

Semantically Enabled Process Synthesis and Optimisation

Antonios Kokossis^b, Claudia Labrador-Darder^a, Franjo Cecelja^a

^aCentre for Process & Information Systems Engineering, University of Surrey, Guildford, GU2 7XH, UK,

^bSchool of Chemical Engineering, National Technical University of Athens, GR-15780, Athens, Greece,

Abstract

This paper introduces a new framework to support synthesis of complex engineering problems and which combines stochastic search optimisation with ontological knowledge modelling. The framework uses Tabu search to generate new solutions and introduces the mechanism of digital certificate to translate between structural information of solutions and semantics of ontology. The solutions are respectively clustered by design features. Tested against complex synthesis of reactor networks, the framework demonstrates noticeable advantage in convergence and superiority in simplifying solutions with favourable comparisons drawn against conventional stochastic optimisation algorithms.

1 Introduction

Most of industrial processes are complex by nature; this is particularly the case with processes involving reaction networks described by highly nonlinear kinetics. Numerous techniques have been developed for design and synthesis of reactor network, i.e. graphical techniques including attainable region technique, simplified approximations in the form of superstructures, and highly analytical representations (Ashley and Linke 2004). In either case selection of options from among large number of alternatives is normally supported by optimisation with numerous optimisation approaches reported. Initially, these were deterministic techniques including non-linear programming (NLP) (Achenie and Biegler 1988) and mixed-integer non-linear programming (MINLP) (Kokossis, A. and Floudas 1990; Raman and Grossmann 1991; Kokossis, A. and Floudas 1994), among others. More recent approaches focus on stochastic search based optimisation with perhaps Simulated Annealing (Marcoulaki, E. and Kokossis 1996; Marcoulaki, E. C. and Kokossis 1999; Mehta and Kokossis 2000; Linke and Kokossis 2003a), Tabu Search (Wang, C et al. 1999; Linke and Kokossis 2003b; Cavin et al. 2004) and Ant Colony (Dorigo et al. 1999; Jayaraman et al. 2000;

Dorigo and Blum 2005) being the most widely applied. The latest advances in stochastic search based on cascading of population and inflection of solutions is particularly attractive as it provides readily access to optimisation solutions at every stage of the process, the search known as the Cascade Algorithm (Labrador-Darder et al. 2009; Kokossis, A et al. 2011; Cecelja, F et al. 2014).

Inherent problems with all optimisation techniques have long been realised, and they include i) tedious analytical considerations, ii) slow convergence and long computational time (Kokossis, A et al. 2011), iii) interpretation of complex and impractical optimisation results (Ashley and Linke 2004), and iv) lack of confidence in the selecting options. As reported, complex analytics was mainly addressed by controlled simplifications, e.g. superstructure representation (Yeomans and Grossman 1999; Linke and Kokossis 2003a), whereas attempts were made to improve slow convergence of deterministic optimisation by simplifying the conceptual content and hence analytics (Raman and Grossmann 1991; Bauer and Stichlmair 1996). Similarly, slow convergence of stochastic searches was addressed by parallelising the execution and hence providing larger number of solutions in shorter time (Talbi et al. 1998; Leite and Topping 1999; Wang, ZG et al. 2005; James et al. 2009; Kokossis, A et al. 2011; Kim et al. 2012; Cecelja, F et al. 2014). Interpretation and assuring the confidence in obtained results, however, still remains a problem. Majority of proposed synthesis solutions were either compared with those obtained using different approaches, or compared with 'similar' solutions already proven in practice.

Application of engineering knowledge in the process synthesis and especially in the process of optimisation helped to both interpret and simplify results, hence to improve the confidence, but also to speed up convergence. Raman et al. attempted to apply engineering insights as additional logical constraints in MILP optimisation (Raman and Grossmann 1991). Similarly, Shah et al (Shah and Kokossis 2001) attempted to further formulate conceptually rich performance model that makes simultaneous use of engineering insights and MILP optimisation and later to systematically formulate overall framework (Shah and Kokossis 2002). To our knowledge, the first attempt to create a knowledge model of a complex process, a reactor network, was reported by Jacobs (Jacobs et al. 1996) who formalised it using production rules. The initial intention was to use knowledge model alone to support selection of reactors in the process of reactor network design, and then to automate

derivation of reactor strategy (Jacobs and Jansweijer 2000). Production rules were also used by Ashley et al to model reactor networks and to better understand the system to guide the optimal search using stochastic optimisation in the form of Tabu search (Ashley and Linke 2004), and consequently to analyse results. Ontological approach to model knowledge was used by Kokossis et al to extract and interpret knowledge generated in the process of optimising reactor network using Simulated Annealing (Kokossis, A. et al. 2008), which was further expanded to cascade solutions and to guide search using Cascade Algorithm (Kokossis, A et al. 2011), as well as to parallel execution (Cecelja, F et al. 2014). An attempt was made by Cecelja et al to use both production rules and ontology for integrated interpretation of solutions and acquisition of optimisation knowledge to guide the search in the process of reactor network synthesis using Cascade Algorithm (Cecelja, F. et al. 2011). The use of production rules and ontology towards using, and in particular integrating of processing technologies, was demonstrated by Raafat et al (Raafat et al. 2013), the process which was then fully elaborated for practical use (Cecelja, F et al. 2015).

This paper presents an approach to model, acquire and beneficially employ knowledge in process synthesis to i) interpret solutions for better understanding of the problem, comparison and increased confidence, ii) to learn from the progress of optimisation and to guide the search towards the optimum solution within predefined and on-the-fly created constraints, and iii) simplify solutions dynamically and in line with problem formulation to speed up the search and to adjust to specification vary. The proposed approach is based on the hypothesis that i) the 'best' solution could always be replaced by 'sufficiently good' solution, and ii) the optimisation based on explicit knowledge is the best way to generate solutions whereas knowledge based on associations, the tacit knowledge, is the best way of selecting solutions. To this end, a widely used, robust and sufficiently adaptable algorithm of Tabu search is used to generate candidate solutions and to optimise. Tacit knowledge about the application and optimisation postulate is modelled using ontologies supported by production rules. While ontologies are used for expanding knowledge base through capturing evolving solution features, for solution interpretation and to share, production rules were employed to model the dynamics of constraints, to simplify solutions and to guide the search.

The reactor network synthesis with superstructure-based optimisation is used to demonstrate the approach with well surveyed single phase Van de Vusse and Denbigh reactions (Kokossis, A. and Floudas 1990) and modified multiphase Denbigh reaction (Mehta and Kokossis 2000) applications. Proposed single phase superstructure representation of reactor network (Mehta and Kokossis 1997) includes Continuous Steering Tank Reactor (CSTR) and Plug Flow Reactor (PFR) reactor units interconnected through mixers and splitters (Figure 1).

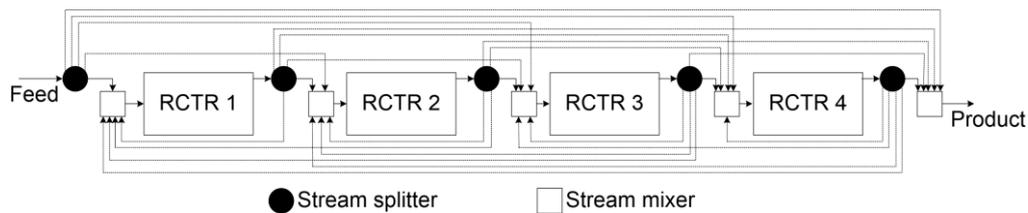


Figure 1 Single phase superstructure representation

The superstructure representation of multiphase reactor networks is built around generic reactor/mass exchanger units (RMX) (Linke and Kokossis 2003a), which enables a flexible and compact representation of fundamental phenomena exploited in the process of design and synthesis, e.g. reaction, mixing, heat transfer and mass exchange (Figure 2).

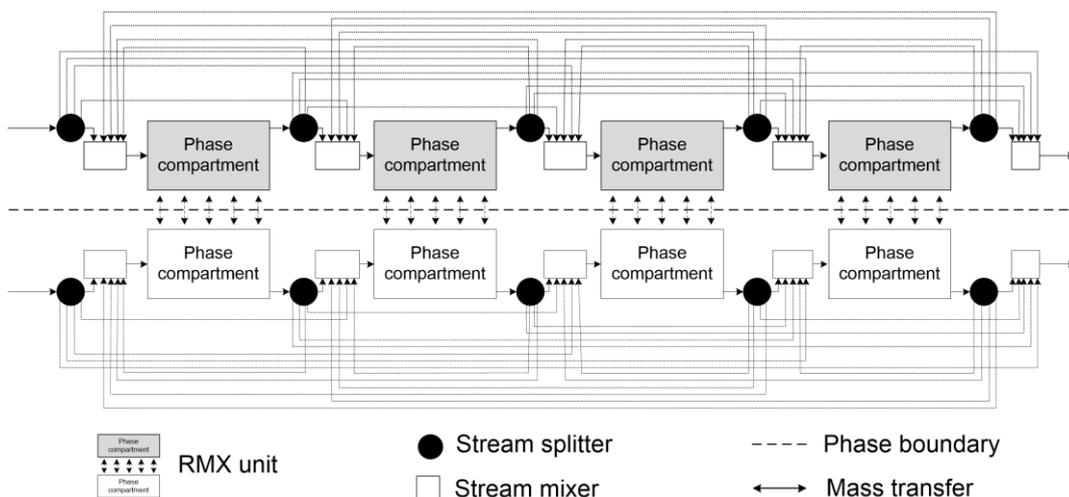


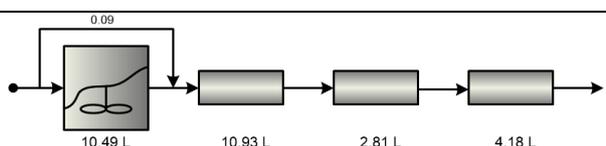
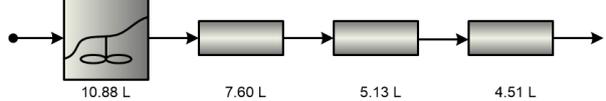
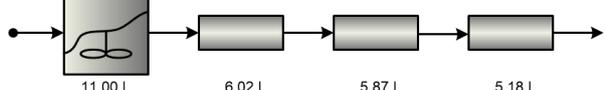
Figure 2 Superstructure representation of a two-phase system

In both cases, the process of synthesis considers type and number of units (max. 4), unit sizes (volumes), mixing pattern, feeding, bypassing and recycling, as well as options for mass exchange to achieve optimal concentrations.

2 Motivating Example

In contrast to deterministic optimisation and in spite of possible converging issues and required computational cost, stochastic search optimisation is increasingly used for synthesis of complex process applications primarily because of their ability to provide readily access to a large number of intermediate or alternative solutions (Marcoulaki, E. and Kokossis 1996; Mehta and Kokossis 2000; Cecelja, F et al. 2014). These solutions, however, often represent competitive designs with similar design performance measures but with different structural and operational characteristics. Minor changes in design parameters may yield major changes in the selected designs. Comparison is cumbersome and optimisation results may be impractical to implement. An example of optimisation guided solution evolution of reactor network for Van de Vusse reaction network (Appendix 1) is presented in Table 1. Apparent as it is, the set of competitive solutions differ in the network layout and operation but share very similar objective values, the product concentration, all within 0.5% difference. By the same token, a number of solutions presented in Table 1 can be simplified to virtually the same configurations by following comparatively simple and fundamental knowledge in chemical engineering; reactor with marginal contribution could be removed, the recycling applied on neighbouring reactors could be simplified into a single unit, etc., and all this without effecting the targeted objective significantly. No design feature, e.g. optimal number and type of reactors or respective volumes, are easily identifiable. Similar level of inconvenience can be experienced with most complex synthesis problems.

Table 1 Excerpt of results generated by the optimisation of reactor network for Van de Vusse reaction

Design Solution	Objective value (mol/L)
	3.6550
	3.6605
	3.6607

	3.6532
	3.6608
	3.6512
	3.6609
	3.6510
	3.6601
	3.6602
Legend: → CSTR → PFR	

3 Semantically Enabled Process Synthesis: Concepts and Definitions

3.1 The Concept

In order to better support the process of decision making in selecting the most appropriate solution in process synthesis, a new multistage knowledge based synthesis framework is proposed. The framework is built around knowledge controlled stochastic search optimisation. The knowledge model is developed upon the understanding of the problem, which includes identification of the design variables enclosed in the appropriate graphical or analytical representations, such as the superstructure, and identification of respective set of structural and operational features. Tabu search is employed to generate new set of candidate design solutions. Along with optimising these variables for specified performance, e.g. product concentration, the solution properties are dynamically captured, grouped by common features into clusters and used to update the knowledge base. The best performing clusters are then selected to set up the next optimisation stage and to customise features of the optimisation search, as illustrated in Figure 3. The features of *limited*

importance are removed through the process of feature customisation. The entire process is repeated until the termination criteria are met.

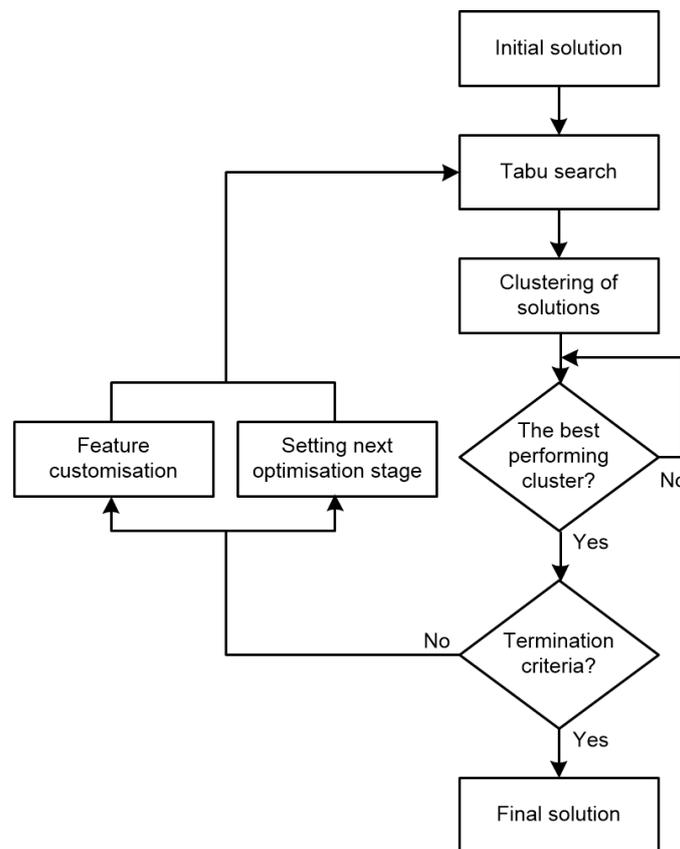


Figure 3 Transition of optimisation stages

The optimisation stage transition is performed in two distinct stages:

Stage 0: the initial stage which takes randomly selected initial solutions and Tabu search optimises for all application variables;

Stage 1 – n: the consequent stages with search performed for all continuous variables and discrete variables identified by the process of clustering. The solutions in the best performing cluster are used as initial solutions for the next optimisation stage.

