Paper IV

Seuranen T., Hurme M., Pajula E., Synthesis of separation processes by case-based reasoning, *Comp. Chem. Eng*. 29 (2005) 1473-1482.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computers and Chemical Engineering 29 (2005) 1473–1482
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

Synthesis of separation processes by case-based reasoning

Timo Seuranen [a], Markku Hurme [a,*], Elina Pajula [b]

[a] *Helsinki University of Technology, Laboratory of Plant Design, P.O. Box 6100, FIN-02015 HUT, Finland*
[b] *KCL Science and Consulting, P.O. Box 70, FIN-02150 Espoo, Finland*

Available online 8 March 2005

Abstract

A new approach to separation process synthesis and selection of single separations is presented. The method is based on the reuse of existing design cases by case-based reasoning (CBR). CBR is a method for finding the most similar existing separation designs and for applying the knowledge of their concept to solve new problems. The method has previously been applied for selecting single separations and simple sequences but has now been extended to cover synthesis of more complicated systems. The method is mainly intended to screening feasible process alternatives in preliminary process design for more detailed study by simulation.

Keywords: Case-based reasoning; Separation process synthesis; Conceptual design

1. Introduction

The paper presents a method for finding feasible separation process sequences by using case-based reasoning (CBR). CBR is a method of reusing existing design cases for making new designs. This means finding most alike existing processes and applying the knowledge of their separation capacity and design for solving new design problems. This is especially important in the early phases of process design when many alternatives should be quickly screened before a more detailed study is done. There is a great need for this kind of screening tool to reduce the number of design options and quicken the process design, since the quick introduction of products into markets is of prime importance (Cordiner, 2001).

When dealing with multicomponent mixtures, the number of possible separation methods, their combinations and process structures to be screened is huge as well as the work involved. The problem has been discussed successfully as a MINLP problem; see for instance Grossmann, Aguirre, and Barttfeld (2004). Very much different type of approach is the CBR-based methodology, which tries to emphasize the practical experience from existing separation systems.

* Corresponding author.
E-mail address: markku.hurme@hut.fi (M. Hurme).

0098-1354/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2005.02.016

2. CBR in process synthesis

Case-based reasoning is one of the non-symbolic AI methods (Watson & Marir, 1994). CBR solves new problems by finding and adapting existing successful designs for solving new problems.

The lack of systematic reuse of existing design experience has been a shortage in process design. The main benefit of CBR approach is that readily available existing knowledge can be utilised systematically also in large and complex problems such as process synthesis and design. In this way, the time-consuming conceptual screening phase of a design project can be fastened. CBR is attracting attention, because it seems to directly address the conceptual process design problems outlined earlier.

Some benefits of the CBR approach are:

1. CBR does not require an explicit domain model and so elicitation becomes a task of gathering case histories.
2. Implementation is reduced to identifying significant features that describe a case, which is an easier task than creating an explicit model.
3. By applying database techniques, large volumes of information can be managed.
4. CBR systems can learn by acquiring new knowledge as cases, thus making maintenance easier.

Because generalisations are not needed in CBR, no data are lost. CBR gives answers to design problems in a straightforward way. The results are dependent on the retrieval parameters and the adaptation applied. The strong interaction with the user makes the flexible and interactive use of existing data and design experience possible. The CBR search can be focused on different aspects by defining new search criteria and weighting retrieval criteria differently. In this way, the same case base can be used for several types of tasks. The system learns by updating the information of the database.

As a general engineering design approach CBR has been used earlier (Maher & de Silva Garza, 1997), the applications to chemical engineering, however, have been quite few as discussed by Avramenko, Nyström, and Kraslawski (2004). The design of separations by CBR has been studied by King, Bañares-Alcántara, and Manan (1999), but they used the residue curve map approach and limited to azeotropic separations. The CBR-based selection of separation methods and the synthesis of simple separation sequences based on the characteristics of separated components has been discussed in earlier papers (Pajula, Seuranen, & Hurme, 2001; Pajula, Seuranen, Koiranen, & Hurme, 2001; Seuranen, Pajula, & Hurme, 2002) but the more comprehensive methodology has not been presented yet.

3. Case retrieval in CBR

A retrieval algorithm using the indices in the case–memory should retrieve the most similar cases to the current problem or situation. The retrieval algorithm relies on the indices and the organisation of the memory to direct the search to potentially useful cases. The issue of choosing the best matching case has been addressed by research on analogy (Watson & Marir, 1994). This approach involves using heuristics to constrain and direct the search.

