are simple linear equations.

Approximations Most of the design problems are underconstrained. In these situations good estimates for the values of some of the unspecified parameters must be supplied to the system through the use of approximate expressions. Figure 3 provides an example of the form used in order to represent such an approximation equation to the system.

Similar to what is the case for the equations in the numerical model, each approximation equation is valid whenever the model fragments and the assumptions contained in its condition slot are valid. In addition, for every parameter approximated the system needs to know how to update its value, in case the estimate supplied by the equation is not good enough. This can range from a simple numerical value (e.g. the number of stages in the column in the current application) to a special heuristic procedure for estimating how much to adjust the initial estimate.

```
(qpe::def-First-Guess-Eqn
 ;;; The name of the equation.
 Gilliland-Fenske-Equation
 ;;; The parameter estimated by the equation. The particular
 ;;; parameter corresponds to the number of stages
 ;;; in the column.
 (Num-of-Stages ?column)
 ;;; The conditions under which the approximation is valid.
 ;;; The first condition is a perspective in the domain theory
 ;;; While the second one is a simplifying assumption for the
 ;;; process.
 ((Steady-State-Column-Design-Features
   ?column ?c-stage ?c-phase ?r-stage ?r-phase
   ?distillate-flow ?bproduct-flow ?substance1
   ?substance2)
  (Constant-Relative-Volatility
   ?substance1 ?substance2 :in ?column))
 ;;; The quantities that take part in the equation.
 ((Mole-Fraction
   ?substance1
   (2-C-S (?substance1 ?substance2) ?c-phase ?c-stage))
  (Mole-Fraction
   ?substance1
   (2-C-S (?substance1 ?substance2) ?r-phase ?r-stage))
  (Constant-Alpha ?column))
 ;;; The numerical form for the approximate expression.
 (= (A (Num-of-Stages ?column))
    (* 2.
       (/ (log
            (* (/ (A (Mole-Fraction
                      ?substance1
                      (2-C-S (?substance1 ?substance2)
                      ?c-phase ?c-stage)))
                  (- 1.
                     (A (Mole-Fraction
                         ?substance1
                         (2-C-S (?substance1 ?substance2)
                         ?c-phase ?c-stage)))))
               (/ (- 1. (A (Mole-Fraction
                            ?substance1
                            (2-C-S (?substance1 ?substance2)
                            ?r-phase ?r-stage))))
                  (A (Mole-Fraction
                      ?substance1
                      (2-C-S (?substance1 ?substance2)
                      ?r-phase ?r-stage))))))
          (log (A (Constant-Alpha ?column)))))))
```

Figure 3: An example of an approximation equation used by the system for estimating the number of stages in a binary column. The function 2-C-S denotes a binary mixture as follows:

2-C-S: (component1, component2) × phase × container ⟶
⟶ Contained-Binary-Mixture

Initial scenario & parameter values This part of the input contains the modelling assumptions that the user is willing to make during the design process, the parameters he is interested in obtaining values for, along with initial values for some of the parameters of the system.

Procedures for modifying the scenario Whenever the parameters that describe the structure of the system (such as the number of stages or the location of the feed in a column) have their values changed, the system needs to modify its current scenario description in order to reflect these changes. This is done using a set of rules that have as conditions the structural features of particular types of columns and result in a set of procedures for updating the current structure.

Design Heuristics These, for the time being, are a set of rules that determine the location of the feed in a column. Each heuristic belongs to a certain class depending on the the kinds of design decisions it guides. For example, all the heuristics that we currently use belong to the same class since they all deal with determining the location of the feed in a column. Only one heuristic per class can be used in the course of deriving a particular design. This solves the problem of using conflicting heuristic rules during the design process. The system tries to apply all of the heuristics in each

class backtracking at each decision point if necessary. In this way a list of alternative designs is generated.

The design method

The design method used by the system is an instance of a general propose-critique-modify (PCM) method [3] and can be described as follows:

1. Instantiate the domain theory and the numerical model that are consistent with the initial scenario description provided by the user.
2. Try to solve for as many as possible unknown parameters using the instantiated numerical model from (1) and the initial parameter values supplied by the user.
3. If there are no more unknown parameters then return with the results, otherwise the problem is underconstrained. In this case the following PCM method is used:

(a) Step 1 (Propose): The system applies its approximate expressions in order to propose values for some of the unspecified parameters. For example the initial specifications for the distillation design problem allow the system to use the Gilliland-Fenske equation in Fig. 3 to estimate the number of stages in a column that satisfy the separation requirements.

(b) Step 2: The system verifies the design proposal from step 1. First it constructs the appropriate scenario description for the proposed design using the procedures described in section 3.1.5. The modified scenario is used to instantiate and solve the appropriate qualitative and numerical models for the column. The solution set is then checked for consistency with the constraints imposed by the qualitative model on the parameter values. In addition, the estimates for the system parameters are checked for consistency with the values computed in (2). In case no errors are found, the system exits with success.

(c) Step 3 (Critique): If step 2 is unsuccesful, the system first tries to apply its design heuristics. In case any modifications to the current scenario or the current parameter values are proposed it proceeds to the next step. For example, in our particular distillation design problem the system applies at this step its design heuristics in order to determine the location of the feed stage in the column.