In the present implementation we use the combination of termination criteria which includes predefined number of clusters and the mean objective value in cluster changes within 1% in two consecutive stages, whichever is reached first.

3.2 Knowledge Modelling

Ontologies combined with production rules are used to build knowledge model (Raafat et al. 2013). While production rules are a simple and robust method of knowledge representation, ontologies offer an informal approach suitable for reusing and sharing.

Used for knowledge representation, ontologies are built upon classes which organise instances with common properties and which have domain dictated relationships between them. In retrospect, ontology O is a 5-tuple (Cecelja, F et al. 2015):

$$O = \langle H^I, H_C, R_i^C, R^C, S_i^I \rangle \quad (1)$$

where:

- 1) H^I is a set of classes S_i^I which themselves are sets of n_C instances s_j sharing common properties p_j , or instances with intensionally equal¹ properties $p_j := p_k$ as:

$$S_i^I = \{s_j\}_{j=1}^{n_C}, p_j := p_k \wedge \forall j \geq 0 \quad (2)$$

- 2) A set of n classes $H^I = \{S_i^I\}_{i=1}^n$ with each class S_i^I having a distinct name N_i^I hence representing a concept with respective semantic (Cecelja, F et al. 2015).
- 3) $H_C = (S_i^I, is - a)$ is a graph forming a subsumption hierarchy. If a classes S_i^I and S_k^I ($S_k^I \subseteq S_i^I$) are characterised by the set of properties P_i and P_k , respectively, then the two classes form a subsumption, if they follow inheritance property such that $P_i \subseteq P_k$ informing that every subclass inherits all the superclass properties.
- 4) R_i^C is a set of bijective relationships $r_{i,j}$ between all instances of domain class S_i^I and range class S_j^I such that

$$R_i^C = \{r_{i,j}(S_i^I, S_j^I) | \forall i \neq j\} \quad (3)$$

Relationships R_i^C can have restriction S_i^D on their domain $S_i^I = \text{dom}(R_i^C)$ defined as partial function $f_D = \text{dom}R_i^C|_{S_i^D}$. As such, f_D establishes the binary relationship between S_i^I and S_j^I based on universal and existential properties P_j of S_i^I , but also on other relationships $r_{i,j}$ of S_i^I ;

¹ Two instances are intensionally equal if they have the same structure of the properties, not necessarily the same property values.

5) R^C is subsumption of n_R relationships R_i^C organised in a property-subproperty hierarchy as

$$R^C = \{r_{i,j}(S_i^I, S_j^I) | \forall i \neq j\}_{i,j=1}^{n_R} \quad (4)$$

6) S_j^I is, as defined by eq. (2), a set of instances. Every class can contain instances. Classes without instances, the empty classes, contain only properties and they are used to enhance semantic aspect of ontology. Attaching instance to a class is the process known as instantiation.

If every class in ontology (1) has a unique name N_i^I associated with a concept, and if the ontology (1) is used to provide hierarchically structured set of causes and effects for understanding the (knowledge) domain, which is an effective means to explicitly describe knowledge in knowledge base, then eq. (1) refers to the domain ontology. In practical terms, domain ontology refers to a collection of interlinked concepts N_i , or names ($N_i \equiv S_i^I$), the concept attributes or properties P_j and functions or logical statements R^C expressing the constraints existing in the domain and restricting the interpretation of vocabulary, all arranged in respective hierarchies H_C and R^C and supplemented by class-attached instances S_j^I .

Ontologies can be inferred by processing specified restrictions f_D and hence reclassifying classes and instances accordingly, the operation which is dominantly used to reclassify solutions in this work. Together with parsing ontology, the process of developing inferences generates new knowledge. For example, if we provide an object property *hasNumberOfZones* which links concepts Network (as the domain) and NumberOfZones (as the range) and if we place a restriction

$$f_D = \exists \text{ hasNumberOfZones NumberOfZones,} \quad (5)$$

the semantic of which is that for some individual to be a member of the concept/class Network, it has to have at least some connections along the property *hasNumberOfZones* to members of the class NumberOfZones, as shown in Figure 4a. By placing additional restrictions on the class ThreeZoneNetwork along the property *hasNumberOfZones* as

$$f_D = \forall \text{ hasNumberOfZones } 3 \quad (6)$$

with semantic interpretation that every member of the class `ThreeZoneNetwork` must have exactly three reactors (Figure 4a), then all experiments which have exactly 3 zones will be inferred to belong to the class `ThreeZoneNetwork`, as shown in Figure 4b.

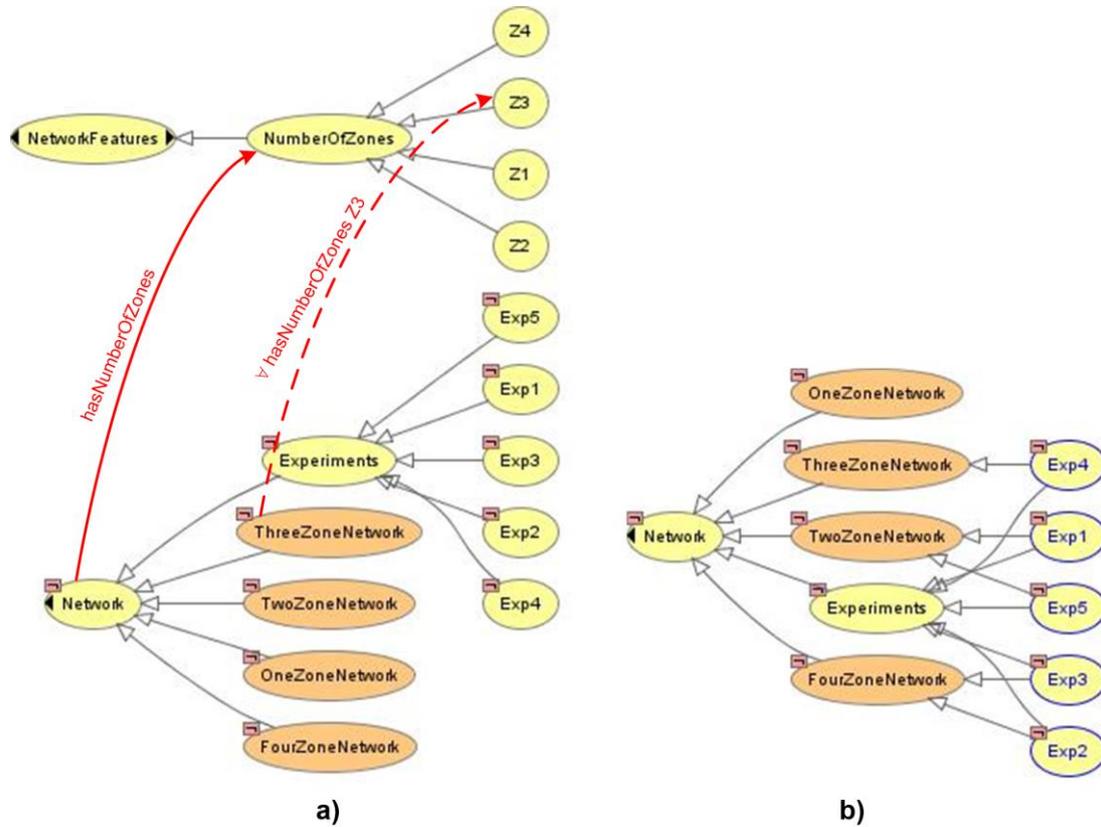


Figure 4 Inferring ontology

In the present implementation, the reactor network ontology has four levels of abstraction: i) meta-level which defines basic ontology structure as referred to by equation (1), ii) upper level providing basic classifications related to the application, including type of networks and structural and operational features and attributes used to characterise concepts, iii) domain level which details the domain of application, i.e. the reactor network synthesis, and iv) instantiation level with individual solutions from the process of optimisation and possibly from real-life experimentation, as illustrated in Figure 5. The meta-level is primarily defined and used for channelling the development of domain ontology level and for ontology sharing and reusing. Other levels will be explained in details in the follow on sections of this paper.

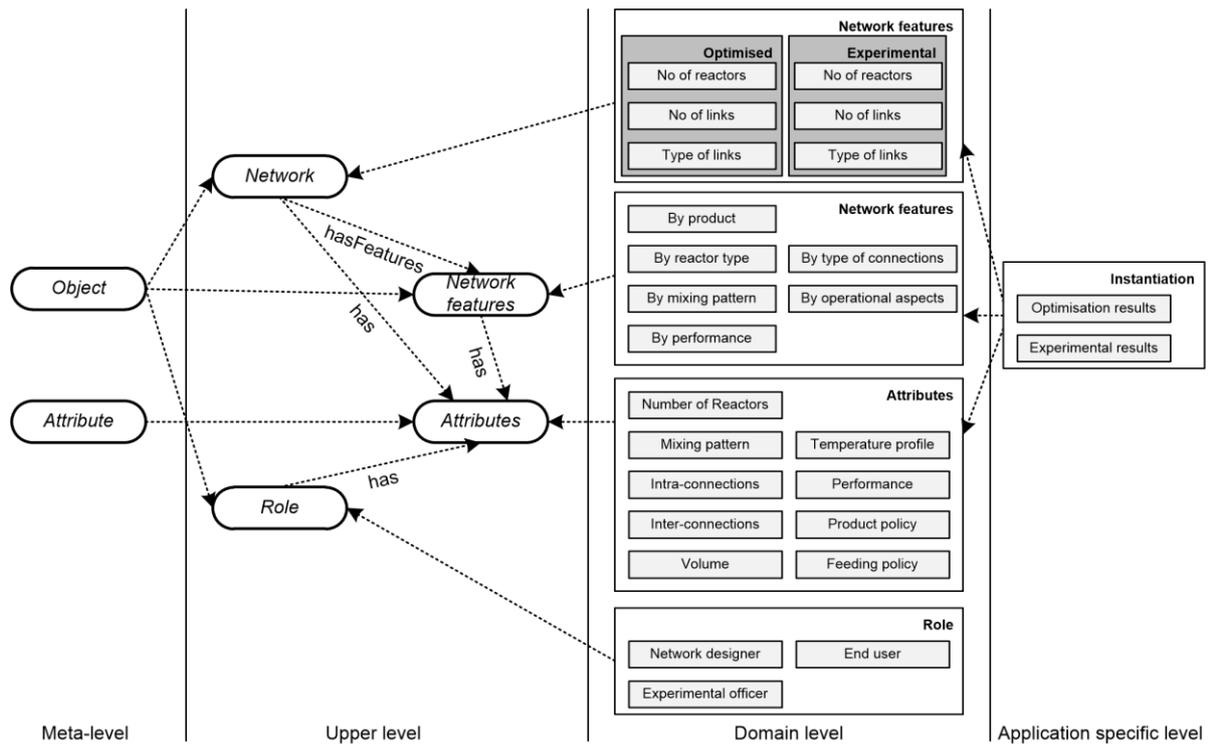


Figure 5 Ontology design

Although one of the earliest forms of knowledge representation, production rules are still in use mainly in combination with other knowledge models, such as ontologies. Conceptually, production rules take the form of an IF-THEN structure:

$$\begin{array}{l}
 \text{IF} \quad (S_1 \ \& \ S_2 \ \& \ \dots \ \& \ S_n) \\
 \text{THEN} \quad (A_1 \ \& \ A_2 \ \& \ \dots \ \& \ A_m)
 \end{array}
 \quad (7)$$

which interprets as IF the combination of statements S_1, S_2, \dots, S_n , called pattern, is true, THAN the set of specified actions A_1, A_2, \dots, A_m is executed. Along with the obvious simplicity, the power of production rules is in their cyclic execution adopted by many expert systems (Figure 6), where the rule patterns are matched against their existence (Guiarratano 2002); the rules with all rule patterns S_n true are placed on the agenda². If more than one rule is on the agenda, the conflict resolution process resolves the order of execution, which could be one of the following:

1. First execute the rule with the most recently updated pattern A_n , which gives priority to newly updated knowledge, or

² Agenda is the list of all the rules with matched patterns ready for execution

2. First execute rules with more complex pattern, that is rules with higher number of patterns A_n , which gives priority to complexity of the problem.

The rules on agenda are executed in order of priority, the knowledge base (KB) is updated according to the specified actions A_m and the whole cycle repeats as long as there is at least one rule on agenda. All respective patterns S_n are stored in the working memory (WM).

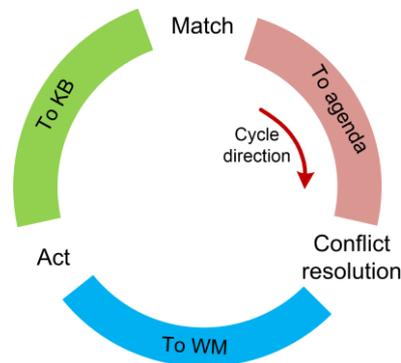


Figure 6 Cyclic nature of production rule execution

In the current implementation, the explicit knowledge on process synthesis is acquired during the ontology instantiation (Section 3.3), whereas the respective tacit knowledge is embedded in the ontology structure. Production rules are used to supplement the ontology inference and to further process explicit knowledge, as explained in Section 4.

3.3 Knowledge Acquisition

Process synthesis knowledge base is updated through the process of acquiring knowledge from optimisation; more specifically it is a process of dynamic instantiation of ontology and developing inferences. A two stage knowledge extraction process is proposed:

1. Translation between candidate solutions and semantic terms of knowledge model,
2. Feature extraction through solution clustering.

Unlike informal semantic form of ontology, the operational and structural information about design solutions are in numerical format referring to specific cases of superstructure. With the superstructure representation of reactor network these include active reactive units combined into reactive zones depending on the mixing pattern, and connection between them. A translation mechanism takes the form of digital certificates (Figure 7), the mechanism which is controlled by production rules.