Methods for case retrieval are nearest neighbour, induction, and knowledge-guided induction and template retrieval. These methods can be used alone or combined into hybrid retrieval strategies. If the nearest neighbour is used, then case features should be able to be weighted and similarity measures customised. If inductive techniques are used, the index tree generated should be open to inspection and alteration by developers. A typical equation for calculating nearest neighbour matching is Eq. (1), where w_i is the importance weighting of a feature i, n the number of features, sim the similarity function, f^{I} and f^{R} are the values for feature i in the input and retrieved cases, respectively.

$$\frac{\sum_{i=1}^{n} w_i \times sim(f_i^{\mathrm{I}}, f_i^{\mathrm{R}})}{\sum_{i=1}^{n} w_i} \quad (1)$$

4. The CBR-based separation synthesis algorithm

The main phases of general CBR-based separation process synthesis algorithm consist of: (1) selection of the methods of single separations; (2) selection of separation sequences; and (3) selection of combined (hybrid) separations. The phases of the algorithm are listed below and discussed in the following sections in more detail.

Proposed CBR-based synthesis procedure:

1. Selection of single separations
   (a) search for the feasibility of conventional distillation based operations;
   (b) search for azeotropes (see subcase; synthesis of azeotropic systems);
   (c) search for suitable mass separation agents (MSA);
   (d) search for other separation methods:
      (i) calculation of relative physical properties (R's);
      (ii) search for separations based on feasible relative properties.

1B. Selection of azeotropic separations
   (a) search for separation in column in isobaric conditions;
   (b) search for separation in columns in non-isobaric conditions;
   (c) separation by using MSA;
   (d) separation by using MSA and non-isobaric pressure;
   (e) separation by other means; reactive, membrane, extraction, etc.;
   (f) separation by hybrid or combined operations.

2. Separation sequencing by using as search criteria:
   - component names or types;
   - relative volatilities of components;
   - VF values of components (see Eq. (4));
   - coefficient of ease of separation (CES) values of components (Liu, McGee, & Epperly, 1987).

   And applying:
   (a) sequences in found cases; or
   (b) sequence heuristics (if they are stored with cases).

   For more details see discussion in Section 8.2.

3. Search for combined separation operations (see Section 9).

5. Selection of single separations

The feasibility of ordinary distillation is studied first in the methodology, since it is the most common way of separating fluid mixtures. This is done by studying the distillation related properties (relative volatilities and lacking chemical reactivities) in the first step. The remaining separation problems are solved with further reasoning, which applies separation methods other than ordinary distillation. In this phase relative properties are calculated and the values that show potential for separation are used as retrieval parameters. The main steps of the approach of selecting single separations are (Pajula, Seuranen, & Hurme, 2001):

Step 1. Conventional distillation is usually applied whenever the relative volatility (α) is large enough. The first search for the solution is made using component names, α and reactivities as retrieval parameters. To make the search simpler, α's can be classified as easy ($\alpha > 1.2$), possible, where mass separating agent (MSA) could be useful ($1.1 < \alpha < 1.2$) and difficult (1.1). A more accurate search is made (capacity and component types also as retrieval parameters), if several alternatives are found. The nearest strategy found is then applied in all the separations, where ordinary distillation is applicable.

Step 2. To be able to compare separation methods, where mass-separating agent is needed, suitable MSA is searched for each binary component pair that cannot be separated by conventional distillation. The retrieval parameters used are, e.g. the types of components, concentrations, relative solubility parameter, dipole moment and dielectric constant. The found MSA is used for defining solubilities and other separation related properties in Step 3.

Step 3. Relative physical property parameters (Jaksland, Gani, & Lien, 1995) are calculated for each component pair that cannot be separated by ordinary distillation. The parameter values are compared to the feasibility limits of different separation methods.

Step 4. Separation methods are searched using the relative parameters (minimum and maximum values), that are within the feasibility limits as retrieval parameters. For example, crystallisation is considered very feasible if the relative melting point is greater or equal to 1.2. Also a more detailed search by using concentration, capacity and component types as retrieval parameters can be defined.

The possibility of combined operations should be checked. This is done in the last phase of the main algorithm and discussed later.

6. Synthesis of azeotropic separations

The presence of azeotropes adds some difficulties to separations and also the synthesis problem becomes much more complex. Therefore, the synthesis of azeotropic distillations is a special case of the general synthesis algorithm.

