CRANFIELD UNIVERSITY

Konstantinos Vergidis

Business Process Optimisation using an Evolutionary Multi-objective Framework

SCHOOL OF APPLIED SCIENCES

PhD Thesis

CRANFIELD UNIVERSITY

SCHOOL OF APPLIED SCIENCES

DEPARTMENT OF MANUFACTURING

PhD Thesis

Academic year 2008-2009

Konstantinos Vergidis

Business Process Optimisation using an Evolutionary Multi-objective Framework

Supervisor: Dr Ashutosh Tiwari

November 2008

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

In response to the increasingly volatile and competitive environment, organisations are examining how their core business processes may be redesigned in order to improve business performance and market responsiveness. However, there is a lack of holistic approaches towards business process redesign through optimisation. The aim of this research is to develop an evolutionary multi-objective optimisation framework for business processes capable of: (i) representing business process designs in a quantitative way, (ii) algorithmically composing designs based on specific process requirements and (iii) identifying the optimal processes utilising evolutionary algorithms.

A literature survey of business process definitions, modelling, analysis and optimisation techniques provides an overview of the current state of research and highlights the gap in business process optimisation. An industry survey within the service sector grounds the research within the industrial context and compares the real-life issues related to business processes with the literature findings. This research proposes a representation technique for business process designs using both a visual and a quantitative perspective. It also proposes the Process Composition Algorithm (PCA) – an algorithm for composing new business process designs. The proposed business process optimisation framework (bpo^F) lies at the heart of this research and employs the representation technique, PCA and a series of state-of-the-art evolutionary optimisation algorithms. The framework is capable of generating a series of alternative optimised business process designs based on given requirements.

A strategy for creating experimental business process scenarios is also proposed by this research. The proposed strategy provides the opportunity of assessing both the capability of the framework in optimising challenging business process scenarios and the performance of the evolutionary algorithms. Finally, a set of real-life business process scenarios is prepared using the proposed representation in order to validate the optimisation framework. Also, a workshop with a series of business process experts assesses the capability of the framework in dealing with these real-life scenarios. In this way, this research proposes a fully tested and validated methodology for capturing, representing and optimising business process designs.

Acknowledgments

Completing this research and writing this PhD thesis was a long and strenuous journey. And as with every journey, and with every path, no one can face the challenges and survive on his own; one needs the strength, virtues and experiences of those that the Universe sends as guides or companions. I am here… I finished this journey safe and sound… Looking back, I now know that it would not be possible without the Divine Grace and all of you who helped me –each in your own unique way. Saying a little 'thank you' is the least I can do, so here it goes…

Dr Ashutosh Tiwari
Thank you for giving me the opportunity to do this PhD. Thank you for being a lot more than a supervisor along the way. Thank you Basim (BT) for all your constructive and creative feedback throughout the course of this research. A lot of gratitude goes to the researchers and staff of the Manufacturing department (building 50) for your support. Special reference is necessary for Teresa, Ruth and Jenny. Thanks for being there!

The friends that came along the way…
Kieron, would I have survived this without our long conversations? Thank you for all the advice and your faith in me. Pedro, your advice is with me along with your positive energy! Dipa, thanks for being so kind and sweet to me. Yoseph, you saw things no one else did… I learnt a lot from you. John-Ahmet, you saw my crazy side and didn't freak out… thanks for being so considerate. Enrique and Chris, the Universe sent you so late that it's almost unfair… all the best for the future! Thank you James, George, Costas and Kostas for all the good times together.

The friends back home…
f.Niko, Xari, Alki, Dimitri, John, Kelly, Konstantina, Sevi… all these years you actively showed your care and support defying the distance. I am looking forward to spend quality time with you once again…

My family
It has been an incredible journey these four years in the UK. It would not be possible without your love and inspiration. Mum, thanks for believing in me. Your love and support kept me safe. Dad, thanks for all your help with my work. You are the best Mathematician! Aphrodite, thanks for being the world's greatest sister… we have serious fun together. Sirish, your perspective on life helped me grow in many ways… I'm going to miss your presence and your wisdom my brother.

Glory to the Father, and to the Son, and to the Holy Spirit
Blessed you are Holy Mother Virgin Mary,
and you Saint John the Russian my loyal friend (thanks for helping with the boring parts!)

As it was in the beginning, is now and ever shall be, world without end

List of Publications

Journal papers

1. Jallow, A.K., Majeed, B., **Vergidis, K.**, Tiwari, A. and Roy, R. (2007), 'Operational Risk Analysis in Business Processes', *BT Technology Journal,* vol. 25, no. 1, pp. 168-177.

2. Tiwari, A., **Vergidis, K.** and Kuo, Y. (2008), 'Computer Assisted Decision Making for New Product Introduction Investments', *Computers in Industry,* vol. 59, pp. 96-105.

3. Tiwari, A., **Vergidis, K.** and Roy, R. (2007), 'Evolutionary Optimisation of Business Process Designs', *Studies in Computational Intelligence,* vol. 49, pp. 513-541.

4. Tiwari, A., **Vergidis, K.**, Lloyd, R. and Cushen, J. (2008), 'Automated inspection using database technology within the aerospace industry, *Proceedings of the Institution of Mechanical Engineers (IMechE), Part B: Journal of Engineering Manufacture,* vol. 222, pp. 175-183.

5. **Vergidis K.**, Turner, C.J. and Tiwari, A. (2008), 'Business Process Perspectives: Theoretical Developments vs. Real-World Practice'*, International Journal of Production Economics*, vol. 114, pp. 91-104.

6. **Vergidis, K.**, Tiwari, A. and Majeed, B. (2006), 'Business Process Improvement Using Multi-Objective Optimisation', *BT Technology Journal*, vol. 24, no. 2, pp. 229-235.

7. **Vergidis, K.**, Tiwari, A. and Majeed, B. (2008), 'Business Process Analysis and Optimisation: Beyond Reengineering', *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*, vol. 38, no. 1, pp.1-15.

8. **Vergidis, K.**, Tiwari, A., Majeed, B. and Roy, R. (2007), 'Optimisation of Business Process Designs: An Algorithmic Approach with Multiple Objectives'*, International Journal of Production Economics*, vol. 109, pp. 105-121.

Conference papers

1. Tiwari, A., **Vergidis, K.** and Majeed, B. (2006), 'Evolutionary Multi-Objective Optimisation of Business Processes', *in Proceedings of IEEE Congress on Evolutionary Computation (CEC'06),* Vancouver, Canada, 16-21 July 2006, pp. 3091-3097.

2. Turner, C.J., **Vergidis, K.** and Tiwari, A. (2007), 'Surveying Business Processes: The Industry Perspective', *in Proceedings of the 5th International Conference on Manufacturing Research (ICMR'07),* Leicester, U.K., 11-13 September 2007, pp. 313-317.

3. **Vergidis, K.** and Tiwari, A. (2008), 'Business Process Design and Attribute Optimization within an Evolutionary Framework', *in Proceedings of IEEE Congress on Evolutionary Computation (CEC'08)*, Hong-Kong, 1-6 June 2008, pp. 668-675.

4. **Vergidis, K.**, Tiwari, A. and Majeed, B. (2007), 'Composite Business Processes: An Evolutionary Multi-Objective Optimisation Approach', *in Proceedings of IEEE Congress on Evolutionary Computation (CEC'07),* Singapore, 25-28 September 2007, pp. 2672-2678.

5. Vergidis, T., **Vergidis, K**. and Tiwari, A. (2008), 'The Evaluation Line: A Posteriori Preference Articulation Approach', *in Proceedings of IEEE Congress on Evolutionary Computing (CEC'08)*, Hong-Kong, 1-6 June 2008, pp. 2699-2705.

Submitted papers

1. **Vergidis, K.** and Tiwari, A. (2008), 'Business Processes Rediscovered' submitted to *International Journal of Business Process Management*.

2. **Vergidis, K.**, and Tiwari, A. (2008), 'An Evolutionary Multi-objective Framework for Business Processes', submitted to *IEEE Transactions on Evolutionary Computation*

3. **Vergidis, K.** and Tiwari, A. (2009), 'bpoF – An Evolutionary Business Process Optimisation Framework', *submitted to IEEE Congress on Evolutionary Computing (CEC'09)*, Norway.

Table of Contents

List of Figures

List of Tables

Frequently used Abbreviations

BPA	Business Process Automation
BPEL	Business Process Execution Language
BPML	Business Process Modelling Language
bpo^F	business process optimisation Framework
DoI	Degree of Infeasibility
EA	Evolutionary Algorithm
EC	Evolutionary Computing
EMOA	Evolutionary Multi-objective Optimisation Algorithm
FES	Framework Evaluation Strategy
IDEF	Integrated DEFinition (modelling)
KPIs	Key Performance Indicators
LSSA	Large Scale Search Algorithm
MILP	Mixed Integer Linear Programming
NSGA2	Non-dominated Sorting Genetic Algorithm 2
PA	Process Attribute
PAES	Pareto Archived Evolutionary Strategy
PCA	Process Composition Algorithm
PESA2	Pareto Envelope-based Selection Algorithm 2
RAD	Role Activity Diagram
SDP	Service Delivery Price
SFT	Service Fulfilment Target
SOA	Service Oriented Architecture
SPEA2	Strength Pareto Evolutionary Algorithm 2
TA	Task Attribute
TAM	Task Attributes Matrix
TRM	Task Resources Matrix
UML	Universal Modelling Language
UTRM	Updated Task Resources Matrix

CHAPTER 1

Introduction

In response to increasingly volatile and competitive environments, organisations are examining how their core business processes may be re-designed to improve business performance and market responsiveness. The design and management of business processes is a key factor for companies to effectively compete in today's volatile business environment. By focusing on the optimisation and continuous improvement of business processes, organisations can establish a solid competitive advantage by reducing cost, improving quality and efficiency, and enabling adaptation to changing requirements.

This research focuses on business process re-design through optimisation and it is carried out as part of the 'Intelli-Process' ('Intelligent Decision Support for Process Re-design and Conformance') project. This project is funded by the Engineering and Physical Sciences Research Council (EPSRC – Grant No.: EP/C54899X/1 – Duration: 2005-08). This chapter provides an introduction to the concepts of business processes and Evolutionary Computing, describes the 'Intelli-Process' project along with the problem statement and motivation for this research. The chapter concludes with the layout of the thesis.

1.1 Introduction to business processes

This section introduces *business processes*, a concept which is the central focus of this research. The following sub-section discusses briefly the various definitions for business processes that exist in literature and presents the definition that is going to be used in this research. The various elements of business processes are then detailed, followed by the current issues related to business processes.

1.1.1 Defining business processes

Business processes are defined by Jacobson *et al.* (1994) as 'the set of internal value-adding activities performed to serve a customer'. Havey (2005) provides a simple definition of business processes as 'step-by-step rules specific to the resolution of a business problem'. Since the 1990's when the first definitions of business processes appeared in literature, many authors attempted to come up with their own improved version of business process definition usually with one purpose: to try and orient business processes towards a particular direction highlighting only specific aspects. Chapter 2 provides a summary of business process definitions (table 2.1) reflecting on the variety and diversity of the different business process definitions that exist in literature.

Although most definitions tend to be similar in the concepts used to express and describe business processes, they have received criticisms for not adequately highlighting the 'business' component and not sufficiently distinguishing from manufacturing or production processes. Volkner and Werners (2000) support that no generally accepted definition of the term business process exists due to the fact that business processes have been approached by a number of different disciplines. Lindsay et al. (2003) report that most business process definitions are limited in depth and the corresponding models are also constrained and confined to a mechanistic viewpoint. The main issue with business process definitions is twofold: either they are too simplistic and basic thus too generic to provide any tangible contribution or they are confined to a very specific application area that prevents them from wide acceptance and applicability. In this research, the author defines business processes as follows:

A **business process** is perceived as a collective set of tasks that when properly connected and sequenced perform a business operation. The aim of a business process is to perform a business operation, i.e. any service-related operation that produces value to the organisation.

1.1.2 Main elements of business processes

Research regarding business processes shows that although there is a wide variety in terms of definitions, when it comes to the structural elements of a business process there is a common ground to build upon. Figure 1.1 presents a proposed business process schema that involves the most common structural elements found in literature. These elements are put together in a hierarchical structure that also reflects the relationships between them. The solid arrows show the main elements of the schema whilst the dashed arrows denote the optional elements. A detailed description of each of the elements of figure 1.1, along with the relevant references, is provided in chapter 2.

Starting from the top, it is necessary to recognise that although business processes are by definition placed within a 'business' context, they are a subclass of generic processes and as such they inherit all of their main properties such as structure, flow, activities, etc. Moving to the second level of the schema, business processes are placed in parallel with workflows.

Workflows (as they are defined in section 2.2.3) are closely linked with business processes and sometimes these terms are interchangeably used thus the bi-directional arrow that links the two concepts. The third level of the business process schema is based on what many authors (see table 2.1) consider as the basic structural elements of a business process: *actors*, *activities* and *resources*.

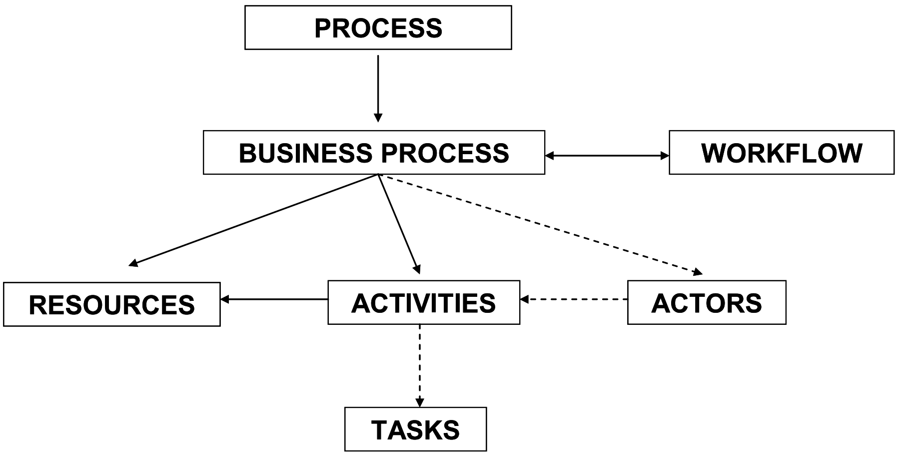


Figure 1.1. Schematic relationship of the main business process elements

These are the three main concepts involved in most business process definitions, although more emphasis is put on activities and resources only. *Actors* are sometimes involved in a business process definition (Lindsay *et al.*, 2003) or sometimes perceived as external entities that enact or execute the process. *Activities* are widely accepted as the central elements that execute the basic business process steps utilising the process inputs in order to produce the desired results. *Resources* are frequently classified as inputs or input resources and they are necessary for activities to be executed. Lastly, *tasks* are perceived as the smallest analysable element of a business process (Orman, 1995). However, they are usually overlooked by most authors or tend to be another synonym for activities.

1.1.3 Towards business process automation

According to Powell et al. (2001), business processes generate most of the costs of any business, so improving organisational efficiency generally requires improving the business processes. Business processes also strongly influence –if not define– the quality of the product and the satisfaction of the customer, both of which are of fundamental importance in the marketplace (Grigori *et al.*, 2001). This has several implications: the business processes should be correctly designed, their execution should be supported by a system that can meet the workload requirements, and the (human or automated) process resources should be able to perform their work items in a timely fashion (Grigori *et al.*, 2004). Zhou and Chen (2003) remark that for systematic and holistic automated business process planning, there must be techniques that support *modelling*, *analysis* and *optimisation* of business processes.

One of the latest trends in literature is towards business process automation. The benefits of such automation are that processes can be executed faster, with lower costs (due to the reduced human involvement) and in a controlled way, according to Castellanos et al. (2004). However, as business processes become automated, the focus of both industry and academia shifts from deployment to effective process modelling, analysis, and optimisation. The focus of this research lies in the area of business process optimisation with reference to modelling and analysis. Business process optimisation will be facilitated by Evolutionary Computing (EC) techniques as the next section discusses.

1.2 Evolutionary multi-objective optimisation

The title of the thesis reveals that the proposed optimisation approach regarding business processes is built on the basis of an evolutionary multi-objective optimisation approach. According to Deb (2001), optimisation refers to finding one or more feasible solutions which correspond to extreme values of one or more objectives.

1.2.1 Why multi-objective optimisation?

Multi-objective optimisation problems and algorithms have received wide attention during the last two decades due to the fact that most real-world problems naturally involve multiple objectives. A multi-objective optimisation problem can involve either minimisation or maximisation and can also be subjected to a number of constraints that limit the problem boundaries. The proposed framework adopts a multi-objective optimisation approach towards business processes for three main reasons:

1. Business process optimisation is inherently a multi-objective optimisation problem due to the variety of factors that a business process can be evaluated with. Dealing with multiple objectives can make this research more appealing and applicable to real-life business process optimisation problems.
2. Evaluation business processes based on a series of relevant factors ensures that this research is versatile in dealing with different objectives for different business goals at a time. The capability of simultaneously addressing a series of customised quantifiable

objectives ensures the generality of the research and its potential applicability in a wider context of business process improvement initiatives.

3. On the contrary, a single-objective optimisation framework focuses on a particular objective (e.g. cost reduction) and thus loosing its generality and its advantages over context-specific business process improvement approaches that target a specific aspect of a business process (e.g. Six Sigma).

1.2.2 Introduction to evolutionary optimisation

In the natural world, evolution has created an unimaginably diverse range of designs, having much greater complexity than mankind could ever hope to achieve. Inspired by this, researchers have started using Evolutionary Computing (EC) techniques that use the principles of evolution to guide the optimisation process. EC is a subfield of Computational Intelligence that involves combinatorial optimisation problems. EC techniques use iterative progress, such as growth or development of a population. This population is then selected in a guided random search using parallel processing to achieve the desired end. Such processes are often inspired by biological mechanisms of evolution.

Usually, an initial population of randomly generated solutions comprises the first generation. The fitness function is applied to the solutions and any subsequent offspring. In selection, parents for the next generation are chosen with a bias towards higher fitness. The parents reproduce by the application of operators such as crossover and/or mutation. *Crossover* acts on the two selected parents and results in one or two children. *Mutation* acts on one solution and results in a new one. These operators create the offspring population of solutions. This process can be repeated until a population of solutions with sufficient quality is found, or a previously defined number of generations is reached.

There are a number of benefits of evolutionary-based optimisation that justify the effort invested in this area. The most significant advantage lies in the gain of flexibility and adaptability to the task in hand, in combination with robust performance and global search characteristics (Coello Coello, 2000). The evolutionary-based optimisation techniques use, in each iteration, a population of solutions instead of a single solution. This enables them, in principle, to identify multiple optimal solutions in their final population. EC techniques often perform well approximating solutions to all types of problems because they ideally do not make any assumption about the underlying fitness landscape; this generality is shown by successes in fields as diverse as engineering, art, biology, economics and operations research.

1.2.3 Why evolutionary optimisation?

The multi-objective optimisation technique employed by the proposed research is a range of Evolutionary Algorithms (EAs). EAs mimic nature's evolutionary principles to guide the optimisation process towards discovering optimal solutions. EAs have become increasingly popular in multi-objective optimisation problems (Coello Coello, 2005). The two main advantages of employing EAs are:

- The outcome of EAs in a single iteration is a *population* of solutions.
- In multi-objective optimisation, EAs treat all the objectives as *equal*.

The lack of preference towards a particular objective in conjunction with the generated population of solutions provides the capability of having a range of optimal solutions that each reflects a different trade-off between the optimisation objectives. This capability enhances the selection process of a single optimal solution by providing a choice of equivalent alternatives.

Both these advantages are essential regarding business process multi-objective optimisation. Moreover, the concept of evolving and improving the population of solutions in the process of evolutionary optimisation is central to business processes for two reasons: (i) business process designs that would otherwise be overlooked by a human designer can be discovered by EAs and (ii) evolving a solution over the generations can transform an infeasible process design to a feasible one.

These characteristics of EAs make them a suitable candidate for optimising business process designs with a series of features. These features include the presence of multiple objectives, composition of feasible business process designs and generation of diverse alternative designs. As a consequence, EAs are better suited to deal with business process optimisation compared to their classical counterparts. This research, therefore, focuses on evolutionary algorithms as optimisers for business processes. The four specific EAs that are selected are detailed in Appendix D and discussed in chapters 6 and 8.

1.3 'Intelli-Process': Parent project of the research

This research is part of a bigger project under the name 'Intelligent Decision Support for Process Redesign and Conformance (Intelli-Process)'. The project aim is to develop an EC-based framework for capturing the business processes in an automated manner, optimising the process design interactively, and identifying the extent of disparity with the optimal in the continuous improvement process. The Intelli-Process project involves two researchers, a principal investigator (Dr. Ashutosh Tiwari) and is funded by EPSRC. The project was launched on September 2005 and its duration extends over a 3-year period. The project investigates areas such business process mining, process design optimisation and process conformance, aiming to cover most of the aspects that build a consistent, efficient and beneficiary business process improvement framework. One researcher is preoccupied with business process mining and conformance while the author of this thesis is focused entirely on business process optimisation.

1.3.1 Project background

Owing to the qualitative nature of business process models, there is a lack of tools for identifying the bottleneck areas in these models. This qualitative nature also explains the difficulty of developing 'parametric' models of business processes. Therefore, although a considerable number of algorithms exist for dealing with process optimisation problems in Manufacturing, there is a lack of algorithmic approaches for the *optimisation of business processes* (Tiwari, 2001). Much of the recent research in the area of business process optimisation has dealt with either selection of a process model from a set of alternatives (Shimizu and Sahara, 2000) or simple single-objective formulation that does not address the strong synergistic/anti-synergistic effects among individual activities that constitute a process design (Hofacker and Vetschera, 2001). Therefore, the current research suffers from serious limitations in dealing with the scalability requirements and complexity of real-life processes. The 'Intelli-Process' project aims to enhance the current work by developing an interactive knowledge-based process improvement tool.

1.3.2 Brief description of 'Intelli-Process'

The aim of this project is to develop an EC-based framework for capturing the business processes in an automated manner, optimising the process design interactively, and identifying the extent of disparity with the optimal in the continuous improvement process. The application scope of this generic framework is in the computer assisted business processes, especially in the Service and IT industries. Examples include order processing and fault handling. To achieve the above aim, the project realised the following objectives:

- To capture the industry best practice and requirements for an intelligent decision support tool for process capture, re-design and conformance.
- To apply EC techniques with automated event monitoring for capturing and mapping the processes.
- To interactively compose a series of the optimised business process designs through the evolution of the current AS-IS process model.
- To mathematically define 'conformity index', as a score of compliance between an individual's/team's process and the optimal model, and to identify the areas of disparity between the two.
- To develop a prototype tool to demonstrate the working of the proposed framework for the Service/IT industry.

1.3.3 'Intelli-Process' challenges

Process improvement is partly based on detailed studies about the processes and partly on subjective decisions involving human judgement. Furthermore, many of these studies are underutilised since it is difficult to check if the people are actually conforming to the suggested optimised process. While research in the process area has mostly focused on manufacturing industry, the application scope of this project is in the *computer-assisted business processes* within the *service industry*. Here, the need is to facilitate automated mapping of the current business processes within an organisation, capture the knowledge elements to partly automate the development of optimised process model, and then monitor people for conformance to that model. The challenges in the above-mentioned areas are briefly discussed below:

- *Capturing the business processes:* Automated capturing of the business processes requires 'sophisticated intelligence' even in a fully computerised work environment. Also, since different individuals/teams inherently perform the same task in different ways, it is difficult to obtain a process model that fits all.

- *Optimising the process design:* Due to their qualitative nature, process designs are hard to characterise in a formal way amenable to analytical methods. Also, there are strong (anti-)synergistic effects among individual activities in a process design.

- *Identifying the extent of disparity:* It is difficult to quantify the conformity of a business process with the suggested optimal. In addition, identification of the areas of disparity between the current and suggested processes is highly subjective in nature.

But these challenges –with emphasis on optimisation– can also be seen as strong arguments for applying EC techniques that use the principles of evolution and provide significant advantages in terms of adaptability and flexibility in combination with robust performance and global search characteristics. The outcome of the proposed project will be a novel and generic EC framework for:

1. automated capturing of business processes,
2. optimising the process design interactively, and
3. identifying the extent of conformance with the optimal.

The focus of this research lies in the optimisation aspect of the 'Intelli-Process' project as discussed in the following sub-section. The industrial collaborator of the 'Intelli-Process' project is British Telecom (BT). The motivation for the 'Intelli-Process' project originated from close collaboration between Cranfield University and the service industry.

1.3.4 'Intelli-Process' and this research

The 'Intelli-Process' project involves two researchers – the author and fellow researcher Christopher Turner. The research detailed in this thesis is carried out solely by the author and is part of his contribution to the optimisation aspect of the 'Intelli-Process' project. The motivation for this research is briefly described in the following section.

1.4 Problem statement & motivation

Based on the 'Intelli-Process' project the problem statement of this research is formulated as follows:

Development of a framework for business process optimisation based on EC techniques that is able to address the main features of business processes and push the existing boundaries in the automated improvement of business processes.

The development of a business process optimisation framework poses a number of challenges. It requires a modelling or *representation technique* that can capture the main features of a business process and express them in a way amenable to optimisation methods. It assumes an algorithmic approach to compose new business process designs and a quantitative measure of evaluating them. Finally, it introduces EC optimisation techniques –that have been successfully applied in other optimisation problems– and applies them in the context of business processes.

The motivation for this research stems from the problem statement and the challenges stated above.

1.5 Thesis Layout

The layout of this thesis follows the main steps of this research as demonstrated in figure 1.2. The main steps of this research also aid in the identification of the individual chapters. A brief description of each chapter is provided below.

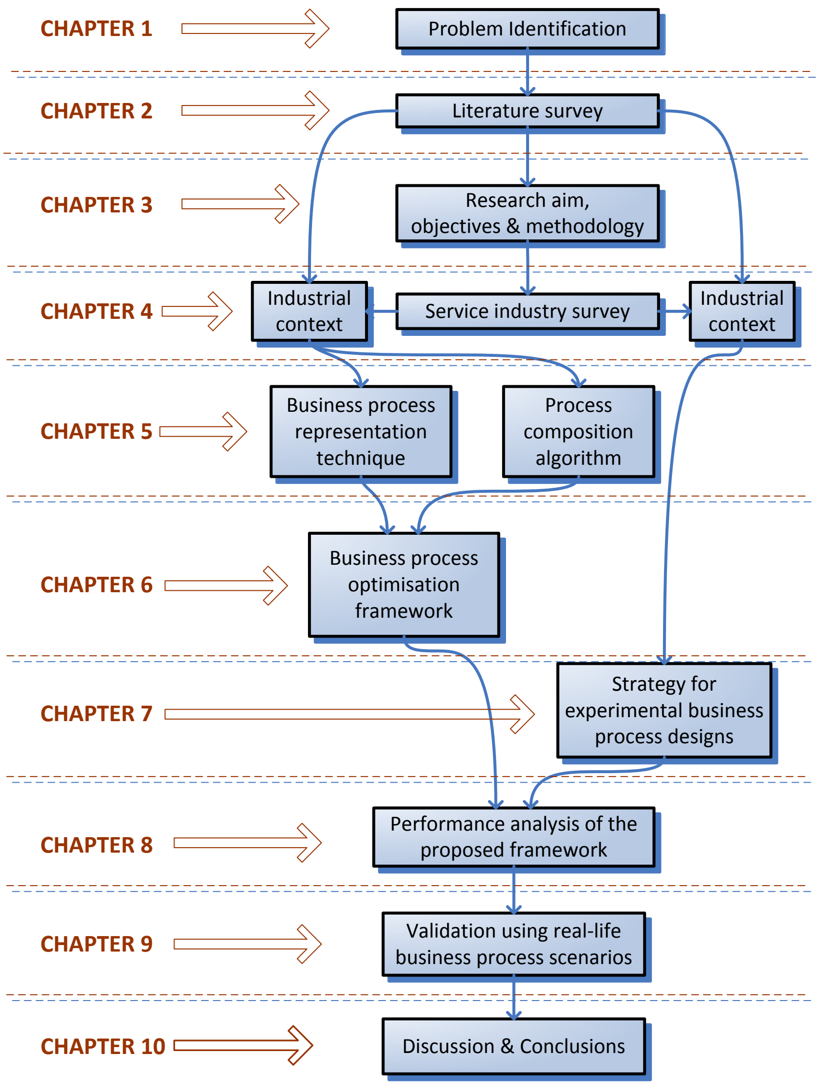


Figure 1.2. Thesis layout and main steps of research

Chapter 1 discusses the background of this research, briefly explaining the aim and objectives of the parent project, 'Intelli-Process'. The problem statement and motivation for this research are described along with the suggested contribution to the parent project.