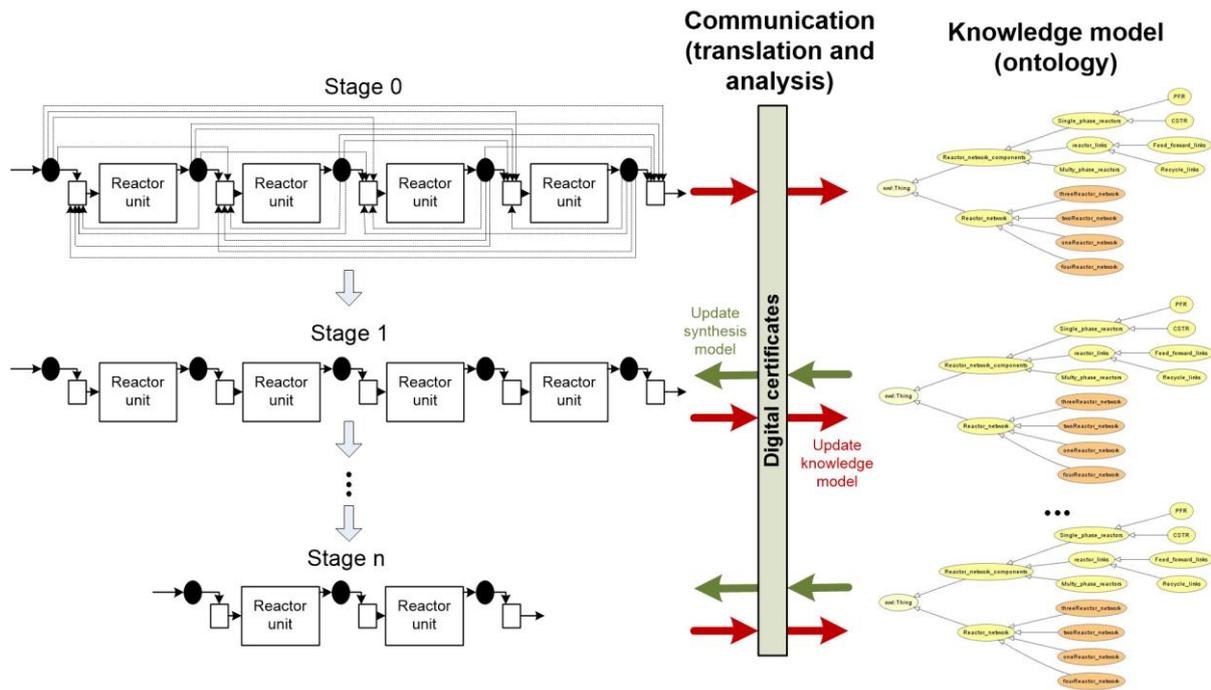


Figure 7 Communication between application optimisation and knowledge model

The ontology is instantiated by the design solutions and design features. Digital certificates are used to extract specific and recognised information about design solutions in the form of a vector, which contain operational and structural features, e.g. those of the superstructure. These features are then converted into instances of respective classes in the ontology. In the current implementation, a set of production rules is used to specify actions to be performed for given conditions, i.e. to establish relationships between the design variables of the superstructure, and to translate between design solutions and semantic terms of the ontology. Rules are also used to dynamically simplify solutions, as fully explained in Sections 4.1 and 4.2. An example of a digital certificate for a single phase reactor network is shown in Figure 10 containing features such as No. of zones, interconnections, intraconnections, feeding policy and product policy.

The solutions are grouped around common features into clusters arranged by their performance to i) instantiate ontology, ii) set up new optimisation stage, and iii) customise features of the optimisation search. We propose clustering of solutions based on pre-set limits of performance, which are narrowed down as optimisation progresses. For example, network volume bounds are reduced from stage to stage. This gradual tuning results in the reduction of the search space. In consequence, knowledge acquired from the analysis of

solutions is used to reduce the degree of freedom for optimisation performed in follow-on stages: i) features enclosed in the selected (the best) cluster are used as the basis for optimisation move selection, and ii) optimisation moves related to features that have been proven irrelevant are cancelled. Search is encouraged into more promising regions and the level of confidence is dynamically adjusted and stringent as the search progresses. Here, confidence is not only based on performance, as is the case with ordinary optimisation, but also on selection of relevant features of the application verified by newly generated solutions, such as number of zones, feeding and product policies etc.

The features used for solutions clustering are those defined by digital certificates. The reverse is possible as dictated by the problem. For a cluster to be considered, it has to have more than one solution. The whole process of solution clustering is summarised in the diagram in Figure 8 where T_{nc} stands for current number of clusters.

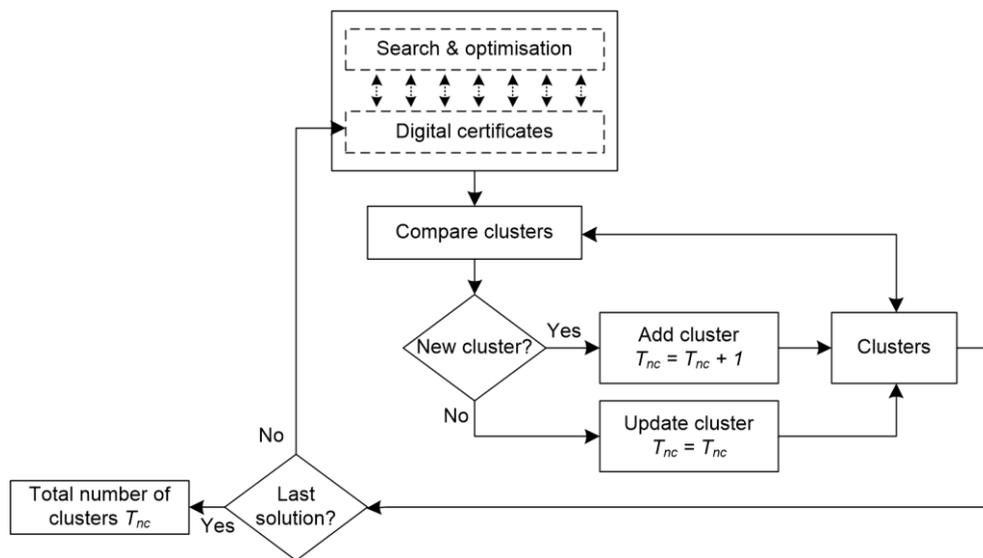


Figure 8 Clustering process

The performance of the clusters is determined as the arithmetic mean of objectives of all solutions in the cluster at the end of the optimisation stage.

3.4 Optimisation and Synthesis

Tabu search was selected for generating candidate design solutions and to perform optimisation because it is a robust and in many ways ordinary local or neighbourhood search widely used with complex engineering problems (Glover and Laguna 2011). Tabu search algorithm has been explained elsewhere (Glover 1989; Linke and Kokossis 2003b;

Cordeau and Laporte 2005; Cecelja, F et al. 2014), but it is suffice to say that it performs transitions from current to new states or solutions in its neighbourhood following a simple descend. We use Markov chain to provide transitions. The search terminates when no improving solution are found. In order to minimise redundant moves, that is to minimise revisiting previously visited solutions and/or cycling in local optimum, the Tabu search algorithm introduces a Tabu list in the form of a short term memory. Moves revisiting the solutions are not allowed. Still, to explore more thoroughly areas in the vicinity of Tabu solutions, aspiration criteria is introduced. The aspiration criteria overrides the Tabu status, if the aspiratration $a_i(s_j, m)$ marks better than a pre-set threshold value $A_i(s_j, m)$:

$$a_i(s_j, m) \in A_i(s_j, m) \quad (i = 1, \dots, n_{EN}) \quad (8)$$

where n_{EN} is the size of neighbouring population. Tabu lists are performed provided that at least one (or some specified number) of conditions (8) are satisfied (Glover et al. 1993).

A long term memory, commonly called the frequency long term memory (FLM) is introduced to enable learning from the past search experience (Glover et al. 1993; Linke and Kokossis 2003b). It records the frequency with which individual moves have been performed over the search history and hence provide bias towards promising areas, the intensification, as much as towards areas that have not been visited before, the diversification. In consequence, intensification generates neighbouring solutions by grafting together good solutions and encouraging search into areas with historically good solutions and returning attractive regions to search them more thoroughly. The diversification, on the other hand, encourages search process to visit unvisited regions and to generate solutions that differ in various significant ways from those seen before (Cecelja, F et al. 2014).

Reportedly, Tabu search uses several termination criteria which include population size with pre-defined number of iterations, progress of optimisation with predefined number of iterations without improvement, and population feature with predefined threshold solution value (Glover 1989; Cecelja, F et al. 2014). The whole process of Tabu search implemented to generate new design solutions is depicted in Figure 9 where S_0 stands for the set of initial solutions at the beginning of search and FLM stands for frequency long term memory.

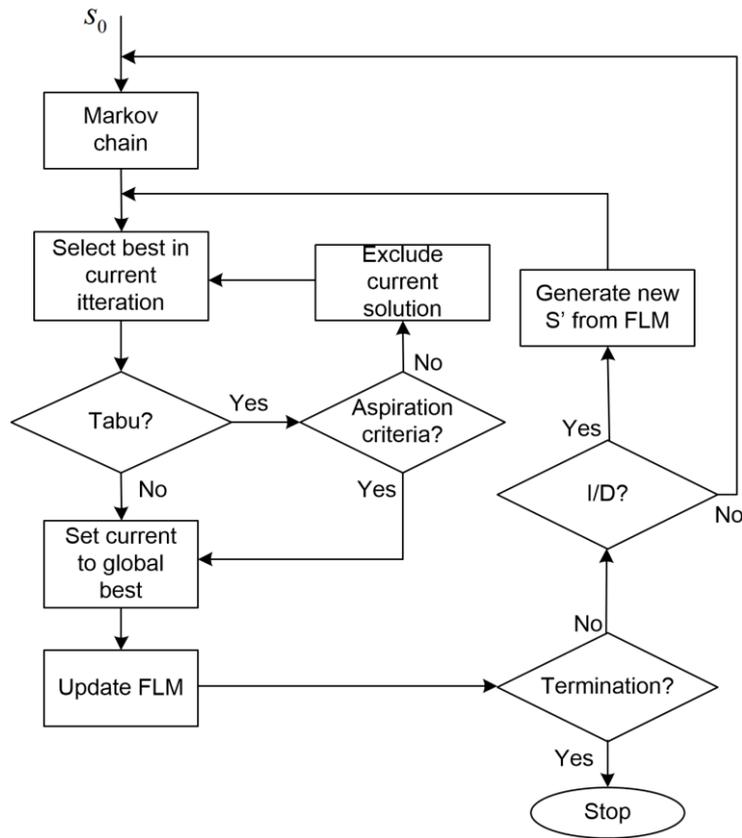


Figure 9 Tabu search algorithm to generate solution population

4 Implementation

Implementation is demonstrated using single phase and multiphase (two-phase) reactor networks represented by superstructures.

4.1 Single Phase Reactor Networks

The single phase reactor network superstructure is given in Figure 1. Each design solution is featured by the number, type and size of units, operational characteristics and the network streams. Operational characteristics include flow rates and temperatures. Active reactor units are always sequentially connected in the way that each reactor receives a minimum from the previous unit. Full mathematical definition of the single phase reactor network superstructure is given in Appendix 4.

The domain ontology developed for single phase systems is modularised to enable reusing, e.g. for multiphase systems given in Section 4.2, and for sharing with other applications in parts or as whole, as shown in Figure 5. The full list of concepts N_i representing classes S_i^I

and organised in subsumption H_C for the top four levels of domain ontology is given in Appendix 5. In Appendix 5 the names of the concepts are self-explanatory, yet with fully explained semantic, and the subsumption *is-a* relationship between classes is indicated by indentation in the left column. The ontology is implemented using Ontology Web Language (OWL).

Relationships R_i^C between classes of the single phase reactor network ontology are used to establish operational static and dynamic associations between them: mixing pattern, number of zones, existence of connections between units, volume, temperature and performance in terms of objective value. All the relationships are listed in Table 2 including respective domain and range classes and names which are self-explanatory. The property subsumption is indicated by indentation in the left-hand side column.

Table 2 Relationships between domain and range classes for single phase ontology

Object property	Domain class S_i^l	Range class S_j^l
<i>hasNumberOfZones</i>	<i>Network</i>	<i>NumberOfZones</i>
<i>hasMixingPattern1</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasMixingPattern2</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasMixingPattern3</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasMixingPattern4</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasFeeding</i>	<i>Network</i>	<i>FeedingPolicy</i>
<i>hasProduct</i>	<i>Network</i>	<i>ProductPolicy</i>
<i>hasConnection</i>	<i>Network</i>	<i>Connections</i>
<i>hasInterConnection</i>	<i>Network</i>	<i>InterConnection</i>
<i>hasIntraConnection</i>	<i>Network</i>	<i>IntraConnection</i>
<i>hasSequentialConnection</i>	<i>Network</i>	
<i>hasSize</i>	<i>Network</i>	<i>Size</i>
<i>hasTemperatureProfile</i>	<i>Network</i>	<i>TemperatureProfile</i>
<i>hasPerformance</i>	<i>Network</i>	<i>Performance</i>

Digital certificates (Figure 10) are created by a set of production rules to specify actions to be performed for given conditions, i.e. relationships between the design variables of the superstructure. The six rules relate to active-reactive units and connections between them and are described in Table 3. The active-reactive units are combined into reactive zones depending on the mixing pattern favoured and connections in the form of recycles and bypasses that relate to these reactive zones are extracted to be represented by the certificates. As a result, the digital certificates are generated for each solutions representing a simplified equivalent structure in terms of their main features, which include the number of reactive zones, the mixing pattern the connections between reactive zones and the

feeding and the product policies, the process previously termed as *feature customisation*. The limiting and simplifying values in Table 3 are determined from logical understanding of the application and as a compromise between required quality of the product and exhaustion and respective computational effort required for the search.

Table 3 Feature customisation and creation of digital certificates

Rule 1:	Reactor network reduction: reactors with marginal contributions are removed if they do not increase/decrease concentration of reference component by more than 5%.
Rule 2:	Reactor network redistribution: change of the sequential unit connection to main flow path according to feed-main and/or bypass-main flow
Rule 3:	Reactive zone classification:
Rule 3a:	<i>Pattern identification:</i> reactor units are classified into three categories by mixing pattern: <i>PFR:</i> PFRs without or with low degree of back mixing – recycle lower than 12% <i>Mid-mixed:</i> PFRs with medium degree of back-mixing – recycle between 12% and 60% <i>Well-mixed:</i> CSTRs and PFRs with high degree of back mixing – recycle over 60%
Rule 3b:	<i>Combination of reactors into reactor zones:</i> <i>Recycling:</i> if neighbouring reactors are enclosed in a recycle – classification is performed according to Rule 3a; <i>Mixing Pattern:</i> if a pair of neighbouring reactors share the same mixing pattern – applies to PFRs only
Rule 4:	Zone network reduction: reactive zones with marginal contribution are removed – zones which do not increase/decrease the concentration of reference component by more than 5%
Rule 5:	Feeding distribution limitation: no feed distribution is considered if the first reactive zone receives between 70% and 95% of the total feeding flow
Rule 6:	Connections: recycles and bypasses are arranged according to the above rules and classified into intra- and inter-connections.

The transition through optimisation of single phase systems is performed in stages (Section 3.1). In the first stage of the optimisation, the Stage 0 (Figure 7), no knowledge is available and a “conventional” stochastic search is executed. The search moves depend on the variables being optimised in each of the two optimisation steps:

- Step 1: Only moves related to the optimisation of discrete variables are considered, which includes all moves except the change in the volume of a single reactive unit and the fraction of a stream;
- Step 2. Both continuous and discrete variables are optimised except for the number and types of units.

In the consequent stages the search moves are customised to clustering process (Figure 8). Feature characterised as irrelevant in the clustering analysis are removed from the moves selection list.