In general to separate azeotropic mixtures various technologies may be used (Hilmen, 2000):

1. Pressure-swing distillation. A series of column operating at different pressures are used to separate binary azeotropes, which change appreciably in composition over a pressure range or where a separating agent, which forms a pressure-sensitive azeotrope, is added to separate a pressure-insensitive azeotrope.
2. In homogeneous azeotropic distillation, a third component is added to modify the components relative volatility.
3. Heterogeneous azeotropic distillation is based on the same principle as homogeneous azeotropic distillation, but the added third component is partially miscible with one of the components, solvent reprocessing is easy by means of a liquid–liquid separation system.
4. Reactive distillation is based on the transformation of one of the components into a component, which does not form an azeotrope with the other components.
5. Salted distillation consists in adding an ionic salt that dissociates in the liquid mixture and changes the azeotrope composition.

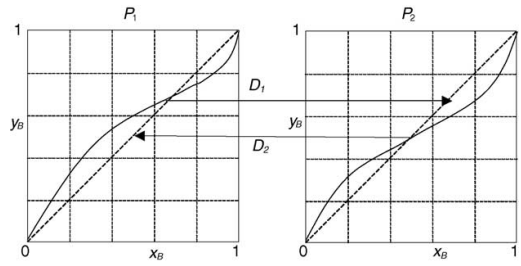

Fig. 1. y–x diagrams at different pressure ($p_1 < p_2$) for minimum-boiling azeotrope.

Pressure changes can have a large effect on the vapor–liquid equilibrium compositions of azeotropic mixtures and thereby affect the possibilities to separate the mixture by ordinary distillation. By increasing or decreasing operating pressures in individual columns the distillation boundaries can be moved in the composition space or the azeotropes can even be made to appear or disappear (Figs. 1 and 2).

The synthesis algorithm for azeotropic separations is a modification of the general synthesis algorithm (Seuranen et

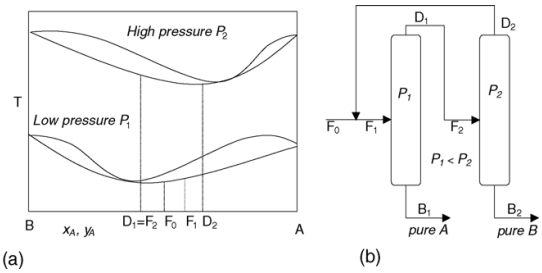

Fig. 2. Pressure-swing distillation: (a) temperature–composition diagram for a minimum-boiling binary azeotrope that is sensitive to changes in pressure and (b) distillation sequence.

al., 2002). The hierarchy of searches in the algorithm is the following:

1. separation in single or multiple columns in isobaric and non-isobaric pressure;
2. separation by using MSA;
3. separation by using MSA and non-isobaric pressure;
4. separation by other means: reactive, membrane, extraction, etc.;
5. separation by hybrid or combined operations.

When azeotropes are present in the mixture, the definition of case description and retrieval parameters is more complex. One idea is to use the relative similarity based on the similarity of feed, product and azeotropic points. To be able to compare separation methods, where mass separating agent is needed, a suitable MSA is searched for each binary component pair that cannot be separated by conventional distillation. The first retrieval parameters used are types of components to be separated. Also, a more accurate search is defined (concentrations, relative solubility parameter, polarity and dielectric constant as retrieval parameters), if several alternatives are found. The found MSA is used for defining solubilities and other separation related properties. If the MSA has not been used earlier for same components, more rigorous studies, simulations and/or experiments are needed to confirm the suitability. An example on searching a MSA is given later.

7. Examples on single separations

7.1. Separation of azeotropic solutions

As an example on single separations the separation of pyridine from water is studied as an example (Pajula, Seuranen, & Koiranen et al., 2001). Since the separation concept is determined by the concentration of product streams required, the composition of the azeotropic point and the solubility of the mixture (i.e. if there is a phase split), these are also the search criteria used in CBR. If the criteria are analogous to the case found in database, the process concepts are similar too and the concept found can be reused.

Problem. Dilute pyridine water solution needs to be separated into products containing 40 and 1 wt% of pyridine.

Search criteria:

(1) azeotropic composition of pyridine with water: 94 °C and 57 wt%;
(2) solubility with water: total;
(3) feed composition: 15 wt% pyridine;
(4) product compositions: 40 and 1 wt% pyridine.