Chapter 2 presents a survey of literature in the areas of business process definition, modelling, analysis and optimisation. This provides an overview of the current state of the main issues of business processes as reported in literature and helps in defining the research gap.

Chapter 3 gives a brief description of this research, outlining its aim, objectives and scope. It also details the methodology that is adopted for ensuring that the aim and objectives of this research are attained.

Chapter 4 grounds the research within the industrial context based on the results of an industry survey within the service sector. This survey offers the opportunity of comparing the literature survey findings and determining the state of business processes in the service industry.

Chapter 5 proposes a specification and a representation technique for business process designs. This helps in capturing and expressing a business process for optimisation. The chapter also proposes a Process Composition Algorithm in order to compose new business process designs.

Chapter 6 presents the proposed framework for business process optimisation. The framework uses the business process representation technique, composition algorithm and evolutionary algorithms in order to generate optimised business process designs.

Chapter 7 introduces a strategy in order to evaluate the proposed optimisation framework. The strategy aims at the generation of experimental business processes so that the performance of the framework in optimising business process designs is assessed in a systematic way.

Chapter 8 generates three experimental scenarios based on the proposed strategy and tests the framework in order to produce optimised results. It presents a discussion of the obtained results with regard to the framework's performance.

Chapter 9 validates this research using three real-life business process scenarios: the 'on-line order placement', 'sales forecasting' and 'fraud investigation' business processes. The proposed optimisation framework is applied on these scenarios and the results are analysed, compared and discussed.

Chapter 10 concludes this thesis with a discussion on the generality of this research, the contribution to knowledge and the limitations of the research methodology, representation technique and proposed optimisation framework. It also discusses the future research directions that stem from this research.

CHAPTER 2

Literature Review

This chapter discusses the main concepts around business processes as they emerge from the problem statement and motivation discussed in chapter 1. It discusses the main findings from a literature survey that is focused on the aspects of *definition*, *modelling*, *analysis* and *optimisation* of business processes. The aim of the chapter is to provide an overview of the existing techniques and approaches and to highlight their strengths and weaknesses. To achieve that, a classification of the types of business process models is proposed. The proposed classification assists in examining the literature and the current practice in terms of business process modelling, analysis and optimisation. The overview of these approaches facilitates in the identification of the research gap thus relevant remarks are drawn that shape the aim and objectives of this research.

2.1 Basic concepts

As discussed in chapter 1, business process automation is one of the latest trends towards business process re-design and improvement. Zhou and Chen (2003a) remark that for systematic and holistic business process planning based on business process automation, there must be techniques that support *modelling*, *analysis* and *optimisation* of business processes. This chapter presents a literature review regarding business processes based on the abovementioned subjects. Figure 2.1 demonstrates a holistic approach towards business processes that starts with an exploration of definitions of business processes and moves towards the investigation of business process modelling approaches, analysis techniques and optimisation (improvement) practices. The chapter structure also follows this approach by dedicating a separate section to each of these aspects as figure 2.1 shows.

Figure 2.1. Chapter structure based on a holistic business process approach

The next section (2.2) discusses how a business process is specified. There are a number of different approaches and definitions in literature originating from different areas that create an ambiguity on the basic concepts. The next section attempts to rationalise and clarify how business processes are perceived, by presenting the most representative definitions based on the proposed business process schema presented in chapter 1. Section 2.3 deals with the various modelling approaches and their level of maturity towards the unique needs and requirements of business processes. Castellanos *et al.* (2004) remark that most of the research focus has been centred around the modelling of business processes without further supporting them with a theoretical basis, analysis techniques and tools (van der Aalst, 1995). What section 2.4 justifies is that business process models should provide the means for quantitative analysis of business processes in order to extract useful performance measures and acquire realistic knowledge about them. After the process analysis, the next step is improvement, but as Grigori *et al.* (2004) report, research in the business process area has been mostly focusing on developing new process models and process automation techniques, whereas little work has been done in the areas of process analysis, prediction, and optimisation. Hofacker and Vetschera (2001) confirm that several approaches exist for formal description and analysis of business processes but only a few for optimisation. Section 2.5 presents how business process optimisation is perceived in literature and classifies the reported optimisation approaches.

2.2 Business process specification

This section discusses how business processes are defined in literature. There are a number of different approaches and definitions originating from different areas. This section attempts to clarify how business processes are perceived, by presenting the most representative definitions for each of the elements of the proposed business process schema (see chapter 1). The aim of this section is to provide an insight towards the main concepts around business processes.

2.2.1 First 'process' then 'business'

Following the latest trends, many authors focus more on the soft aspects of business processes and tend to view them under a sociotechnical perspective as opposed to the 'mechanistic viewpoint' (Lindsay *et al.*, 2003) that the established approaches have been accused of. However, as the term 'business process' has received such a wide acceptance it cannot be ignored that it interprets a business operation as 'process' thus passing the key attributes of a process to a particular business function.

According to Bal (1998), a process is a sequence of activities which are performed across time and place. A process also has a well defined beginning and end with identifiable inputs and outputs. Similar definitions are provided by Davenport and Short (1990) who emphasize on the defined outcome, and Aldowaisan and Gaafar (1999) who highlight the structured nature of process. Havey (2005) identifies that a process involves movement, work and time; it performs actions over some interval of time in order to achieve, or to progress to, some objective. Li *et al.* (2003) provide more details on what is involved in a process definition apart from the participating activities, and mentions elements such as the criteria to indicate the start and termination of the process, and information about the individual activities, the main participants, and associated IT applications and data. These definitions of process underline the important attributes that are inherited by any type of process.

2.2.2 Definitions of business processes

Since the 1990's when the first definitions of business processes appeared in literature, many authors attempted to come up with their own improved version of business process definition usually with one purpose: to try and orient business processes towards a particular direction highlighting only specific aspects. However, in almost every reference in this area, the authors cite particular business process definitions by Hammer and Champy (1993) and Davenport (1993). Also there are references such as Lindsay *et al.* (2003), Melao and Pidd (2000) and Tinnila (1995) that provide compilations of the various business process definitions. Table 2.1 provides a summary of these definitions. It reflects on the variety and diversity of the different business process definitions that exist in literature.

As is evident from the definitions in table 2.1, most of the authors use the concepts of activities, sequence, inputs and outputs to describe a business process. This proves that most definitions are similar. Significant differences lie on the emphasis on particular aspects of business process. Agerfalk *et al.* (1999), for example, focus on the necessity of the activities to be organised and structured in a specific way within a business process. Castellanos (2004) and Fan (2001) underline on the goal orientation of a business process. Davenport (1993), Gunasekaran and Kobu (2002) and Hammer and Champy (1993) offer more customer oriented definitions while Irani *et al.* (2002) move the focus on the necessity of clear inputs and outputs.

Author(s)	Business process definitions
Agerfalk (1999)	A *business process* consists of activities ordered in a structured way with the purpose of providing valuable results to the customer.
Castellanos *et al.* (2004)	The term *business process* is used to denote a set of activities that collectively achieve certain business goal. Examples of these processes are the hiring of a new employee or the processing of an order.
Davenport and Short (1990)	*Business process* is a set of logically related tasks performed to achieve a defined business outcome.
Davenport (1993)	*Business process* is defined as the chain of activities whose final aim is the production of a specific output for a particular customer or market
Fan (2001) Shen *et al.* (2004)	*Business process* is a set of one or more linked procedures or activities that collectively realise a business objective or policy goal, normally within the context of an organisational structure defining functional roles and relationships.
Gunasekaran and Kobu (2002)	A group of related tasks that together create value for a customer is called a *business process.*
Hammer and Champy (1993)	A *business process* is a collection of activities that takes one or more kinds of inputs and creates an output that is of value to the customer. A business process has a goal and is affected by events occurring in the external world or in other processes.
Irani *et al.* (2002)	A *business process* is a dynamic ordering of work activities across time and place, with a beginning, an end, and clearly identified inputs and outputs.
Johanson *et al.* (1993)	A business process is a set of linked activities that takes an input and it transforms it to create an output. It should add value to the input and create an output that is more useful and effective to the recipient.
Pall (1987)	Business process is the logical organisation of people, materials, energy, equipment and procedures into work activities designed to produce a specified end result.
Soliman (1998)	*Business process* may be considered as a complex network of activities connected together.
Stock and Lambert (2001)	A *business process* can be viewed as a structure of activities designed for action with focus on the end customer and the dynamic management of flows involving products, information, cash, knowledge and ideas.
Stohr and Zhao (2001)	A *business process* consists of a sequence of activities. It has distinct inputs and outputs and serves a meaningful purpose within an organisation or between organisations.
Volkner and Werners (2000)	*Business process* is defined as a sequence of states, which result from the execution of activities in organisations to reach a certain objective.
Wang and Wang (2005)	*Business process* is defined as a set of business rules that control tasks through explicit representation of process knowledge.

Table 2.1. Business process definitions

Zakarian (2001) emphasises that any transformation occurring in the business process should add value to the inputs and create an output that is useful to a downstream recipient. Others, such as Davenport (1993), Johanson *et al.* (1993), Shen *et al.* (2004) and Stohr and Zhao (2001) provide definitions that involve most of the above issues. There are also some distinctive definitions, such as the ones from Volkner and Werners (2000) and Wang and Wang (2005). Although Volkner and Werners (2000) involve activities in their definition, they emphasise more on *states* as the basic structural elements of business process. This approach provides a different insight into business processes as evolving series of states that change as a result of execution of activities. This definition of business processes can be attractive for using Petri-nets as a business process modelling technique as Petri-nets take into account the different process states. Petri-nets are discussed later in this chapter. Wang and Wang (2005) define business process as a set of business rules that control tasks although they do not sufficiently clarify who executes these tasks and if they are structured in some way.

Although most definitions tend to be similar in the concepts used to express and describe business processes, they have received criticisms for not adequately highlighting the 'business' component and not sufficiently distinguishing from manufacturing or production processes. Volkner and Werners (2000) support that no generally accepted definition of the term business process exists due to the fact that business processes have been approached by a number of different disciplines. Lindsay *et al.* (2003) report that most business process definitions are limited in depth and the corresponding models are also constrained and confined to a mechanistic viewpoint. According to these authors, whereas the production processes focus on the activities being performed, the business processes focus on the goal that needs to be attained and on the people who enact the process. These authors emphasise that business processes are carried out by human operators; they are a balancing act between learning from the past and experimenting with and adapting to the future, and between rules and constraints versus freedom and flexibility. Smith (2003) also refers to business processes as 'human-centred phenomena' that are long lived, persistent, consisting of system-to-system, person-to-system and person-to-person interactions. Volkner and Werners (2000) consider the *flow* as the basis of business process, suggesting that business processes are characterised by the fact that the activities of the flow are executed repeatedly.

Finally, three perspectives or approaches to business processes are identified by Tinnila (1995). The first considers IT as an enabler of business processes to improve operative efficiency. The second discusses the potential of business processes in redesigning of organizations. The third recognises business processes as units of strategic planning and therefore acknowledges the need to connect them more closely to business strategies. Similar to this classification, Chen *et al.* (2001) distinguish between operational, supportive and managerial business processes.

2.2.3 How are workflows different?

As with most concepts, business processes emerged from a related concept: the workflows. The concept of workflow existed before business processes and still is widely used. Workflows are not limited to the business context only, although it is one of their popular applications. Although workflows are precisely defined by the Workflow Management Coalition (WfMC, 1995) in table 2.2, the emergence of business processes created a mismatch between these two concepts. Van der Aalst and ter Hofstede (2002) attribute this lack of consensus to the variety of ways in which business processes are described. Table 2.2 provides a key to the most common workflow definitions and show that there are still different perspectives used by various authors.

Author(s)	Definitions of workflow
Basu and Blanning (2000)	A *workflow* is a particular instantiation of a process. Because a process may include decision points that can cause the process to branch in different ways during execution, a process can contain several possible workflows, each corresponding to a particular set of values for all relevant branching conditions.
Li *et al.* (2004a)	A *workflow* specification is a formal description of business processes in the real world.
Stohr and Zhao (2001)	A *workflow* is a specific kind of process, whose transitions between activities are controlled by an information system (workflow management system).
WfMC (1995)	A *workflow* is defined as the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules.

Table 2.2. Workflow definitions

A review of table 2.2 makes apparent the different approaches to workflows and their relationship with business processes. The definition provided by WfMC relates workflows with business processes and emphasises automation according to a set of procedural rules. Basu and Blanning (2000) support that workflow is only a particular instance of a business process, depicting each time one of the alternative process paths. This approach is not in-line with WfMC and limits workflows to simple business process instances without the ability to demonstrate process patterns such as decision boxes. According to van der Aalst

(1998a), workflows are case-based, i.e., every piece of work is executed for a specific case (e.g. an insurance claim, a tax declaration, or a request for information). Stohr and Zhao (2001) specify workflows as specific kind of processes that are software assisted and enacted. Finally, Li *et al.* (2004a) consider workflows as formal descriptions that rationalise real-world business processes. Business process definitions lack formality and workflows can provide the semantics to push business processes into more structured approaches and specifications. Workflows enable better management and control of the process (Wamelink *et al.*, 2002). However, Wang and Wang (2005) compare traditional workflow approaches concluding that they are too rigid and exact to match complex and dynamic business activities due to the lack of flexibility and adaptability.

2.2.4 Activities & resources as structural elements of business processes

The majority of business process definitions cited in table 2.1 involve the concepts of *activities* and *resources* to describe a business process. Actors are overlooked in most business process definitions since many authors perceive actors as human resources thus omitting any explicit reference. According to van der Aalst (1995), the objective of a business process is the processing of cases and to completely define a business process two things need to be specified: the activities, i.e. partially ordered sets of tasks, and the allocation of resources to tasks. This section discusses *activities* and *resources* as the two main structural elements of business processes and identifies the different perspectives related to these that exist in literature.

Activities

Activities are perceived by the majority of authors as a central element that defines business processes. They are the executable part of a process that is enacted by the actors utilising the resources; therefore activities provide the link between the actors and resources. Table 2.3 provides an overview of the definitions of activity presented in literature in the context of business processes.

Van der Aalst (1998a) provides the simplest definition of activity as a transaction. He also specifies the properties that an activity –similar to a transaction– should satisfy:

- *Atomicity:* An activity is executed successfully or is rolled back completely, i.e., a task cannot be partially completed.
- *Consistency:* The execution of an activity leads to a consistent state.

- *Isolation:* The effect of the execution of an activity in parallel with other activities is equal to the effect of the execution of one activity in isolation.

- *Durability:* The result of a committed activity cannot get lost.

Author(s)	Activity definitions
Van der Aalst (1998a)	One can think of an *activity* as a *transaction*.
Aldowaisan and Gaafar (1999)	An *activity* is defined as a set of operations commonly performed by a single employee type without forced interruptions.
Basten and van der Aalst (1999)	*Activities* are assumed to be atomic entities without internal structure
Kiepuszewski *et al.* (2003)	*Activities* in elementary form are atomic units of work, and in compound form they modularise an execution order/
Li *et al.* (2003)	An *activity* identifies an action which can be characterised by a verb and an object upon which the action applies.
Stohr and Zhao (2001)	An *activity* is a discrete process step performed either by a machine or human agent. An activity may consist of one or more tasks.
Zakarian (2001)	A process model includes a set of *activities* arranged in a specific order, with clearly identified inputs and outputs. Each *activity* in a process takes an input and transforms it into an output with some value to a customer.

Table 2.3. Definitions of activity

Aldowaisan and Gaafar (1999) attempt to classify activities and assign them to particular employee types (i.e. actors), highlighting also the need for their uninterrupted operation. The perception that activities have no internal structure ((Basten and van der Aalst, 1999), (Kiepuszewski *et al.*, 2003)) and are simply atomic units or entities contrasts with Stohr's and Zhao's (2001) hypothesis that an activity may consist of one or more tasks. Usually, the decomposition –or not- of activities depends on the author's perspective of business processes and the details required. Li *et al.* (2003) attempt to identify activities using verbs and objects and Zakarian (2001) claims that like processes, activities transform inputs to value-adding outputs and thus an activity is a process miniature.

Along with these definitions, there are also classifications of activities according to different criteria. Li *et al.* (2003) separate activities to *manual* and *automated*, depending on whether they are realised by a human or a software system and to *primary* or *final* depending on whether they can be refined at a certain stage or not. Zakarian and Kusiak (2001) distinguish between three types of activities:

- *value-adding activities:* activities that are important to the customer;

- *work flow activities:* activities that move work flow across boundaries that are functional, departmental, or organisational; and
- *control activities:* activities that are created to control value-adding and work flow activities.

There are also different perspectives among authors on whether an activity can be decomposed into tasks. According to table 2.3, Basten and van der Aalst (1999) and Kiepuszewski *et al.* (2003) view activities without any internal structure. But Orman (1995) claims that an activity can be further decomposed into tasks that are the smallest identifiable units of analysis. Similar opinions are expressed by van der Aalst (1998a), Biazzo (2000), Li *et al.* (2003) and Stohr and Zhao (2001). As an example Van der Aalst (1995) communicates an inclusive description tying up activities and tasks: 'Business processes are centred around activities. Each activity specifies the set and the order of tasks to be executed in order to achieve the business process goal'. However, many authors tend to use the terms 'activity' and 'task' as equivalent in the context of business processes.

Resources

The second central element of business processes are the resources. According to van der Aalst and van Hee (1996), the allocation of resources to activities, schedules the business process. Many authors refer to resources simply as inputs and others classify resources to input and output (Hofacker and Vetschera, 2001). Resources are used by activities and transformed to create the process output. A number of authors provide different definitions for resources. Li *et al.* (2003) along with van der Aalst and van Hee (1996) consider a resource as any human and/or machine supporting the fulfilment of activities. In a later reference, van der Aalst (1998a) limits resources to human only, stating that 'in most environments where workflow management systems are used the resources are mainly human'. Biazzo (2000) comes up with a more generic definition claiming that resources include everything that is either used or modified by the tasks. While in most business process definitions (table 2.1) the activities are utilising the resources, Castellanos *et al.* (2004) suggest that the resources execute the activity, implying that resources are mostly humans or machines. Hofacker and Vetchera (2001) also classify resources into information and physical according to their nature.

2.2.5 Summary

This section discussed the basic concepts around business processes and can be summarised with the following remarks:

- There is an abundance of definitions and specifications related to business processes in relevant literature.
- The standardisation of the business process definition can have an impact on the business process community and can contribute towards the integration and homogenisation of the approaches towards business process modelling.
- Business processes are inherently a type of process, but they are different in the sense that they perform a business operation.
- Workflows are a similar concept; however, traditional workflow approaches are too rigid and exact to match complex and dynamic business activities due to the lack of flexibility and adaptability.
- An activity is considered as the executable part of a process that is enacted by the actors and utilises the available resources. The term 'task' is either used as a synonym to 'activity' or to denote the smallest identifiable unit in a business process.
- Resources are used by activities and transformed to create the process output. They can be classified as input and output resources for either a business process or a particular activity.

2.3 Business process modelling

Business process modelling is directly related with the perception and understanding of business processes. In most of the cases, a business process is as expressive and as communicative as is the technique that has been used to model it. Therefore the elements and the capabilities of a business process model play a significant role in describing and understanding a business process. IDEF and Petri-nets are frequently encountered in business process modelling literature and this section starts by briefly discussing why it is the case. Then, existing classifications of business process models in literature are presented before the classification proposed by the author is detailed.

2.3.1 Popular modelling techniques: IDEF and Petri-nets

IDEF and Petri-nets are two of the most widely acknowledged and adopted business process modelling techniques according to literature research. Both of them are families of constructs and both have been widely extended and applied to a range of different contexts. Below is an overview of IDEF and Petri-nets in the context of business process modelling.

IDEF (Integrated DEFinition) process modelling

The development of Integrated Definition (IDEF) models for overview and analysis of business processes has been motivated by the initial desire to increase productivity by improving the communication and structure of manufacturing systems (Gunasekaran and Kobu, 2002). The IDEF family of modelling techniques has been popular in companies to model diverse processes and it is also used by many authors because it allows for a systematic and a well-defined representation of processes (Zakarian, 2001). The IDEF family is used in different platforms and applications. The most important types are: IDEF0, IDEF1, IDEF1X, IDEF2, IDEF3, IDEF4 and IDEF5. However, for process modelling, the most widely used techniques are IDEF0 and IDEF3. IDEF3 is the most popular and widely used method in the business process context. One of the major advantages of IDEF3 representation is its simplicity and its descriptive power. These models are also easy to extend. Kusiak and Zakarian (1996b) remark that the essence of IDEF3 methodology is its ability to describe activities and their relationship at various levels of detail, because an initial model includes parent activities that can be decomposed into lower level activities. According to Zakarian and Kusiak (2000) IDEF3 offers several important characteristics for successful process representation:

1. process description in the form of activities,
2. structure of the underlying process, and
3. flow of objects and their relationship.

IDEF3 models have been used by a number of authors as the starting point for further exploitation of models. Kusiak and Zakarian (1996a) perform reliability analysis to identify critical activities in an IDEF3 model, improve process performance, and decrease the operating cost of the process. Zakarian (2001) applies fuzzy reasoning to efficiently model the incomplete information about process variables using an IDEF3 model as a basis. Kusiak *et al.* (1994) performs observational analysis of business processes to demonstrate the current use of IDEF models and Badica *et al.* (2003a) propose a novel business process modelling approach combining IDEF0 and IDEF3 concepts. Lastly, Zhou and Chen (2002) use a combination of IDEF3 and AON (Activity on Node) graphs to formally describe a business process. However, as IDEF is a diagrammatic approach to business process modelling, it has some disadvantages. Zakarian in a number of references ((2001), (2000) and (2001)) and Peters and Peters (1997) highlight the major drawbacks of the IDEF approach:

- *The amount of time required for a process to be completed* – Time is ignored. IDEF diagrams are, basically, like plumbing layouts. They show where everything comes from and goes to without indicating when or how long such a traversal will take.
- *The costs associated with the process* – Being dataflow oriented, IDEF ignores this issue which is often a key motivation for process reengineering.
- *The utilisation of resources during the process* - Not including time makes it impossible to compute what percentage of the total process resources (e.g. people, machines, communications lines) are being utilised.
- *The possibility of company policy being violated* - IDEF, like other static analysis techniques, assumes a rather benign environment. One in which everything and everyone will follow the rules. The possibility for unauthorised detours around company guidelines cannot be checked because no dynamic or simulated events can be examined.
- *The frequency at which time limits are exceeded* - Again, dynamic analyses can demonstrate how often a process will fail to meet time limits.
- The methodology is static and qualitative which is a drawback for the analysis of processes.
- Activities in a model are at a relatively high level of abstraction, making it difficult to associate exact quantitative data for the process variable of interest.
- It is based on informal notation that lacks mathematical rigour. If mathematical definitions are to be applied, these have to be specified for each particular process and each activity separately.

Petri-nets

A Petri-net is a graphical language that is appropriate for modelling systems with concurrency (van der Aalst, 1998a). Petri-nets have been modified and extended by various researchers to allow for more powerful modelling capabilities. Some of their variations include Timed Petri-nets, Stochastic Petri-nets, Coloured Petri-nets and Hierarchical Petri-nets. A Petri-net is a suitable model for a wide variety of applications (e.g. modelling and analysis of concurrent and parallel systems, communication protocols and manufacturing control systems). Figure 2.2 depicts a sample Petri-net of an insurance process claim.

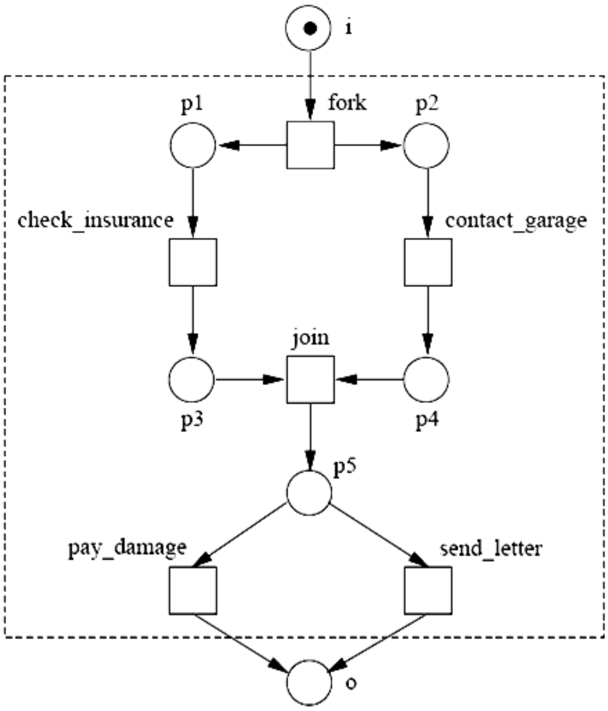


Figure 2.2. A Petri-net modelled business process for an insurance claim (van der Aalst, 1995)

A Petri-net is a kind of directed graph with an initial state called initial marking. The underlying graph of a Petri-net is a directed, bipartite graph consisting of two kinds of nodes, called places and transitions. Arcs (or arrows) represent connections between nodes. An arc can only connect from a place to a transition or from a transition to a place. Connections between two nodes that are of the same kind are not allowed. In graphical representation, places are drawn as circles and transitions as bars or boxes. A marking (state) is an assignment of tokens to the places of the net. A transition is enabled if each place connected to the transition input arc (input place), contains at least one token. According to van der Aalst (1998b) Petri-nets are unique as they cover different perspectives of business process modelling and as such they have three distinctive advantages:

1. They have formal semantics despite the graphical nature.

2. Unlike most of the modelling techniques, they are state-based instead of (just) event-based.

3. They allow the application of analysis techniques.

Li *et al.* (2004b) support Petri-nets due to the above advantages pointing that 'Petri-nets are a naturally selected mathematical foundation for the formal performance analysis of workflow models'. Van der Aalst (1998a) considers Petri-nets as powerful analytical tools that are essential for formally modelling and analysing workflow processes for correctness and consistency (Stohr and Zhao, 2001). Zakarian and Kusiak (2000) highlight that Petri-nets are concurrent, asynchronous, distributed, parallel, nondeterministic, and have a stochastic nature. There are a number of applications of Petri-nets reported in literature. Hofacker and Vetschera (2001) report that, most of the approaches for formal description and analysis of business process designs are based on graphs or Petri nets. Donatelli *et al.* (1995) use Generalised Stochastic Petri Nets (GSPN) and Performance Evaluation Process Algebra (PEPA) to study qualitative and quantitative behaviour of systems in a single environment and identify as strength of Petri-nets their causality, conflict and concurrency clearly depicted within a model. Raposo *et al.* (2000) use a Petri-net based approach to model the coordination mechanisms in multi-workflow environments.

Apart from their wide acceptance, Petri-nets have also received criticisms. Peters and Peters (1997) sum up the essential process modelling elements that the initial form of Petri-nets lack, although most of these have been dealt with in later Petri-net extensions:

1. Time has been left out.
2. The tokens (used to mark conditions) are anonymous.
3. Transitions always behave the same way; people and other systems do not exhibit this property.

Other deficiencies have also been identified. Basu and Blanning (2000) claim that Petri-nets are primarily oriented to analysis and conflict resolution considerations, rather than workflow component connectivity and interactions. Two serious drawbacks are also mentioned by Aguilar-Saven (2004): (i)Petri-nets do not have data concepts and (ii) there are no hierarchy concepts, hence the models can become excessively large. Although the Petri-net techniques can capture system dynamics and physical constraints, they are not adequate to solve optimisation problems (Lee *et al.*, 2001) due to their inability to compose new designs based on an existing one and also quantitatively assess a business process.

Also, Petri-nets are not suitable for someone seeking to understand the flow of a business process due to their focus on states and transitions in a process.