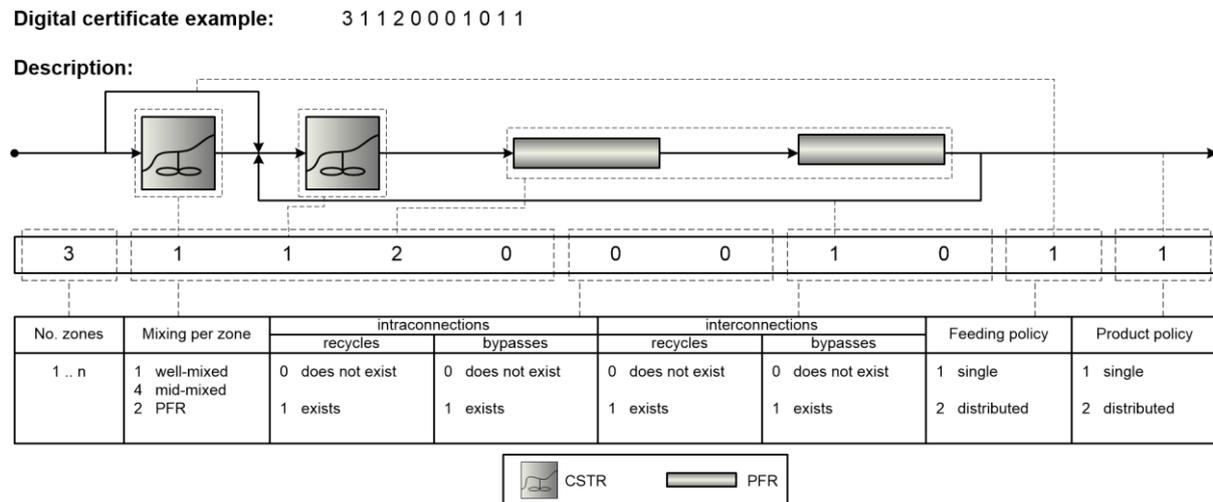


Figure 10 A digital certificate for a solution of a single phase reactor network

4.2 Multiphase Systems

A superstructure representation of multiphase systems is proposed which follows on simplification introduced by (Linke and Kokossis 2003a) and adoption of generic reactor/mass exchanger (RMX) units by (Mehta and Kokossis 1997; Mehta and Kokossis 2000).

RMX units allow for a compact representation of all possible design options resulting from the possible contacting and mixing patterns that exist between different phases. They consist of compartments belonging to each of the two phases with each phase compartment having corresponding shadow phase compartments. With a single RMX unit, a reactor, a mass exchanger, a reactive mass exchanger or the combination of them can be represented. In each phase compartment, different reciprocally excluding mixing patterns are introduced. Mixing options include well-mixed and plug flow. The plug flow behaviour typically modelled with differential equations is approximated by a series of equal volume CSTR units following approach proposed by (Kokossis, A. and Floudas 1990): complex differential equations are all replaced by simplified algebraic equations. Shadow compartments are presented in the additional phases and may be linked to their matching compartments through mass transfer. The mass exchange is limited to a pair of

compartments that belong to different phases. As a result of the different mixing patterns considered in each phase and of the links between phases, for a two phase system, a single unit can represent conventional designs, i.e. mechanically agitated reactors (CSTR/CSTR), bubble column reactors (CSTR/PFR), co-current or counter-current packed bed reactors (PFR/PFR) and spray column reactors (PFR/CSTR). Mass is exchanged across boundaries as a result of phase equilibrium or diffusional transport and represents the only way of interaction between contacting phases. Each phase compartment is fed with a stream (Figure 11). The outlet stream can leave the compartment, can be recycled within the compartment or it can be sent to another compartment across the boundary when physically and technically possible. e.g. state change operations. The existence of mass exchange between the compartments of RMX units and the existence of reaction inside a compartment are design variables included in the unit representation. This feature allows for the deactivation of reaction and/or mass transfer, making in some cases the introduction of additional phases not necessary in some parts of the network. The possible existence of catalysed reactions in the reactive phase is also considered in the RMX representation.

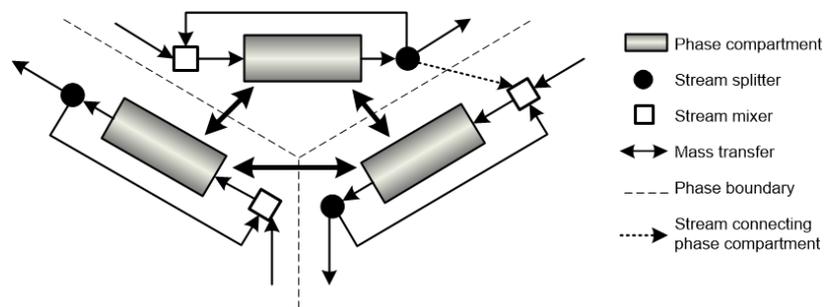


Figure 11 RMX unit composed of three phase compartments

The superstructure representation employed here involves RMX units, raw material sources, product sinks and all the physically possible connections between RMX units by means of a stream network, as shown in Figure 2. The process of synthesis involves decisions on the number of units (up to four), unit sizes, feeding, bypassing, recycling, mixing patterns and flow patterns. As illustrated in Figure 2, in each phase, the arrangement of compartments in series/parallel is possible through mixers and splitters. For a given phase, the stream network that allows for connections between compartments (interconnecting streams, recycles and bypasses) does not interact with the network stream of its contacting phase. However, the outlet stream of a phase compartment may be sent to another phase

compartment across the phase boundaries when physically and technically possible. The arrows linking a pair of compartments across the phase boundary in Figure 2 stand for the mass transfer that takes place between a sub-unit and its matching shadow sub-unit contained in the shadow phase compartment.

A formal ontology O in the domain of multiphase reactor networks synthesis is created by reusing the single phase reactor network ontology (Table 1). It is directly linked to the multiphase superstructure representation employed. The concepts N_i of the first four levels of abstraction are given in Appendix 6. The explanations of concepts, although the names are self-explanatory, are also given in Appendix 6, whereas the subsumption class-subclass relationship H_C is given by the indentation in the left-hand column. It is apparent that the concepts N_i identified for multiphase reactor networks and which represent respective classes S_i^f relate to the design variables included in the superstructure employed for this application, including the definition of: i) the reactive/non-reactive nature of the phases involved in the network (NetworkPhase), ii) the physical state of the phases (PhaseState), iii) additional features of the network (NetworkFeatures) such as combined MixingPatterns, OverallFlowArrangement and MassTransfer, and iv) features specific to each phase such as Loops (for both phases) and Recycles (for the non-reactive phase). Besides, some features such as FeedingPolicy and ProductPolicy, which were represented as NetworkFeatures for single phase systems, have been relocated to account for existence of more than one phase in the system.

Relationships R_i^C between classes of the multiphase ontology mainly refer to establish operational static and dynamic associations between them: mixing pattern, number of zones, existence of connections between units and between phases, volume, temperature and performance in terms of objective value. All relationships are listed in Table 4 where the names are self-explanatory and respective subsumption is indicated by indentation in the left-hand column: relationship *hasConnection* has sub-relationships *hasInterConnection*, and *hasIntraConnection*.

Table 4 Relationships between domain and range classes of multiphase systems

Object property	Domain class S_i^I	Range class S_j^I
<i>hasNetworkPhase1</i>	<i>Network</i>	<i>NetworkPhase</i>
<i>hasNetworkPhase2</i>	<i>Network</i>	<i>NetworkPhase</i>
<i>hasPhaseState1</i>	<i>Network</i>	<i>PhaseState</i>
<i>hasPhaseState2</i>	<i>Network</i>	<i>PhaseState</i>
<i>hasNumberOfUnits</i>	<i>Network</i>	<i>NumberOfUnits</i>
<i>hasMixingPattern1</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasMixingPattern2</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasMixingPattern3</i>	<i>Network</i>	<i>MixingPattern</i>
<i>hasOverallFlowPattern</i>	<i>Network</i>	<i>OverallFlowPattern</i>
<i>hasMassTransfer1</i>	<i>Network</i>	<i>MassTransfer</i>
<i>hasMassTransfer2</i>	<i>Network</i>	<i>MassTransfer</i>
<i>hasMassTransfer3</i>	<i>Network</i>	<i>MassTransfer</i>
<i>hasSize</i>	<i>Network</i>	<i>Size</i>
<i>hasTemperatureProfile</i>	<i>Network</i>	<i>TemperatureProfile</i>
<i>hasPerformance</i>	<i>Network</i>	<i>Performance</i>
<i>hasLoop1</i>	<i>Network</i>	<i>Loops</i>
<i>hasLoop2</i>	<i>Network</i>	<i>Loops</i>
<i>hasFeeding1</i>	<i>Network</i>	<i>FeedingPolicy</i>
<i>hasFeeding2</i>	<i>Network</i>	<i>FeedingPolicy</i>
<i>hasProduct1</i>	<i>Network</i>	<i>ProductPolicy</i>
<i>hasProduct2</i>	<i>Network</i>	<i>ProductPolicy</i>
<i>hasRecycles</i>	<i>Network</i>	<i>Recycles</i>
<i>hasConnections</i>	<i>Network</i>	<i>Connections</i>
<i>hasInterConnection</i>	<i>Network</i>	<i>InterConnections</i>
<i>hasIntraConnection</i>	<i>Network</i>	<i>IntraConnections</i>

Digital certificates for multiphase systems are created to acquire knowledge on:

i) Features specific to each phase:

- Mixing pattern of each phase compartment, to include well-mixed, mid-mixed and PFR;
- Connections between phase compartments, to include:
 - Sequential connections included only for non-reactive phase to distinguish between configurations in series and in parallel;
 - Loops as sequential connections between two non-consecutive phase compartments to ensure the connectivity of some of the units of the network, i.e. loops ensure the connection of its source unit or/and its sink unit to another unit. Sequential connections and loops allow for the description of the arrangement of the compartments (series or parallel);
 - Recycles and bypasses classified as intra- and interconnections and are represented;

- Feeding policy:
 - In the non-reactive phase, it refers to the feed streams connected to a compartment;
 - In the reactive phase, it has the same representation as for single phase networks;
 - Product policy:
 - In the non-reactive phase, it refers to the product streams from compartments;
 - In the reactive phase, as for the feeding policy determining single or distributed product strategy;
- ii) General features:
- Number of RMX units;
 - Existence of mass transfer in each of the RMX units;
 - Flow pattern of each RMX unit with options of co-current and counter-current flow.

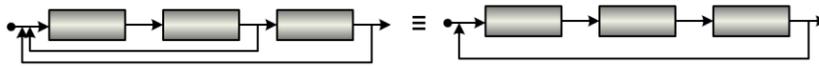
As with the single phase network, the digital certificates are created by production rules which relate to active RMX units and connections between them. Rules are applied to each of the phases provided that the nature of the contacting phenomena between phases is not altered, i.e. the mass transfer distribution and flow arrangement of the RMX units remain unchanged. The set of rules used for multiphase systems are summarised in Table 5.

Table 5 Feature customisation and creation of digital certificates for multiphase systems

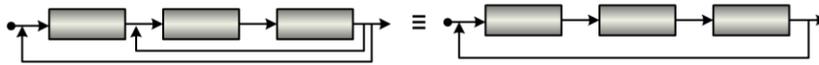
Rule 1:	RMX network reduction: Active RMX units with marginal contribution, i.e. change of the desired component concentration lower than 2% respect to its maximum value in the network, are removed from the network provided that the RMX unit does not function as a reactive mass exchanger and when the unit is not involved either in recycles or bypasses.
Rule 2: [†]	Phase compartment classification - pattern identification: Phase compartments are classified in three categories depending on the mixing pattern they present: <i>PFR:</i> PFRs without or with a low degree of back-mixing (recycle fraction lower than 12% of the outlet flow); <i>Mid-mixed:</i> PFRs with a medium degree of back-mixing (recycle fraction between 12% and 60%); <i>Well-mixed:</i> CSTRs and PFRs with high degree of back-mixing (recycle fraction over 60%).
Rule 3:	Replacement of multiple connections by an overall connection.
Rule 3a:	<i>Replacement of multiple recycles by an overall recycle:</i> the existence of multiple recycles involving more than one compartment in one phase may be replaced by a total recycle with source being the last active phase compartment and with sink being the first active phase compartment. The overall recycle maintains the dilution

effect achieved by the set of recycles keeping down the degree of complexity of the layout by which the compartments are displayed and interconnected, as summarised in the following scenarios:

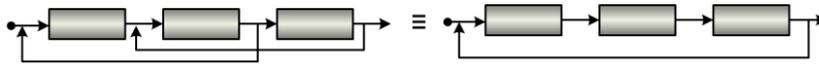
Scenario A



Scenario B



Scenario C

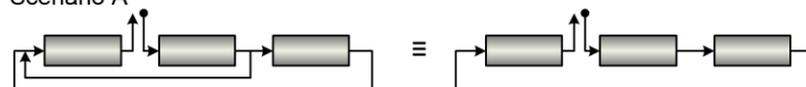


This rule is applicable when all active phase compartments are involved in at least one recycle.

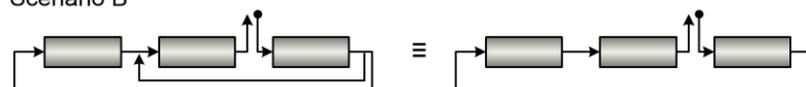
Rule 3b: *Replacement of multiple loops by an overall loop*: Loops are sequential connections between two non-consecutive phase compartments. Unlike recycles, their removal could result in unfeasible structures. Loops are necessary streams that ensure the connectivity of some of the units of the network, i.e. loops ensure the connection of its source unit or/and its sink unit to another unit. The existence of multiple loops connecting more than one compartment in one phase may be replaced by a total loop. The overall loop maintains the sequential connection achieved by the set of loops keeping down the degree of complexity of the layout by which the compartments are displayed and interconnected. This rule is applicable when all active phase compartments are involved in at least one loop, which can be summarised in the following situations:

Open loops:

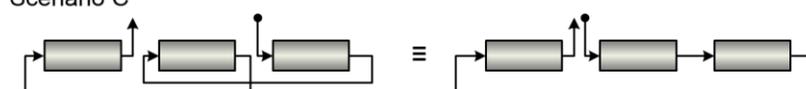
Scenario A



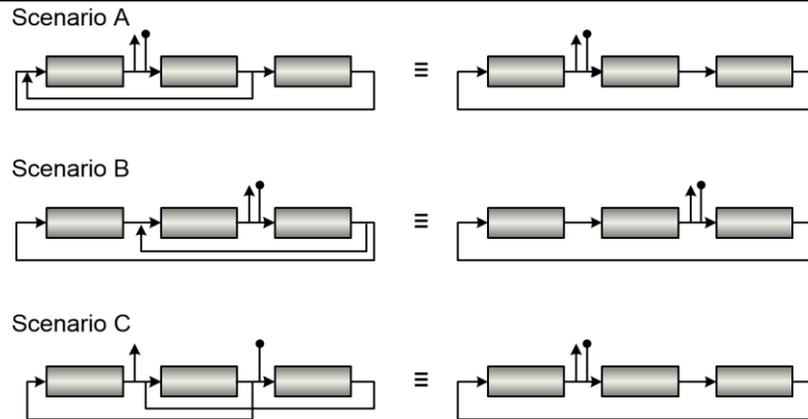
Scenario B



Scenario C



Closed loops:



- Rule 4: **Feeding distribution limitation:** no feed distribution is considered if a phase compartment receives between 70%-95% (upper optimisation bound) of the total feeding flow fed to that phase. This rule is applicable provided that:
The removal of the feed distribution does not imply the addition of new connections between phase compartments.
No side-product exists in between the phase compartments involved.
- Rule 5: **Connections:** Stream connections are arranged according the previous rules for each of the phases. They are classified into sequential connections (including loops) and recycles and bypasses, which are also classified into intra- and interconnections.