Table 1
Query and nearest cases in the THF/water problem

	Query	Found 1	Found 2
Component 1 type	Water	Water	Water
Component 2 type	Ether	Ether	Acetate
Component 1	Water	Water	Water
Component 2	THF	Diethyl ether	Ethyl acetate
Solubility parameter	9.9	7.4	9.1
Dipole moment/*D*	1.75	1.3	1.7
Dielectric constant	7.6	4.34	6.02
MSA's		*n*-Hexane, benzene, toluene	*n*-Pentane, 2,2-dimethyl-butane, dichloromethane
Similarity		0.92	0.85

Using these parameters the nearest case found is:

Components:	Tetrahydrofuran (THF) and water
Azeotrope:	At 64 °C 96 wt%, THF
Solubility:	Total (at 70 °C)
Feed:	20 wt% THF
Products:	55 and 3 wt% THF
Separation:	Distillation in a single column without entrainer

The found case is analogous in the azeotropic behaviour and in relative stream concentrations (especially relative to the azeotropic point). In neither case, the azeotropic concentration is crossed. Based on this, it can be reasoned that distillation in a single column is applicable also in the pyridine case searched and no entrainer is required.

7.2. Selection of mass transfer agent

Problem. Find suitable mass separation agent for THF/water separation.

The search is made using following retrieval parameters: component type, solubility parameter, dipole moment and dielectric constant. These three describe solvent's separation capability (Pajula, Seuranen, & Hurme, 2001). The nearest cases are shown in Table 1.

The found MSA is used for defining solubilities and other separation related properties for Step 3. If the MSA has not been used earlier for exactly the same components, more rigorous studies, simulations and/or experiments are needed to confirm the suitability. In this case *n*-pentane has been reported for THF/water separation (Smallwood, 1993).

7.3. Separation in non-isobaric distillation system

Problem. Separate tetrahydrofuran (15 wt%) from water. Purity requirement for tetrahydrofuran product is 99 wt%.

This cannot be reached with ordinary distillation, because tetrahydrofuran and water form an azeotrope at 64 °C with 96 wt% THF, which has to be crossed (Smallwood, 1993).

Search criteria:

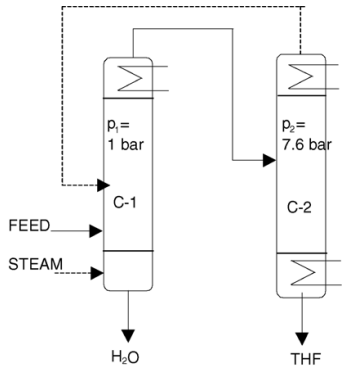

Fig. 3. Separation of THF/water mixture in two-column system.

(1) feed composition: 15 wt% THF;
(2) product composition: 99 wt% THF;
(3) use pressure–swing distillation;
(4) one or more columns;
(5) no MSA.

Using these parameters the nearest cases are shown in Table 2.

It is found that the THF/water system can be separated by shifting the azeotropic concentration by changing the system pressure in one or two-column systems (Fig. 3).

8. Synthesis of separation sequences

8.1. Approaches to synthesis of separation sequences

In the synthesis approach presented the selection of the separation sequence starts from choosing of a single separation as shown before. There are several alternatives for the selection of a separation sequence:

(1) Finding all possible separation sequence combinations. This is feasible only in small cases, because the combinatorial explosion takes place quickly when the number of products to be separated increases.
(2) Using optimisation algorithms such as or genetic algorithms (Hurme, 1996) or mixed integer programming (MINLP) to find the most feasible separation sequence. Both approaches require the selection of a superstructure. Other disadvantage is the limitation on the size of the problem handled because of the combinatorial explosion, which is problematic especially with MINLP.
(3) The case-based reasoning approach can be used through an ‘upper level’ CBR. This is discussed in more detail in the following section.

8.2. Separation sequence synthesis by CBR

The case-based reasoning approach can be used through an ‘upper level’ CBR for finding out a separation sequence. This is possible since a database can also be used for storing feasible separation sequences—not only information on single separations.

The search can be done:

1. directly with *component names*;
2. *component types* (e.g. aliphatic alcohol); or
3. in a more creative way by using analogies through *characteristic properties* of the components to be separated.

There are two alternative ways to interpret the search results:

(i) directly as feasible separation *sequences*; or
(ii) as feasible separation *sequence heuristics*, which can to be applied on the design of the new sequence. In this approach, the heuristics are stored together with the cases and the sequence is adapted from the strategy of the nearest existing case found (Pajula, Seuranen, & Hurme, 2001).