If no modifications have been suggested by the heuristics, the system assumes that the estimates for some of the system parameters were inacurrate. For every parameter whose value was found to be inconsistent, the system tries to find a causal path in the domain theory that relates it with one of the estimated parameters. It then suggests possible modifications for the estimate of the guessed parameter. For example, step 1 in our distillation design example, produced an estimate of 13 equilibrium stages for the column. After the qualitative and numerical models for a 13-stage column were solved, the proposed mole fraction for the volatile component at the reboiler was found to be less than the one computed in (2). After traversing the causal path that connects the number of stages with the mole fraction of the liquid at the reboiler, the system suggested to reduce the number of stages in the column in order to increase the mole fraction of the volatile component at the reboiler.

(d) Step 4 (Modify): The system uses the suggestions from step 3 to modify parameter values and/or the scenario description for the proposed design and returns to step 2.

In its current instantiation the system is allowed to estimate the value of only one parameter. More complex schemes for estimating multiple parameters can be found in [11].

Output

The system generates a list of designs that consist of a set of values for the column parameters and a scenario description that is consistent with these values. This scenario description can then be used with SIMGEN [6], [7] to generate simulators for the dynamic behavior of the proposed column.

Elapsed Real Time for the Design system (approximate)	1 hr 58′
Elapsed Real Time for Qualitative Analysis (approximate)	1 hr 45′
Number of scenario models tested in producing the final design proposal	4
Number of Quantities in the proposed design	314
Number of active Processes in the proposed design	26
Number of instantiated Views in the proposed design	139

Figure 4: Statistics for the distillation design example. The times shown correspond to the time spent in finding the first acceptable design in the problem.

Results

The program runs on an IBM RS/6000, Model 320, with 64MB of RAM running Lucid Common Lisp. The system was succesfully used to solve the design problem for binary distillation columns described in [9]. The results from the program (a column with 12 equilibrium

stages) were consistent with the ones found in [9]. As Figure 4 indicates most of the running time was spent in the qualitative analysis of the various column models that were proposed during the design process. The system uses the same type of qualitative analysis with SIMGEN [7]. The particular design example that we used is quite simple. Due to the fact that the system used only one heuristic for the location of the feed in the column, only one column design was proposed. We are currently testing the system in a series of other examples that are contained in the same book and are related to binary column design.

Discussion

Although our research is still in progress, we believe that there are some very interesting conclusions that can be drawn from it concerning the relations between QR and design tasks. Our current design system demonstrates that:

- Qualitative models can be succesfully used in guiding and monitoring design decisions.
- The ability to organise and access multiple, alternative models of physical systems based on the simplifying or operating assumptions that were introduced in each one of them gives QR significant advantages as a way of modelling physical systems for design tasks.
- Qualitative domain theories represent the physical principles involved in modelling a device using a task independent formalism. This makes them suitable for supporting multiple tasks (e.g. design, control etc).

We are currently extending our qualitative and numerical models in order to cover multicomponent distillation columns. We hope to use this expanded theory along with the current design system in a system able to do process synthesis and in particular synthesis of separation sequences for multicomponent mixtures. We believe that this task will help us investigate ways of efficiently combining heuristic and first-principles knowledge in design.

Acknowledgements

Thanks to Ken Forbus for useful comments. This research has been supported by the NASA Johnson Space Center.

References

[1] Addanki, S., Cremonini, R. and Penberthy, J. S. 1990. Reasoning about Assumptions in Graphs of Models. In *Readings in Qual. Reasoning about Physical Systems* by Weld, D. S. and deKleer. J., Morgan Kaufmann Inc.

[2] Catino, C. A., Grantham, S. D., Ungar, L. H. 1991. Automatic generation of qualitative models of chemical process units. *Computers chem. Engng.* 15: 583-599.

[3] Chandrasekaran B. 1990. Design Problem Solving: A Task Analysis. *AI Magazine*, vol. 11, no. 4.

[4] deKleer, J. and Brown, J. S. 1990. A Qualitative Physics Based on Confluences. In *Readings in Qual. Reasoning about Physical Systems* by Weld, D. S. and deKleer. J., Morgan Kaufmann Inc.

[5] Falkenhainer B. and Forbus, K.D. 1991. Compositional Modelling: Finding the Right Model for the Job. *Artificial Intelligence*, 51.

[6] Forbus, K. D. and Falkenhainer, B. 1990. Self-Explanatory Simulations: An integration of qualitative and quantitative knowledge. In *AAAI-90.*

[7] Forbus, K. D. and Falkenhainer, B. 1992. Self-Explanatory Simulations: Scaling up to large models. In *AAAI-92.*

[8] Forbus, K. D. 1984. Qualitative Process Theory. *Artificial Intelligence*, 24.

[9] King, C. J. 1971. *Separation Processes.* McGraw-Hill, Inc.

[10] Sgouros, N. M., Building a self-explanatory simulator for an ideal binary distillation column, Unpublished manuscript.

[11] Westerberg, A. W. et al. 1979. *Process flowsheeting.* Cambridge University Press.