[†] Unlike single phase systems, RMX units are not allowed to be merged due to the action of a recycle or when sharing the same mixing pattern. For merging to be possible, a minimum of two consecutive units should be sequentially connected and without interconnecting streams. Besides, due to the existence of an additional phase, the units should share the same mass transfer characteristics and flow direction configuration. Such scenario has been found to be very rare for multiphase applications as usually several connections between units exist making merging units unlikely.

The link between the synthesis representation and the respective digital certificates is illustrated in Figure 12, where the top phase of the structure corresponds to the non-reactive phase and the lower phase is the reactive phase.

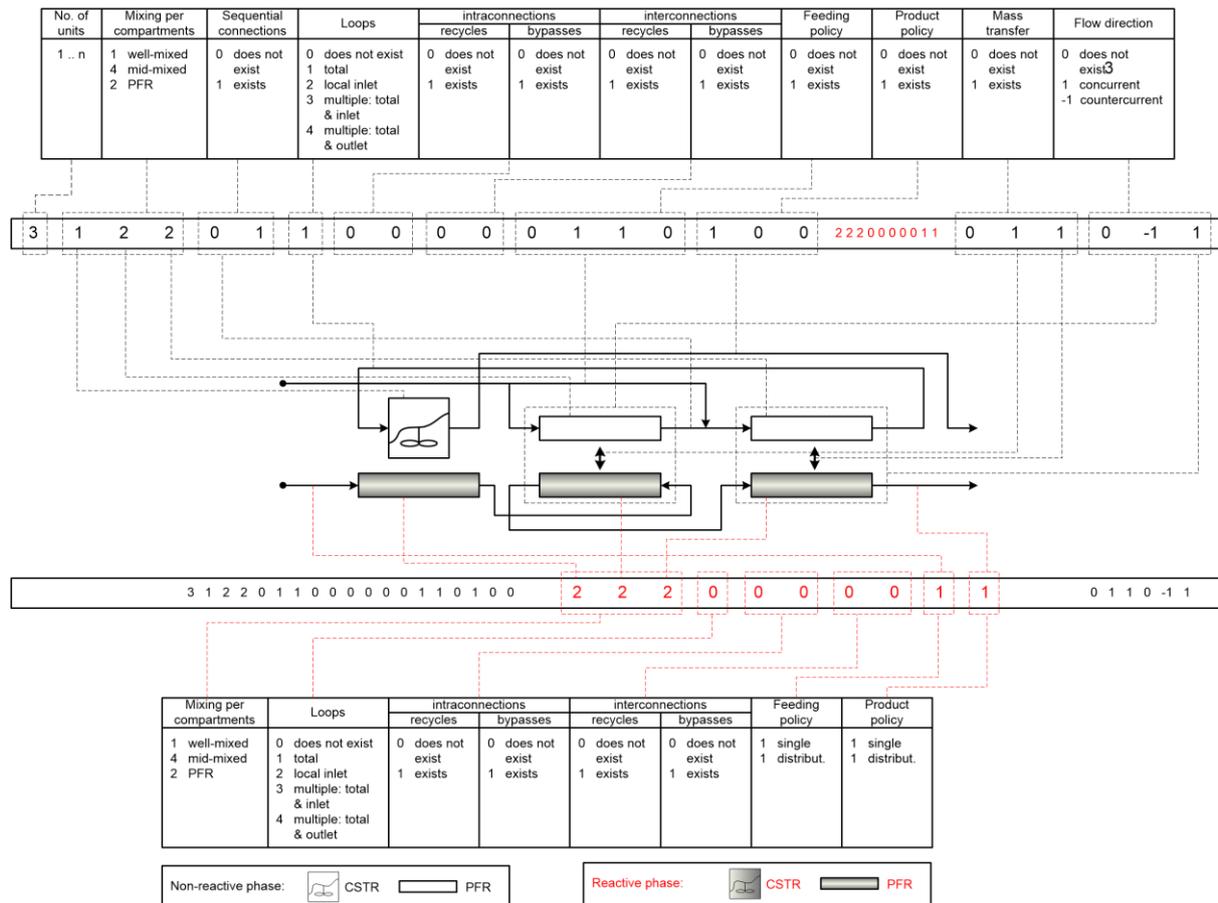


Figure 12 Digital certificate for solutions of a multiphase system

Optimisation transition between stages for multiphase systems involves new options which arise from the existence of additional phase, i.e. mass transfer and the contacting pattern between phases. In the first stage of the optimisation, the Stage 0 (Figure 7), no knowledge is available and a “conventional” stochastic search is performed. The moves considered in it, depend on the variables being optimised in each of the two optimisation steps:

- Step 1: Only moves related to the optimisation of discrete variables are considered, which includes all moves except the change in the volume of a single reactive unit and the fraction of a stream;
- Step 2. Both continuous and discrete variables are optimised except for the number and types of units.

In the consequent stages the moves are customised to clustering process (Figure 8). Those moves related to features that have proven to be irrelevant in the clustering analysis are removed from the moves selection list. However, with multiphase systems some exceptions

apply. The activation/deactivation of mass transfer and the change of the contacting flow pattern of a RMX unit are allowed throughout the whole optimisation to avoid constraining these contacting phase features to the structural decisions made in early stages of the search. In a similar way, feeding and product policies are degrees of freedom throughout the process to enable a progressive accommodation of temporary design features in order to make the transition between solutions possible. Also, we simplify the process of optimising continuous variables as proposed by (Vanderbit and Louise 1984; Mehta and Kokossis 1998). Changes in the continuous variables, i.e. volume and stream split fractions, are possible by performing step changes of variable size within the acceptable regions of operation from lower to upper bounds. Split fractions are modified continuously within the lower and upper bounds range $[f_{min}, f_{max}]$. The magnitude of the change for RMX unit volumes is probabilistically selected and the change is based on a random multiplier. Three different changes are considered for RMX unit volumes depending on the three probabilities (P_1, P_2, P_3) that are user defined such as they add up to 1:

$$\begin{aligned}
 V_i^{new} &= \max\{V_{min}, \min\{rnd_1 \cdot V_i^{current}, V_{max}\}\} && \text{with } P_1 \\
 V_i^{new} &= \min\{rnd_2 \cdot V_i^{current}, V_{max}\} && \text{with } P_2 \\
 V_i^{new} &= \max\{V_i^{current}/rnd_2, V_{min}\} && \text{with } P_3
 \end{aligned} \tag{9}$$

where $V_i^{current}$ is the current volume of a RMX unit on which the change is applied, V_i^{new} is the resulting volume after the change, rnd_1 is a random number in the interval $[0.5, 1.5]$, rnd_2 is a random integer from the interval $[2, 10]$, and V_{min} and V_{max} are the volume lower and upper bounds, respectively. In final stages, smaller changes on RMX unit volumes are introduced to allow for the increase or decrease of the current volume $V_i^{current}$ within a 10% range, provided the resulting volume falls within the bounds. As a result, the optimisation is intensified reducing the number of final optimisation stages.

5 Experimental Evaluation

5.1 Illustrative Examples

The synthesis examples are selected to demonstrate reasonable complexity and they are sufficiently demanding to justify implementation of proposed knowledge supported optimisation framework. Their formulations include two experiments for single phase

reactor network and one for multiphase reactors network. Exact formulations of experiments are presented in the Appendix 1 – Appendix 3. A comparative analysis of results with conventional optimisation algorithms, namely Tabu Search and Simulated Annealing, as well as with published ones is also presented.

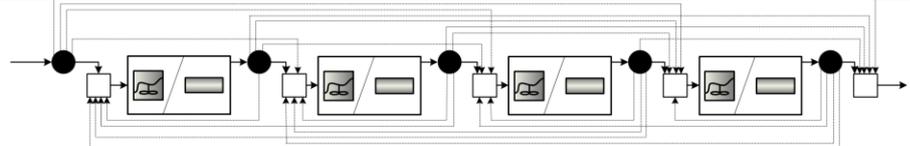
5.2 Synthesis of Single Phase Reactor Network

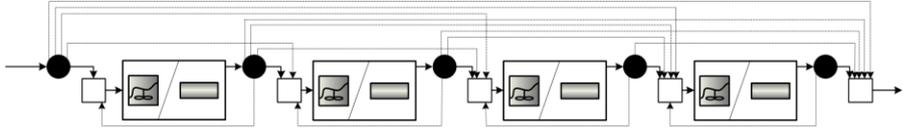
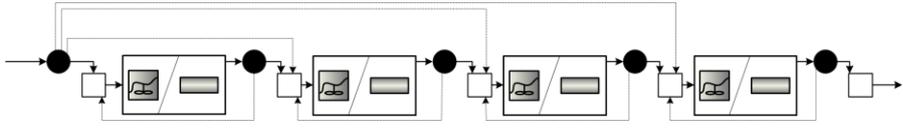
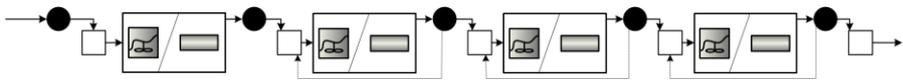
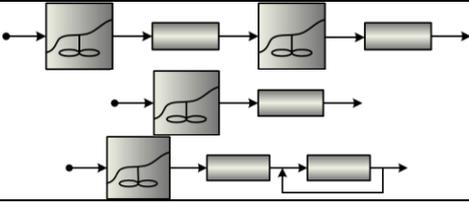
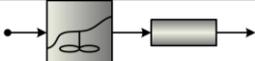
5.1.1 Van de Vusse Reaction

Van de Vusse reaction system consists of the four reactions (Kokossis, A. and Floudas 1990) and the complete kinematics is given in Appendix 1. The respective evolution of the superstructure, the selected clusters and their maximum objective values are illustrated in Table 6. At Stage 0 ten different initial structures are optimised for 20 iterations each. The maximum objective obtained for each of them is after selecting the best 50% of performing clusters. The solutions presented become the initial solution for the next stage, the stage 1, as shown in Table 6.

Stage 1 eliminates the possibility of well-mixed behaviour in the final reactive zone and removes connections of type recycle between reactive zones. Stage 2 removes connections of type bypass between reactive zones and fixes the superstructure to three reactive zones with plug flow/mid-mixed behaviour. Feed distribution is also removed. The analysis performed at Stage 3 results in three near-optimal designs and at Stage 4, two similar structures are obtained. They both fall into the best 2% performing solutions classification. The small differences in the objectives are due to the presence of a minor recycle in the first structure. In Stage 5, a single PFR with volume range between 25.70 L and 25.86 L results as the optimal solution.

Table 6 Evolution of the superstructure and results for Van de Vusse reaction

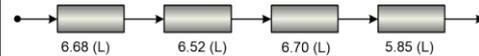
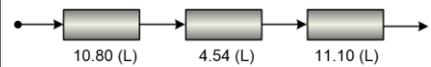
Superstructure	Digital certificates	Max obj (mol/L)
Stage 0		
	-	-
Stage 1		

	21100000021 31110000021 12000000011 32410000011 22400000021 32120000011 22100000021 11000000011 21200000011 14000000011 24200000021 31120000021 31140000021 21200000111 41212000011 41212100011 21200100011 12000100011 41214000011 21400000011	3.4695 2.7397 3.5513 2.3991 3.4356 2.5033 3.5322 3.2645 3.6488 3.5357 3.6004 3.6166 3.5749 3.6193 3.6445 3.6085 3.6395 3.5852 3.6298 3.5942
Stage 2		
	12000000011 22200000021 22100000011 21200000021 14000000011 12000100011 24200000011 31120000021 31140000011 41212000011 41214000011 31240000011 21400000011 21200100011	3.5568 3.5621 3.5322 3.6527 3.5408 3.5474 3.6013 3.6178 3.5919 3.6445 3.6369 3.6402 3.5942 3.6040
Stage 3		
	21200000011 41212000011 41214000011 31240000011 21200100011	3.6527 3.6445 3.6369 3.6413 3.6447
Stage 4		
	41212000011 21200000011 21240000011	3.6445 3.6527 3.6413
Stage 5		
	21200000011	3.6527

The final results compared to the conventional Tabu Search as well as results from the literature are presented in Table 7. With comparatively similar optimal objective value, the concentration of the product (mol/L), the simplicity of the structure with only one PFR

reactor is apparent. In addition, approved with more than 10 different configurations, the proposed methodology converges faster saving more than 30% of computation time for the same configuration.

Table 7 Van de Vusse reaction results comparison

Van de Vusse	Status	Functions evaluations	Objective	Volume (L)	Optimal structure ⁽⁴⁾
This work	Total ⁽¹⁾	4502	0.4326	25.86	
Tabu Search	ave ⁽²⁾	974	0.4324	25.84	
	max ⁽²⁾	1878	0.4326	26.71	
	min ⁽²⁾	465	0.4318	25.53	
Simulated Annealing (Marcoulaki, E. C. and Kokossis 1999)	ave ⁽³⁾	-	0.4293	26.97	
	max ⁽³⁾		0.4304	28.97	
	min ⁽³⁾		0.4284	25.08	

⁽¹⁾ Total number of function evaluations for all stages. At Stage 0, 10 initial runs are considered. For the next stages, the number of runs is equal to the number of clusters selected in the previous stage as these are used as initial structures for the runs to be performed in the next optimisation stage

⁽²⁾ Summary statistics based on 10 runs

⁽³⁾ Summary statistics based on 5 runs

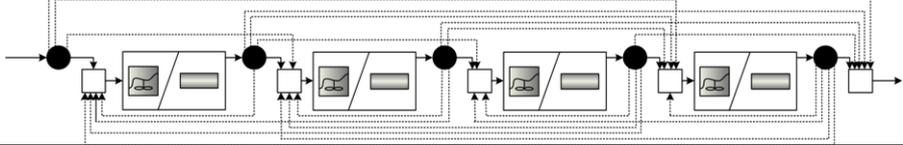
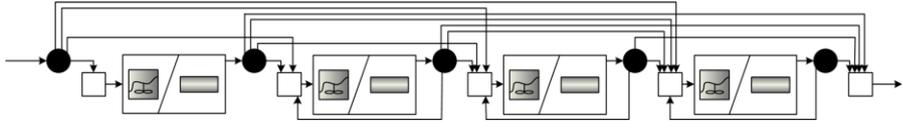
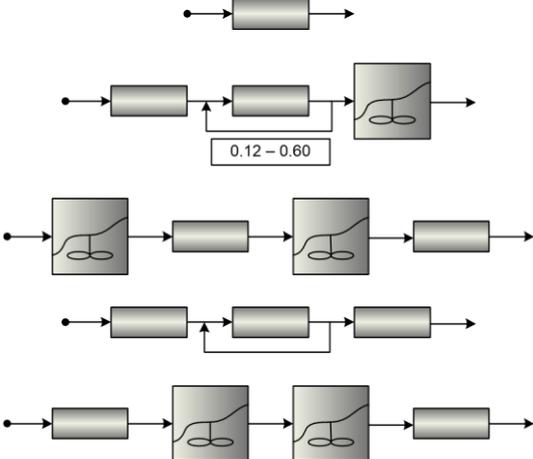
⁽⁴⁾ Best for 10 runs for Tabu Search and 5 runs for Simulated Annealing based on different initial structures

5.1.2 Denbigh Reaction Results

The Denbigh reaction scheme is defined by four reactions that involve five components, as shown by the kinetics in Appendix 2 (Kokossis, A. and Floudas 1990). The respective evolution of the superstructure, the selected clusters and their maximum objective values are illustrated in Table 8.