In the first sequencing strategy, it is possible to make searches not by names but criteria, which are related to the separation properties of the component pairs. The criteria should describe the difficulty of the separation by using properties such as boiling points, relative volatilities or coefficients of ease of separation (CES) (Liu et al., 1987). However, in the end we are interested in the costs of the required separation tasks. The cost of separation is not only dependent on the physical separation properties of the components but also on their concentrations in the feed and in the required products. Therefore, the use of CES is a possible way of making searches. However, the value of CES is dependent on the separation sequence and consequently its use as search criteria is difficult. Therefore, a new more straightforward search criterion is developed in the following.

It is well known that the column operating and capital costs are related to the vapor flow of the column. Porter and Momoh (1991) have suggested Eq. (2) as an approximate method of calculating the vapor flow V in a column, which can also serve as a simple estimate on the both costs:

$$V = D\left[1 + \frac{R_F}{\alpha - 1}\frac{F}{D}\right] \quad (2)$$

where V: vapor flow; D: distillate flow rate; F: feed flow rate; α: relative volatility.

$$R_F = \frac{R}{R_{\min}} \quad (3)$$

where R: reflux rate; $R_{\min}$: minimum reflux rate.

Eqs. (2) and (3) can be simplified to a search criterion VF, which is calculated for all the component pairs to be

Table 2
Query and nearest cases for non-isobaric separation

	Query	Found 1	Found 2
Component 1	Water	Water	Water
Component 2	THF	THF	THF
Feed (wt%)	15	10	12
Products	99 wt% THF	99 and 3 wt% THF	95 and 4 wt% THF
No. of columns	–	2	1
Comments	–	Distillation in a two-column system operating at different pressure (p_1 = 1 bar, p_2 = >6 bar)	Distillation in a one column system operating at different pressure (top 1 bar, bottom 8,0 bar)
Similarity		0.95	0.85

separated:

$$\mathrm{VF} = \left(D + \frac{1.1F}{\alpha - 1}\right) N \tag{4}$$

R_F was substituted with a typical value of 1.1. D and F are not the molar flows but percentage concentrations of the component pair. N is the number of components in the separation problem. N is used for scaling to make the values comparable between different separation problems. VF is an indicator of separation difficulty. It is a vapor flow estimate for a binary separation normalized to the whole problem scale by multiplying by the number of components in the problem.

If methods other than a conventional distillation is used, the values of relative volatilities are scaled to give a correct view of the economic feasibility. For extraction and extractive distillation the method of Sounders (1964) is employed for cost scaling. Other types of cost comparison charts are available elsewhere (Liu et al., 1987).

The method of finding separation sequences by CBR is the following:

1. Single separation methods are determined as described earlier.
2. VF values are calculated for all component pairs to be separated by using percentage concentrations in the feed of the whole system. The VF is scaled with the number of components in the feed N to make the values comparable with problems of unequal number of components.
3. A search is made in the database, which has known separation sequences stored with the component VF values. Since several sequences are often nearly as good at least in the theoretical sense (i.e. they give nearly similar values for the objective function) (Hurme, 1996; Liu et al., 1987), it is useful to store also near optimal sequences for a problem into the database. Also it is beneficial to store subsequences formed by removing components from the longer sequences at the end or beginning of sequence. This is possible especially if the amounts of those components are small.
4. The search is made on subsequences of different lengths. It is also useful to split the problem into subproblems, since the solutions can be combined from parts as seen later.
5. The similarities between input and a retrieved case is calculated for each separation. The similarity s between values a and b is defined by Eq. (5):

$$s = 1 - \frac{|a - b|}{a} \tag{5}$$

The similarity s can be calculated based either on absolute or relative values of properties. For instance, the VF values can be scaled to unity for a certain base component to calculate relative similarities. For absolute similarities there is no scaling.

6. The similarity for the whole separation subsequence is calculated as the average of the similarities of the single separations.
7. When all acceptable (i.e. similarity > tolerance) subsequences are found, total sequences are combined from the found subsequences starting from one end of the sequence (lightest or heaviest). The combining is done so that all separations are made only once in the sequence. Infeasible sequences are sequences that include same separations more than once or lack separations. As a result from the combining alternative sequences can be received, which then can be studied by simulation in more detail.

The approach for separation sequence synthesis is demonstrated by the following sample problem.

8.3. Example on separation sequence synthesis

Problem. Separate the mixture of light hydrocarbons shown in Table 3 into 'pure' components (Hartmann & Kaplick, 1990).

The problem is solved in two ways to present the different solution alternatives.