Peters and Peters (1997) examine the possibility of using Petri-nets together with IDEF0 and express their concern on how well these two techniques can match each other. Bosilj-Vuksic *et al.* (2000) also investigate the suitability of IDEF diagrams (IDEF0 and IDEF3) and Petri Nets (DES-nets) for modelling business processes and present a comparative evaluation of their features. According to these authors the comparison reveals that these two methods complement each other and that they can be used together for modelling business processes for better results. Due to their simplicity, it seems appropriate to develop IDEF diagrams during the preliminary phases of business process modelling projects in order to develop 'AS-IS' models and in later phases, when 'TO-BE' models are developed, IDEF diagrams could be transformed into Petri-nets that add formal semantics.

2.3.2 Existing classification of business process models

According to van der Aalst *et al.* (2003), business process modelling is used to characterise the identification and specification of business processes. Business process modelling includes modelling of activities and their causal and temporal relationships as well as specific business rules that process enactments have to comply with. There is an abundance of business process modelling techniques with approaches that capture different aspects of a business process, each having distinctive advantages and disadvantages. Before presenting existing classifications of modelling techniques, the aim, usability and benefits of business process modelling are briefly discussed.

Lindsay *et al.* (2003) describe business process modelling as a snapshot of what is perceived at a point in time regarding the actual business process. The *objective* of business process modelling is, according to Sadiq and Orlowska (2000), the high-level specification of processes, while according to Biazzo (2002), it is the representation of relationships between the activities, people, data and objects involved in the production of a specified output. According to Volkner and Werners (2000) and Aguilar-Saven (2004), business process modelling is essential for the analysis, evaluation and improvement of business processes as it is used to structure the process, such that the existing and alternative sequence of tasks can be analysed systematically and comprehensively. Business process modelling is a useful tool to capture, structure and formalise the knowledge about business processes ((Guha *et al.*, 1993), (Abate *et al.*, 2002)). Aguilar-Saven (2004) suggests that business process models are mainly used to learn about the process, to make decisions on the process, or to develop business process software. For each of these purposes particular business process models are better suited depending on their particular constructs.

Authors such as Kettinger *et al.* (1997), Melao and Pidd (2000) and Aguilar–Saven (2004), have provided frameworks for presenting and classifying different business process modelling techniques. Kettinger *et al.* (1997) conducted a thorough study of business process reengineering methodologies (25), techniques (72) and tools (102) that are adopted by 25 international consultancy firms. The study revealed that in every stage of the reengineering process there is a variety of approaches followed. Kettinger *et al.* (1997) report a widespread use of process capture and modelling techniques. They also present a comprehensive list of the appropriate software tools and the techniques (e.g. process flowcharting, data flow diagramming) that each of the tools supports. However, there is not much emphasis on process modelling itself as it is viewed merely as a technique among others that constitute the wider picture of business process reengineering.

Melao and Pidd (2000) focus exclusively on business processes and their modelling. They adopt four different perspectives for understanding the nature of business processes and then identify the most common modelling approaches for each perspective. The first perspective views business processes as *deterministic machines*; a fixed sequence of well–defined activities that convert inputs to outputs in order to accomplish clear objectives. For this perspective static process modelling is sufficient, with techniques such as Integrated Definition methods (IDEF0, IDEF3) and Role Activity Diagrams (RADs). The second perspective views business processes as *complex dynamic systems*, assemblies of interchangeable components. This second viewpoint focuses on the complex, dynamic and interactive features of business processes. The authors suggest discrete event simulation (discussed later in this chapter) as a suitable way to model the dynamic behaviour of this approach. The third perspective of business processes is *interacting feedback loops* that highlight the information feedback structure of business processes. System dynamics modellers are recommended for this perspective. The last perspective of business process is *social constructs* and emphasises more on the people side. It is the people who make and enact business processes, people with different values, expectations and roles. This soft side of business processes can be modelled with soft unstructured illustrative models. However a real–life business process involves elements for all the four perspectives and therefore it is evident that there is no such modelling technique that can embrace all this variety of characteristics that constitute a business process.

Another notable review regarding business process modelling classification comes from Aguilar–Saven (2004). The author presents the main process modelling techniques and classifies them based on two dimensions: The first dimension is concerned with four different purposes of use and classifies the business process models based on whether they are (i) descriptive for learning, (ii) enable decision support for process development/design, (iii) enable decision support for process execution or (iv) allow IT enactment support. The second dimension distinguishes between *active* and *passive* models. As *active* are considered those models that allow the user to interact with them (dynamic model) while *passive* are those that do not provide this capability. It is important to note that Aguilar–Saven (2004) provides an extensive list of software tools that are associated with all the process modelling techniques presented in their paper. As seen from the references described above each of the authors provides a different modelling framework according to his or her focus on specific directions.

2.3.3 Proposed classification of business process models

The author proposes *three sets* to classify business process modelling techniques as demonstrated in figure 2.3. The first set (i.e. diagrammatic models) involves business process models that sketch a business process using a visual diagram. The second set (i.e. mathematical models) corresponds to models in which all the elements have a mathematical or a formal underpinning. Finally, the third set (i.e. business process languages) contains software–based languages that support business process modelling and most of the times process execution. The classification of the most representative modelling techniques is demonstrated using a Venn diagram (see figure 2.3). Each of the techniques is further discussed later in this section. Table 2.4 presents the classification of figure 2.3 and also cites a selection of references for each of the key techniques. The remaining of this section discusses the main features of these process modelling techniques based on the set (or sets) that they belong to.

Diagrammatic models

The first techniques that were used for business process modelling were plain graphical representations (i.e. flowcharts) that were initially developed for software specification ((Knuth, 1963), (Chapin N., 1971)). These simplistic diagrams depicted a business process but most of the times without using a standard notation (Havey, 2005). These techniques are useful for fast and informal process representation but they lack the necessary semantics to support more complex and standardised constructs. This led to the development of standard methodologies such as IDEF and Unified Modelling Language (UML) for process modelling and/or software development. Business process modelling benefited from these standardised diagrammatic approaches since they are simple and easy to use. However, they have also received a series of criticisms from various authors. The central point of argument is that these modelling approaches are based on graphical notations only (Zakarian, 2001), thus lacking formal semantics (Valiris and Glykas, 1999). They also lack quantitative information that obstructs any further analysis and development of analysis methods and tools (van der Aalst and van Hee, 1996); there is no formal underpinning to ensure consistency across models (Valiris and Glykas, 1999). Phalp (2000) notes that any analysis attempt using these types of models often consists solely of inspection of diagrams and the conclusions are heavily dependent upon the skills of the analyst.

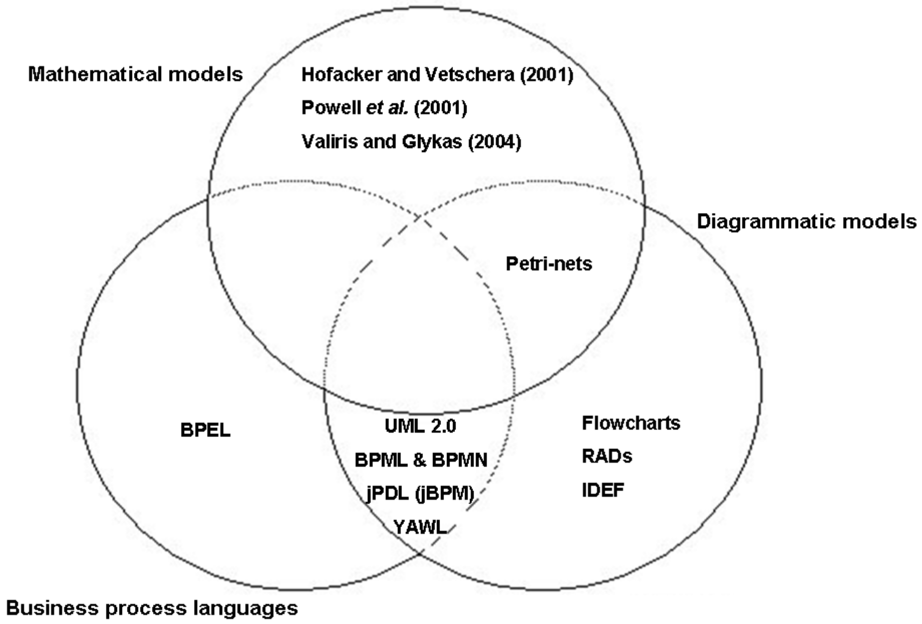


Figure 2.3. Classification of business process modelling techniques

Although visual inspection of diagrams tends to be highly subjective, these diagrams are still widely used in business process environments. The advantage to visually depict the flow of a business process in a way that no technical expertise is required is very appealing to the business analysts. Even advanced and more sophisticated modelling techniques are influenced by this perspective and they support a visual representation of the modelled processes.

Modelling techniques	Modelling set(s)	Selected references
–Flowcharts	–Diagrammatic models	– (Knuth, 1963) – (Chapin, 1971) – (Chapin, 1974) – (Feldman, 1998) – (Lakin *et al.*, 1996)
–IDEF	–Diagrammatic models	– (Mayer *et al.*, 1994) – (Menzel and Mayer, 1998) – (Peters and Peters, 1997) – (Zakarian and Kusiak, 2001) – (Zakarian and Kusiak, 2000) – (Zakarian, 2001) – (Badica *et al.*, 2003a) – (Shimizu and Sahara, 2000) – (Zhou and Chen, 2002)
–RADs	–Diagrammatic models	– (Ould, 1995) – (Holt, 2000) – (Phalp and Shepperd, 2000) – (Badica *et al.*, 2003b)
–UML	–Diagrammatic models –Business process language	– (Quatrani, 2001) – (Kim *et al.*, 2003) – (Wohed *et al.*, 2004)
–Petri–nets	–Diagrammatic models –Formal/mathematical models	– (van der Aalst, 1998a) – (Li *et al.*, 2004b) – (Donatelli *et al.*, 1995) – (Raposo *et al.*, 2000) – (Peters and Peters, 1997)
–Business process models based on mathematical or algorithmic models	–Formal/mathematical models	– (Hofacker and Vetschera, 2001) – (Powell *et al.*, 2001), – (Valiris and Glykas, 1999)
–BPEL –BPML	–Business process language	– (Reimer *et al.*, 2000) – (Havey, 2005) – (Grigori *et al.*, 2004) – (Smith, 2003)
–jPDL (jBPM)	–Diagrammatic models –Business process language	– (Koenig, 2004)

Table 2.4. Main modelling techniques, corresponding sets and selected references

Formal/mathematical models

The necessity for formal semantics for business process modelling led to a second generation of formal models. Formal models are the ones in which process concepts are defined rigorously and precisely, so that mathematics can be used to analyse them, extract knowledge from them and reason about them. An advantage of formal models is that they can be verified mathematically, and can be checked for consistency and other properties (Koubarakis and Plexousakis, 2002). These models are in line with van der Aalst et al. (2003) suggestion that business process models 'should have a formal foundation' because formal models do not leave any scope for ambiguity and increase the potential for analysis. However, there is a lack of formal methods to support the design of processes (Hofacker and Vetschera, 2001) because business process elements and constraints are mostly of qualitative nature and it is hard to characterise them in a formal way amenable to analytical methods (Tiwari, 2001). This explains the difficulty of developing 'parametric' models of business processes and the fact that only a few practical examples are found in relevant literature (e.g. Hofacker and Vetschera, 2001). Petri–nets are an example of a business process modelling technique that combines visual representation using standard notation with an underlying mathematical representation.

Coming to the approaches that use mathematical models only, there is no widely accepted model. This results into different authors presenting their individual approaches towards mathematical business process modelling. An approach that has a mathematical basis is proposed by Hofacker and Vetschera (2001). They describe a business process using a series of mathematical constraints (that define the feasibility boundaries of the business process) and a set of objective functions (that consist of the various objectives for business process design). Their approach can only handle sequential processes and cannot model complex modelling constructs. Also, although there is no emphasis on the diagrammatic representation, this approach can be subject to quantitative analysis and improvement as it is based on a mathematical model. A similar approach is presented by Powell *et al.* (2001). They describe a mathematical model that has the main ingredients of a generic business process. Valiris and Glykas (1999) also propose the use of formal mathematical notations as a way of introducing business rules and verifying the logical consistency of diagrammatic models.

Despite their advantages over simple diagrammatic approaches, criticisms for formal/mathematical business process models have also been reported. Building a formal business process model can prove much more complex and demanding compared to traditional techniques where a process diagram is sufficient (Hofacker and Vetschera, 2001). These authors also show that the representation of real–life processes using mathematical models may be complex and sometimes not possible as these include complex features such as decision points, feedback loops and parallel or hierarchical flow.

Koubarakis and Plexousakis (2002) note that the use of complex mathematical notations might discourage the business analyst since 'it is a lot of work to create, maintain a formal business process and retain its consistency'. However, as a diagram can lead to ambiguity about the process, the formal model ensures that the process is described accurately and analysis tools can be used to extract quantitative information about the process. This is the main advantage of formal business process modelling techniques.

Business process languages

The third –and most recent– generation of business process modelling techniques comes as an attempt to tackle the complexity of the formal models but retain their consistency and potential for further analysis. As the first generation of business process modelling techniques was strongly influenced by the ones used in software development; so is this generation. It is the dynamic, complex and rapidly evolving nature of business process models that makes them similar to software development techniques. The third set takes business process modelling a step further as it uses *process languages* –usually XML–based– to model and execute a business process. These context–specific executable languages are the latest trend is business process modelling, a trend that has already produced a number of different semantic packages, with Business Process Execution Language for Web Services (BPEL4WS –also known as BPEL) and Business Process Modelling Language (BPML) being the most distinctive. Van der Aalst *et al.* (2003) remark that process languages with clear semantics are useful as they can express business process models and contribute to the analysis of their structural properties.

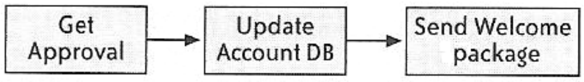
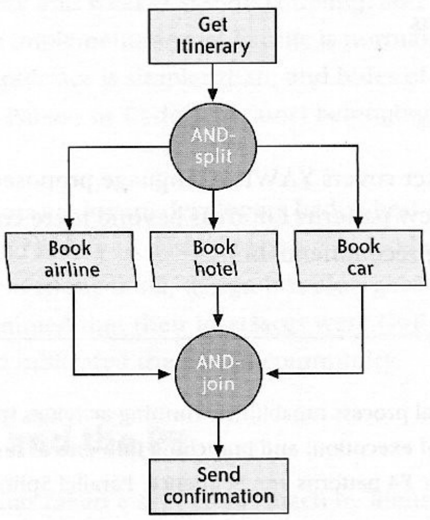
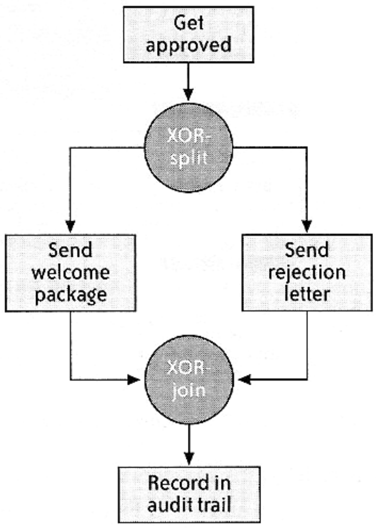
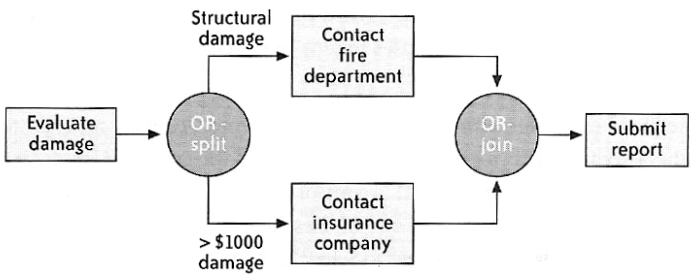
Havey (2005) claims that BPEL is the most popular as it is supported by IBM, Microsoft and BEA. BPEL is not a notational language but it is also XML–based and as such it inherits XML attributes such as programmability, executability and exportablility. BPML is a product of the Business Process Modelling Initiative (www.bpmi.org). It is also an XML–based language that encodes the flow of a business process in an executable form. BPML is accompanied by BPMN (Business Process Modelling Notation), a graphical flowchart language that is able to represent a business process in an intuitive visual form (Havey, 2005). Each BPML process has a name, a set of activities and a handler; it also supports subprocesses. YAWL (Yet Another Workflow Language) is another –as the name itself says– graphical process language created by van der Aalst and ter Hofstede (2003). YAWL is a Petri–net based language that was built with the primary target to support a wide range of business process patterns. It has received criticism for being inadequate in terms of expressiveness and system integration capabilities (Havey, 2005). JBoss Business Process Management (JBPM) execution language named jPDL (Koenig, 2004) is also a novel approach to business process modelling and execution. This new approach facilitates the natural transition from declarative input by the business analyst to the programming logic needed to implement a business process, thus simplifying business process development and allowing even non–programmers to develop business processes using visual tools. jBPM engine is based on open source software, providing infrastructure to developers who have access to a variety of supplementary software tools with which they can easily design and analyse business processes in a graphical environment.

2.3.4 Handling complexity of business process patterns

Any business process modelling technique should be able to support a range of patterns. According to Riehle and Zuillinghoven (1996) a pattern 'is the abstraction from a concrete form which keeps recurring in specific non-arbitrary contexts'. Wohed *et al.* (2002) refer to patterns as 'abstracted forms of recurring situations in processes'. Havey (2005) is more specific about *business process patterns*: 'they are inherently spatial and visual. A process pattern is a cluster, or a constellation of process activities arranged in just the right way to solve a difficult problem'. Zapf (2000) supports the pattern construction for specific application domains as this allows a detailed analysis.

Patterns enable the standardisation of solutions to commonly recurring problems within business processes and the reuse of these standardised process parts across different process models. Identifying the basic process constructs is necessary for any business process modelling approach to be able to consider several complex dependencies between the activities (Scheer, 1994). Authors such as Kiepuszewski *et al.* (2003), van der Aalst and ter Hofstede (2002) and Zhou and Chen (2002) refer to sequence, choice, parallelism, and synchronization as the basic patterns for modelling and controlling a business process.

Similar constructs are mentioned by Volkner and Werners (2000) as AND, inclusive-OR, exclusive-OR and their combinations. Van der Aalst and ter Hofstede (2002) also introduce a comprehensive list of 20 workflow patters. These patterns have been compiled from an analysis of existing workflow languages and they capture typical control flow dependencies encountered in workflow modelling (Wohed *et al.*, 2002). In (van der Aalst and ter Hofstede, 2002), the functionality of 15 workflow management systems is compared. The results of this experiment revealed two problems: (i) current workflow systems do not have significant expressive power and (ii) they do not support a consistent range of patterns. Table 2.5 presents a selection of patterns from van der Aalst and ter Hofstede (2002) that were considered as the basic constructs for any business process model. These patterns are provided with a brief explanation while the pattern images are taken from Havey (2005).

<table>
<tr><td colspan="2">1. Sequence
Demonstrates the process steps that are performed sequentially.

</td></tr>
<tr><td>2. Parallel split & Synchronisation (AND-split & join)
The branch from a single activity to multiple parallel paths and their convergence to a single activity, which waits for the completion of all paths before starting.

</td><td>3. Exclusive choice & Simple merge (XOR-split & join)
The branch from a simple activity to exactly one of several paths, based on the evaluation of a condition and the convergence on a simple activity which starts when one of the chosen paths completes.

</td></tr>
<tr><td colspan="2">4. Multi-choice & Synchronising merge (OR-split & join)
The choice of one or more parallel branches, in which each branch is taken only if it satisfies a particular condition and the branch join when all of the active parallel paths are complete.

</td></tr>
<tr><td colspan="2">5. Discriminator (N-out-of-M join)
Multiple parallel patterns join but exactly one (discriminator) or N (out of M) are allowed to continue in</td></tr>
</table>

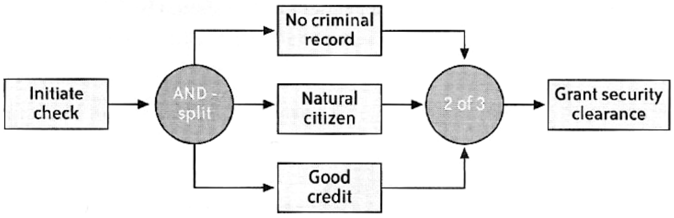
the process, based on a condition evaluated at runtime. The remaining branches are blocked. 
6. Arbitrary cycles (GOTO or feedback loop) The repetition of an activity or a set of activities by cycling back to it in the process.
7. Cancel activity (Kill activity) To stop the execution of a particular process activity on a cancellation trigger.
8. Cancel case (Kill process) To stop the execution of the entire process on a cancellation trigger.

Table 2.5. Main business process patterns (images from Havey, 2005)

Pattern	IDEF3	UML 2.0	Petri-nets	Math. model	BPEL	BPML	jPDL
1. Sequence	✓	✓	✓	✓	✓	✓	✓
2. AND-split & join	✓	✓	✓	✗	✓	✓	✓
3. XOR-split & join	✓	✓	✓	✗	✓	✓	✓
4. OR-split & join	✓	✓/✗	✓	✗	✓	✗	✓
5. Discriminator	✗	✗	✗	✗	✗	✗	✗
6. Arbitrary cycles	✗	✓	✗	✗	✓	✓	✓
7. Cancel activity	✗	✓	✗	✗	✓	✓	✓
8. Cancel case	✗	✓	✗	✗	✓	✓	✓

Table 2.6. Process patterns supported by modelling techniques and languages

Most of the main business process patterns are inspired by software specifications. Table 2.6 identifies which business process modelling techniques support these patterns. The modelling techniques are selected across all the three modelling dimensions. IDEF3 supports only the basic patterns (Zhou and Chen, 2002). UML provides support for almost all the patterns presented here apart from OR-join and Discriminator (Wohed *et al.*, 2004). Petri-nets and IDEF support the same patterns according to van der Aalst (1998a). However, most of the business process patterns are covered by the various Petri-net extensions. The mathematical model (Hofacker and Vetschera, 2001) –although praised for its formality and optimisation capabilities– illustrates a simplistic approach towards business processes thus no pattern is implemented apart from the sequential flow of activities. This is due to the complexity of the mathematical model development. Most of the business process languages are implemented based on the process patterns. For example, YAWL supports all the patterns in table 2.6 since it was created primarily for this purpose (van der Aalst and ter Hofstede, 2002). BPEL also supports most patterns (Wohed *et al.*, 2002), (Havey, 2005) and also BPML (Havey, 2005). According to Koenig (2004), jBPM's jPDL was also implemented to cover all the patterns presented here.

2.3.5 Summary

This section discussed the main approaches towards business process modelling and proposed a novel classification of the existing modelling techniques. The following remarks summarise the section:

- Business process modelling is an essential aspect of business processes as –in the majority of cases– a business process is as expressive and as communicative as is the technique that is used to model it.
- IDEF and Petri-nets are still popular techniques for business process modelling; however they have some drawbacks which make them inappropriate for optimisation.
- A proposed classification of the existing business process modelling techniques involves three sets: diagrammatic models, mathematical models and business process languages. This classification of business processes contributes to visually highlighting a number of interesting observations.
- Despite the existence of many formal process modelling notations, the majority of the business process community still uses simple diagrammatic modelling techniques.
- There is an increasing need for formal methods and techniques to support both the modelling and the analysis of business processes.
- Business process languages provide diagrammatic depiction of business processes and associated analysis techniques which can be used for investigating properties of processes. However, a disadvantage of the business process languages that aim at automating business processes are the limitations of their modelling concepts.
- Business process models from all the three sets are analysed to show that only the business process languages fully support a wide range of process patterns.

2.4 Business process analysis

According to Irani *et al.* (2002) businesses should not be analysed in terms of the functions in which they can be decomposed to or in terms of the products they produce, but in terms of the key *business processes* that they perform. Due to the complexity of process design and control encountered in modern businesses, there is a need for the development of suitable analysis techniques (van der Aalst, 1998a). However, *business process analysis* is a term used with a broad meaning including a range of different tactics such as simulation and diagnosis, verification and performance analysis of business processes. Van der Aalst and ter Hofstede (2003) underline that business process analysis should aim at investigating properties of business processes that are neither obvious nor trivial. Boekhoudt *et al.* (2000) justify the necessity for analysis of business process models in order to clarify the business process characteristics, identify possible bottlenecks and compare any potential process alternatives. Yet most of business process analysis approaches are based on subjective rather than objective methods (Valiris and Glykas, 1999). In line with van der Aalst and ter Hofstede (2003), Boekhoudt *et al.* (2000) also report that among the modelling techniques, those that have formal semantics and mathematical basis are the most suitable for analysis. Irani *et al.* (2002) citing Davenport (1993) state that to understand and analyse a business process helps to recognise the sources of problems and ensure that they are not repeated in the new process thus providing a measure of value for the proposed changes. This approach opposes the radical attitude towards business process redesign introduced by Hammer and Champy (1993). This section presents the different types of business process analysis and presents a variety of representative approaches found in literature.

2.4.1 From observational analysis to performance evaluation

There are different *types* of analysis related to business process. Figure 2.4 presents these different analysis types in a Venn diagram. It matches the types of process analysis to each of the three business process modelling sets introduced in the previous section (2.3). For the diagrammatic models (first set) only *observational analysis* can be applied. Observational analysis, which primarily entails altering the process structure via inspection of the diagrams (Aldowaisan and Gaafar, 1999), is the most common analysis approach using visual models of business processes (Phalp and Shepperd, 2000). The observational analysis technique offers a set of options to redesign a process that includes eliminating non–value–added activities (e.g. redundant, rework and supervisory activities), simplifying activities, combining activities, increasing the concurrency of activities and automating activities (Kusiak *et al.*, 1994). However, this analysis approach can be time consuming and heavily dependent upon the experience of the modeller whose conclusions are frequently based upon his/her knowledge of the particular business domain and his/her skills (Ould, 1995). Zakarian (2001) recognised that diagrammatic process models have qualitative notation and this results in the lack of analysis tools thus making the application of quantitative methods unusual (Volkner and Werners, 2000) and unattractive. Making business process analysis meaningful and attractive is not only linked to the construction of ever–more detailed maps, which use increasingly sophisticated representational techniques, but also the willingness to combine seemingly irreconcilable strategies for analysis (Biazzo, 2000).

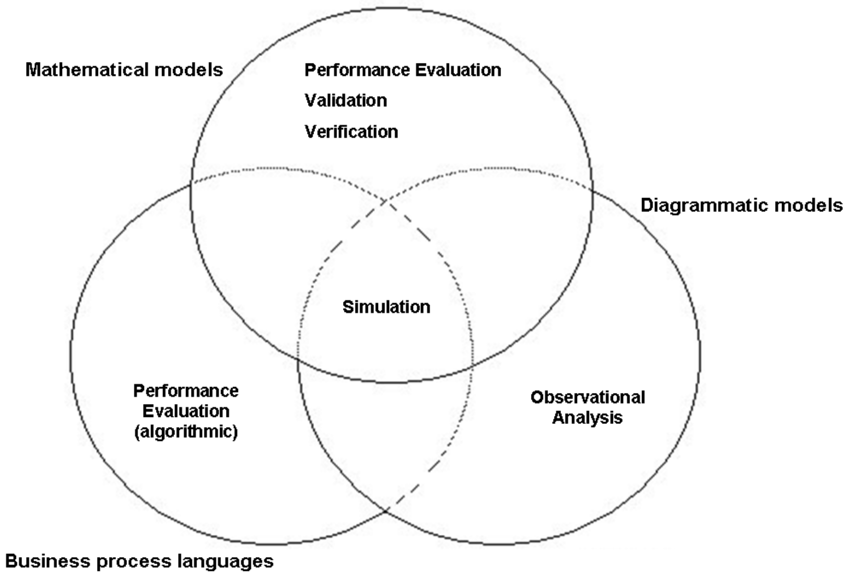


Figure 2.4. Types of process analysis for the business process modelling sets

When analysing a business process it is necessary to have mechanisms more sophisticated than simple qualitative analysis of static diagrammatic models. Authors such as Aguilar–Saven (2004) and Zakarian (2001) stress the need for formal techniques for analysis of process models, in order to make process modelling methodologies more attractive and meaningful. The need for quantitative analysis of the business process models is one of the major reasons for the evolution of process models with formal underpinning (i.e. mathematical models). These formal approaches to modelling of business processes provide a sound basis for setting performance indicators that measure the attainment of strategic goals and objectives by relating these goals and objectives to the core processes (Lewis, 1993). For these to occur, analyses types that present both dynamic and functional aspects of the process are required. According to van der Aalst (1998a), most of the techniques that are used for the analysis of formal business process models, originate from operations research.