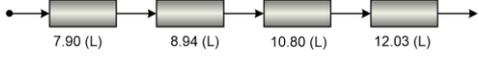
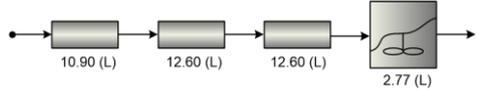
As with Van de Vusse reaction, at Stage 0 ten different initial structures are optimised for 20 iterations each. The maximum objective obtained for each of them after selecting the best 50% of performing clusters. The solutions presented become the initial solution for the next stage, the stage 1. At Stage 1 (Table 8), connectivities of type recycle between reactive zones and mid-mixed pattern in the first and the latest unit are eliminated. At Stage 2 five clusters are selected. It appears that those structures with no well-mixed behaviour in any of the units perform better. As a result, a single PFR is obtained as the best solution in Stage 3.

Table 8 Evolution of superstructure for Denbigh reaction

Superstructure	Digital certificates	Max obj (mol/L)
<p>Stage 0</p> 	-	-
<p>Stage 1</p> 	<p>1200000011 0.2525 2210000011 0.2016 3212000011 0.2002 12000010011 0.2460 22100010011 0.1995 21200000011 0.2233 22400000011 0.2303 32410000021 0.2455 42421000021 0.2173 31410000021 0.2279 41412000011 0.1946 41121000011 0.2144 41212000111 0.2339 41112000111 0.2174 21200000111 0.2278 32420000011 0.2507 32440000011 0.2358 41211000111 0.2295 42112000111 0.3257 31210000011 0.2352 31420000011 0.2363 41212000011 0.2438 41112000011 0.2398 41211000011 0.2374 42112000011 0.2445</p>	
<p>Stage 2</p> 	<p>12000000011 0.2526 32410000011 0.2455 41212000011 0.2438 32420000011 0.2507 42112000011 0.2445</p>	
<p>Stage 3</p> 	12000000011	0.2527

The results obtained in this work compare well to those obtained in the literature (Marcoulaki, E. C. and Kokossis 1999) and with those obtained by conventional Tabu Search (Table 9). It is apparent that results from Simulated Annealing optimisation offer a small CSTR as the final reactor, which is equivalent to having an extra sub-CSTR at the end of the previous PFR, hence performing the same as a single PFR. The discretisation into 28 sub-CSTRs of the single PFR found in this work delivers a higher quality solution.

Table 9 Comparison of results for Denbigh reaction

Denbight	Status	Functions evaluations	Objective	Volume (L)	Optimal structure ⁽⁴⁾
This work	Total ⁽¹⁾	10932	0.2578	45.84	
Tabu Search	ave ⁽²⁾	916	0.2539	39.77	
	max ⁽²⁾	1077	0.2539	40.28	
	min ⁽²⁾	603	0.2537	39.35	
Simulated Annealing (Marcoulaki, E. C. and Kokossis 1999)	ave ⁽³⁾	-	0.2524	39.50	
	max ⁽³⁾		0.2528	40.24	
	min ⁽³⁾		0.2519	38.87	

⁽¹⁾ Total number of function evaluations for all stages. At Stage 0, 10 initial runs are considered. For the next stages, the number of runs is equal to the number of clusters selected in the previous stage as these are used as initial structures for the runs to be performed in the next optimisation stage

⁽²⁾ Summary statistics based on 10 runs

⁽³⁾ Summary statistics based on 3 runs

⁽⁴⁾ Best for 10 runs for Tabu Search and 3 runs for SA based on different initial structures

5.2 Synthesis of Multiphase Reactor Network: Modified Denbigh Reaction

Modified Denbigh Reaction is a gas-liquid multiphase version of the Denbigh reaction proposed by (Mehta and Kokossis 2000) and with kinetics which is in detailed explained in Appendix 3. The objective is to minimise the yield of the target component B. The selected probabilities for moves related to the optimisation of discrete variables in the first and consequent optimisation stages are given in Table 10.

Table 10 Perturbation probabilities for multiphase Denbigh reaction

Move [†]	Probability	
	Stage 1	Other stages
Phase change: phase 1, phase 2	0.7, 0.3	0.7, 0.3
RMX unit, stream	0.3, 0.7	0.3, 0.7
RMX unit moves: add, delete, modify, change type	0.35, 0.35, 0.0, 0.3	0.0, 0.0, 0.9, 0.1

Streams moves: add, delete, modify	0.5, 0.5, 0.0	0.6, 0.2, 0.2
Network expansion: in series, in parallel	0.5, 1.0	0.5, 1.0
RMX unit modifications: volume, switch mass transfer	0.5, 0.5	0.5, 0.5
RMX unit type changes: mixing pattern, flow direction	0.5, 0.5	0.0, 1.0
RMX unit volume modification (P_1, P_2, P_3, P_4)	0.6, 0.2, 0.2, 0.0	0.6, 0.2, 0.2, 0.0 ^{††}
Mixing pattern selection: type 1, type 2, type 3, type 4, type 5	0.25, 0.25, 0.25, 0.25, 0.25	0.0, 0.0, 0.0, 0.0, 0.0
CSTR selection: phase 1, phase 2	0.5, 0.5	0.0, 0.0
PFR selection: phase 1, phase 2	0.5, 0.5	0.0, 0.0
Flow direction: co-current, counter-current	0.5, 0.5	0.5, 0.5
Mass transfer switches: deactivated, activated	0.3, 0.7	0.3, 0.7

[†] P_1 : probability to change volume within $\pm 50\%$; P_2 : probability to increase RMX unit volume up to $\min(V_{max}, 10 \cdot volume)$, P_3 : probability to decrease RMX unit volume up to $\max(V_{min}, volume/10)$; P_4 : probability to change volume within $\pm 10\%$. Mixing patterns type 1: CSTR/CSTR; type 2: CSTR/PFR; type 3: PFR/CSTR; type 4: PFRs co-current; type 5: PFRs counter-current.

^{††} In final stages this probability becomes 1.0 and P_1, P_2 and P_3 are set to 0.0.

The summary of the optimisation results, the clusters selected and the maximum objective values achieved, are shown in Table 11.

Table 11 Results of the optimisation for Denbigh multiphase reaction

Stage	Selection criteria (%)	Clusters selected				Clusters generated	Maximum objective (kmol/h)
		One unit	Two units	Three units	Total		
1	50	3	20	11	34	46	11.4472
2	10	3	19	3	25	342	11.8905
3	5	3	15	3	21	318	12.2965
4	2	3	3	3	9	228	12.4265
5	1	3	3	3	9	63	12.7823
6	1	3	3	3	9	10	12.7823
7	1	3	3	3	9	9	12.7852
8	1	3	3	3	9	9	12.7906
9	1	3	3	3	9	9	12.8034
10	1	3	3	3	9	9	12.8034

To facilitate understanding of the results presented in Table 11, the results obtained for three unit network are taken as representative and are explained in details. From early stages (Table 12), the mixing tendency for all units is PFR behaviour ($mixing\ per\ unit = \{4,5\}$), which becomes the only option from intermediate stages, i.e. from Stage 3. In all stages a sequential arrangement of units with single feeding and product policy for the reactive phase (phase 2) is preferred ($feeding\ policy = 1, product\ policy = 1$). Neither loops nor recycles are identified throughout the stages.

Table 12 Evolution of superstructure and results for Denbigh multiphase reaction

Stage	Cluster ID	Compact digital certificate															Optimum volume (m ³)	Maximum objective (kmol/h)		
		No. of units	Mixing per unit	Overall flow arrangement	Phase 1						Phase 2								Mass transfer	
					Loop	Recycle	Feeding policy			Product policy	Loop	Recycles	Bypasses	Intracon		Feeding policy				Product Policy
							Single	Distributed	Bypass					Recycles	Bypasses					
1	17	3	4 5 5	1	0 0 1 0 0 2	0 0 0 0 0 1 1	0 0 0 0 0 1 1	1 1 1	150.00	7.3333										
	16	3	4 3 5	1	0 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	150.00	7.4410											
	15	3	5 3 5	1	0 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	150.00	7.8866											
	34	3	4 4 5	2	3 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1391.05	8.8777											
	32	3	1 4 4	2	3 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1391.05	8.9293											
	2	3	1 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 1 1	450.00	9.0077											
	3	3	3 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 1 1	450.00	9.0451											
	33	3	4 4 4	2	3 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1391.05	9.1370											
	18	3	2 5 2	1	0 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1250.00	9.4832											
	4	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 1 1	450.00	10.0438											
	22	3	2 4 2	1	0 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1100.00	10.0495											
2	25	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 1 1	450.00	10.0438											
	24	3	2 4 2	1	0 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	1100.00	10.0495											
	23	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 2	1 1 1	480.70	10.2831											
3	21	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 0 1	480.70	10.0514											
	20	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 1	1 0 0	260.31	10.2144											
	19	3	5 4 4	1	0 0 1 0 0 1	0 0 0 0 1 1 2	1 1 1	480.70	10.2831											
4	9	3	5 4 4	3	1 1 0 1 0 3	0 0 0 0 1 1 1	1 1 0	651.30	11.6208											
	8	3	5 4 4	3	1 0 0 1 0 3	0 0 0 0 1 1 1	1 1 0	648.86	11.6804											
	7	3	5 5 4	3	1 1 0 1 0 3	0 0 0 0 1 1 1	1 1 0	488.79	11.6868											
5	9	3	5 4 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	479.26	12.5538											
	8	3	5 4 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 0	529.28	12.6919											
	1	3	5 5 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 0	508.80	12.7823											
6	9	3	5 4 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	497.55	12.5755											
	8	3	5 4 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 0	529.28	12.6919											
	1	3	5 5 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 0	508.80	12.7823											
7	9	3	5 4 4	2	1 0 1 0 0 1	0 0 0 0 0 1 1	1 1 1	497.55	12.5755											

	8	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	512.15	12.7011
	1	3	5	5	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	510.73	12.7852
8	9	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	1	497.55	12.5755
	8	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	512.15	12.7011
	1	3	5	5	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	506.61	12.7906
9	9	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	1	497.55	12.5755
	8	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	512.15	12.7011
	1	3	5	5	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	506.96	12.8034
10	9	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	1	497.55	12.5755
	8	3	5	4	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	512.15	12.7011
	1	3	5	5	4	2	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	506.96	12.7034

At Stage 4, a change in the overall flow pattern is observed evolving from co-current configuration (*overall flow* = 1) to intermediate (*overall flow* = 3). This is due to the appearance of a total loop connecting the extreme compartments (*loop* = 1), which in some cases appears as a closed loop presenting a dilution effect (*loop* = 1 and *recycle* = 1). Besides, changes on the feeding strategy from single to distributed (*feeding policy: distributed* = 1) and on the product strategy from single (*product policy* = 1) to distributed towards the last two units (*product policy* = 3) are observed, which appear to accommodate appearance of new loop. These changes correspond to intermediate steps of the network evolution in order to evolve from a co-current flow arrangement at Stage 3 towards an overall counter-current flow arrangement (*overall flow* = 2) found from Stage 5 onwards, where both the feeding strategy and the product policy become single again. As for the gas phase, the layout of the reactive phase remains constant from Stage 5.

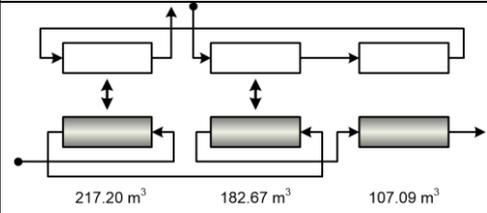
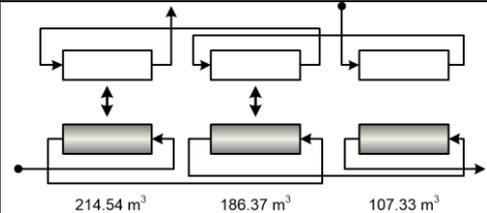
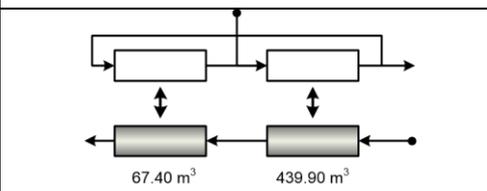
In the gas phase, partial loops connecting the final units (*loop* = 3) identified at Stage 1, are removed at Stage 2. In the reactive phase, bypasses are removed as a consequence of the further optimisation of the gas feeding strategy (transition from Stage 3 to Stage 5).

Variations in the mass transfer for the two latest compartments are observed at Stages 3 and 4 but appear to be fixed in later stages with preference of mass transfer in all compartments but the final one. Such configuration gives better performance than transferring mass along all compartments.

Counter-current arrangement (*mixing* = 5) for the first compartment is early identified (Stage 3). A co-current arrangement (*mixing* = 4) with a preference for no mass transfer in the third compartment is preferred from Stage 5. From Stage 5 onwards it is observed that by considering counter-current display in the middle compartment instead of the co-current option, higher OFVs are obtained.

Final results compared to optimisation by conventional Simulated Annealing and the published one are presented in Table 13.