Table 3
Problem feed compositions, adjacent relative volatilities, VF values and corresponding sequences found

	Component	mol%	α	VF	Sequence	
1	Propane	13.7	2.53	224	y	u
2	*i*-Butane	11.7	1.26	588	y	u
3	*n*-Butane	5.5	2.39	126	y	u, w
4	*i*-Pentane	9.9	1.30	1000	x	u, v, w
5	*n*-Pentane	26.4	2.16	400	x	v, w
6	*i*-Hexane	5.6	1.31	854	x	v
7	*n*-Hexane	27.2				

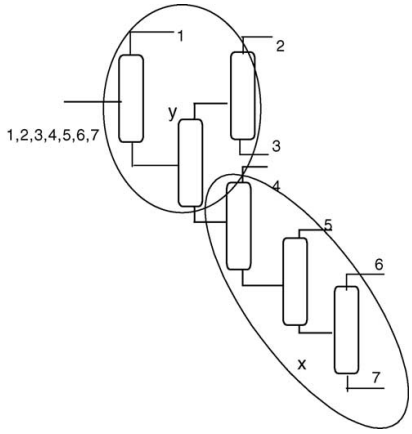

Fig. 4. The synthesised separation sequence (Solution 1).

Solution 1. The values of search criterion VF are calculated for adjacent separations. A search is made among the database sequences by trying to find similarity to any subset of query, which includes ≥ 3 separations. The search can be made by using either absolute or relative similarity as criterion, but the relative similarity has been used here.

The most similar subsequences found are shown in Table 4. The corresponding subsequences in query problem are marked with x and y also in Table 3. It can be seen that the relative similarities *s* are of 75% magnitude, which we have found generally acceptable.

The sequence synthesised based on the combination of the two found best cases (Hartmann & Kaplick, 1990; Smith, 1995) is shown in Fig. 4. The three lowest separations are from the first case (x) and the three uppermost from the latter case (y). Note that not all separations of the latter case were used, but from the y sequence the lightest component has been removed. The synthesis result found corresponds to the optimum result reported (Hartmann & Kaplick, 1990).

Solution 2. The synthesis problem given earlier can also be solved by applying other cases found. The cases used are shown in Table 5. The corresponding subsequences in Table 3 are marked with u, v and w. Note that this time the sequences used overlap each other and therefore add more credibility to the result of reasoning.

From u and w sequences the heaviest component has been deleted. In fact in w sequence, it can be seen that many sequences presented by Liu et al. (1987) become similar, if the heaviest component is removed. The removing of components is possible from the ends of sequences (i.e. the lightest or heaviest component) without much interfering the optimum sequence especially, if the quantity of the component is small and the separation is easy. The sequence found (Fig. 5) is the same as previously. In fact also other sequences could be also found by the method, therefore the method works mostly as a tool for finding feasible alternatives for more detailed studies by simulation.

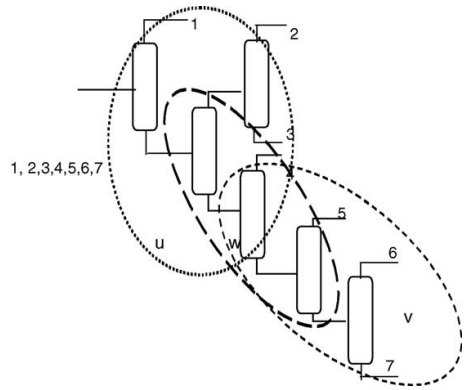

Fig. 5. The synthesised separation sequence (Solution 2).

9. Combined operations

After the separation sequence synthesis the possibility of combined operations should be studied (Pajula, Seuranen, & Koiranen et al., 2001). For example, a single column can separate several products using side streams. The approach for this is first to consider conventional separation sequences and then try to combine single separations one by one. Alternative approach is to conclude possible combination operations from the retrieved cases.

Therefore, the approaches are:

1. Combine two sequential separations together.
2. Search if analogous combined separation can be found from database by using relative volatilities, etc.

Or

1. Calculate VF values of separations.
2. Search if combined separation cases can be found in the database for those VF values.

Example. Consider distillation for a four component mixture (A/B/C/D), where A is a very light component. The originally synthesised separation sequence is shown in Fig. 6.

Searching for possible combination operations would give a case for hydrocarbon separations with noncondensable gases where noncondensables are taken out from the condenser as a third stream. Applying this case to the problem would give a combined system shown in Fig. 7.