Figure 2.4 demonstrates that the three different types of business process analyses proposed by van der Aalst (2004) (having workflows –and in particular Petri–nets– in mind) belong to the mathematical models set:

- *validation*, i.e. testing whether the business process behaves as expected in a given context,
- *verification*, i.e. establishing the correctness of a business process and
- *performance analysis* (or performance evaluation), i.e. evaluating the ability to meet requirements with respect to throughput times, service levels, and resource utilisation or other quantitative factors.

None of the above analysis types can be applied to a visual diagram only; formal underpinning of the process model is required. Validation checks whether the system behaves as expected in a particular context, while verification checks whether the business process model is free of logical errors (van der Aalst, 1998a). Verification, unlike validation, is context independent; it detects, for example, deadlocks in process designs a logical error independent of the purpose of the process. Performance evaluation aims to describe, analyse, and optimise the dynamic, time–dependent behaviour of systems (Hermanns *et al.*, 2002), (Raposo *et al.*, 2000). Validation can be done by interactive simulation: a number of fictitious cases are fed to the system to see whether they are handled well. However, verification and performance analysis require more advanced analysis techniques (van der Aalst, 2004).

Li *et al.* (2004b) present another classification of business process analyses. This classification is very similar to that proposed by van der Aalst (2004). According to Li *et al.* (2004b), workflow model analysis is conducted mainly at three levels –the logical, the temporal, and the performance levels which deal with different aspects of a workflow model. Logical level focuses on the correctness of the various process events (i.e. verification) and temporal level focuses on the interval dependency relations of a workflow model with imposed timing constraints (i.e. validation). The logical and temporal levels ensure only a functionally working workflow but not its operational efficiency. The performance level focuses on evaluating the ability of the workflow to meet requirements with respect to some key performance indicators. Although performance analysis of business processes is recognised as a significant step towards quantitative analysis of business processes, it has not captured the attention of many researchers (Salimifard and Wright, 2001).

The concept behind business process languages is to make a process executable and hence amenable to quantitative analysis. However, for business process languages set, only simulation is proposed in literature explicitly. Simulation is a software–assisted technique for analysing business process; it is discussed below in a separate sub-section. Although formal languages have been exploited in order to define and model business processes, the use of formal languages to handle the performance evaluation of workflows has received little coverage (Abate *et al.*, 2002). However, some process languages have associated analysis techniques which can be used for investigating process properties. These techniques can then be relied upon to provide insight into the behaviour and characteristics of a business process model specified in the language (van der Aalst *et al.*, 2003). According to the authors' opinion this level of modelling and execution of business processes (i.e. using a process language) is the most suitable for the application of any analysis technique. These can be in the form of algorithmic expressions that can be expressed using the process language and thus be integrated within the process model. However, it is not sufficient to just develop these techniques. It is important to look at methods and tools to make them applicable in the practical context (van der Aalst, 1998a).

2.4.2 Simulation of business processes

Simulation is a popular technique for analysing business processes and it can involve other types of analyses mentioned above. According to Volkner and Werners (2000) many problems of business processes have similarities to problems in project management or production process planning which have already been analysed successfully using simulation. Simulation provides a structured environment in which one can understand, analyse, and improve business processes (Gunasekaran and Kobu, 2002). Business process simulation is used to assist decision making by providing a tool that allows the current behaviour of a system to be analysed and understood. It can also predict the performance of the system under a number of scenarios determined by the decision maker (Greasley, 2003). Process simulation facilitates process diagnosis (i.e. analysis) in the sense that by simulating real–world cases, what–if analyses can be carried out (van der Aalst, 1998a). The advantage of simulation is that it is a flexible technique (van der Aalst, 2001) because it can be used to obtain an assessment of the current process performance and/or to formulate hypotheses about possible process redesign (Abate *et al.*, 2002).

Modern simulation packages allow for both the visualisation and performance analysis of a given process (van der Aalst, 2001) and are frequently used to evaluate the dynamic behaviour of alternative designs (Aldowaisan and Gaafar, 1999). Visualisation and graphical user interface are important in making the simulation process more user–friendly. According to Fathee *et al.* (1998) simulation is most useful for the analysis of stable business processes and less useful for dynamic systems that do not reach equilibrium. The main advantage of simulation–based analysis is that it can predict process performance using a number of quantitative measures such as lead–time, resource utilisation and cost (Greasley, 2003). As such, it provides a means of evaluating the execution of the business process to determine inefficient behaviour (Ferscha, 1998). Thus business process execution data can feed simulation tools that exploit mathematical models for the purpose of business process optimisation and redesign (Abate *et al.*, 2002). Dynamic process models can enable the analysis of alternative process scenarios through simulation by providing quantitative process metrics such as cost, cycle time, serviceability and resource utilisation (Gunasekaran and Kobu, 2002). These metrics form the basis for evaluating alternatives and selecting the most promising scenario for implementation (Levas *et al.*, 1995). However, these analytical models (mostly mathematical), according to Gunasekaran (2002), have not received much attention due to their complexity despite their ability to play a greater role in measuring performance and in conducting experiments.

The advantages of applying simulation are: (i) the possibility for the quantitative analysis of business processes with consideration to their dynamic characteristics, (ii) the possibility for a systematic generation of alternatives by modifications in identified weak points and (iii) the high flexibility in modelling as well as an adequate consideration of stochastic influences (Volkner and Werners, 2000). However, simulation has some weak points as well. Some authors ((Greasley, 2003), (Volkner and Werners, 2000)) report the large costs involved and the large amount of time to build a simulation model due to the complexity and knowledge required in building such models. Van der Aalst (2001), underlines that simulation supports only 'what–if' analysis and does not suggest any process improvements. Basu and Blanning (2000) also claim that while process simulation can provide useful insight into process behaviour, it does not address questions about the interrelationships among process components.

2.4.3 A compilation of approaches for business process analysis

After identifying the main analysis types for business process, the most relevant approaches found in literature are discussed. Table 2.7 presents the analysis types and approaches for a selection of business process modelling techniques. For each process modelling technique the table cites the modelling set(s) it belongs to (based on figure 2.3), the types of analyses applicable based on these sets (based on figure 2.4) and a selected number of related approaches (references). According to table 2.7, most analysis approaches reported in the literature are based on models that belong to the diagrammatic models set. Also, no analysis approach is reported for the business process languages set.

business process MODEL	modelling SET(S)	business process analysis TYPES	business process analysis APPROACHES
IDEF	–Diagrammatic models	–Observational –Simulation	– (Kusiak and Zakarian, 1996a) – (Kusiak and Zakarian, 1996b) – (Zakarian and Kusiak, 2001) – (Zakarian and Kusiak, 2000) – (Zakarian, 2001) – (Badica *et al.*, 2003a) – (Peters and Peters, 1997) – (Shimizu and Sahara, 2000)
RADs	–Diagrammatic models	–Observational –Performance analysis	– (Phalp and Shepperd, 2000) – (Badica *et al.*, 2003b)
Petri–nets	–Diagrammatic models –Mathematical/formal models	–Observational –Validation –Verification –Performance analysis –Simulation	– (van der Aalst, 1998) – (van der Aalst *et al.*, 1994) – (van der Aalst and van Hee, 1996) – (van der Aalst, 1995) – (van der Aalst, 2003) – (Kiepuszewski *et al.*, 2003) – (Li *et al.*, 2004b) – (Donatelli *et al.*, 1995) – (Gao *et al.*, 2003) – (Raposo *et al.*, 2000) – (Peters and Peters, 1997)
Mathematical models	–Mathematical/formal models	–Performance analysis –Simulation	– (Powell *et al.*, 2001) – (Valiris and Glykas, 2004)
Business process languages	–Business process languages	–Performance analysis (algorithmic) –Simulation	(none reported in literature)

Table 2.7. Business process analysis approaches based on modelling sets and analysis types

IDEF models have been a starting point for business process analysis for authors such as Kusiak and Zakarian that have published a series of papers (refer to table 2.7 for the references) exploring and analysing various aspects of IDEF models. The most representative is Zakarian (2001) where the author is using an IDEF3 model attempting to model and analyse/quantify a business process using a combination of fuzzy logic and rule–based reasoning. Using –although not explicitly mentioned– observational analysis, he extracts IF–THEN fuzzy rules from the IDEF3 model and defines a number of linguistic variables. The linguistic variables are categorised into fuzzy sets which are defuzzified by assigning precise boundaries. The process is accurately executed and its output is quantified and predicted by assigning values to each variable. Combinations of different values for each variable can be applied to analyse and test the process and its outputs. Peters and Peters (1997) also present a tool to simulate an IDEF0 model by making dynamic transformations. Other IDEF–based analysis approaches come from Badica *et al.* (2003a) and Shimizu and Sahara (2000). Another group of analysis approaches is related to the quantification of Role Activity Diagrams (RADs). Phalp and Shepperd (2000) attempted to quantify RADs. The authors extracted a metric (coupling ratio) to measure the correlation between actions (sole activities of a role) and interactions (involvement of another role). By reducing coupling, roles can become more autonomous within the process because they do not need to synchronise. Badica *et al.* (2003b) attempted to map and quantify RADs using a similar approach.

When it comes to process models with formal underpinnings, two main approaches are identified: those built around Petri–nets and those that use mathematical models of business processes. Van der Aalst has produced a series of papers focusing on different aspects of Petri–nets and workflow analysis (refer to table 2.7 for references), but he tends to focus more on validation, verification and correctness of workflows rather than performance analysis. Other analysis approaches include: Donatelli *et al.* (1995) that involves Process Algebra and Stochastic Petri–nets, and Gao *et al.* (2003) that applies fuzzy–reasoning to Petri–nets. In terms of mathematical models, Powell *et al.* (2001) propose a series of mathematical formulations and ratios to measure, analyse and control business processes. Valiris and Glykas (2004) propose a framework that contains a series of metrics for business processes. As mentioned previously in this section, as of yet there are no reported analysis approaches explicitly for business process languages.

2.4.4 Summary

This section discussed the main types of business process analysis and mapped these types to the proposed sets of business process modelling techniques. The following remarks are drawn from the extension of the proposed classification to cover business process analysis and they summarise this section:

- Analysis of business processes includes a range of different tactics such as simulation and diagnosis, verification and performance analysis.
- The proposed classification of business process modelling techniques was used in this section to demonstrate the available analysis types for each of the proposed sets.
- For the diagrammatic models (first set) only *observational analysis* can be applied. However, this analysis approach can be time consuming and heavily dependent upon the experience of the modeller. Nonetheless, most analysis approaches reported in the literature are based on models that belong to the diagrammatic models set.
- The need for quantitative analysis of the business process models is one of the major reasons for the evolution of process models with formal underpinning (i.e. mathematical models).
- Performance analysis can be directly used for decision–support and further improvement of the process. The knowledge extracted from performance analysis should be fed back to the process in order to improve it. However, the proposed analysis classification demonstrates a lack of reported performance analysis approaches.
- Simulation is a popular technique for analysing business processes and it can involve other types of analyses mentioned. However, simulation supports only 'what–if' analysis and does not suggest any process improvements.
- There are no reported analysis approaches explicitly for business process languages.

2.5 Business process optimisation

As the previous sections discussed, business process modelling does not add much value without further inspection and analysis of the business processes model. Likewise, process analysis has little value, unless it helps in improving or optimising a business process (van der Aalst *et al.*, 2003). Process improvement can occur through associated formal techniques (van der Aalst *et al.*, 2003) that support both the modelling and the analysis of business processes (van der Aalst and van Hee, 1996). A holistic approach towards business processes should capture a business process (business process modelling), provide the necessary means for bottleneck identification and performance analysis and – eventually– generate alternative improved business process(es) in terms of specified objectives. But often this last part (business process optimisation) is overlooked –if not completely neglected in business process literature. This section discusses the difference between process improvement and optimisation, and provides a classification of the current business process optimisation approaches.

2.5.1 Improvement is not enough

Business process improvement started as part of business process redesign and/or reengineering efforts that promised exceptional results. Gunasekaran and Kobu (2002) claim that a business process has to undergo fundamental changes to achieve significant performance improvements. According to Soliman (1998), the objectives of business process re–engineering are to improve the business processes and reduce costs. However, although most of the business process re–engineering (or re–design) attempts in literature claim to support business process improvement, there are scarce cases that describe with sufficient details the actual improvement steps that need to be undertaken. Jaeger *et al.* (1995) is a typical case where business process improvement is limited to a broad description of steps:

1. specify the system.
2. identify the performance bottleneck(s).
3. choose among the possible modifications to resolve the performance bottlenecks.

These guidelines are not sufficient for a structured process improvement as they do not provide the necessary insight and level of detail for the actions that lead to process improvement. Another similar approach is presented by Aldowaisan and Gaafar (1999) and it is based on observational analysis. Their technique has a set of options to redesign a process. This includes: eliminating non–value–added activities (e.g. redundant, rework and supervisory activities), simplifying activities, combining activities and increasing the concurrency of activities; but again the improvement process is not transparent. This approach does not guarantee an optimum redesign as it manually derives alternative process maps starting from the current process map.

A methodology for business process improvement is only as good as the tools and techniques that support it (Bal, 1998). However, the literature restricts itself to descriptions of the 'situation before' and the 'situation after', giving very little information on the redesign process itself (Reijers and Liman-Mansar, 2005). Valiris and Glykas (2004) criticise this perspective, stating that most of these re–engineering methodologies lack the formal underpinning to ensure the logical consistency of the generation of the improved business process models. This leads to a lack of systematic approach that can guide a process re–designer through a series of (repeatable) steps for the achievement of process redesign (Valiris and Glykas, 2004). While there are several methodologies for structuring business process redesign projects, the task of developing optimal designs of business processes is left to the designer's intuition (Hofacker and Vetschera, 2001). *Business process optimisation* is the *automated* improvement of business processes using pre–specified quantitative measures of performance (objectives) and as discussed in the next section it is the appropriate systematic approach to fill in this gap.

2.5.2 Two perspectives for business process optimisation

Business process optimisation can espouse techniques from relevant disciplines. Gunasekaran and Kobu (2002) claim that, within the business process context, there is a need for a wider use of Decision Support Systems based on Artificial Intelligence and Expert Systems. They also support the need for developing queuing, linear programming and simulation models to represent business processes and to select the optimal design. In this section we discuss and relate two other disciplines with business processes: *scheduling* and *evolutionary computing*. Scheduling shares a range of common topics with business processes and evolutionary computing is an already successful optimisation approach in other areas.

Business processes and Scheduling

Scheduling problems are similar to business process optimisation problems. Both disciplines share common topics such as the optimal allocation of resources to tasks (van der Aalst, 1996). Having this in mind, a range of already successful optimisation approaches from scheduling can become available to business processes taking into account what Ernst *et al.* (2004) claim that optimisation capabilities are generally targeted at a specific application area and cannot be easily transferred to another discipline.

According to Bellabdaoui and Teghem (2006), the development of optimisation models for planning and scheduling is one of the most useful tools for improving productivity in a large number of companies. There is a range of review papers about scheduling optimisation approaches. Mathematical programming, especially Mixed Integer Linear Programming (MILP) has become one of the most widely explored methods for process scheduling problems because of its rigour, flexibility and extensive modelling capability. Floudas and Lin (2005) present an overview of the developments of MILP-based approaches for scheduling and observe increasing application of the formal MILP optimisation framework to real scheduling problems in process and related industries. Kallrath (2002) gives an overview of the current state-of-the-art of planning and scheduling problems and reaches to similar conclusions. According to this author the state-of-the-art technology based on mathematical, especially mixed-integer optimisation for planning is advanced and appropriate for solving real world planning problems. The reason is that mixed integer optimisation can provide a quantitative basis for decisions and it has proven itself as a useful technique to reduce costs and to support other objectives. Rommelfanger (2004) presents another scheduling optimisation approach that involves fuzzy mathematical programming. While in the case of classical models the vague data is replaced by 'average data', fuzzy models offer the opportunity to model subjective judgement of a decision maker as precisely as the decision maker is able to describe it. In contrary to classical systems, in fuzzy systems combined with an interactive solution process the information can be gathered step by step. Another advantage of fuzzy models is the fact that mixed integer programming problems can be solved easily because the boundaries are not crisp.

These scheduling problems are inherently combinatorial in nature because of the many discrete decisions involved, such as equipment assignment and task allocation over time. Shah (1998) examines different techniques for optimising production schedules with an emphasis on formal mathematical methods. Pinto and Grossmann (1998) also present an overview of assignment and sequencing models used in scheduling with mathematical programming techniques. A recent review comes from Mendez *et al.* (2006) that present an extensive classification of scheduling problem types that demonstrates their diversity. Addressing this diversity, these authors also present a general classification of optimisation models as a framework for describing the major optimisation approaches that have emerged over the last decade regarding scheduling.

From the above, one can conclude that scheduling optimisation is an established research area reporting successful approaches. These approaches can inspire relevant applications in business process optimisation. However, business processes involve other elements not covered by scheduling problems, such as decisions, business rules, etc. that are hard to be expressed mathematically. Ernst *et al.* (2004) reports that mathematical programming formulations can only be applied when constraints and objectives can be expressed mathematically. Hence relevant approaches can be applied to simplified versions of business processes. As it is later discussed, there are optimisation approaches on mathematically formulated business processes from authors such as Hofacker and Vetschera (2001). These approaches, although consistent, are overly complicated and still deal with simplistic sequential business processes. Taking into account that scheduling is solely based on mathematical models, it is questionable whether business process optimisation should follow the same path or investigate alternative ways that express a business process using a variety of components.

Business processes and Evolutionary Computing

Evolutionary Computing (EC) techniques use the principles of evolution to guide the optimisation process and they have been successfully applied to several combinatorial problems. Genetic algorithms (GAs), for example, have already been used to find solutions to scheduling problems and their variants (Hofacker and Vetschera, 2001). Hart *et al.* (2005) present a review of applied evolutionary computing methods to scheduling problems and they report the existence of evolutionary algorithms that are capable of tackling large and hard real-world problems and are competitive with traditional techniques. There are a number of benefits in using evolutionary optimisation. One significant advantage lies in the gain of flexibility and adaptability to the task in hand, in combination with robust performance and global search characteristics (Back *et al.*, 1997). According to Moon and Seo (2005) the most attractive feature of evolutionary algorithms is the flexibility of handling various kinds of objective functions with few requirements on mathematical properties. Wang *et al.* (2004) note that process optimisation is a difficult task due to the non–linear, non–convex and often discontinuous nature of the mathematical models used.

Regarding business processes, the evolutionary approaches reported are rather limited. Hofacker and Vetschera (2001) have attempted to transform and optimise a business process model using GAs but they report non satisfactory results. The model is based on a series of mathematical formulations and is highly constrained thus making it hard for the algorithm to locate solutions. Tiwari *et al.* (2006) and Vergidis *et al.* (2006) extended their mathematical model and applied multi–objective optimisation algorithms, such as the Non-Dominated Sorting Genetic Algorithm 2 (NSGA2) and the Strength Pareto Evolutionary Algorithm 2 (SPEA2,) and report satisfactory results that provide encouraging opportunities for further investigation. These and other approaches towards business process optimisation are further discussed in the following section.

In general, evolutionary optimisation could benefit business processes by discovering process designs that are perhaps overlooked by a human designer. Also these techniques can evaluate a significant number of alternative designs based on the same process and determine the fittest based on specific objectives. Genetic algorithms could also be related with a new concept: *automated process generation*. A process design could be either generated or modified in an automated way based on different paths of execution and different objectives each time. It is a new and intriguing area of process optimisation where evolutionary techniques can significantly contribute to. The focus of this research lies in the area of business process optimisation using EC techniques.

2.5.3 Current business process optimisation approaches

Zhou and Chen (2003b) suggest that business process optimisation should aim at reducing lead–time and cost, improving quality of product, and enhancing the satisfaction of customer and personnel so that the competitive advantage of an organisation can be retained. Reijers (2002) suggests that the goals of business process optimisation are often the reduction of cost and flow time. However, Hofacker and Vetschera (2001) underline that the concept of 'optimality' of process designs is not trivial and the quality of processes is defined by many, often conflicting criteria. Both in application and theory, great importance is attached to the optimisation of business processes, mostly without explaining the criteria and the alternatives considered for optimisation (Volkner and Werners, 2000). But Zhou and Chen (2003a) remark that there is still no systematic optimisation methodology for business processes. Figure 2.5 classifies the improvement and the optimisation capabilities of business process models using the same proposed sets as used in figures 2.3 and 2.4.

As mentioned previously, optimisation is not an option for diagrammatic process models. This is because optimisation requires quantitative measures of process performance that cannot be produced in diagrammatic models. However, there are many qualitative improvement approaches applied to diagrammatic process models such as Zakarian (2001) and Phalp and Shepperd (2000). But these techniques are limited as they develop the existing diagrammatic models based on trial–and–error. Graph reduction technique is another systematic approach for business process optimisation applicable to models that have elements from both the diagrammatic and the mathematical models. Current optimisation approaches are related almost exclusively to the formal modelling techniques on the mathematical models set. This is because the formality and quantitative nature of these models allows for systematic optimisation. Quantitative criteria are considered essential in order to evaluate the improvements in a business process through modifications to the basic structure (Volkner and Werners, 2000). Business process languages set could accommodate executable models of process optimisation but to the author's knowledge there is no literature reference in this area.

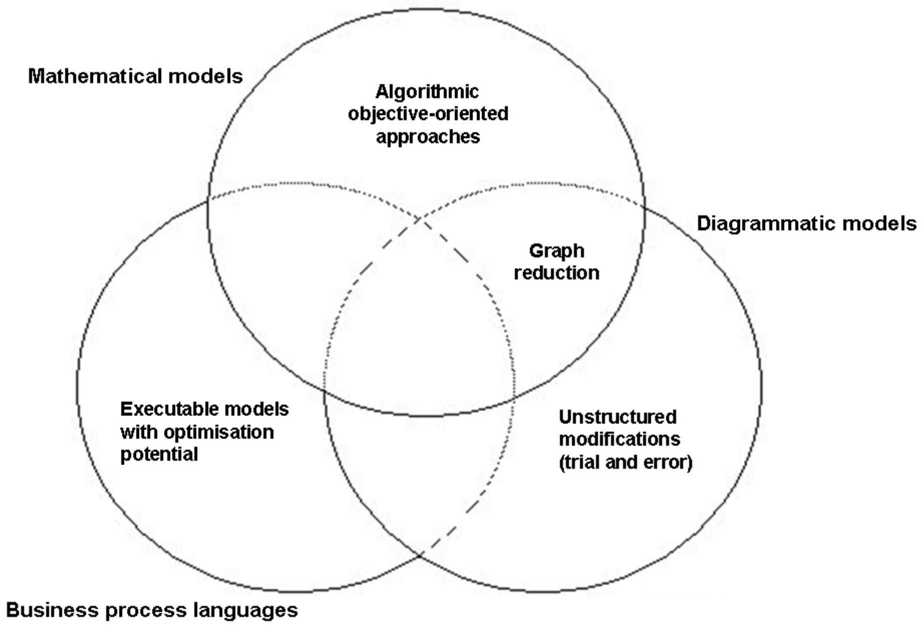


Figure 2.5. Improvement/Optimisation capabilities of the business process modelling sets

Table 2.8 summarises the main business process optimisation approaches found in literature, mostly related to Petri–nets and mathematical process models. Taking into consideration the emphasis that has been put on Petri–nets for their analysis capabilities, one would expect that they would also fit for optimisation purposes. But according to Lee (2004) Petri–nets are not adequate to solve optimisation problems except when using graph reduction techniques. Although they can capture system dynamics and physical constraints, they are not suitable for optimisation problems with combinatorial characteristics and complex precedence relations.

Li *et al.* (2004a) suggest that another way of analysing and improving a business process, is graph reduction technique for structural conflict identification or deadlock removal. Graph–reduction techniques have also engaged the attention of a series of authors. These are algorithmic techniques that modify a diagrammatic model of a process. Sadiq and Orlowska (2000) identify and try to analyse and resolve two structural conflicts in process models: deadlock and lack of synchronisation. Van der Aalst *et al.* (2002) regard the previous approach as incomplete and propose a new algorithm. A similar approach is also followed by Lin *et al.* (2002) who present a complete and minimal set of rules and a novel algorithm to implement the identification of structural conflicts in process models. In this case, the correctness and completeness of the algorithm are proved. Again, graph reduction techniques are not related with quantifiable performance measures although they have algorithmic foundation.

MODEL of business process	modelling SET(S)	TYPES of business process optimisation	APPROACHES to business process optimisation
Petri–nets (and workflows)	–Diagrammatic models –Mathematical/formal models	–Graph reduction techniques	– (Sadiq and Orlowska, 2000) – (van der Aalst *et al.*, 2002) – (Lin *et al.*, 2002)
Mathematical models	–Mathematical/formal models	–Algorithmic approaches	– (Han, 2003) – (Gutjahr *et al.*, 2000) – (Jaeger *et al.*, 1995) – (Hofacker and Vetschera, 2001) – (Soliman, 1998) – (Tiwari *et al.*, 2006) – (Vergidis *et al.*, 2006) – (Volkner and Werners, 2000) – (Zhou and Chen, 2002) – (Zhou and Chen, 2003a) – (Zhou and Chen, 2003b)
		–Activity/Task consolidation	– (Dewan *et al.*, 1998) – (Rummel *et al.*, 2005)

Table 2.8. Optimisation approaches for formal business process models

The majority of optimisation techniques are related to algorithmic approaches. Soliman (1998) provides a typical description of an optimisation problem. According to this author, business processes may be considered as a complex network of activities connected together with decision variables and an objective function subject to a number of constraints. Similar approach to the optimisation problem is proposed by Hofacker and Vetschera (2001) who provide analytical support for optimising the design of (mainly administrative) business processes. Their paper introduces formal models of the business process design problem, which can be used to analytically determine optimal designs with respect to various objective functions subject to a number of constraints. It is perceived to be the most complete paper in the area of business process optimisation because along with the formal business process model, three different optimisation techniques are examined: mathematical programming, a branch and bound method, and genetic algorithms. Tiwari *et al.* (2006) present an extension of the same formal model by applying multi–objective optimisation for business process designs and Vergidis *et al.* (2006) demonstrate the optimised alternatives. Optimisation of a business process under multiple criteria is attractive since business processes often have conflicting criteria (Hofacker and Vetschera, 2001).

Gutjahr *et al.* (2000) present a stochastic branch and bound approach for solving hard combinatorial business process related problems. Jaeger *et al.* (1995) also provide an optimisation framework based on performance evaluation that makes both resource and process changes to improve a system's performance. Han (2003) develops an algorithmic framework to design business processes using decision models. The aim of this methodology is to reduce the total cost of implementing decisions by creating a quantitative model using four design change patterns: (1) simple automation for process streamlining, (2) linear sequencing, (3) re–sequencing involving process parallelisation and (4) radical process integration that is implemented algorithmically.

Zhou and Chen ((2002), (2003a) and (2003b)) have published three papers regarding business process optimisation. Zhou and Chen (2003b) introduced the concept of assignment quality and developed multi–objective evaluation, combining optimisation models for intra– and inter–enterprise business processes; they use the NSGA to solve this problem. Zhou and Chen (2002) focus more on time, cost and resource constraints of a business process model and attempt to optimise it by utilising a genetic algorithm to minimise the process cost. Lastly, Zhou and Chen (2003a) develop a systematic design methodology for business process optimisation from strategic, tactical and operational perspectives using structured and quantitative methods that support the design. This optimisation optimally assigns resource capabilities, organisational responsibilities and authorities, and organisational decision structure.