Table 13 Comparison of results for multiphase Denbigh reaction

Denbight	Status	Functions evaluations	Objective	Volume (L)	Optimal structure ⁽⁴⁾
This work	Total ⁽¹⁾	14821	12.8034	506.96	
Simulated Annealing	ave ⁽²⁾	3880	12.8023	506.64	
	max ⁽²⁾	5306	12.8036	513.63	
	min ⁽²⁾	3069	12.8006	499.70	
Simulated Annealing (Mehta and Kokossis 1998)	ave ⁽³⁾ max ⁽³⁾ min ⁽³⁾	-	12.2840 12.7800 11.7700	-	

⁽¹⁾ Total number of function evaluations for the runs performed from Stage 0 until last stage. In Stage 0, 10 initial runs are considered. For the next stages, the number of runs is equal to the number of clusters selected in the previous stage as these are used as initial structures for the runs to be performed in the next optimisation stage

⁽²⁾ Summary statistics based on 10 runs

⁽³⁾ Summary statistics based on 16 runs

⁽⁴⁾ Best for 10 runs for SA and 16 runs for SA based on different initial structures

The optimal solution suggested by Simulated Annealing agrees with those obtained in this work. The pair of loops found in the optimal solution for Simulated Annealing corresponds to the total loop identified in this work. The flow arrangement from the last unit is equivalent in both cases (the counter-current/co-current concept loses significance with no

mass transfer). Computational efforts are reduced by 62% when using the proposed method, if compared with a classical Simulated Annealing approach. The presented results outperform those found in the literature both in quality and maximum objective identified.

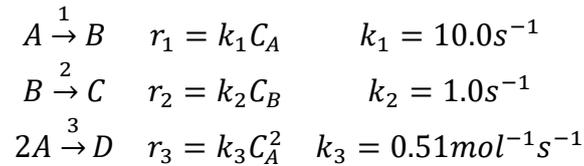
6 Conclusion

A novel framework for engineering synthesis and design is presented. The proposed approach benefits from knowledge modelling, acquisition and processing to interpret optimisation solutions, to learn from the progress of optimisation and to guide the search respectively, and to simplify solutions dynamically in line with the problem formulation. Ontologies supported by production rules are used to model knowledge and hence to systematise solution features by their performance and contribution. Tabu search is used to generate new solutions and to optimize. Digital certificate are proposed for representation of design features are translate between structural information about solutions and semantic of ontology. The framework is demonstrated using superstructure representations of single phase and multi-phase reactor network optimised by the concentration of the product and network structure.

Three complex experiments of reactor network synthesis are used to assess the performance of proposed framework; Van de Vusse and Denbight reaction for single phase networks and Modified Denbigh reaction for multiphase network. Exhaustive experimentation and comparison with Tabu Search and Simulated annealing shows noticeable advantage in convergence and superiority in simplification of solutions. The main advantage is, however, in readily available interpretation of solutions and understanding of effects of design parameters.

Appendix 1: The Van de Vusse reaction

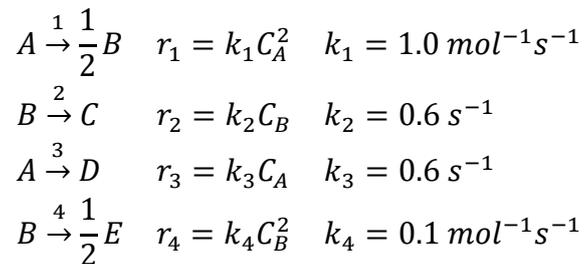
The¹ Van de Vusse reaction consists of a combination of parallel and serial reactions (Kokossis, A. and Floudas 1990; Marcoulaki, E. C. and Kokossis 1999). In this reaction scheme, *A* is the reactant, *B* is desired product, and *C* and *D* are by-products (Ashley and Linke 2004):



The feed flows at 100 l/s and contains pure A with the feed concentration of 5.8 mol/l. The objective of this reaction is to maximise the outlet concentration of desired product B while minimising by-products C and D. As a matter of reference, the best reported solution (Labrador-Darder et al. 2009), to our knowledge, is the outlet concentration of 3.66 with CSTR + PFR reactors with no recycle and by-pass flows and with total volume of 28.77 l (11.64 + 17.13 l).

Appendix 2: Denbigh Reaction

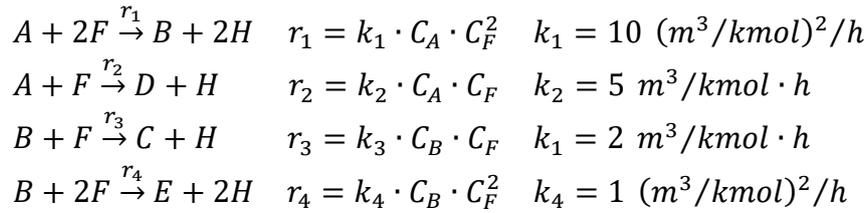
The Denbigh reaction scheme is defined by four reactions that involve five components (Kokossis, A. and Floudas 1990):



where B is the desired product. The feed flow rate is 100 L/s and consists of 6.0 mol/L of component A and 0.6 mol/L of component D. The objective is to maximise the product yield of component B (C_B/C_A^0).

Appendix 3: Modified Denbigh Reaction

The example presented here involves a gas-liquid multiphase version of the Denbigh reaction (Mehta and Kokossis 2000). The liquid phase contains five non-volatile components A, B, C, D and E. The gas feed is made of an inert component G, which is not soluble in the liquid at the process conditions, and component F that diffuses in the liquid phase and produces the volatile product H. The reactions that take place in the liquid phase are:



The liquid feed consists of pure A with a flow rate of 100 kmol/h. A constant flow rate of 200 kmol/h of F and 800 kmol/h of G is assumed as fresh gas feed. The volumetric mass transfer coefficients and the phase holdups for the liquid phase are assumed to be constant for all possible types of mixing patterns considered as shown in xxx along with the molar specific volumes. The bound values for the volume and the split fractions employed are also shown. The objective is to maximise the yield of component B.

Table 14 Data for multiphase Denbigh reaction system

Parameter	Value
Volumetric mass transfer coefficients (1/h)	$k_{1aF} = 500, k_{1aH} = 500, k_{1aG} = 0$
Molar specific volumes for liquid phase (m^3/kmol)	$V_{spA} = V_{spB} = 1.667$ $V_{spC} = V_{spD} = V_{spE} = 1.667$ $V_{spF} = V_{spH} = 0$
Henry's constants ($\text{bar} \cdot \text{m}^3/\text{kmol}$)	$H_F = H_H = 10.0$
Liquid phase holdup	$\Phi_L = 0.5$
Operating pressure (bar)	$P = 10.0$
Volume bounds V_{max}, V_{min} (m^3)	2000, 10
Split fractions bounds f_{max}, f_{min}	0.95, 0.05

Appendix 4: Single Phase Process Representations

The superstructure representation employed for single phase systems involves reactor units, raw material sources, product sinks and all the physically possible connections between them through mixers and splitters, as illustrated in Figure 1. The reactor network superstructure representation employed in this work follows analysis provided by (Mehta and Kokossis 1997) which involves the following variables and parameters.

RU	Reactor units	MI	Mixers
F	Sources of raw material	CP	Components
P	Products	SK_{ru}	Well-mixed sub-units
SP	Splitters	RX	Reaction

Partitions of the previous basic sets include the following subsets:

RU^A	Active reactor unit	MI^A	Active mixer
F^A	Active raw material	$MI_{ru,sk}^{RU}$	Mixer prior to sub-unit
P^A	Active product	MI_{ru}^{PRU}	Final product mixer
SP^A	Active splitter	MI_p^P	Mixer of product
SP_f^F	Splits a raw material	RX_{ru}^A	Active reaction
SP_{ru}^{RU}	Splits outlet of unit	CP_p^P	Component in product
$SP_{ru,sk}^{IRU}$	Splits the outlet of sub-unit	CP_f^F	Component in the inlet
SP_p^P	Splits a product stream	CP_{ru}^{RU}	Component in the outlet

Based on the previous sets, the variables employed in the formulation of the superstructure are defined. The following set of variables includes the flow rates of each component through splitters and mixers:

$\{FD_{f,cp}\}$	Component flow rate through splitters	$\{OUTSK_{ru,sk,cp}\}$	Component flow rate at the outlet of subunit
$\{INR_{ru,sk,cp}\}$	Component flow rate at the inlet of sub-unit	$\{INP_{p,cp}\}$	Component flow rate through mixers
$\{INPR_{ru,cp}\}$	Component flow rate through mixers	$\{OUTP_{p,cp}\}$	Component flow rate through splitters
$\{OUTR_{ru,cp}\}$	Component flow rate through splitters		

The next set of variables includes the split fractions of streams connecting splitters to mixers of the superstructure:

$\{SFR_{f,ru,sk}\}$	Fraction of $FD_{f,cp}$ entering mixer	$\{SRP_{r,p}\}$	Fraction of $OUTR_{ru,cp}$ entering mixer
$\{SFP_{f,p}\}$	Fraction of $FD_{f,cp}$ entering mixer	$\{SPR_{p,ru,sk}\}$	Fraction of $OUTP_{p,cp}$ entering mixer
$\{SRR_{r,ru,sk}\}$	Fraction of $OUTR_{ru,cp}$ entering mixer	$\{SKP_{ru,sk}\}$	Fraction of $OUTP_{p,cp}$ entering mixer

Variables related to the reaction rates are:

$\{RXR_{rx,ru,sk}\}$	Specific reaction rate	$\{V_{ru}\}$	Volume of reaction unit
$\{\vartheta_{rx,cp}\}$	Stoichiometric coefficient of component	$\{\epsilon_{ru,sk}\}$	Holdup in sub-unit which is expressed as a fraction

The mathematical formulation for the superstructure is defined as follows. Initially, the balance equations around the mixers are considered. Mixers $MI_{ru,sk}^{RU}$ are fed by streams

from any splitter SP_f^F and $SP_{jru \neq iru}^{RU}$ and mixers MI_{ip}^P receive streams from any splitter SP_{ru}^{RU} . Initially, the balance equations around mixers $MI_{ru,sk}^{RU}$ prior to reactor are defined as:

$$\sum_{f \in PA} FD_{f,cp} \cdot SFR_{f,ru,sk} + \sum_{r \in RU^A} OTR_{r,cp} \cdot SRR_{r,ru,sk} + PREV_{sk} + \sum_{p \in PA} OUTP_{p,cp} \cdot SPR_{p,ru,sk} - INR_{ru,sk,cp} = 0 \quad (A3.1)$$

$$\forall cp \in CP, ru \in RU^A, sk \in SK_{ru}$$

$$PREV_{sk=1} = 0$$

$$PREV_{sk \in SK_{ru} \setminus \{1\}} = OUTSK_{ru,sk-1,cp} \cdot (1 - SKP_{ru,sk-1})$$

For outlet mixers of reactor units MI_{ru}^{PRU} :

$$\sum_{sk \in SK_{ru}} OUTSK_{ru,sk,cp} \cdot SKP_{ru,sk} - OTR_{ru,cp} = 0 \quad (A3.2)$$

$$\forall cp \in CP, ru \in RU^A$$

For network product mixers MI_p^P :

$$\sum_{f \in FA} FD_{f,cp} \cdot SFP_{f,p} + \sum_{ru \in RU^A} OTR_{ru,cp} \cdot SRP_{ru,p} - OUTP_{p,cp} = 0 \quad (A3.3)$$

$$\forall cp \in CP, p \in P^A$$

The balance equations describing the well-mixed cells SK_{ru} of the reactor units are presented:

$$INR_{ru,sk,cp} + \sum_{rx \in RX_{ru}^A} \vartheta_{rx,cp} \cdot RXR_{rx,ru,sk} \cdot \frac{\varepsilon_{ru,sk} \cdot V_{ru}}{|SK_{ru}|} - OUTSK_{ru,sk,cp} = 0 \quad (A3.4)$$

$$\forall cp \in CP, ru \in RU^A, sk \in SK_{ru}$$

The following constraints for the splitters need to be satisfied for a structure to be feasible. Structural feasibility is principally related to the active connections between the mixers and splitters and is checked after a move is applied. For the split fractions associated with splitters SP_f^F :

$$\sum_{ru \in RU^A} \sum_{sk \in SK} SFR_{f,ru,sk} + \sum_{p \in PA} SFP_{f,p} - 1 = 0 \quad (A3.5)$$

$$\forall f \in F^A$$

For the split fractions associated with splitters SP_{ru}^{RU} :

$$\sum_{ru \in RU^A} \sum_{sk \in SK} SPR_{r,ru,sk} + \sum_{p \in P^A} SRP_{r,p} - 1 = 0 \quad (A3.6)$$

$$\forall r \in RU^A$$

For the split fractions associated with splitters SP_p^P :

$$\sum_{ru \in RU^A} \sum_{sk \in SK_{ru}} SPR_{p,ru,sk} - 1 = 0 \quad (A1.7)$$

$$\forall p \in P^A$$

The set of non-linear equations is solved using the NEQLU routine.

Appendix 5: The single phase reactor network domain ontology

Concepts	Description
Network	A network consists of one or combination of reactors that are interconnected in any physically possible way
OneZoneNetwork	Any network that has only one, two, three or four zones, respectively
TwoZoneNetwork	
ThreeZoneNetwork	
FourZoneNetwork	
Experiments	A network that has been generated during the specific stage of optimisation. They might be expressed in the form of the digital certificate or general solution form when the respective property are defined at application specific level (Figure 5).
NetworkFeatures	Characteristics that define the network in terms of structural and operational aspects.
NumberOfZones	Number of reactive zones that network consists of (max. 4)
Z1	Used to represent respective number of zones, 1, 2, 3, 4, respectively.
Z2	
Z3	
Z4	
MixingPattern	Describes the mixing of the zone.
Well-mixed	CSTR or PFR with reactor recycle ratio fraction >0.60.
Mid-mixed	PFR with reactor recycle ratio fraction <0.60.
PFR	PFR with reactor recycle ratio fraction <0.12.
FeedingPolicy	Refers to how the fresh feed stream is fed to the network.
Single	Flow connected from a source to a sink through one stream: fresh feed stream is fed to a single zone
Distributed	Flow split into/made from more than one stream: the fresh stream is split amongst different zones.
ProductPolicy	Refers to how the product is removed from the network
Single	Flow connected from a source to a sink through one stream: the product is removed from a single zone
Distributed	Flow split into/made from more than one stream: the product is removed from more than one zone.
Connections	Refers to the existence of streams that connect zones
IntraConnections	Connection within a zone
Recycle	Backward connection
Bypass	Forward connection
InterConnections	Connect different zones
Recycle	Backward connection

	Bypass	Forward connection
Size	Low	Volume of the network Lower than defined threshold
	Medium	Around given threshold
	High	Above given threshold
OperationalAspects	TemperatureProfile	Features that refer to operational features of the network Temperature profile along the network
	Isothermal	Constant temperature profile along the network
	Nonisothermal	Variable temperature profile along the network
Performance		Objective function value of the optimisation solution: function of the outlet concentration and the size of the network
	Low	Lower than 33% of a range
	Medium	Between 33% and 66% of a range
	High	Above 66% of a range.