Table 4
Solution 1; found feed compositions, adjacent relative volatilities, VF values and similarities

	Component	mol%	α	VF	s	Sequence	Reference
1	A	57.6	1.47	820	0.5	x	Hartmann and Kaplick (1990, p. 180)
2	B	5.5	1.72	164	1	x	
3	C	17.7	1,42	456	0.75	x	
4	D	19.2					
Average similarity		0.75					
1	Benzene	31.0	1.90	545	–		Smith (1995, p.140)
2	Toluene	33.0	1.76	450	0.74	y	
3	EB		6.6	1.06	2930	0.54	y
4	Xylenes	25.0	1.76	340	1	y	
5	C9s	4.9					
Average similarity		0.76					

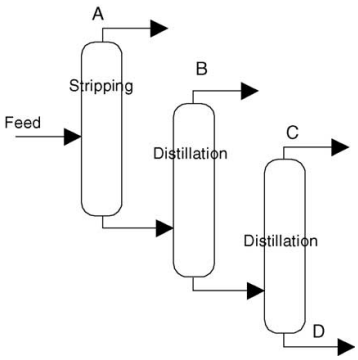

Fig. 6. Train of single distillation separations.

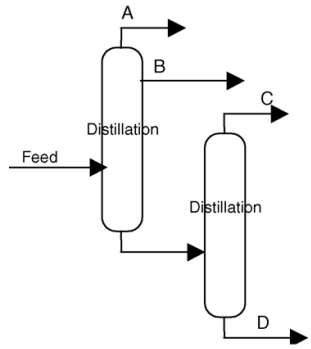

Fig. 7. The train with a combined operation.

Table 5
Solution 2; found feed compositions, adjacent relative volatilities, VF values and similarities

	Component	mol%	α	VF	s	Sequence	Reference
1	Ethane	20	3.50	210	0.45	u	Liu et al. (1987, p. 250)
2	Propylene	15	1.20	1240	0.97	u	
3	Propane	20	2.70	258	1	u	
4	1-Butane	15	1.21	1032	0.50	u	
5	*n*-Butane	15	3.00	186	–	–	
6	*n*-Pentane	15					
Average similarity		0.73					
1	Propane	1.5	2.45	70	0.26	w	Liu et al. (1987, p. 157)
2	1-Butene	14.7	1.18	2060	1	w	
3	*n*-Butane	50.3	1.70	860	0.96	w	
4	2-Butenes	27.6	2.50	260	–	–	
5	*n*-Pentane	5.9					
Average similarity		0.74					
1	A	20	2.1	280	0.87	v	Gruhn, Hartmann, Kardos, Dietzsch, and Kauschus (1977, p. 176)
2	B		30	28	128	1	v
3	C		30	3.7	200	0.73	v
4	D		20				
Average similarity					0.86		

10. Creativity and learning aspects

An important aspect in process design is creativity. A design system should not only be capable of modifying existing designs included in the database but also able to create new designs especially by proposing new separation alternatives. One possible way of including creativity into synthesis is to use analogies (Pajula, Seuranen, & Koiranen et al., 2001). Analogies can be included by using 'generalisations' and structural features such as proper hierarchy. The generalisations introduced may include general level categories in database of separation cases such as type of separation, type of components or their physical properties.

11. The CBR implementation

Since the presented approach is based on the use of existing knowledge, it is obvious that the method requires necessarily a database of single separations and separation sequences to store the data. The database can be collected from the literature and the earlier designs made. Separation process synthesis has been widely discussed in the public literature (e.g. Liu et al., 1987; Smith, 1995), which made it possible to create our prototype database include large variety of separation processes. The prototype database includes less than 100 separations and separation sequences, which allowed solving the case study problems presented. It should be noted that because several sequences are often nearly equally good in the economic sense, it is useful to store several feasible sequences for the same problem into the database. Finally the database can form a valuable databank of design experience and know how of the company, if also extensive in-house information is included.

The first CBR prototype was made in MS Excel and the computing time was negligible. The actual CBR application was made using a commercial CBR development tool, CBR-Works 4.0. The hierarchical structure of the database is based on classification of separations presented by Wankat (1990). In the database structure, each process document contains general data for separation (feed composition, purification requirements, capacity, etc.) and the separation methods that exists in the separation process train. Every separation process consists of several pieces of process equipment, which have to be selected. Therefore, detailed equipment specific information needs to be stored as well. The computing time depends on the retrieved case varying with our prototype database and Mobile Pentium III processor at 500 MHz about 5–20 s. However, it should be noted that the computation time depends much on the size of the database, the way of coding and the search algorithm used.