Another approach to optimisation is the consolidation of the activities (or tasks) of a business process. Rummel *et al.* (2005) propose a model that focuses on shortening the cycle time of a business process by consolidating activities –assigning multiple activities to one actor– thereby eliminating the coordination and handoff delay between different activities. As this approach is activity (or task) focused, it ignores interactivity delay which may contribute significantly to overall process cycle time. Dewan *et al.* (1998) claim that there is no systematic methodology to determine the optimal re–bundling of information–intensive tasks. They present an approach to optimally consolidate tasks in order to reduce the overall process cycle time. The authors present a mathematical model to optimally redesign complex process networks but a limitation of the paper is that it refers to business processes with information flows only. Its main contribution is the effective business process re–structuring and the reduction of the overall task time using handoff delay reduction or elimination as a result of a unified methodology applicable to multiple task–based business processes.

Although formal languages have associated analysis techniques that can be used for investigating properties of processes (van der Aalst *et al.*, 2003), an optimisation approach based on executable process languages was not observed in literature. Since most of the optimisation approaches –as discussed above– are based on algorithmic approaches, these could be easily translated to executable software programs. Analysis and optimisation of business processes can be done best using an approach based on explicit and executable process models. Such models would allow evaluating performance in terms of flows, calculating costs against objectives, recognising constraints and evaluating the impact of internal and external events (Reyneri, 1999). The idea is that, by being able to assess the process execution quality and costs, it is possible to take actions to improve and optimise process execution (Castellanos *et al.*, 2004).

2.5.4 Summary

This section discussed the main approaches for business process optimisation and classified them based on the proposed three sets of business process modelling techniques. The following remarks summarise this section:

- The notion of business process improvement is usually limited to a broad description of steps thus does not provide the necessary insight and level of detail required.

- ⊖ Business process optimisation is the *automated* improvement of business processes using pre–specified quantitative measures of performance (objectives).
- ⊖ Business process optimisation is a difficult task due to the non–linear, non–convex and often discontinuous nature of the mathematical models used.
- ⊖ Scheduling problems are similar to business process optimisation problems. However, business processes involve other elements not covered by scheduling problems that are hard to be expressed mathematically.
- ⊖ Petri–nets are not adequate to solve optimisation problems with combinatorial characteristics and complex precedence relations.
- ⊖ The majority of business process optimisation techniques are related to algorithmic approaches.
- ⊖ Evolutionary techniques have been successfully applied to several combinatorial problems. There is a range of approaches reported in literature regarding business process optimisation using evolutionary techniques.

MODEL of business process	modelling SET(S)	TYPES of business process analysis	TYPES of business process optimisation
Flowcharts	–Diagrammatic models	–Observational	
IDEF	–Diagrammatic models	–Observational –Simulation	
RADs	–Diagrammatic models	–Observational –Performance analysis	
Petri–nets	–Diagrammatic models –Mathematical/formal models	–Observational –Validation –Verification –Performance analysis –Simulation	–Graph reduction
Mathematical models	–Mathematical/formal models	–Performance analysis –Simulation	–Algorithmic approaches –Activity/Task consolidation
Business process languages	–Business process languages	–Performance analysis (algorithmic) –Simulation	

Table 2.9. Overview of business process models, sets, analysis and optimisation types

2.6 Research gap

This chapter presented and classified the references regarding business process definition, modelling, analysis and optimisation. The review of modelling, analysis and optimisation approaches was based on a proposed classification of the types of business process models based on three sets. These classifications resulted in visually highlighting a number of interesting observations and especially the lack of certain approaches. Table 2.9 summarises the main business process models that were discussed in this paper along with their associated modelling, analysis and optimisation capabilities.

It is evident from table 2.9 that business process optimisation has not received as much attention as business process modelling and analysis techniques. Business process modelling has always attracted the attention of researches from a variety of fields. This resulted in a variety of modelling approaches that are used for business processes. Each of these diverse modelling approaches has distinctive advantages but still what is missing is a holistic approach that will involve elements from all the three sets presented in this chapter. There is a need for defining operational and reusable business process models within different types of enterprises, in different contexts and at the required level of detail. These models should be able to address the complexity of the design and identify problems encountered in modern business processes. Therefore, there is an increasing need for formal methods and techniques to support both the modelling and the analysis of business processes. However, despite the existence of many formal process modelling notations, the majority of the business process community still uses simple diagrammatic modelling techniques that have little potential for performance analysis and/or optimisation.

Table 2.9 demonstrates this gap in the lack of reported performance analysis and optimisation approaches. For most of the business process models there is no structured and repeatable improvement technique reported. In terms of process analysis, there should be a trend to focus on performance analysis as it can be directly used for decision–support and further improvement of the process. Performance evaluation needs to be integrated into the design process from the very beginning so that the objectives of the process can be rationalised from an early stage. Performance indicators are critical for the control and monitoring of a business process. The knowledge extracted from performance analysis should be fed back to the process in order to improve it. However, there are very few attempts reported in literature to combine performance evaluation and process optimisation. Regarding the latter, there are some successful attempts reported, but they are highly complicated and yet address only simple sequential business processes.

With process modelling techniques such as IDEF and Petri–nets still popular, what is missing is a modelling technique that involves elements from all the three modelling sets and thus supports analysis and optimisation. This hybrid modelling technique could (i) support a visual diagrammatic representation of the process (thus having all the advantages of visualisation), (ii) have a formal mathematical underpinning so that quantitative measures can be extracted and (iii) can be expressed using a software–based process language and thus allow optimisation extensions. Business process optimisation has a potential growth with direct benefit to the business process community and there are still a lot remaining to be done. This is why the focus of this research attempts to address the gap in business process modelling and optimisation.

2.7 Summary

This chapter examined the basic aspects regarding business process definition, modelling, analysis and optimisation. The standardisation of the business process definition can have an impact on the business process community and will boost the integration and homogenisation of the approaches towards business process modelling. A proposed compilation of three sets provided a classification for business process models based on their *mathematical*, *diagrammatic* and *language* characteristics. The advantage of this classification is that it allows a modelling technique to be positioned based on several sets simultaneously. These three sets also provided a basis for the classification of modelling, analysis and optimisation approaches.

The following remarks highlight the research gap:

- The current trend in business process modelling is the use of diagrammatic models that visualise the business process but do not provide the necessary quantitative constructs for performance analysis and optimisation.
- The proposed classification demonstrated a lack of support by most business process modelling techniques for structured process improvement
- The few business process optimisation approaches reported in literature are highly complicated and yet address only simple sequential business processes.

As mentioned in chapter 1, this research attempts to provide a contribution to the area of business process optimisation embracing the distinctive features of business processes. This chapter provided an overview of the current modelling, analysis and optimisation approaches and highlighted the lack of a holistic and formal approach towards business process optimisation. This literature survey enables the identification of the research aim and objectives in the next chapter.

CHAPTER 3

Research Aim, Objectives & Methodology

This research aims at contributing to the area of business process optimisation using formally defined business process models and existing state-of-the-art optimisation algorithms. This chapter specifies and discusses the aim and objectives of this research. Based on these, the research scope and methodology are also elaborated and discussed.

3.1 Research Aim

The aim of this research is to develop and propose a new framework for business process optimisation capable of: (i) representing business processes in a quantitative way, (ii) algorithmically composing business process designs based on specific requirements and (iii) identifying the optimal processes utilising state-of-the art evolutionary multi-objective optimisation algorithms.

3.2 Research Objectives

The research issues involved in the fulfilment of the aim are broken down into specific objectives. The research objectives, which address these issues, are:

1. To investigate and establish the state-of-the-art regarding business process modelling, analysis and optimisation.
2. To explore the industrial context of this research through a survey that identifies the main issues regarding business processes in the service industry.
3. To provide a formal specification and a representation technique for modelling business processes quantitatively so that they can be plugged to evolutionary optimisation techniques.
4. To develop an algorithmic technique that composes new business process models based on specific requirements.
5. To construct an evolutionary multi-objective optimisation framework for business processes.
6. To identify the basic features of the problem and suggest a strategy for generating tuneable business process scenarios in order to systematically evaluate the performance of the optimisation framework.

7. To validate the business process representation technique, composition algorithm and optimisation framework using a set of real-life business process scenarios.

3.3 Research Scope

Based on the objectives stated above, the scope of this research can be summarised as follows:

- *Context:* The issues and solutions proposed in this research are with regard to business processes in the service industry.
- *Domain:* The main focus of this research is business process optimisation of designs that are quantitatively represented and composed based on a proposed algorithmical approach.
- *Business processes:* The research is focusing on business processes composed of a set of discrete steps with identified inputs and outputs. Business processes that satisfy these criteria can be modelled using a basic flowchart and can be used as input to the framework.
- *Literature survey:* The literature survey in this research concentrates on business process modelling, analysis and optimisation in order to identify the recent developments in the area and establish a clear understanding about where any further contributions should be made.
- *Service industry survey:* The industry survey within the service industry focuses on the level of adoption of the business process perspective and the potential benefits or issues that this perspective raises.
- *Optimisation algorithms:* This research is focusing on existing state-of-the-art Evolutionary Multi-objective Optimisation Algorithms (EMOAs) due to their capability of handling multiple objectives, constraints and their global search characteristics.
- *Areas of customisation of optimisation algorithms:* In this research, the EMOAs are customised for handling business process designs that are composed based on a dedicated algorithm and represented using a proposed quantitative technique.
- *Areas of development of test business process designs:* This research focuses on the development of test business process designs for performing controlled and systematic investigation on designs with different features such as number of tasks and attribute values.

- *Validation:* In this research, the validation takes place using the development of real-life scenarios of business processes. The performance of the framework based on these scenarios is assessed by a group of experts in order to assess the generality and the contribution of this research

3.4 Research Strategies

According to Robson (2002), research strategies include:

- *fixed design strategies*, that require tight pre-specification before data collection; also known as *quantitative strategies* and
- *flexible design strategies*, that evolve during data collection; also known as *qualitative strategies.*

Quantitative research is often referred to as the traditional scientific research approach. It is considered as a pervasive, scientific mode of enquiry, characterised by objectivity, reliability, and prediction. Much of the data collected and used is of a numerical format. The most common form of this research approach is within laboratory settings, where the environment and experimental conditions can be closely controlled. The main strengths of the quantitative approach lie in precision and control. Control is achieved through the sampling and design; precision is achieved through quantitative and reliable measurement. The main limitation, with respect to 'real world enquiries', is that human beings are far more complex than the 'narrow' view imposed by a quantitative approach (Burns, 2000).

Qualitative research is primarily based on an investigative approach, where much of the data collected is through interviews, surveys, and observation, and is in the form of words (Robson, 2002). Qualitative researchers tend to be personally involved with their study. As a result, the research questions and design tends to 'evolve' over time as more information is collected. Sociologists, psychologists, anthropologists, and more recently business and industry, tend to use a qualitative research approach (Gummesson, 1991). The main strengths of the qualitative research approach are the insights gained from an inside view of the world under investigation and the researcher's personal involvement. This enables the researcher to derive unexpected and striking observations to examine further. The main limitations and criticisms are validity and reliability. Data collection methods are time consuming, subjective and prone to interpretation bias. The fact that the researcher is present causes bias during the collection of data. It is difficult to replicate studies; furthermore, it is difficult to make generalisations from the research findings.

There are many research strategies or methods that can be used to collect the data necessary to answer the research question. The method of research chosen depends on the nature of the enquiry. Robson (2002) presents three traditional research methods widely used and recognised: *Experiments*, *Surveys* and *Case Studies.* The characteristics of these are presented in table 3.1.

Research strategy	Main characteristics
Experimental	**Description:** Measuring the effect of manipulating one variable on another variable. **Features:** ⊖ Selection of samples of individuals from known populations ⊖ Allocation of samples to different experimental conditions ⊖ Introduction of planned change on one or more variables ⊖ Measurement on small number of variables ⊖ Control of other variables ⊖ Usually involves hypothesis testing
Surveys	**Description:** Collection of information in standardised form from groups of people. **Features:** ⊖ Selection of samples of individuals from known populations ⊖ Collection of relatively small amount of data in standardised form from each individual ⊖ Usually employs questionnaire or structured interview
Case studies	**Description:** Development of detailed, intensive knowledge about a single case, or of a small number of related cases. **Features:** ⊖ Selection of a single case or a small number of related cases of a situation, individual or group of interest or concern. ⊖ Study of the case in the context. ⊖ Collection of information via a range of data collection. ⊖ Techniques include observation, interview and documentary analysis.

Table 3.1. Characteristics of research strategies (Robson, 2002)

3.5 Research Methodology

Business process optimisation falls under the category of quantitative research (fixed design strategy). To assess the capabilities of the proposed framework and the extent that these contribute to business process research, the performance of the framework and the generated business process designs need to be *quantitatively* measured and evaluated. Fixed design strategies are theory-driven and therefore they require sound theoretical justification. A traditional approach to quantitative research is the *experimental strategy*. Experimental design and analysis is an essential part of scientific methodology; it entails the specification of the conditions in which experimental data will be observed (Greenfield, 2002). Experimental design can involve response experiments that investigate the effect of several variables at different levels. Therefore, this research requires an experimental strategy to be undertaken.

The research methodology stems from the hypothesis that:

Business process optimisation using an evolutionary multi-objective framework can produce a number of alternative optimised business process designs for a range of experimental and real-life business process scenarios.

Based on the guidelines that result from the nature of this research and the hypothesis, the main steps of the research methodology that have guided the main activities of this research are identified and depicted in figure 3.1.

Problem identification

As mentioned in chapter 1, this research is part of the 'Intelli-Process' project. The problem statement for this research is derived based on the objectives of the 'Intelli-Process' project. This research shares the vision of the 'Intelli-Process' project to progress toward intelligent methods and evolutionary techniques for tackling issues related to business process optimisation.

Literature survey

An extensive literature survey is carried out as part of this research in order to present and discuss the current state of research related to business processes. In particular, based on the primary focus on representation and optimisation of business process designs, the literature survey is carried out with respect to business process *modelling*, *analysis* and *optimisation* approaches. These issues are selected based on the business process automation trend a discussed earlier in this chapter. Once the main subjects are selected the literature research involves the investigation of books, peer reviewed journals and on-line articles in order to obtain in-depth knowledge. This assists in attaining a clear understanding of the existing work and the level of any related business process optimisation approaches along with their strengths and weaknesses.

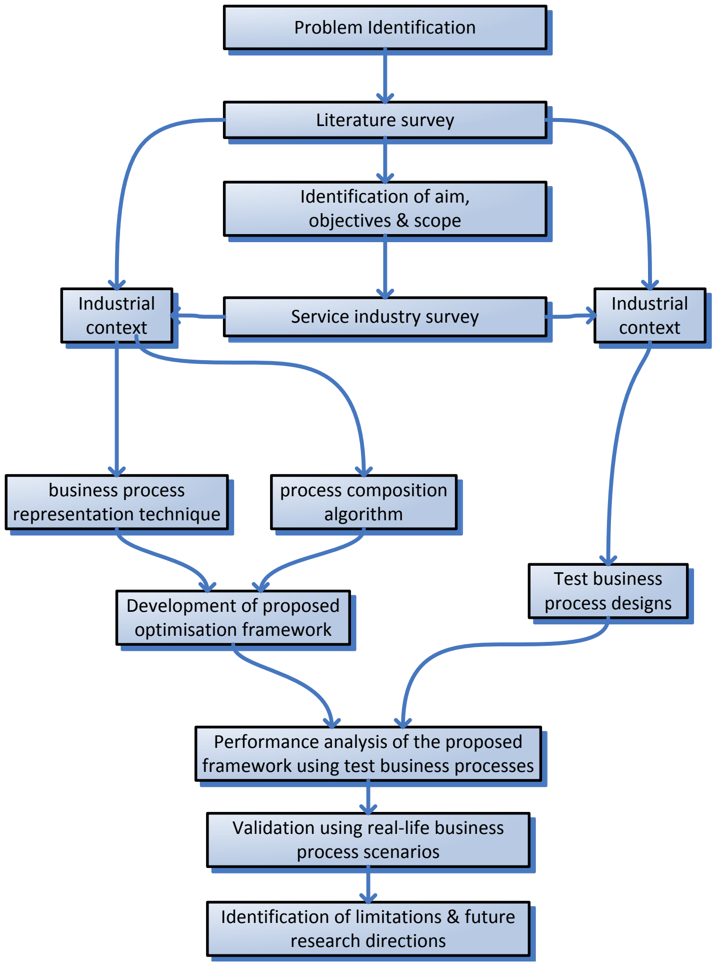


Figure 3.1. Main steps of the research methodology

Identification of aim, objectives and scope

Along with the problem statement, the literature survey provides evidence of the main research issues that need to be addressed in order to push forward the domain knowledge and provide potential solutions to persisting issues in the area of business process optimisation. This enables the precise definition of the aim and objectives that this research seeks to address. The scope provides the boundaries that the research is restrained to. The attempt to address the aim and objectives does not necessarily guarantee a complete solution to the issues that are raised by this research. However, defining the aim and objectives provides a solid guide with specified outcomes that significantly assist the course of the research.

Service industry survey

The context of this work is the service industry and a relevant survey helps in grounding the research within the industrial context. Companies that belong to the service industry are surveyed in order to investigate their business process related activities with main focus any improvement/optimisation initiatives. The survey is carried out through industry visits and on-line questionnaires for collecting information from related experts. The detailed survey methodology is discussed in the next chapter.

Development of business process representation technique

The representation technique is developed in this research to address the gap in existing business process modelling approaches for multi-objective optimisation. First, the main features of a business process than need to be captured are identified. Based on these and the optimisation focus of this research, the aim and objectives of the proposed representation are defined. Using them as a starting point, each aspect of the representation technique is developed. The technique is developed with two things in mind: (a) capturing and preserving the main elements and features of a business process design and (b) providing the capability of optimising the captured design using state-of-the-art EMOAs.

Development of the process composition algorithm

The process composition algorithm is developed in this research to address the lack of similar approaches for automated process composition based on specific process requirements. This algorithm is developed based on the proposed representation technique. It is implementing the necessary steps that create a diagram of a business process design using the proposed representation. It is also constructed in a way that it can be plugged into an optimisation framework such as the one proposed by this research.

Development of the proposed optimisation framework

The proposed optimisation framework employs existing state-of-the-art evolutionary multi-objective evolutionary algorithms (EMOAs) to achieve the optimal generation of business process designs based on specific process requirements. Due to the nature of the problem, the framework provides a customisation to these algorithms in order to encode a solution using the proposed representation technique and generate solutions using the process composition algorithm. The framework is using state-of-the-art EMOAs aiming at high quality results that deliver optimal processes.

Experimental business process scenarios & performance analysis

This research proposes a strategy for generating tuneable experimental business process scenarios for the business process optimisation problem. The reason is that it is difficult to locate real examples of business processes that possess all the elements of the proposed business process representation. Therefore, in order to assess the optimisation framework in a systematic way, it is essential to devise a strategy for generating experimental business process scenarios. The strategy is based on the features of the business process optimisation problem. Based on these features the corresponding problem parameters are identified and classified appropriately. Based on this classification, experimental scenarios of varying complexity can be generated. The performance of the framework is evaluated on experimental scenarios based on systematic and controlled variation of the identified parameters. This also helps in assessing the performance of the optimisation algorithms and identifying the strengths and weaknesses of the proposed approach.

Validation using real-life business process scenarios

Furthermore, a small set of real-life business process scenarios reported in the literature are tested within the proposed framework. These scenarios are converted to the proposed representation approach, subjected to the composition algorithm and optimised within the framework. These scenarios are adopted as indicators on whether the proposed research can have direct applicability to current business process improvement initiatives. In this way, this research proposes a fully tested and validated methodology for dealing with the representation, composition and optimisation of business processes.

Identification of limitations & future research directions

Finally, the limitations of the proposed research are identified and acknowledged. Based on these limitations, the generality of this research along with its contributions are established. Moreover, future research directions are proposed to enhance and further elaborate this research.

3.6 Summary

This chapter presented the aim of this research as the development of a business process optimisation framework capable of representing, composing and generating optimal business process designs. The aim is elaborated in specific objectives which detail the main actions of the research. Also, the research methodology is discussed in order to ensure the methodical approach that is followed. The next chapter discusses the industrial context of this research and presents the findings of an industry survey in the service sector.

CHAPTER 4

Industrial Context & Focus

This chapter grounds the research within the industrial context based on a survey within the service industry. It also determines the research focus by discussing the findings of the literature and service industry surveys. The aim, objectives and methodology that guided the survey are presented along with the main findings in the areas of business process definition, modelling, analysis and optimisation. The main remarks from these areas shape the industrial context and research focus and guide the course of action that is followed in this research.

4.1 Service industry survey

An industry survey within the service sector is carried out for grounding the research within the industrial context. This targeted survey was conducted within the service industry in order to investigate the current state of practice regarding key aspects of business processes. The survey involved the participation of 25 respondents working in service industry sectors such as finance, public sector and consultancy.

4.1.1 Aim & objectives of the survey

The aim of the service industry survey is to contrast the theoretical issues as those were discussed in chapter 2 with the real-world practical problems regarding business process definition, modelling, analysis and optimisation in order to determine the focus for this research. Examining the real-world problems can assist relevant research to recognise and address more practical problems related to business processes and can also highlight the requirements and solutions that the service industry is seeking. The survey objectives that lead to this aim are stated as follows:

- To identify the level of perception and adoption of business processes in the participating organisations,
- To establish the current industry practice in capturing and modelling a business process, and
- To investigate the existence of any quantitative analysis and optimisation initiatives.

4.1.2 Methodology of the survey

To satisfy the aim and objectives of the service industry survey, figure 4.1 shows the methodology that was followed. Each of the main steps of the methodology is briefly discussed below:

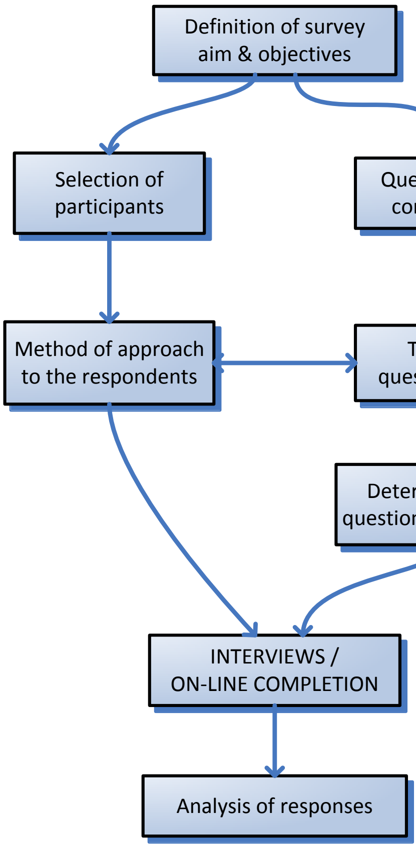


Figure 4.1. Main steps of the survey methodology

Organisation type	No. of participants
Finance & Banking	7
University & Public Sector	7
Consultancy	3
Other Service-based Organisations	8
TOTAL:	**25**

Table 4.1. Number of surveyed respondents based on organisation type

Selection of participants

The survey presented targeted service industry practitioners engaged in business process related activities in a range of service organisations. The selection was based on Cranfield University's existing list of contacts. The total number of participants was 25 from various types of service-oriented organisations. Table 4.1 presents a classification of the overall participants based on the type of the organisation. Appendix A provides further details for each participant such as his/her job role, organisation and years of experience.

Method of approach to respondents

It was decided that the participants would be approached in two ways: (a) face-to-face interview or (b) request for an on-line survey completion. This resulted in two different types of questionnaires as discussed below. In total, 5 interviews were conducted through industry visits. Prior to the visit, information regarding the research was sent to the main contact in the company. Furthermore, the interviews were preceded by a presentation, which introduced the research and explained the purpose of the visit. After guaranteeing confidentiality, the researcher interviewed the participant and wrote the responses given, during the interview. The information collected from the interview was used for preparing the final transcript. This approach was followed for the companies visited. Complementary to the visits was an on-line survey which resulted in a different type of questionnaire. The potential participants were initially invited by e-mail. Those who agreed to participate were sent the on-line survey. In total, 20 responses were received from this type of approach.

Types of questionnaires

It was decided that the best form of data capture for this survey would be provided by the questionnaires. This was due to the nature of the subject being researched and the need to be consistent and precise in the questions. The questionnaires are able to deliver a more accurate view of the overall trends in the service industry with reference to business process issues that this research is concerned with. The survey was conducted using two types of questionnaires: (a) semi-structured questionnaire for face-to-face interviews and (b) on-line questionnaire (containing mostly multiple-choice questions). Both questionnaires can be found in Appendix A.

The semi-structured questionnaire for face-to-face interviews was partly survey-based and partly fully-structured. A fully structured questionnaire has predetermined open response questions and differentiates from a survey questionnaire in which questions are more likely to be closed (Robson, 2002). In total 5 face-to-face interviews were conducted, each lasting approximately one hour and containing a mixture of open and closed questions. This included a multiple-choice section which repeated some key questions for confirmation in order to address bias. Participants from face-to-face interviews were selected from service sector based companies that deploy the concept of business processes. Results from the face-to-face interviews were analysed to develop the on-line version of the questionnaire.

The on-line version of the questionnaire was a self-completion survey. The purpose of the on-line questionnaire was to capture a wider audience and ensure generality. It allowed 20 respondents to answer a range of multiple choice and short answer questions via a web-based fill-in form. The average time of completion was 15 minutes. The on-line questionnaire, while lacking in detail in some areas, was able to deliver a quantitative view of the overall trends in the service industry with reference to business process modelling, analysis and optimisation. In this case, the reliance on the interviewee to interpret the questions correctly and provide responses made it even more important to trial the questions. Both questionnaires were piloted by three people from a service-based organisation.

Determination of questionnaire sections

Oppenheim (1992) puts forward the concept of dividing the questionnaire into modules with each module concentrating on one concept or variable. This notion has been incorporated into the design of the questionnaire used in this research, breaking it down into three main sections:

1. *Business process basics*, that investigates the understanding of the participants about the notion of business processes,

2. *Modelling techniques*, that seeks to capture the business process modelling techniques used in industry, and

3. *Analysis and improvement*, that focuses on quantitative analysis approaches and improvement initiatives related to business processes.

Both the interview and online questionnaires were structured in a similar way and contained these three sections. These sections illustrate the areas of business processes that this work is focusing on. For each of these sections, the theoretical developments based on the literature research findings are summarised and contrasted with the service industry survey results.

Analysis of responses

For each of the questions in the questionnaire, the responses given by the respondents were compiled and analysed. Those that revealed significant similarities or differences with the literature survey are discussed in the following sections.

4.2 The service industry perception of business processes

The first issue regarding business processes is the understanding about the concept itself and the benefits it can bring to an organisation. This section demonstrates in service industry there is a generic –and sometimes vague– understanding about business processes. This is one of the main reasons for the diversity of approaches that seek to address issues vital to the elaboration of the business process perspective.

As discussed in chapter 2, the main issue with business process definitions is twofold: either they are too simplistic and basic thus too generic to provide any tangible contribution or they are confined to a very specific application area that prevents them from wide acceptance and applicability. For this reason, the survey respondents were first asked about their perception of business processes. The motive was not only to ask for a definition but also investigate the structure of their organisation, the flow of the business processes and the use of any business process related software tools.

Based on the answers provided, the understanding of business processes is characterised by three distinctive points of view: (i) as structured processes similar to production processes, (ii) as methodologies to achieve a business goal and (iii) as complex sociotechnical constructs that involve elements from both the above categories but with more emphasis on human interactions and relationships. The participants were asked to rate the above three approaches. Their ratings focused on the structured process view (with clearly identified inputs and outputs) and the view of process as enactment of actors aiming to achieve a business goal (softer and unstructured perspective). The definition regarding the sociotechnical perspective received little coverage. 21 of the respondents (84%) feel more comfortable in dealing with business processes in a structured way rather than a softer approach that involves non-quantifiable business goals from strategic level. Service industry practitioners want a clear and concise view of the business processes within the organisation and a solid understanding about their flow rather than discussing about the social interactions and effects that the process triggers. The need for rationalising business processes is one of the major drivers for business process modelling and the reason for the abundance of existing modelling techniques.