Appendix 6: The multiphase reactor network domain ontology

Concepts	Description
Network	A Network consists of one or a combination of RMX units that may be interconnected in any physically possible way through connections and that by operating under certain operation conditions gives a performance.
OneUnitNetwork	Any network that has only one RMX unit.
TwoUnitNetwork	Any network that has only two RMX units.
ThreeUnitNetwork	Any network that has only three RMX units.
Experiments	A network that has been generated during the optimisation at a specific optimisation stage. Each of these networks is expressed in the form of a compact digital certificate and represents one of the clusters selected to launch the next optimisation stage. As each stage represent a group of solutions that share the same features, each experiment can be seen as a class that defines a group of individuals (i.e. optimisation solutions).
NetworkPhase	Type of the phases (reactive / non-reactive) involved in the multiphase system
NonReactivePhase	Phase in which no reaction takes place.
ReactivePhase	Phase in which reaction takes place.
PhaseState	State of the phases involved in the multiphase system.
Gas	State of matter without a definite volume or shape.
Liquid	State of matter with a definite volume but not a definite shape.
NetworkFeatures	Characteristics that define the network in terms of structural and operational aspects.
NumberOfUnits	Number of RMX units the network consists of. The maximum number of RMX units allowed is three.
U1	Used to represent the number of one RMX unit of the network.
U2	Used to represent the number of two RMX units of the network.
U3	Used to represent the number of three RMX units of the network.
MixingPattern	Describes mixing of the RMX unit.
BackMixedBackMixed	Combinations of (phase 1/phase 2): well-mixed/well-mixed, well-mixed/mid-mixed, mid-mixed/mid-mixed or mid-mixed/well-mixed.
BackMixedPFR	Combinations of (phase 1/phase 2): well-mixed/PFR or mid-mixed/PFR.
PFRBackMixed	Combinations of (phase 1/phase 2): PFR/well-mixed or PFR/mid-mixed.
PFRCoCurrent	Combination of (phase 1/phase 2): PFR/PFR arranged in co-current.
PFRCounterCurrent	Combination of (phase 1/phase 2): PFR/PFR arranged in counter-current.

OverallFlowArrangement	Direction of the main flow of a phase respect to the other.
CoCurrent	Overall flow arrangement for a pair of phases flowing in the same direction.
CounterCurrent	Overall flow arrangement for a pair of phases flowing in opposite direction.
Intermediate	Overall flow arrangement for a pair of phases flowing in the same direction in some parts of the network and in opposite direction in some other parts of the network.
MassTransfer	Mass flow as the result of a species concentration difference in a mixture.
Active	Presence of mass transfer.
Inactive	Lack of mass transfer.
Size	Volume of the network.
Low	Magnitude lower than the 33% of a range.
Medium	Magnitude between the 33% and 66% of a range.
High	Magnitude higher than the 66% of a range.
OperationalAspects	Features that refer to operational issues of the network.
TemperatureProfile	Temperature profile along the network.
Isothermal	Constant temperature profile along the network.
NonIsothermal	Variable temperature profile along the network.
Performance	Objective function value of the optimisation problem. The performance measure is a function of the outlet compositions and the size of the network; it may be, for instance, the yield of a given product, the selectivity between products, or the overall profitability of the process.
Low	Magnitude lower than the 33% of a range.
Medium	Magnitude between the 33% and 66% of a range.
High	Magnitude higher than the 66% of a range.
PhaseFeatures	Characteristics that define each phase of the network in terms of structural and operational aspects.
Loops	Loops are sequential connections between two non-consecutive phase compartments. Unlike recycles, their removal would result in an unfeasible structure. Loops are necessary streams that ensure the connectivity of some of the units of the network (i.e. loops ensure the connection of its source unit or / and its sink unit to another unit).
LocalInlet	Refers to the front part of a network where a stream or a group of streams are connected to. When applied to the feeding policy, it means that the feeding of a phase is split amongst the inlet compartments. When applied to the product policy, it means that the product of a phase is removed from the inlet compartments. For loops and recycles, it means that they connect the inlet compartments
LocalOutlet	Refers to the end part of a network, where a stream or a group of streams are connected to. When applied to the feeding policy, it means that the feeding of a phase is split amongst the outlet compartments. When applied to the product policy, it means that the product of a phase is removed from the outlet compartments. For loops and recycles, it means that they connect the outlet compartments of a phase. However, for recycles, it can also refer to a recycle around the last compartment of a phase for a network consisting of more than one unit. Due to being three the maximum number of units allowed in this work, the outlet compartments refer to the two last compartments of a three unit network.
Total	Refers to all the units of a network, where a stream or a group of streams are connected to. When applied to the feeding policy, it means that the feeding of a phase is split amongst all the compartments. When applied to the product policy, it means that the

		product of a phase is removed from all compartments. For loops and recycles, it means that they connect the extreme compartments of a phase and therefore all units are affected by the stream. This concept applies to networks made of more than one unit.
FeedingPolicy	Single	Refers to how the fresh feed flow of a phase is fed to the network. Flow connected from a source to a sink through one stream. When applied to the feeding policy, it means that the fresh feed stream in a phase is fed to a single compartment. When applied to the product policy, it means that the product in a phase is removed from a single compartment.
	Distributed	Flow split into/made from more than one stream. When applied to the feeding policy, it means that the fresh feed stream in a phase is split amongst different compartments. When applied to the product policy, it means that the product in a phase is removed from more than one compartment.
	LocalInlet	Refers to the front part of a network where a stream or a group of streams are connected to – as above.
	LocalOutlet	Refers to the end part of a network, where a stream or a group of streams are connected to – as above.
	Total	Refers to all the units of a network, where a stream or a group of streams are connected to – as above.
Bypassed ProductPolicy		Fresh feed stream connected to the product.
	Single	Refers to how the product flow from a phase is removed from the network. Flow connected from a source to a sink through one stream. When applied to the feeding policy, it means that the fresh feed stream in a phase is fed to a single compartment. When applied to the product policy, it means that the product in a phase is removed from a single compartment.
	Distributed	Flow split into/made from more than one stream. When applied to the feeding policy, it means that the fresh feed stream in a phase is split amongst different compartments. When applied to the product policy, it means that the product in a phase is removed from more than one compartment.
	LocalInlet	Refers to the front part of a network where a stream or a group of streams are connected to – as above.
	LocalOutlet	Refers to the end part of a network, where a stream or a group of streams are connected to – as above.
NonReactivePhaseFeatures Recycles	Total	Refers to all the units of a network, where a stream or a group of streams are connected to – as above.
	LocalInlet	Refers to the front part of a network where a stream or a group of streams are connected to – as above.
	LocalMiddle	Refers to the middle part of a network, where a stream or a group of streams are connected to. Due to being three the maximum number of units allowed in this work, this concept can only be applied to recycles around the second compartment of a phase for a three unit network.
	LocalOutlet	Refers to the end part of a network, where a stream or a group of streams are connected to – as above.
	Total	Refers to all the units of a network, where a stream or a group of streams are connected to – as above.
ReactivePhasefeatures Connections		Phase features specific to the non-reactive phase of a network.
		Backward connection.
ReactivePhasefeatures Connections		Phase features specific to the reactive phase of the network.
	IntraConnections	Refers to the existence of streams that connect the compartments in the reactive phase. Connections that take place within a reactive phase compartment.

Recycle	Backward connection.
Bypass	Forward connection.
InterConnections	Connections that connect different compartments of the reactive phase.
Recycle	Backward connection.
Bypass	Forward connection.

References

- Achenie, L. and L. Biegler (1988). Developing Targets for the Performance Index of a Chemical Reactor Network: Isothermal Systems. *Ind. Eng. Chem. Res.* 27: pp 1811-1821.
- Ashley, V. M. and P. Linke (2004). A novel approach for reactor network synthesis using knowledge discovery and optimization techniques. *Chem.Eng.Res.Des.* 82: pp 952-960.
- Bauer, M. and J. Stichlmair (1996). SUPERSTRUCTURES FOR THE MIXED INTEGER OPTIMIZATION OF NONIDEAL AND AZEOTROPIC DISTILLATION PROCESSES. *Computers for Chem. Rng.* 20: pp S25-S30.
- Cavin, L., U. Fisher, F. Glover and K. Hungerbuhler (2004). Multiobjective process design in multi-purpose batch plants using a Tabu search algorithm. *Comput.Chem.Eng.* 28: pp 459-478.
- Cecelja, F., A. Kokosis, D. Du and S. Yang (2014). Asynchronous Optimisation with the Use of a Cascade Search Algorithm. *Computers & Chemical Engineering* 66: pp 276 - 289.
- Cecelja, F., A. Kokosis and D. Du (2011). Integration of ontology and knowledge-based optimization in process synthesis applications. *Computer Aided Chemical Engineering* 29: pp 427-432.
- Cecelja, F., N. Trokanas, T. Raafat and M. Yu (2015). Semantic Algorithm for Industrial Symbiosis Network Synthesis. *Journal of Computers & Chemical Engineering On-line*: in press, accepted manuscript: pp.
- Cordeau, J. F. and G. Laporte (2005). TABU SEARCH HEURISTICS FOR THE VEHICLE ROUTING PROBLEM. *Metaheuristic Optimization via Memory and Evolution - Tabu Search and Scatter Search.* 30: pp 117-144.
- Dorigo, M. and C. Blum (2005). Ant colony optimization theory: A survey. *Theoretical Computer Science* 344: pp Theoretical Computer Science.
- Dorigo, M., G. DiCaro and L. M. Gambardella (1999). Ant algorithm for discrete optimisation. *Artificial Life* 5: pp 137-172.
- Glover, F. (1989). Tabu Search - Part 1. *ORSA Journal on Computing* 1: pp 190-206.
- Glover, F. and M. Laguna (2011). TABU SEARCH. *Handbook of Combinatorial Optimization*: pp.
- Glover, F., E. Taillard and D. Werra (1993). A user's guide to tabu search. *Annals of Operations Research* 41: pp 3-28.
- Guiarratano, J. (2002). "CLIPS - User's Guide." 2012, from <http://clipsrules.sourceforge.net/index.html>.
- Jacobs, R. and W. Jansweijer (2000). A knowledge-based system for reactor selection. *Computers and Chemical Engineering* 24: pp 1781-1801.
- Jacobs, R., W. Jansweijer and P. Iedema (1996). A KNOWLEDGE BASED SYSTEM FOR REACTOR SELECTION. *Computers and Chemical Engineering* 20: pp S165-S170.
- James, T., C. Rego and F. Glover (2009). A cooperative parallel tabu search algorithm for the quadratic assignment problem. *European Journal of Operational Research* 195: pp 810-826.
- Jayaraman, V. K., B. D. Kulkarni, S. Karale and P. Shelokar (2000). Ant colony framework for optimal design and scheduling of batch plants. *Comput.Chem.Eng.* 24: pp 1901.
- Kim, J., M. Kim, M. Stehr, H. Oh and S. KHa (2012). A parallel and distributed meta-heuristic framework based on partially ordered knowledge sharing. *Journal of Parallel and Distributed Computing* 72: pp 564-578.

- Kokossis, A., F. Cecelja, C. Labrador-Darder and P. Linke (2008). An Ontology Approach to Represent and Extract Process Synthesis Knowledge in High-Throughput Optimization. AIChE Annual Meeting in Philadelphia: pp.
- Kokossis, A. and C. A. Floudas (1990). Optimization of complex reactor networks-I. Isothermal operation. Chemical Engineering Science 45(3): pp 595-614.
- Kokossis, A. and C. A. Floudas (1994). Optimization of complex reactor networks-II. Nonisothermal operation. Chemical Engineering Science 49: pp 1037-1051.
- Kokossis, A., P. Linke and S. Yang (2011). The Cascade Optimization Algorithm: A New Distributed Approach for the Stochastic Optimization of Engineering Applications. Industrial & Engineering Chemistry Research 50: pp 5266-5278.
- Labrador-Darder, C., F. Cecelja, A. Kokossis and P. Linke (2009). Integration of superstructure-based optimization and semantic models for the synthesis of reactor networks. Computer Aided Chemical Engineering 26: pp 865-870.
- Leite, J. P. B. and B. H. V. Topping (1999). Parallel simulated annealing for Structural Optimisation. Comput.Struct. 73: pp 545-564.
- Linke, P. and A. Kokossis (2003a). Attainable Reaction and Separation Processes from a Superstructure-Based Method. AIChE Journal 49(6): pp 1451-1470.
- Linke, P. and A. Kokossis (2003b). On the robust application of stochastic optimisation technology for the synthesis of reaction/separation systems. Computers and Chemical Engineering 27: pp 733-758.
- Marcoulaki, E. and A. Kokossis (1996). STOCHASTIC OPTIMISATION OF COMPLEX REACTION SYSTEMS. Computers chem.Engng 20: pp S231-S236.
- Marcoulaki, E. C. and A. C. Kokossis (1999). Scoping and Screening Complex Reaction Networks Using Stochastic Optimization. AIChE Journal 45(9): pp 1977-1991.
- Mehta, V. and A. Kokossis (1997). Development of Novel Multiphase Reactors using a Systematic Design Framework. Computers chem.Engng 21: pp S325-S330.
- Mehta, V. and A. Kokossis (1998). New Generation Tools for Multiphase Reaction Systems: A Validated Systematic Methodology for Novelty and Design Automation. Computers them.Engng 22: pp S119-S126.
- Mehta, V. and A. Kokossis (2000). Nonisothermal synthesis of homogenous and multiphase reactor networks. AIChE Journal 46(11): pp 2256-2273.
- Raafat, T., F. Cecelja, N. Trokanas and B. Xrisha (2013). An Ontological Approach Towards Enabling Processing Technologies Participation in Industrial Symbiosis. Computers & Chemical Engineering 59(1): pp 33-46.
- Raman, R. and I. Grossmann (1991). RELATION BETWEEN MILP MODELLING AND LOGICAL INFERENCE FOR CHEMICAL PROCESS SYNTHESIS. Computers them. Engng 15(2): pp 73-84.
- Shah, P. and A. Kokossis (2001). Knowledge based models for the analysis of complex separation processes. Computers and Chemical Engineering 25: pp 867-878.
- Shah, P. and A. Kokossis (2002). New Synthesis Framework for the Optimization of Complex Distillation Systems. AIChE Journal 48(3): pp 527-550.
- Talbi, E. G., Z. Hafidi and J. M. Geib (1998). A parallel adaptive tabu search approach. Parallel Comput. 24: pp 2003-2019.
- Vanderbit, D. and S. Louise (1984). A Monte Carlo Simulated Annealing Approach to Optimisation Over Continuous Variables. Journal of Computational Physics 56: pp 259-271.
- Wang, C., H. Quan and X. Xu (1999). Optimal design of multiproduct batch chemical processes using tabu search. Computers & Chemical Engineering 23(3): pp 427-437.
- Wang, Z., Y. Wong and M. Rahman (2005). Development of a parallel optimization method based on genetic simulated annealing algorithm. Parallel Computing 31: pp 839-857.
- Yeomans, H. and I. Grossman (1999). A systematic modeling framework of superstructure optimization in process synthesis Computers & Chemical Engineering 23(6): pp 709-731.