12. Conclusions

The case-based reasoning method for separation process synthesis is applicable especially in conceptual process design for screening options to be studied by simulation in more detail. The advantage of CBR compared to rule-based methods is that all the existing knowledge is available as detailed cases and can be utilised in a non-reduced form. The method is flexible, since the user can extract various types of information even from the same cases by using different types of search criteria, weights and similarity measures depending on the requirements and point of view. The method aims not to substitute the process designer or process simulation but the goal is to give a design tool to support the designer.

The method compares with the optimization based methods (e.g. MINLP) in the following ways: optimization requires an explicit objective function to be defined. However, many design criteria such as safety or operability are difficult or impossible to quantify and therefore include in the objective function. The optimization approaches are not interactive; the optimization cannot be guided easily by the user as in CBR. Therefore, optimization often finds impractical or infeasible optima. The use of optimization requires that the optimization alternatives and area have to be restricted. This tedious task called 'generation of superstructure' has to be done by the user. If done improperly, it has a detrimental effect on the result. If the superstructure selected is too large, a combinatorial explosion will take place, if too narrow, feasible options are cut off.

The characteristic (and problem) of CBR is its dependency on the database and the engineering skills of the designer. On the other hand, the method can utilise the huge amount of data available in the literature, if the information can only be extracted. Therefore, methods of data mining become of importance. The general quality of design can be improved, since CBR enhances the systematic reuse of existing design experience, especially if the design cases are stored with the feedback and practical experience gained from existing engineering designs. Ultimately, the database generated can form a valuable 'institutional memory' of the company.

References

Avramenko, Y., Nyström, L., & Kraslawski, A. (2004). *Computers and Chemical Engineering, 28*, 37–44.

Cordiner, J. L. (2001). *European symposium on computer aided process engineering, Vol. 11* (pp. 27–40). Amsterdam: Elsevier.

Grossmann, I. E., Aguirre, P. A., & Barttfeld, M. (2004). *European symposium on computer aided process engineering, Vol. 14* (pp. 53–74). Amsterdam: Elsevier.

Gruhn, G., Hartmann, K., Kardos, J., Dietzsch, L., & Kauschus, W. (1977). Systemverfahrenstechnik II, Deutscher Verlag fur Grundstoffindustrie, Leipzig.

Hartmann, K., & Kaplick, K. (1990). *Analysis and synthesis of chemical process systems*. Amsterdam: Elsevier.

Hilmen, E. (2000). Separation of azeotropic mixtures: tools for analysis and studies on batch distillation operation. Dr. Thesis. Trondheim: NTNU,.

Hurme, M. (1996). *Proceedings of the second nordic workshop on genetic algorithms and their applications* (pp. 219–224). University of Vaasa.

Jaksland, C., Gani, R., & Lien, K. M. (1995). *Chemical Engineering Science, 50*, 511–530.

King, J. M. P., Bañares-Alcántara, R., & Manan, Z. A. (1999). *Environmental Modelling & Software*, *14*, 359–366.

Liu, Y. A., McGee, H. A., & Epperly, W. R. (1987). *Recent developments in chemical process and plant design*. New York: John Wiley.

Maher, M. L., & de Silva Garza, A. G. (1997). *IEEE Expert*, *12*(2), 34–41.

Pajula, E., Seuranen, T., & Hurme, M. (2001). *European symposium on computer aided process engineering, Vol. 9* (pp. 469–474). Amsterdam: Elsevier.

Pajula, E., Seuranen, T., Koiranen, T., & Hurme, M. (2001). *Computers and Chemical Engineering*, *25*, 775–782.

Porter, K. E., & Momoh, S. O. (1991). *Chemical Engineering Journal*, *46*, 97–108.

Seuranen, T., Pajula, E., & Hurme, M. (2002). *European symposium on computer aided process engineering, Vol. 10* (pp. 343–348). Amsterdam: Elsevier.

Smallwood, I. (1993). *Solvent recovery handbook*. New York: McGraw-Hill.

Smith, R. (1995). *Chemical process design*. New York: McGraw-Hill.

Sounders, M. (1964). *Chemical Engineering Progress*, *60*(2), 75–82.

Wankat, P. C. (1990). *Rate-controlled separations*. Barking, England: Elsevier.

Watson, I., & Marir, F. (1994). *The Knowledge Engineering Review*, *9*(4), 50–60.