For the business processes treated as structured processes within the organisations, the issues regarding the control of the process flow and the organisational structure from a business process perspective were discussed. Business process pioneers envisioned business-process-centric organisations, where all the resources are organised around an organisation's business processes. This would remove the necessity of reinforcing an explicit process flow across the different departments because the departments would function *a priori* in a business process oriented fashion. However, this survey suggests that this is far from being realised in the service industry where the traditional departmental segmentation dominates the organisational structure.

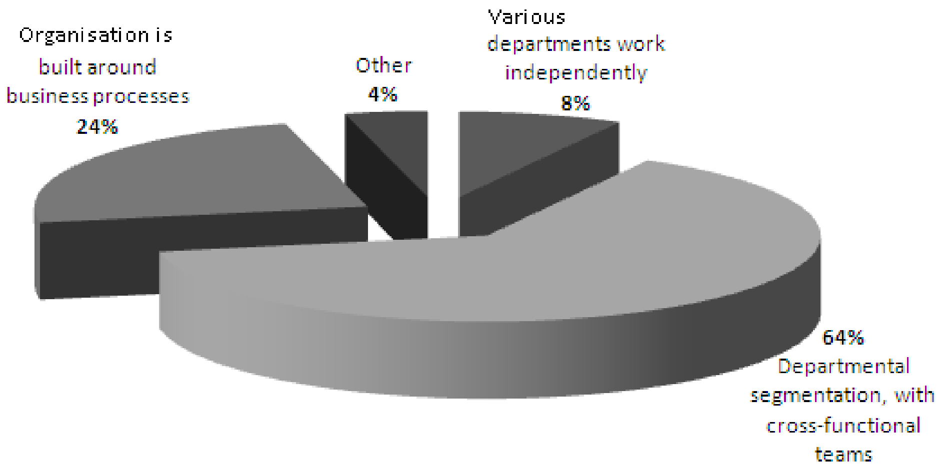


Figure 4.2. Business processes and organisational structure

Figure 4.2 demonstrates the responses of the 25 participants regarding the organisational structure. 64% provide a recognisable common practice regarding business processes. This perspective, although preserves the traditional departmental segmentation, recognises the need for cross-departmental co-ordination and co-operation for processes to be effectively enacted. 24% responded that their organisation has moved away from traditional structures and operates around the main business processes.

In the majority of the cases when the organisation is not built around business processes, controlling the process flow becomes a crucial issue for uninterrupted and immaculate process enactment. In order to control the process one has to own and manage it. The results shown in Figure 4.3 communicate the responses about the current practice of process ownership and process flow knowledge within an organisation. It is encouraging that 44% of the respondents (11) claim that their organisation appoints specific process owners responsible for each business process. Central co-ordination and understanding of the complete process is essential in order to manage its enactment efficiently. Another 32% state that the process knowledge is shared among the main participants of the process. This usually results in the lack of concrete understanding and central co-ordination in business process enactment. There is also an 8% (2 participants) who responded that no-one has explicit knowledge about the complete process flow. This percentage can be easily matched with the 8% of the previous question which stated that the various departments work independently in the organisations (see figure 4.2). Departmental segmentation without established interdepartmental communication results in isolation of the different operations and this can have detrimental effects on the processes at stake.

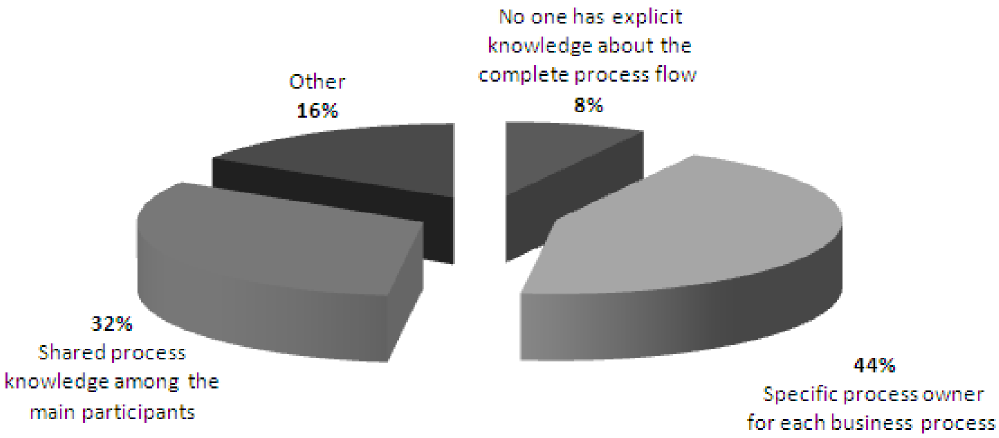


Figure 4.3. Business process ownership within organisations

Business process automation (BPA) is one of the main trends regarding business processes and involves the automation (part or complete) of the process using software enactment tools. Automating a business process gives the opportunity to collect real execution data continuously from which information about the process performance can be obtained and can be used for monitoring, work balancing and decision support (Abate *et al.*, 2002). The benefits of such automation are that the processes can be executed faster, with lower costs (due to the reduced human involvement), and in a controlled way, since the enactment system can detect exceptions or delays in process executions (Castellanos *et al.*, 2004). For the organisations that participated in the survey, there is at least a 50:50 split between automated and manual processes. The overall percentage of automated processes was identified as 66% for banking and finance organisations which was the highest of any sector. The lowest level of automation was found in the university and public sector with only 30% of processes automated, followed by the consultancy sector at 33%. It is important to note that, as more and more processes become automated, the focus of both service industry and academia shifts from deployment to process analysis and optimisation (Castellanos *et al.*, 2004).

The last question related with business process basics examined if the organisations that participated in the survey make use of the capabilities of the contemporary business process management tools (e.g. ARIS, SAP, Tibco). Over half the respondents stated that their organisation uses business process management software, a fact that is encouraging for the development and elaboration of business process oriented software tools. Among the software packages named by the respondents, SAP NetWeaver and various components from ARIS Platform were the most commonly implemented solutions. Other organisations reported customised tools where the process management elements consisted of custom built code integrated with corporate databases or existing process/task specific software. Customised tools aim to address critical software integration issues between specific software solutions that have been implemented in the organisation. The other half of the respondents claimed that either their organisation does not use any software or they are not aware of a business process suite being used within their company. These findings indicate that the market for business process management software is growing and it is finding its way to the corporate environment. Based on the survey answers, the requirements from a business process suite are: visual editor for process modelling (requested by 28% of the respondents), customised specification of Key Performance Indicators (KPIs) (16%), support for process simulation (12%), and generation of optimised process models (8%). Other requirements include: process execution/enactment capability, different process detail levels (i.e. hierarchical processes) and capability to model services.

4.3 Business process modelling

After investigating the basic notions of business processes, the second major part of the survey sought to identify the service industry practices regarding business process modelling. Business process modelling is concerned with depicting and representing adequately a business process emphasising its aspects that need to be communicated and dealt with; it is extensively discussed in literature (see chapter 2) but the service industry practice seems to have fallen behind as the survey suggests.

Business process modelling is a useful tool to capture, structure and formalise the knowledge about business processes. However, there is an abundance of business process modelling techniques that capture different aspects of a business process with some being better suited depending on their particular constructs. Given the abundance of process modelling techniques, this part of the survey sought to identify whether there is a common practice in capturing and depicting business processes within service organisations. The participants were provided with a list of the most common business process modelling techniques identified from the relevant literature research: Flowcharts, IDEF models, Petri-nets and documentation (textual description of the business process). The participants were asked to identify and rank how frequently each of these techniques is used within their organisation

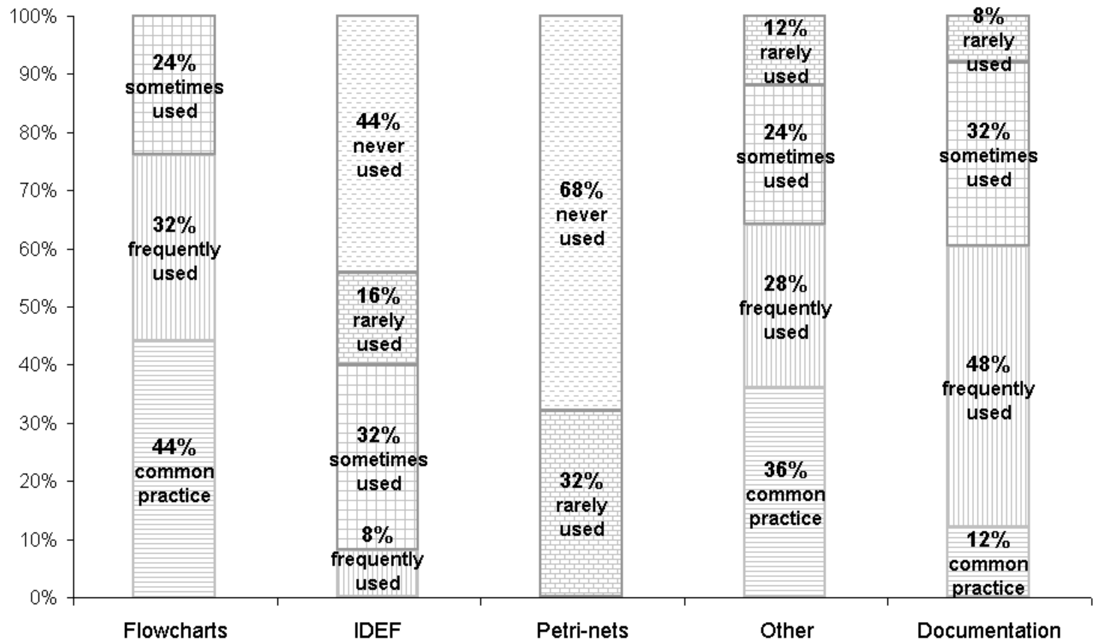


Figure 4.4. Business process modelling techniques used within the service industry

The results, demonstrated in figure 4.4, show that the majority of participants (11) use basic flowcharts with informal notation. A vast 44% responded that flowcharts are 'common practice', 32% responded 'frequently used' and 24% responded 'sometimes used'. IDEF models are 'sometimes' used by 32% while 60% 'rarely' or 'never used' this modelling technique for business process modelling, although relevant literature seems to strongly favour it. For Petri-nets –that are strongly supported by authors related with workflows– 31% responded that they are 'rarely used' while the vast majority (68%) have not used them at all. Process documentation is a common practice for 12% while another 78% claim that it is 'frequently' or 'sometimes' used. While process documentation is ignored in business process literature –as it can be argued whether it is a modelling technique– it is well established as a technique for describing and detailing business processes. Another 36% of the respondents (9) claim to use other modelling techniques for business process modelling without providing further details.

Discussions about business process modelling often involve the patterns that are observed and supported by the various business process models. A modelling pattern is 'the abstraction from a concrete form which keeps recurring in specific non-arbitrary contexts' (Riehle and Zuillinghoven, (1996). A process modelling pattern is a cluster, or a constellation of process activities arranged in just the right way to solve a difficult problem. The survey participants were asked to select from a predefined list one or more of the patterns that appear in their business process models. The business process patterns that were listed were: (i) sequential flow, (i) parallel flow (AND), (iii) decision (OR) and (iv) feedback loops (LOOP). These were considered to be the most profound, based on the list of patterns that Havey (2005) cites. All these business process patterns were recognised by the vast majority of the respondents (22) as frequently occurring in their organisational business processes. Some participants reported other process patterns that they have encountered such as asynchronous flow, event- and rule-driven flow. The majority of participants stressed that business processes have complex constructs that need to be taken into account when modelling.

4.4 Business process analysis and improvement

The final section of the survey dealt with process analysis and improvement. These are two essential aspects of business processes that can justify their potential for optimisation. Business process analysis is a term used with a broad meaning including simulation and diagnosis, verification and performance analysis of business processes. Process improvement can also occur through formal techniques. However, as chapter 2 discussed, the current state of analysis of business processes often consists only of simple inspection of process diagrams. A holistic approach towards business processes should capture a business process (business process modelling), provide the necessary means for bottleneck identification and performance analysis and –finally– generate alternative improved business process(es) based on specified objectives.

Business process analysis can significantly contribute to the organisations by locating any bottlenecks in their processes. Analysis can be carried out in a qualitative or quantitative fashion. The service industry is focused on quantitative analysis as only this can result in measurable business process improvement (i.e. optimisation). 80% of survey participants (20) responded that quantitative analysis of business processes occurs in their organisation and it is performed using the concept of Key Performance Indicators (KPIs). The KPIs are measurable factors that assess either direct results of business processes or aspects that are directly affected. 36% of the respondents (9) use simulation packages to simulate and extract quantifiable results from business process executions. Simulation packages allow the definition of various KPIs but the results are not of the same value as the ones based on real process execution data. Other approaches to quantitative business process analysis involve manual processing, customer feedback and measurements provided by business process management suites. Since KPIs are widely used within service industry, the most widely adopted KPIs in terms of analysing business processes were identified from the participants. The main responses are grouped and summarised in Table 4.2, along with the percentage of respondents who reported using the same KPI.

Business Process KPIs	% of Respondents (no. of Respondents)
Lead time/cycle time	20% (5)
Balanced scorecard	16% (4)
Client acceptance/appreciation	12% (3)
Process cost	8% (2)
KPIs customised for each particular process	8% (2)
KPIs mapped to strategic/business goals	8% (2)
Benchmarking	4% (1)
Profitability	4% (1)
Financial/stock measures	4% (1)

Table 4.2. Responses about the most widely used KPIs

Five of the survey participants indicated that the time it takes to complete and produce or satisfy the business outcome (lead time) is an established measure of a business process. The use of balanced scorecard to measure/evaluate processes is common among 4 participants. The scorecard contains a range of KPIs such as the ones cited in table 4.2. Three participants evaluate business process performance based on client acceptance and appreciation. This occurs by filling customer satisfaction surveys that are subject to quantification and further analysis. Process cost has also been mentioned as an important factor for evaluating the business processes by 2 participants. Apart from specific KPIs, two different generic perspectives were mentioned by 4 respondents. Two of them use process-customised KPIs while the other 2 use KPIs that are mapped to strategic and business goals thus relating the processes directly to the corporate strategies. Other KPIs (or approaches to process measurement) reported were: benchmarking, profitability and other financial/stock measures. One respondent also reported that the findings of the process evaluation are fed back into the execution of the process. In relation to business process optimisation, the various KPIs can be considered as potential optimisation objectives.

Business Process Improvement Techniques	% of Respondents (no. of Respondents)
Six Sigma	16% (4)
Software assisted – ARIS Platform, PRISM, Intelicorp	12% (3)
Corporate internal methodologies/projects	12% (3)
Lean techniques	8% (2)
Total Quality Management – TQM	8% (2)
Observational analysis/process reuse	4% (1)

Table 4.3. Responses about business process improvement techniques

Regarding business process improvement/optimisation approaches, a dominating 56% of the respondents (14) stated that there is no improvement initiative for the organisation's business processes. The remaining 44% reported process improvement techniques that are grouped and summarised in table 4.3. The most popular approach is Six Sigma as stated by 4 participants. Three stated that their processes are improved using software tools, such as ARIS Platform, PRISM, Intelicorp, and another three claimed that there are internal projects or improvement methodologies within their organisations, without providing further details. Lean techniques and Total Quality Management (TQM) were reported as process improvement initiatives by 2 participants. Another 2 reported observational analysis and process reuse as part of improving the organisation's business processes. It is important to note that none of the improvement approaches cited in table 4.3 is exclusive to business processes. Most of the techniques are borrowed from management or manufacturing related disciplines. The lack of a consistent optimisation technique created exclusively and customised for business processes is evident based on the results.

The idea of developing a dedicated optimisation technique for business processes appealed to the vast majority of the participants. They were also asked to rank the importance of four different factors about the development of such an optimisation technique. The most important factor was 'resource allocation' as 56% of the respondents ranked this factor as 'very important' for a business process optimisation approach. Second with 40% responses is the 'activities reduction/consolidation' element and third with 32% is 'company policy/rules'. Finally, 24% of the respondents would take into account 'external environment/competitors' in a business process optimisation framework.

4.5 Industrial context of the research

The service industry survey enables the identification of the current practice of aspects related to business processes. Based on the three main sections of the service industry survey, the main observations reveal the industrial context of the research and along with the literature survey shape the research focus as discussed in the next section. The industrial context of this research can be summarised as:

- Both the academic researchers and service industry practitioners feel more confident in dealing with structured and defined business processes. This is justified by the fact that a structured process with expected (or predefined) inputs and outputs is subject to quantification and measurable evaluation.
- The majority of organisations are still operating under the traditional departmental structure. This results in a continuous challenge for effective business process enactment.
- The lack of concrete process management results in vague understanding about the process, its main elements and its flow. Business processes without explicit ownership and management become fragmented within the various departments and their scope and outcomes become unclear.
- Investigation of business process modelling both in literature and in service industry proved that simple diagrammatic techniques such as flowcharts still dominate the area. This reflects the need for a simple, communicative and effective illustration of business processes.
- In literature there are plenty of advanced modelling techniques and methods for business processes. Advanced –and perhaps more complex– modelling methods do not guarantee a more formal and structured approach towards business processes; they might even discourage the industry practitioners.

- Process analysis is still largely perceived as the manual inspection of diagrams. Due to the qualitative nature of business process modelling, quantitative analysis and process evaluation are hard to apply.
- Manual or qualitative analysis approaches, such as diagram inspection, overshadow techniques that can be used for performance analysis to aid process improvement initiatives.
- The survey participants have a clear focus on the quantitative KPIs they would like business processes to be evaluated with. But this quantitative evaluation currently takes place only in a small number of the participating organisations.
- Business process improvement in an automated fashion is perhaps the most attractive potential that can grant a valuable advantage to business processes and secure them with a new direction for future development.
- The majority (58%) of respondents are not using a structured methodology for improving their business processes; thus there is a large gap and a potential for a methodology for automated improvement (process optimisation) based on a standard process model.
- Functionalities such as business process analysis and optimisation are largely lacking in most commercial software systems available in the market today. A business process suite could be potentially developed addressing the above elements in a holistic way with the aim of providing a truly beneficial solution for the requirements of business processes.

4.6 Research focus

Chapter 2 discussed that theoretical research is dealing with sophisticated issues around business processes. However, the service industry survey demonstrated that the service industry is reluctant to adopt a similar perspective and still uses simple and manual techniques in dealing with business processes. The main reason is that the service industry is not convinced that a business process approach could bring significant tangible and measurable benefits. This is due to the fact that as of today there is no comprehensive and systematic solution in terms of a fully functional business process framework.

The literature and service industry surveys provided a comprehensive view of the main issues related to business processes, i.e. definition, modelling, analysis and optimisation. As mentioned in chapter 3, this research primarily aims at developing a business process optimisation framework. Along with the primary aim, there are some prerequisites that will also be addressed by this research in accordance with the issues highlighted above. The research focus entails the aspects summarised below:

- Business processes will be defined and their main elements will be specified in accordance with the context of this research as denoted by the literature and industry surveys (chapter 5)
- The proposed quantitative representation approach will use as input a simple business process model as indicated by the service industry survey. Both the model and the representation technique will be fully specified in terms of structural elements based on the business process schema that was presented in the literature survey (chapter 5).
- A dedicated composition algorithm will be presented and elaborated in chapter 5. This algorithm will compose alternative business process designs based on specific requirements. Although the concept was not explicitly discussed in the industry survey, it is a necessary step for creating optimised alternative business process designs.
- The process indicators will not be explicitly investigated by this work. The optimisation framework is kept generic and not oriented towards specific targets (e.g. cost reduction). Therefore, there will be no selection or preference towards particular process indicators as optimisation objectives. This will occur only later, in the validation of the framework with specific business process scenarios from the service industry.
- Finally, this research proposes an optimisation framework for business processes pushing the existing boundaries of business process improvement initiatives as reported in the literature and service industry surveys. The optimisation framework is capable of generating optimised alternative business process designs based on multi-criteria evaluation using state-of-the-art evolutionary algorithms (chapter 6).

4.7 Summary

This chapter presented a service industry survey in order to determine the industrial context and focus of this research. The survey investigated the current state of business processes within the service industry. It highlighted that business processes still need to demonstrate clear and tangible benefits in order to gain wider acceptance. Researchers have attempted to develop sophisticated techniques for tackling business process issues but the service industry uses basic and mostly manual techniques for dealing with issues such as process modelling and analysis. The potential advantages of business process optimisation have neither been clearly demonstrated as of yet nor been supported by existing software solutions. The survey highlighted the need for a framework that adequately and efficiently addresses the issues related with business process specification, modelling and optimisation. Chapter 5 addresses the issues of specification and modelling by proposing a definition and a formal representation for business processes.

CHAPTER 5

Proposed Business Process Representation

This chapter introduces a representation for business processes. Before the representation is presented, the concept of 'business process' is specified within the context of this research. Business process specification is considered as necessary due to the variety of definitions and approaches identified in the literature and service industry surveys. The proposed representation is then presented in detail. The aim of the representation, in accordance with the aim of this research, is to make business process designs amenable to evolutionary multi-objective optimisation techniques. In this chapter, the various aspects of the proposed representation such as the process diagram and the mathematical parameters are elaborated and discussed in detail. The proposed representation also encompasses an algorithmic procedure for ensuring the correctness of the business process design. This algorithm is an essential part of the representation and is also detailed in this chapter.

5.1 Specification of business processes

This section specifies the notion of *'business process'* within the context of this research. Chapter 2 examined the various definitions in literature and chapter 4 surveyed the different perceptions in the service industry. Based on these, the author proposes a definition and specification that is straight-forward and grounds business processes within the context of this research. The purpose of the specification is to provide a definition of business processes, identify the domain and their features, and finally specify the business process elements that are considered essential by this research. The sub-sections below detail the various definitions and specifications of the business process as defined by the author within the context of this research based on the literature and industry survey.

5.1.1 Definitions

In this research, the definition of business process is as follows:

A **business process** is perceived as a collective set of tasks that when properly connected and sequenced perform a business operation. The aim of a business process is to perform a business operation, i.e. any service-related operation that produces value to the organisation.

The *design* of a business process is the means of communicating it, hence defined as:

A **business process design** is the representation of a business process depicting the participating tasks and their connectivity patterns that determine the flow of the process. The aim of the design is to capture, visualise and communicate a business process.

This chapter proposes a *representation technique* as a means to construct and communicate business process designs. As such it is defined as:

A **representation technique** provides the means to construct a business process design. A design representation can be visual (to communicate the design as a diagram) and/or quantitative (to communicate the design in a way amenable to quantitative analytical methods). The elements involved in the representation define the capabilities of the design in terms of clarity, accuracy and lack of ambiguity.

5.1.2 Domain / Context

This research focuses on business processes in the service industry. This means that the business process itself is considered as a service and its outcomes are non-material equivalents of goods based on the service definition. Examples of such business process involve order processing and fault/complaint handling.

5.1.3 Features

In terms of features, a business process is:

- *More than customer oriented.* Unlike Hammer and Champy's (1993) customer oriented definition of business processes, the proposed specification covers back-end or internal business processes. It provides a general perspective for all the value-adding business operations performed within an organisation.

- *Service/functionality oriented.* The proposed specification examines business processes from a perspective of the functionalities that are involved –not the steps that need to be executed. This perspective emphasises more on the flow and connectivity of the participating functionalities rather than on execution details.

- *Hierarchical structure.* Having this perspective of identifying the main functionalities included in the business process means that a strategic process and an operational process can be similarly perceived. Therefore, a functionality identified in a strategic level can itself be a business process at a lower level. The proposed specification allows for hierarchical structuring of the organisation's business processes.

5.1.4 Elements

The elements that are involved in the business process and consequently represented in the business process design are based on the business process schema (figure 2.1) that was presented in chapter 2. These are:

1. The participating tasks,
2. The resources of a task / business process,
3. The attributes of a task / business process and
4. The connectivity patterns.

Essentially, the main elements involved are the *tasks* and *resources* of the business process. As a direct result from the optimisation focus of this research, the *attributes* of the tasks and the process are also taken into consideration in order to provide the capability of evaluating a business process design. Finally, the *patterns* that interconnect the tasks are also included, as they are identified by this research as one of the key characteristics that distinguish business processes. The proposed representation –presented later in this chapter– provides constructs for each of the identified process elements. Each of these elements is further discussed.

Participating tasks

The participating tasks are the main elements of the process: each task represents a specific functionality. The tasks are considered as 'black-box' functionalities, that are joined together to utilise the aim of the process (i.e. perform the business operation). They are similar in nature and characteristics but different in terms of the core operation they perform.

Resources of a task / business process

The resources are considered as the input and output products of the participating tasks and the business process. The resources are the process elements that flow and are transformed through the process to produce the final outcome. This research does not assume any specific nature or type for the resources. The resources of the participating tasks are considered as either *task input* or *task output* resources. They connect the tasks based on common inputs and outputs and help in shaping the connectivity patterns that occur in the process design. The resources of the business process are considered as *process requirements* and are classified as the required *process input* and expected *process output* resources.

Attributes of a task / business process

The attributes are considered as the measurable (quantitative) characteristics of the participating tasks and the business process. It is assumed that the attributes are common across the participating tasks (e.g. task cost) and can be mapped to the corresponding process attributes (e.g. process cost) using a suitable aggregation function. The attributes are used to evaluate a business process design and consequently compare it with others and optimise it. Examples of attributes involve task/process cost, duration and reliability.

Connectivity patterns

The importance of connectivity patterns and the need to support them when representing a business process was discussed in chapter 2. The patterns are constructs that help in expressing recurring paths in a process and are largely responsible for shaping the process design. Involving the patterns in the proposed representation approach is important, as patterns are able to consider several complex dependencies between the tasks (Scheer, 1994). Authors such as Kiepuszewski *et al.* (2003), van der Aalst and ter Hofstede (2002) and Zhou and Chen (2002) refer to:

- Sequence,
- Parallel execution (AND),
- Multi-choice (OR), and
- Arbitrary loops (GOTO),

as the basic patterns for modelling and controlling a business process. These patterns are identified by this research as essential to be involved in the business process specification and proposed representation approach.

5.2 Proposed representation approach

This section introduces the proposed representation approach. It starts by stating the aim and objectives that the approach needs to satisfy. Based on these, the proposed representation consists of two perspectives: visual and quantitative. The visual representation communicates the business process design as a diagram and the quantitative representation captures the design in a way amenable to analytical methods. Both perspectives aim to capture and represent the business process elements identified in the previous section.

5.2.1 Aim and objectives of the proposed approach

The aim and objectives of the proposed representation approach are as follows:

Aim of the representation

The aim of the proposed representation approach is to capture, visualise and express a business process design in a quantitative way that allows Evolutionary Multi-objective Optimisation Algorithms (EMOAs) to generate a series of alternative optimised designs.

Objectives that the representation needs to achieve

To achieve the aim, the proposed approach needs to:

1. Provide a visual communication of the business process design,
2. Express the elements of the process using mathematical parameters,
3. Provide quantitative means to evaluate the business process design,
4. Form a basis for generation of alternative designs based on existing ones.

5.2.2 Visual representation of a business process design

Based on the literature and service industry surveys, the visual representation of a business process design is based on the principles of simple flowchart. The reason is that flowchart receives wide recognition and familiarity across researchers and industry practitioners related to business processes. The flowchart can accommodate visually all the process elements that were identified in sub-section 5.1.4, except the task attributes, and can communicate the basic flow of the business process design. The proposed notation for a flowchart depicting a business process design is:

- Two rounded boxes marked as 'START' and 'END' appear in every design and denote the beginning and the end of the process.
- The participating *tasks* are sketched as boxes.
- The *resources* are the connecting arrows that link the tasks
- The *patterns* are depicted as follows:
    - Sequence is sketched as the connecting arrow between two tasks

- Parallel flow (AND) is sketched as box
- Multi-choice (OR) is sketched as rhombus
- Arbitrary loops (GOTO) are sketched as arrows pointing backwards

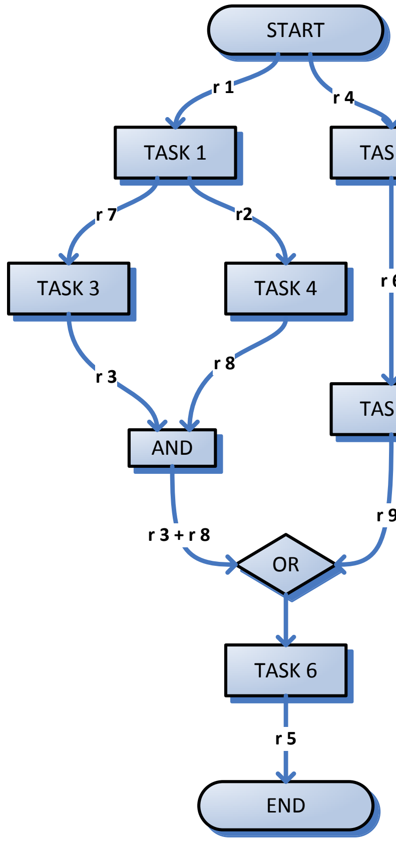


Figure 5.1. Example of the visual representation of a generic business process design

Figure 5.1 shows an example flowchart for a generic business process design based on the proposed notation. The visual demonstration of a business process design is an essential part of the proposed representation. It provides the means of communicating the framework results (optimised business process designs) to the stakeholders (e.g. business analysts) by showing the main steps of the process design and the way they are interconnected. The next challenge is the conversion of the visual representation to the equivalent quantitative representation using mathematical constructs.

5.2.3 Mathematical parameters of the business process elements

This sub-section shows the mathematical parameters of the main process elements as identified in section 5.1.4. Expressing these concepts using mathematical notation assists in the introduction of more complex constructs that can represent a business process design in a quantitative way. Table 5.1 shows the encoding of the main process parameters.

Parameter	Description	Parameter	Description
n_d	Number of tasks in the design	N_d	Set of the n_d tasks
r_d	No. of resources in the design	R_d	Set the r_d resources
t_{in}	No. of task input resources	I_i	Set of the t_{in} resources for a task *i*
t_{out}	No. of task output resources	O_i	Set of the t_{out} resources for a task *i*
r_{in}	No. of process input resources	R_{in}	Set of the r_{in} resources
r_{out}	No. of process output resources	R_{out}	Set of the r_{out} resources
p	No. of task/process attributes	TA_i	Set of the task attribute values for a task *i*
		PA	Set of the *p* process attribute values

Table 5.1. Main process parameters

The set of n_d tasks that belong to a particular process design is $N_d = \{t_1, t_2, t_3, \ldots, t_{nd}\}$. The set of r_d resources in the design $R_d = \{r_1, r_2, r_3, \ldots, r_{rd}\}$ accommodates the subsets R_{in} and R_{out} that store the process input resources and process output resources respectively. The business process design utilises all the resources in R_{in} and produces all the resources in R_{out}. Also, each task *i* in the design has t_{in} input resources stored in $I_i \subseteq R_d$ and t_{out} output resources stored in $O_i \subseteq R_d$. Finally, each task *i* has *p* attribute values stored in the TA_i set and the corresponding *p* process attributes are stored in the PA set.

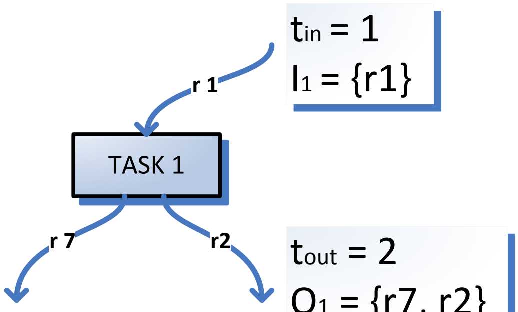


Figure 5.2. Mathematical parameters and visual representation of a task

Figure 5.2 focuses on 'TASK 1' from figure 5.1 and shows how the parameters from table 5.1 relate to its visual illustration. 'TASK 1' has one input resource {r1} and two output resources {r7, r2}. The only task related elements that are not visualised are the task attribute values. Figure 5.3 shows the business process design in figure 5.1 related to the parameters of table 5.1. The resources that flow from the 'START' node are the process input requirements and the resources that conclude to the 'END' node are the process output requirements. The process design has 6 participating tasks and 9 different resources that flow though the tasks. The process attribute values that are calculated from the participating task attribute values are not depicted in the visual representation of the business process design.

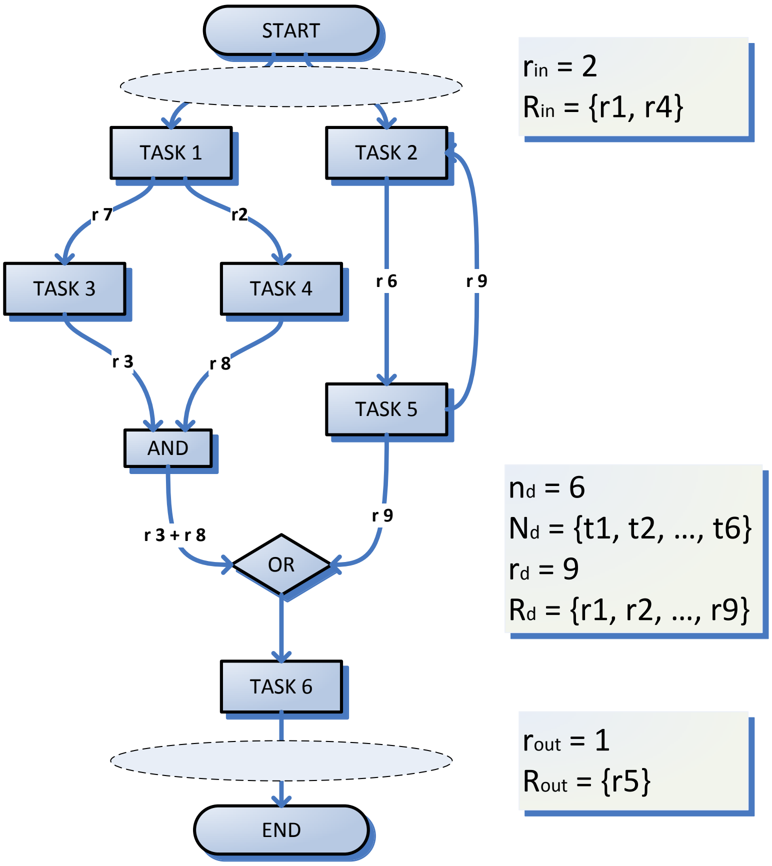


Figure 5.3. Mathematical parameters and visual representation of a process design

5.2.4 Task Attributes Matrix (TAM)

Having defined the initial parameters for the main elements of the business process, more complex constructs can be elaborated. The first construct is a matrix that aims at capturing the *attribute values of the tasks* in the design thus helping in calculating the process attribute values and evaluating the design. It is a two-dimensional matrix called *Task Attributes Matrix* (TAM) with dimensions $n_d \times p$. The rows in TAM accommodate the n_d tasks in the process design and the columns accommodate the values of the p attributes per task. TAM contributes to the calculation of the process attribute values for a specific process design assuming that a process attribute is calculated based on the participating task attribute values. Using TAM, a process attribute *j* (PA_j) can be calculated as an aggregate of the corresponding task attributes for all the n_d tasks in the process design:

$$PA_j = \sum_{i=1}^{n_d} TAM_{ij}$$ (Equation 5.1)

Table 5.2 provides an example of TAM for the generic business design in figure 5.1 assuming two attributes (A_1 and A_2). The process attributes are calculated based on the aggregation equation 5.1. TAM does not store the process attributes but only the task attribute values.

Tasks \ Attributes	A_1	A_2
Task 1	100	300
Task 2	120	302
Task 3	117	324
Task 4	178	308
Task 5	145	356
Task 6	157	389
PROCESS	*817*	*1979*

Table 5.2. Example of Task Attributes Matrix (TAM) and process attributes calculation

5.2.5 Task Resources Matrix (TRM)

The second construct –similar to TAM– is also a matrix that aims at capturing the *task sequencing* and the *patterns* formulated in the process design. To achieve this, the matrix maps the input and the output resources of the tasks in the process design. This matrix is called *Task Resources Matrix (TRM)* and it is two-dimensional matrix with size $n_d \times r_d$. The rows in TRM accommodate the n_d tasks in the process design and the columns accommodate the r_d resources in the design. Each cell in TRM shows the relationship between the task and the resource. For a task $i \in N_d$ and a resource $j \in R_d$:

- If $r_j \in I_i$ then $TRM_{ij} = 1$
  (If the resource belongs to the set of input resources of the task then their relationship is flagged as '1')

- If $r_j \in O_i$ then $TRM_{ij} = 2$
  (If the resource belongs to the set of output resources of the task then their relationship is flagged as '2')

- If $r_j \notin I_i$ and $r_j \notin O_i$ then $TRM_{ij} = 0$
  (If the resource is neither in the input nor in the output resources of the task then their relationship is flagged as '0')

Figure 5.4 shows the TRM mapping for 'TASK 1' (figure 5.2) and the 9 resources in the business process design (figure 5.3). Table 5.3 demonstrates the TRM for all the tasks in the process design based on the proposed mapping of the relationship between the tasks and resources.

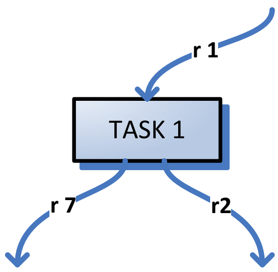
Visual representation	TRM mapping
	**$TRM_{11} = 1$** **$TRM_{12} = 2$** $TRM_{13} = 0$ $TRM_{14} = 0$ $TRM_{15} = 0$ $TRM_{16} = 0$ **$TRM_{17} = 2$** $TRM_{18} = 0$ $TRM_{19} = 0$

Figure 5.4. Example of TRM mapping based on 'TASK 1'

Resources / Tasks	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9
Task 1	1	2	0	0	0	0	2	0	0
Task 2	0	0	0	1	0	2	0	0	1
Task 3	0	0	2	0	0	0	1	0	0
Task 4	0	1	0	0	0	0	0	2	0
Task 5	0	0	0	0	0	1	0	0	2
Task 6	0	0	1	0	2	0	0	1	1

Table 5.3. Example of Task Resources Matrix (TRM)

TRM can capture the task sequencing of a business process design and also provide a basis for reproducing one based on the business process requirements. However, the process patterns are not mapped explicitly in TRM but are formulated based on a set of rules. The next sub-section discusses how patterns are captured using the TRM mapping.

5.2.6 Mapping the process patterns

The mapping of a business process design in TRM provides the capability of setting rules in order to capture and map the various patterns that can occur in the design. The rules are based on the tasks and the way the resources flow. The resources connect the tasks and shape the various patterns that occur. Table 5.4 demonstrates the rules for each pattern in TRM mapping and the corresponding visual representation.

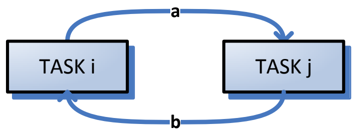
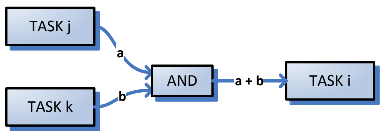
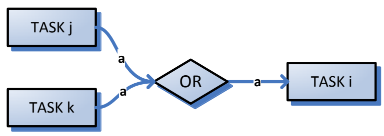
PATTERN & TRM MAPPING	DESCRIPTION	VISUAL REPRESENTATION
SEQUENCE *occurs if:* $TRM_{ia} = 2$ and $TRM_{ja} = 1$	Two tasks (*i, j*) are placed in *sequence* if a resource (a) is the output of one task (*i*) *and* the input to the other (*j*). That means that resource *a flows* from task *i* to task *j*.	
LOOP *occurs if:* $TRM_{ia} = 2$ and $TRM_{ib} = 1$ *AND* $TRM_{jb} = 2$ and $TRM_{ja} = 1$	*Loop* occurs when two tasks (*i, j*) are connected with two resources (*a, b*) in a way that resource *a* flows from task *i* to *j* and resource b flows from task *j* to *i*.	
AND *occurs if:* $TRM_{ja} = 2$ and $TRM_{ia} = 1$ *AND* $TRM_{kb} = 2$ and $TRM_{ib} = 1$	*Parallel execution* (AND) occurs when a task (*i*) accepts two (or more) *different* input resources (*a, b*) from *different* tasks (*j, k*).	
OR *occurs if:* $TRM_{ja} = 2$ and $TRM_{ka} = 2$ *AND* $TRM_{ia} = 1$	*Multi-choice* (OR) occurs when a task (*i*) accepts the same resource (*a*) or an equivalent from (two or more) tasks (*j, k*).	

Table 5.4. Rules using TRM mapping to capture business process patterns

For two tasks *i, j* to be placed in sequence, at least an output resource of task *i* needs to be input resource of task *j*. A loop is created in the process design when for two tasks *i, j* one resources flows from *i* to *j* and one from *j* to *i*. Parallel execution (AND or AND-join) occurs when a task *i* requires different resources from (two or more) different tasks (*j, k*). If two (or more) different resources flow from one task to another then this is mapped as sequence. Also, if two different resources flow from one task to a different task, it is mapped as sequence and not as AND-split. The reason that the pattern is omitted is the assumption that both tasks are necessary in the design. This also contributes to leaner designs that are not loaded with patterns. Multi-choice (OR or OR-join) occurs when the same resource is produced by two (or more) different tasks. This pattern provides the capability of selecting one or more paths in non-exclusive way (as opposed to exclusive choice - XOR pattern). The OR-split again is omitted from the supported patterns as it is considered redundant. This is based on the assumption that a task's output resource is utilised only in the case that there is an available task to use it and thus there is no need to place an OR-split among the output resources of a task.

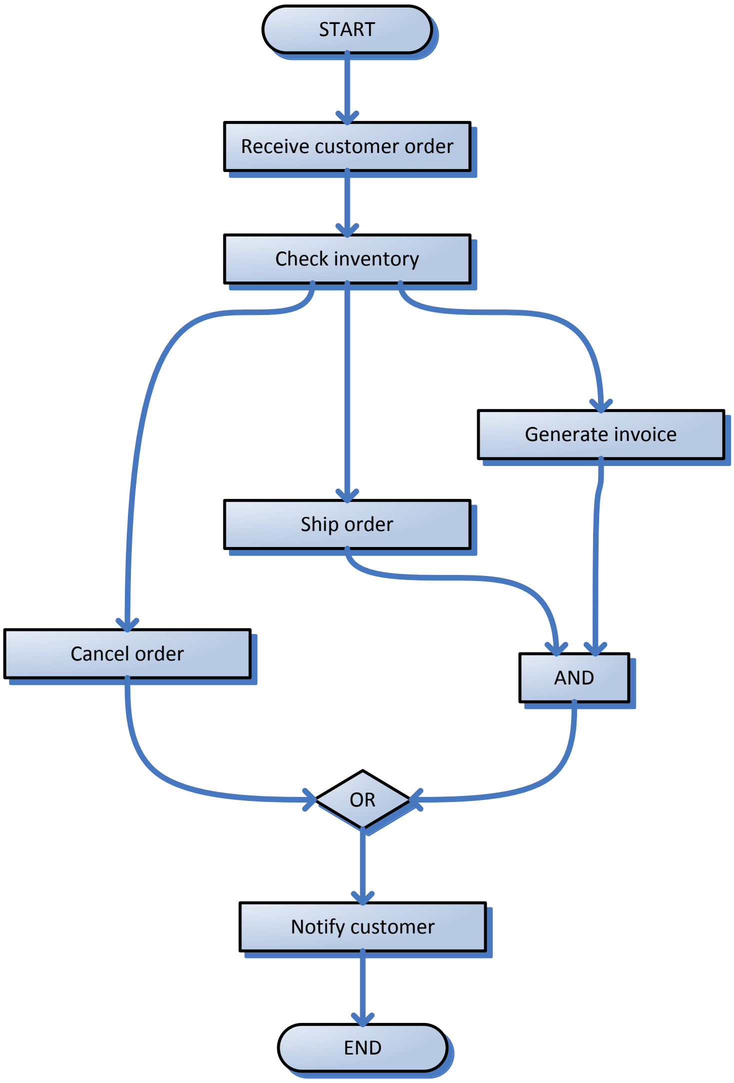


Figure 5.5. Sales order business process (source: Havey, 2005)

5.2.7 Representation example: Sales order

This sub-section provides an example business process mapped with the proposed representation approach. The business process demonstrated in figure 5.5 show a sales order business process captured in a flowchart as described in Havey (2005). The process starts by receiving the customer order and checking the inventory. This task implements a loop to show that all company inventories are checked. In the case that the order cannot be fulfilled it is cancelled and the customer is notified accordingly. In case that the order can be fulfilled, two parallel activities take place: The order is shipped to the customer and the appropriate invoice is generated. Finally, the customer is informed with the order details. The flowchart that maps the business process consists of 6 different tasks and involves one loop, one OR and one AND pattern.

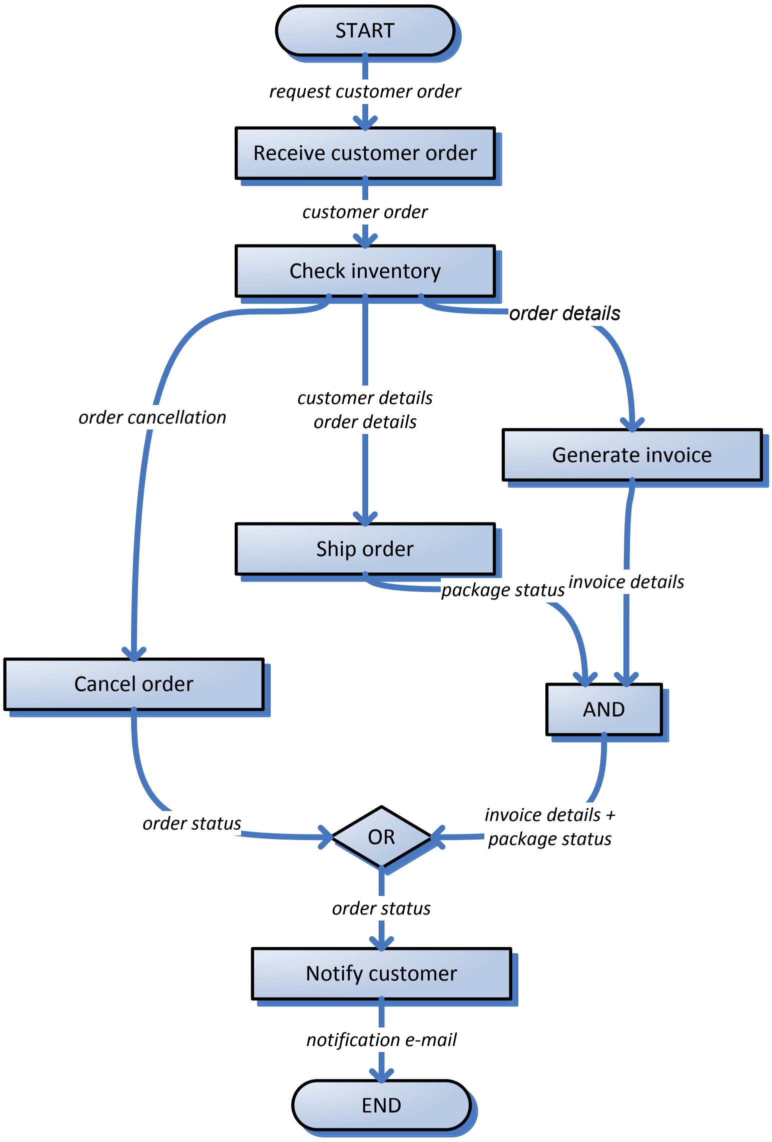


Figure 5.6. Sales order business process design with resources

Figure 5.6 shows the business process design of figure 5.5 enhanced with possible input and output resources per task in order to match the proposed representation. Table 5.5 shows an example of TAM for the tasks in the particular example with two task/process attributes: cost and duration. The attribute values of each task are hypothetical. Table 5.6 shows the TRM as created based on the relationships between the tasks and resources in figure 5.6. These two tables can be used to evaluate (TAM) and reconstruct (TRM) the business process design of the sales order process.

Tasks \ Attributes	Cost (£)	Duration
Receive customer order	100	300’
Check inventory	120	302’
Ship order	117	324’
Generate invoice	178	308’
Cancel order	145	356’
Notify customer	157	389’

Table 5.5. TAM for the sales order example

Tasks \ Resources	request customer order	customer order	order cancellation	customer details	order details	package status	invoice details	order status	notification e-mail
Receive customer order	1	2	0	0	0	0	0	0	0
Check inventory	0	1	2	2	2	0	0	0	0
Ship order	0	0	0	1	1	2	0	0	0
Generate invoice	0	0	0	0	1	0	2	0	0
Cancel order	0	0	1	0	0	0	0	2	0
Notify customer	0	0	0	0	0	1	1	1	2

Table 5.6. TRM for the sales order example

5.2.8 Visual vs. quantitative perspective

The previous sub-sections demonstrated the two perspectives of the proposed representation; the visual and the quantitative perspective. The visual perspective communicates the business process design through a diagram while the quantitative perspective maps the design to make it suitable for analytical methods according to the optimisation focus of this research. The visual perspective can be used to capture or describe the AS-IS situation of a business process. Once the visual perspective is at hand, its transformation to the quantitative aspect of the representation is straight forward. The task attributes are recorded in TAM and the relations of the task and the resources are mapped in TRM according to the proposed mapping and the pattern rules.

However, in the case that a business process design is expressed or captured based on the quantitative perspective, the transformation to visual diagram poses a challenge. This occurs because the quantitative perspective does not ensure that the business process design is feasible and thus an algorithmic process is necessary. The algorithm to construct the business process diagram can be based on (i) the information stored in TRM and (ii) the process input and output requirements so that the start and end point of the process are known. Figure 5.7 provides a basic pseudo-code to construct the visual perspective based on the quantitative representation perspective. The diagram starts with the 'START' node and the process input resources. The tasks that –based on the TRM mapping– accept those resources as inputs are attached. The output resources of the tasks are again sketched based on the information stored in TRM. If those resources coincide with the process output resources, then the process design is completed and the 'END' node is sketched. In any other case, any remaining tasks are attached until the process is completed.

1. START with the process input resources

2. Attach the tasks that accept the resources as inputs

3. Draw the output resources of the attached tasks

4. IF the process output resources are produced,

THEN the process design is complete (END)

ELSE GOTO to step 2

Figure 5.7. Pseudo-code for constructing the visual representation perspective

The pseudo-code in figure 5.7 is not robust as it is based on the assumption that the process design mapped in TRM can be constructed as a diagram and therefore is *feasible*. However, this might not be the case as the tasks in TRM might have been selected in an arbitrary or any other way. Therefore, for the representation to be complete there is a strong need for an algorithm that can compose the visual perspective of a business process design and ensure its feasibility based on the quantitative perspective. This algorithm, called the 'Process Composition Algorithm', is presented in the next section.

5.3 The Process Composition Algorithm (PCA)

This section presents the algorithm of the proposed representation approach, the Process Composition Algorithm (PCA). PCA is an essential part of the proposed representation as it provides the bridge between the visual and quantitative perspective. Moreover, the algorithm by composing a design ensures that the design captured by both representation perspectives is feasible.

5.3.1 Purpose of the algorithm

One of the main objectives of the proposed representation is to provide the capability of generating alternative designs. Based on the representation, a business process design can be in the form of TRM and checked whether it corresponds to a feasible business process design. The previous section stressed the need for an algorithm to perform this operation. Based on this, the purpose of the algorithm introduced in this section is two-fold:

a. To produce the visual representation of a business process design given the quantitative representation, and
b. To check whether the captured design corresponds to a *feasible* business process.

The proposed algorithm is called *Process Composition Algorithm* (PCA). This algorithm attempts to *compose* a business process design as a diagram, given its quantitative representation, and check whether the final outcome corresponds to a feasible business process. The concept of feasibility regarding a business process design is discussed later in this section. The concept of *task library* as a repository of tasks that can potentially participate in a business process design is introduced next.

5.3.2 Library of tasks

The aim of this research is the generation of alternative optimised business process designs. This cannot happen without having a range of available tasks which in different combinations can shape a variety of equivalent business process designs. It is assumed, that for the composition of a business process design, there is a library of available tasks. The introduction of task library affects the problem parameters and introduces two new parameters. Table 5.7 presents the updated parameters for the composition of business process designs.

Parameter	Description	Parameter	Description
n	Number of tasks in the library	N	Set of the n tasks
n_d	No. of tasks in the design	N_d	Set of the n_d tasks
r	No. of available resources	R	Set of the r resources
t_{in}	No. of task input resources	I_i	Set of the t_{in} resources for a task i
t_{out}	No. of task output resources	O_i	Set of the t_{out} resources for a task i
r_{in}	No. of process input resources	R_{in}	Set of the r_{in} resources
r_{out}	No. of process output resources	R_{out}	Set of the r_{out} resources
p	No. of task/process attributes	TA_i	Set of the task attribute values for a task i
		PA	Set of the p process attribute values

Table 5.7. Updated parameters for composition of business process designs

The number of tasks in the library is n and the tasks are stored in the set $N = \{t_1, t_2, t_3, \ldots, t_n\}$. Subsequently the n_d tasks that participate in the design are a subset of N, $N_d \subseteq N$. Also, the number of resources in the design (r_d) is now replaced by the number of resources of all the tasks in the library (r) and its corresponding set R. Consequently the size of TRM becomes now $n_d \times r$ in order to show the relationships of the tasks in the design with all the available resources. The task library provides a starting point for PCA to compose a feasible process design from a range of available tasks.

5.3.3 Infeasibility of process designs

The pseudo-code presented in figure 5.7 constructs the visual representation of a business process assuming that the design stored in TRM is feasible. A business process design is considered as *feasible* when:

1. $TRM_{ij} = 1, \forall j: r_j \in R_{in}, i: t_i \in N_d$
   (*All* the process input resources are utilised by one or more tasks that participate in the process design)

2. $TRM_{ij} = 2, \forall j: r_j \in R_{out}, i: t_i \in N_d$
   (*All* the process output resources are produced by one or more tasks that participate in the process design)

3. Each task in the design is connected either with the process inputs, the process outputs or another task in the design.

Figure 5.8 shows how a business process design can be elaborated based on the pseudo-code in figure 5.7. A feasible business process design is one that starts with the resources in R_{in} and by properly connecting the tasks in TRM produces the requested R_{out} resources. Having discussed the concept of task library, the challenge in producing a feasible design is to obtain a TRM matrix with those tasks from the library that satisfy the feasibility constraints. However, the three feasibility constraints yield a significant number of infeasible cases:

1. One or more process input resources *cannot be utilised* from the tasks in TRM,

2. One or more process output resources *cannot be produced* from the tasks in TRM,

3. There is a *broken link* in the design; there is no task in TRM that can be attached to the process diagram based on its input and output resources.

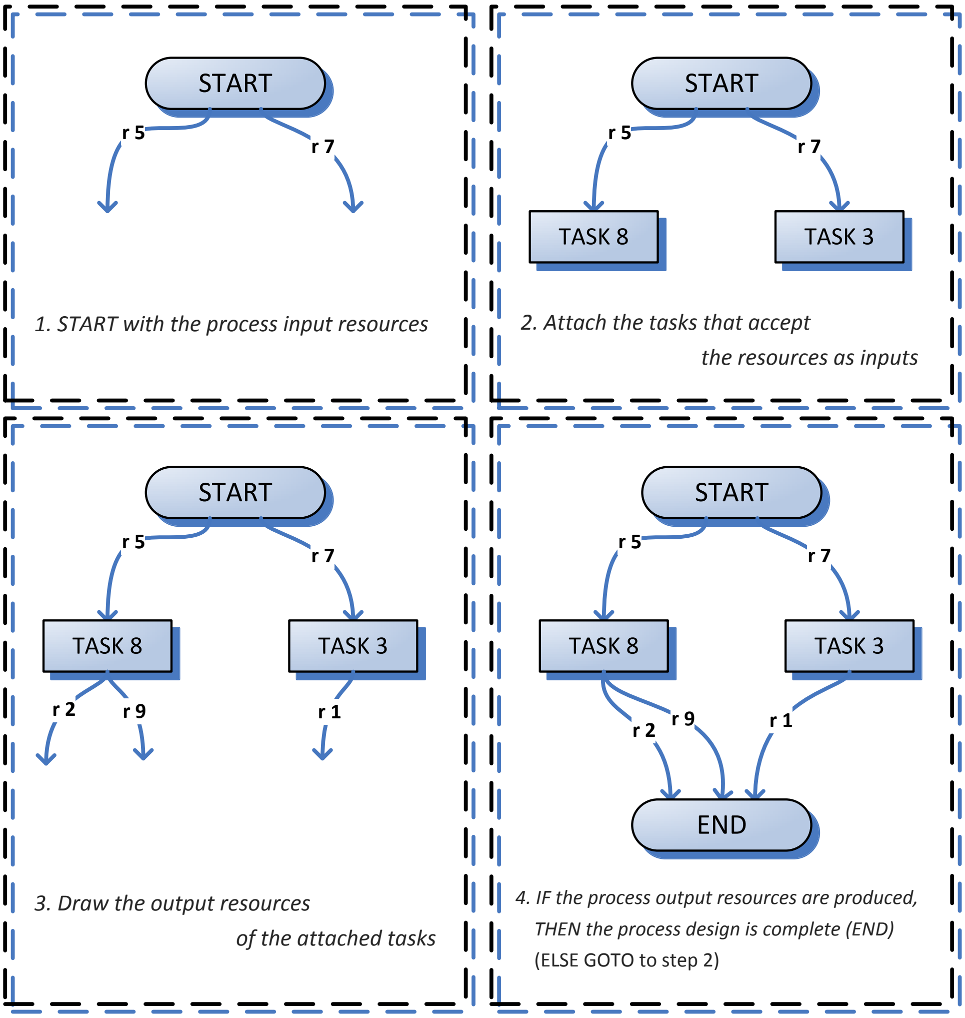


Figure 5.8. Business process graph elaboration based on pseudo-code (figure 5.7)

These cases of infeasibility result in the business process design not being able to elaborate when algorithmically composing it using PCA. Each of these three cases can result in high probability of infeasible solutions even for a large size of the task library. In the case that one resource of R_{in} is not utilised or one resource of R_{out} is not produced, the process design is considered as infeasible. Even in the case that R_{in} and R_{out} are utilised and produced, ensuring that *all* tasks in between are connected through their input and output resources has a low probability. Appendix B demonstrates this high infeasibility ratio and shows the complexity and constraints of algorithmically composing a feasible business process design. PCA attempts to address these challenges and provide a viable solution in composing feasible business process designs.

5.3.4 Requirements and outcomes of PCA

PCA is required to tackle the infeasibility issues and construct a feasible process diagram. Figure 5.9 shows the requirements (or inputs) of the algorithm. The process requirements in the form of the process input and output resources are required as the termination conditions. The algorithm adds tasks to the process design until the process inputs are utilised and the process outputs are produced. The second requirement is TRM that contains the tasks that form the design. PCA will add tasks to the design from this matrix and check whether they correspond to a feasible design. Finally, task library is essential in order to modify or repair the design. As previously discussed, the composition of a process design has high probability of infeasibility, therefore the tasks in the library can help in repairing the design and thus making it feasible with minor alternations. Also, the task library can be used to improve a feasible process design by replacing tasks with better attribute values; this optimisation approach is discussed in the next chapter.

1. Process input and output requirements (R_{in} and R_{out})

2. Participating tasks in the design (TRM)

3. Task library (N)

Figure 5.9. Requirements for the Process Composition Algorithm

Figure 5.10 shows the three outcomes of PCA. The main outcome of the algorithm is the business process design. The design is composed and represented as a directed graph. The nodes of the graph are the tasks of the business process design and the edges represent the connecting resources. The graph is directed, which means that the edges are directed to show the flow of the resources between the tasks. Cycles are allowed in the graph on the basis of the LOOP pattern. PCA uses a graph to compose the business process diagram due to the availability of the various graph elaboration and traversal strategies (e.g. breadth-first, depth-first) which facilitate the main operation of PCA. The traversal strategies followed at different stages of the PCA are explained later in this section.

1. Business process design (process graph)

2. Updated set of tasks in the design (N_d)

3. Degree of Infeasibility (DoI)

Figure 5.10. Outcomes of the Process Composition Algorithm

The second outcome of the PCA is the updated set of tasks in the design (N_d). Based on the proposed representation, PCA translates TRM into a process diagram. However, there is a need for updating TRM itself during the execution of PCA for two main reasons:

1. The *elimination* of any tasks in TRM that have not been added to the process diagram during its composition, thus do not contribute to fulfilling the process requirements, and
2. The *replacement* of any tasks in TRM with tasks in the library that ensure the feasibility of the composed design.

Based on the modifications in TRM during the PCA execution, the second outcome of PCA is the updated set of tasks (N_d) that participate in the process design based on the execution of the algorithm. Finally, the third outcome of PCA is the *Degree of Infeasibility* (DoI). DoI is suggested by the author as a measure of the extent to which a process design is infeasible. Measuring the infeasibility of a design means that different designs can be compared and evaluated. As it will be discussed in chapter 6, the proposed optimisation framework operates with a population of solutions. These solutions are evaluated based on their process attribute values. However, not all solutions might be feasible at any generation during the optimisation process. The DoI helps in selecting the 'less' infeasible solutions and preserving them in the population with the hope that they have a better chance of evolving towards feasible solutions during the optimisation process. DoI based on three main factors of infeasibility (examined in section 5.3.3) and is calculated as:

$$DoI = 1 \cdot n_{in} + 5 \cdot (\sim r_{out}) + 3 \cdot (\sim r_{in}) \qquad \text{(Equation 5.2)}$$

DoI assigns a different weight to each infeasibility case. These weights are selected in a way that reflects the relative importance and frequency of each infeasibility case. For every task *inserted from the library* in the process design, DoI is increased by 1 (n_{in} = total number of tasks inserted from the library) as it is considered a frequently occurring case during the design composition. For every output resource *not produced*, DoI is increased by 5 ($\sim r_{out}$ = total number of output resources not produced) as it is considered an important condition for the feasibility of the design. Finally, for every input resource *not utilised*, DoI is increased by 3 ($\sim r_{in}$ = total number of input resources not utilised). The weight here is less that the output resources because for the output resources to be produced it means that at some point all input resources were utilised. For one or more input resources to be missing it means that corresponding task(s) were omitted during the last stage of the PCA and thus the penalty is less. As each process design carries a DoI, it is straight-forward to compare the feasibility of the designs generated by the PCA. A feasible process design has zero DoI.

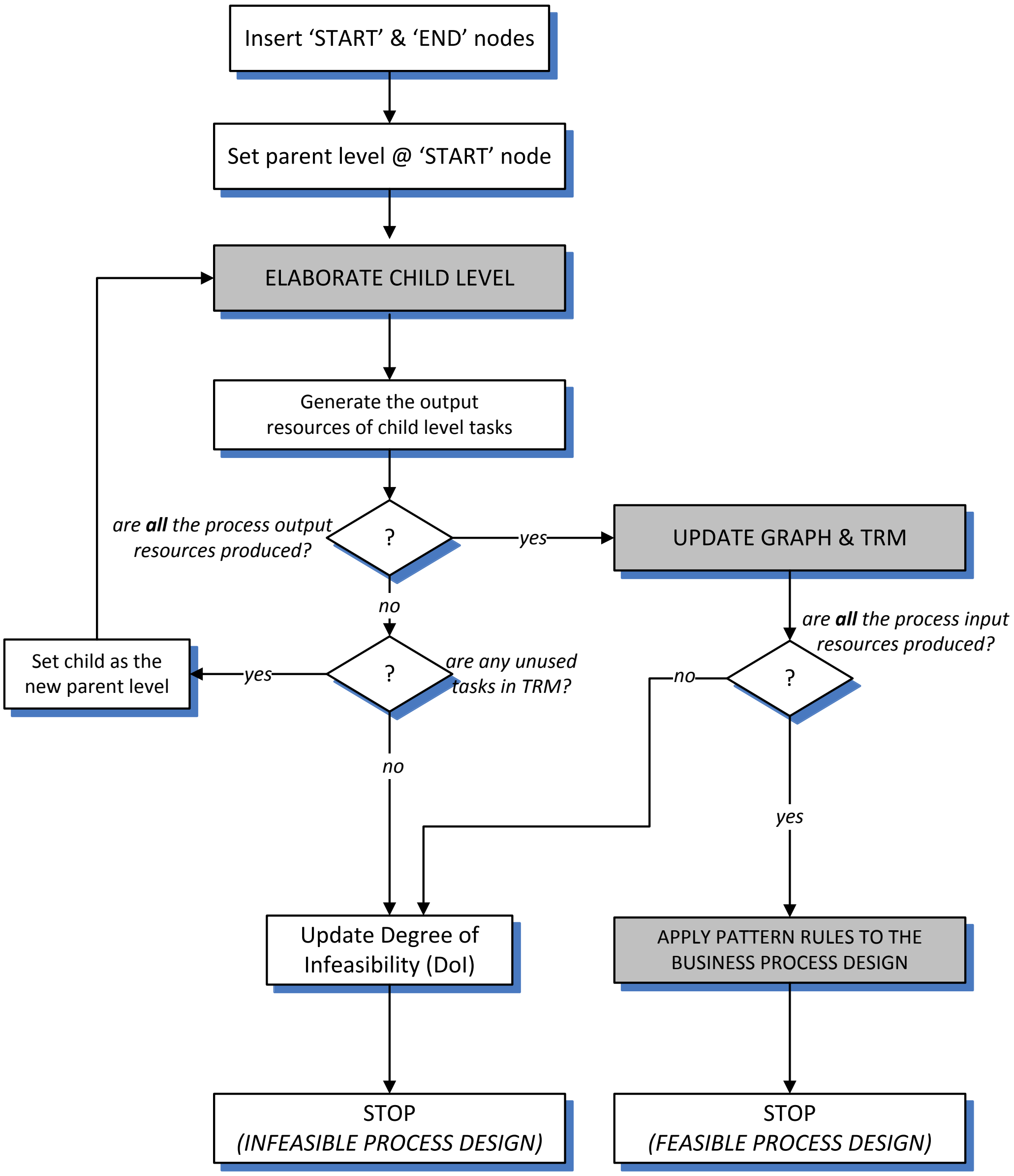


Figure 5.11. Main steps of the Process Composition Algorithm (PCA)

5.3.5 Main steps of PCA

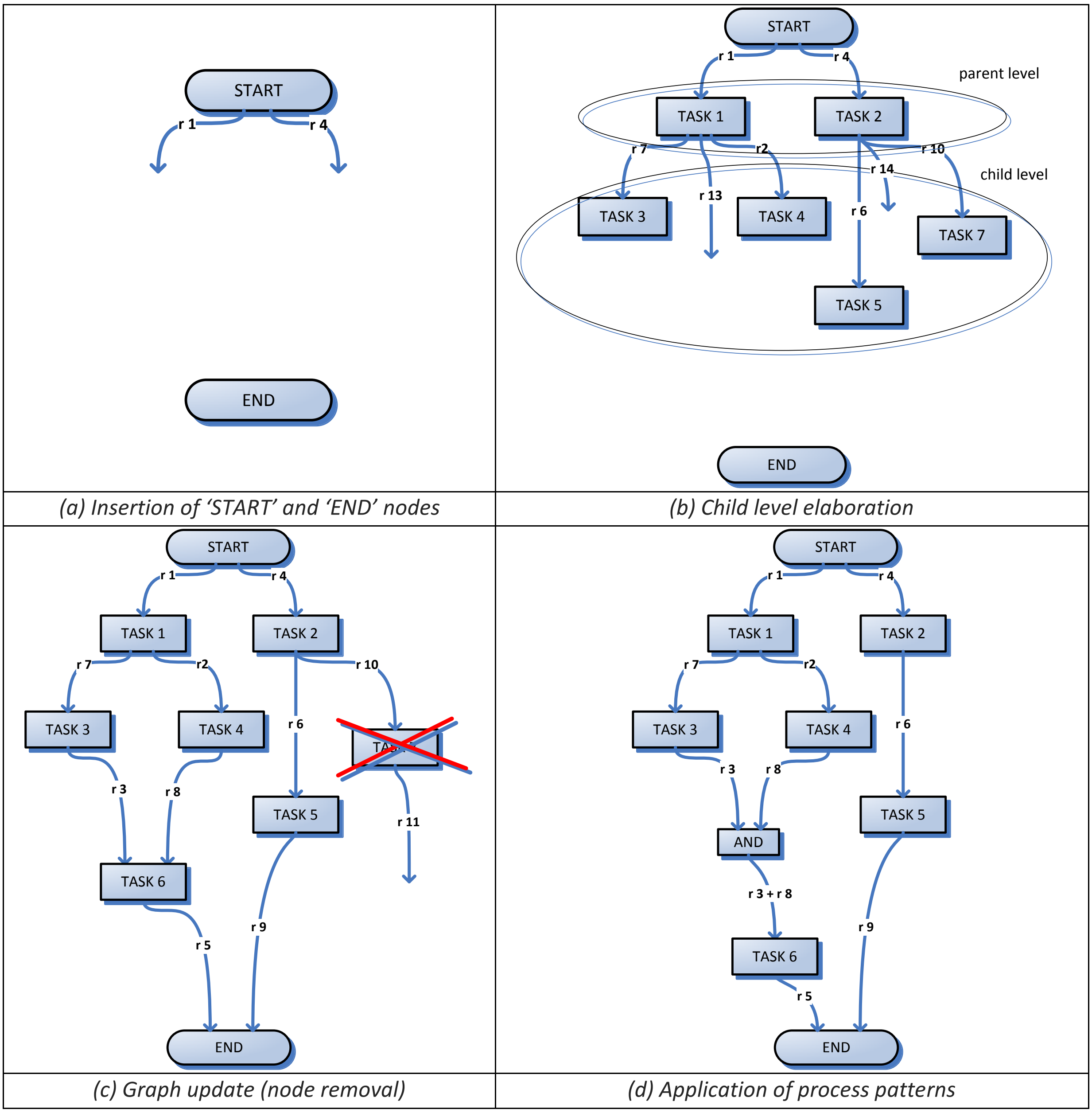
Figure 5.11 displays the main steps of the Process Composition Algorithm (PCA). PCA constructs a process graph and traverses it to ensure that it meets the process requirements. In the graph, each task is represented as a node and there are two artificial nodes, the 'START' node with the process input resources and the 'END' node with the process output resources. These nodes facilitate the connection of the process input and output resources with the participating tasks in order to produce a process design that meets the process requirements. The graph is elaborated with the breadth-first strategy using the concepts of 'parent' and 'child' levels. The 'parent' level consists of the nodes already inserted in the graph and the 'child' level is the one where the new tasks are added in the design based on the output resources of the tasks in the 'parent' level. Once the elaboration of all the tasks in 'child' level is completed, it becomes 'parent' level for the graph elaboration to proceed.

PCA starts by inserting the artificial nodes 'START' and 'END' to an empty graph. The 'START' node is initially marked as the 'parent level'. Then, the algorithm visits all the nodes in parent level one by one in order to elaborate the child level. Graph elaboration requires a small algorithm that is discussed in the next sub-section. Once the child level elaboration is completed, the output resources of the recently attached tasks along with the unlinked output resources of previous tasks are checked to find out whether they contain the process output resources. In the case that not all the output resources are produced and there are unused tasks in TRM, the tasks in 'child' level become the new 'parent' level and the elaboration process is repeated. If there are no unused tasks in TRM then for every output resource that has not been produced there is a penalty attached to the design and DoI is updated accordingly.

In the case that –at some stage of the elaboration process– all the process output resources are produced, TRM and the graph are updated. The update process involves two parts: (i) the elimination from TRM of any tasks that have not been inserted in the process design, and (ii) the elimination of graph nodes (tasks) that do not contribute to the production of the process outputs. The update process is discussed as a separate algorithm in the following sub-section. After the update, PCA checks whether all the process input resources are produced. Some of the tasks that were utilising the process inputs might not have contributed to the process outputs and therefore are removed from the design. In the case that one or more process inputs are not utilised, there is a penalty attached to the design and DoI is updated accordingly. In the case that all the process inputs are produced, the design is marked as *feasible* and it is traversed for the patterns to be applied on the design based on the pattern rules presented in sub-section 5.2.6. The steps of the pattern application algorithm are also discussed later.

Figure 5.12 shows the main stages of a business process design composition based on the main steps of the PCA. Initially, the 'START' and 'END' nodes are inserted (figure 5.12.a).

In the child level elaboration phase, tasks are inserted in the graph and attached to the output resources of the parent level nodes (figure 5.12.b). There might be resources for which there is no matching task. When all the process output resources are produced, PCA removes from the graph any nodes that haven't contributed to the process outcome (figure 5.12.c). Finally, in the pattern application phase (figure 5.12.d), the graph is traversed and the appropriate patterns are inserted based on the pattern rules.



(a) Insertion of 'START' and 'END' nodes

(b) Child level elaboration

(c) Graph update (node removal)

(d) Application of process patterns

Figure 5.12. The main stages of business process design composition

The result of PCA is, in the best case, a feasible business process design in which all the tasks in the design are linked together utilising the process input resources and producing the process output resources. PCA does not scrap the infeasible solutions but it repairs them (utilising the task library) or attaches a penalty to demonstrate their Degree of Infeasibility. In the evolutionary optimisation approach presented in the next chapter, infeasible solutions can lead to feasible ones as they evolve over the optimisation generations.

5.3.6 Algorithm to 'Elaborate child level'

Figure 5.13 shows the algorithm for the 'elaborate child level' operation of PCA. The purpose of this algorithm is to elaborate the process graph by adding tasks to the 'child' level that can be attached to tasks from the 'parent' level. In doing so, this algorithm also marks the tasks from TRM that are added to the process graph as 'used'. This is useful later in the execution of PCA to determine on whether to continue the graph elaboration. In the case that all tasks from TRM are marked as 'used' and the process output resources are not produced, then the design is infeasible.

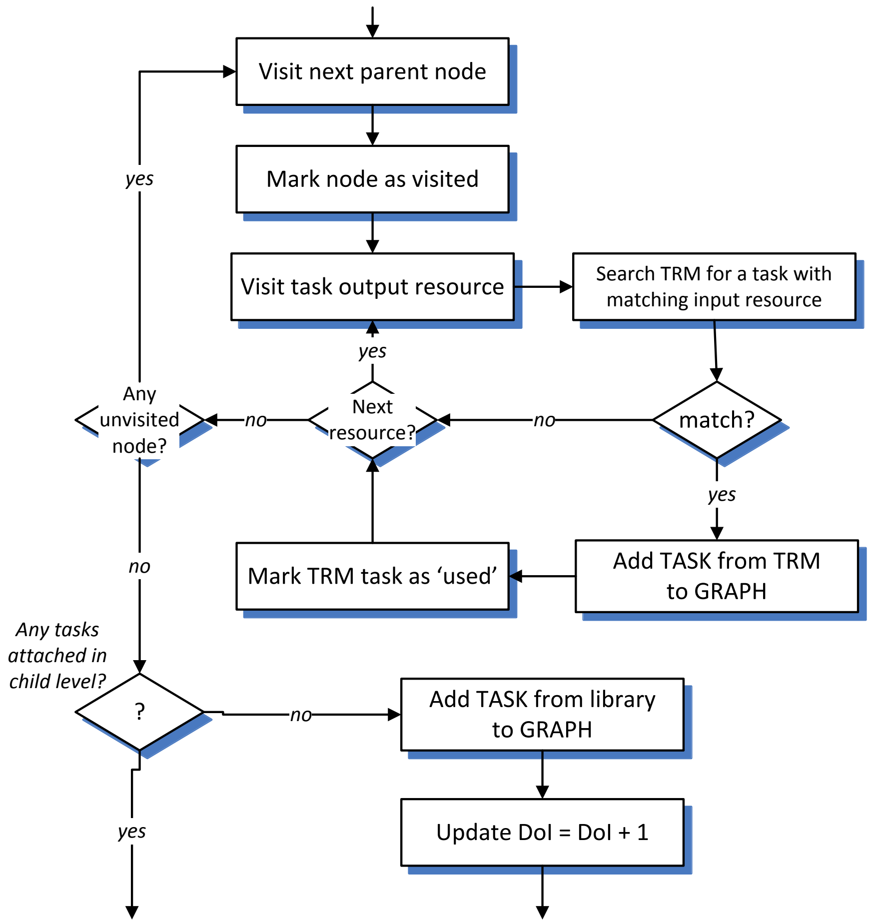


Figure 5.13. Algorithm for the 'ELABORATE CHILD LEVEL' operation of PCA

The basic steps of this algorithm are described as follows. For every node in the 'parent' level', all the output resources are visited. For every output resource the algorithm checks in TRM to find a task with at least one matching input resource. If a task with common resource is found, it is inserted in the graph, linked with the parent task and added to the 'child' level set.

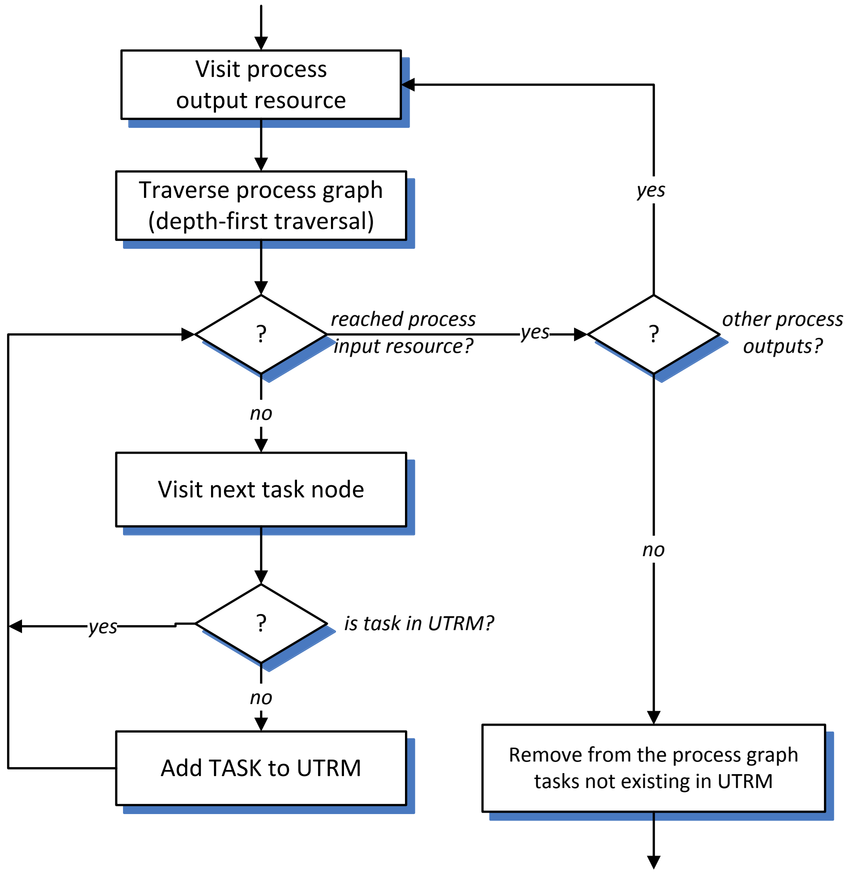


Figure 5.14. Basic steps for the 'UPDATE GRAPH & TRM' operation of PCA

In case that there is no matching task, the algorithm proceeds to the next output resource of the parent level. When the algorithm reaches the last output resource of the last 'parent' level task, it checks whether there are any tasks attached in the 'child' level. In the case that there are no tasks inserted in the 'child' level, the algorithm attaches a matching task from the task library in order to continue with the graph elaboration process. As a result, a penalty is attached to the design and DoI is updated. Every task that is added to the design is linked not only with the parent task but also with any task with which it has a matching resource.

5.3.7 Algorithm to 'Update graph & TRM'

Figure 5.14 shows the steps of the 'UPDATE GRAPH & TRM' operation of PCA. This is an important operation in PCA that occurs only in the case of all the process output resources being produced during the graph elaboration. The purpose of this operation is two-fold: (i) to create an Updated TRM (UTRM) and N_d that contain the actual tasks in the graph that starts from a process input and concludes to a process output, and (ii) to update the graph based on the updated N_d by removing the tasks that do not contribute to the process outputs. The graph elaboration starts with a top-down approach from the artificial 'START' node. The graph elaboration follows the breadth-first strategy in a blind approach and all the nodes added do not necessarily contribute to the production of the process outputs. Therefore the first termination criterion is the check for the output resources.

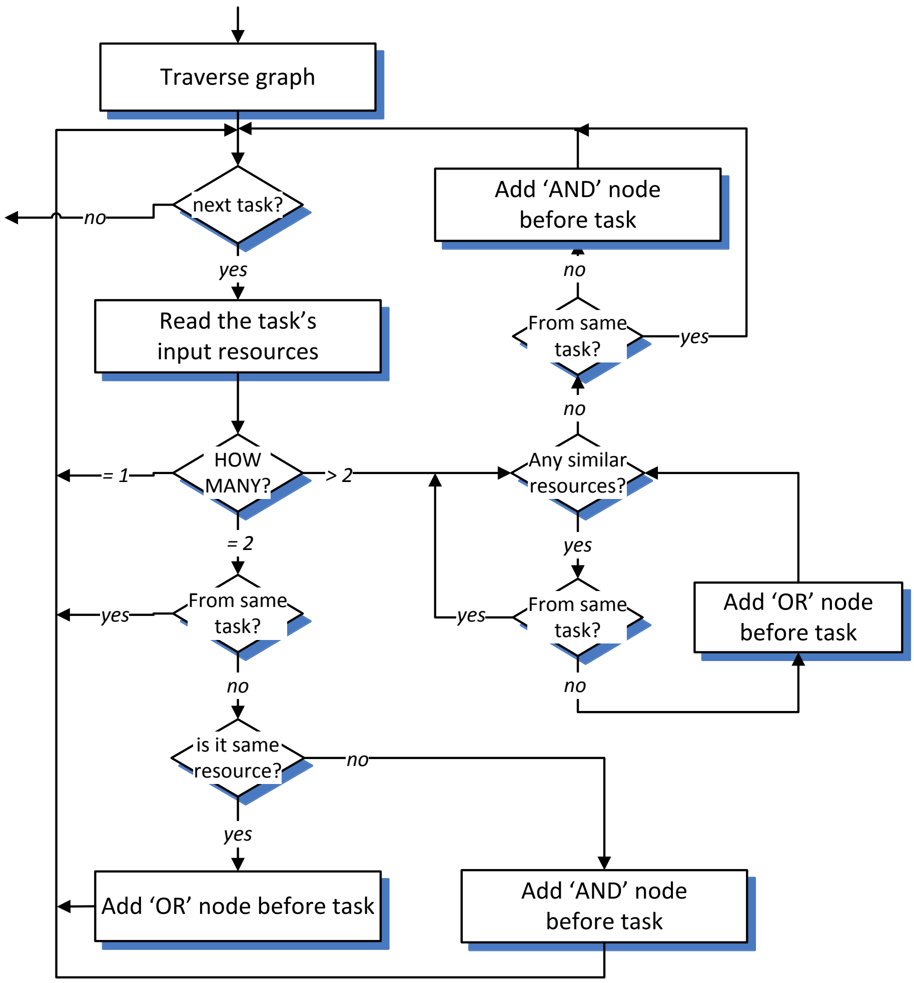


Figure 5.15. Algorithm for the 'pattern application' operation of PCA

Provided that the process output requirements are produced, this algorithm traverses the graph backwards, following the depth-first strategy. Starting from each output resource, the algorithm locates the shortest path from the process output resource to the process input resources. The tasks that are visited during this process are inserted in UTRM. When this operation is repeated for all the output resources, the tasks in the process graph that are not visited are eliminated since they do not link the process inputs with the process outputs. The result of this algorithm is the Updated TRM and N_d that contain only the tasks that are necessary and an updated process graph with redundant nodes removed.

5.3.8 Algorithm for 'Pattern application'

The last step of PCA –when a design is feasible– is to insert the patterns in the process design based on the pattern rules discussed in sub-section 5.2.6. The patterns 'AND' and 'OR' are inserted as nodes in the graph. Sequence and loops are already shaped during the graph elaboration through the resources flow, thus not mapped explicitly using a node. Figure 5.15 shows the steps of the 'pattern application' operation of PCA. The algorithm traverses the graph and reads the input resources of each visited task. This traversal can be performed following either depth-first or breadth-first strategy as the only target is to visit all the graph nodes regardless of sequence. For each task that is visited, its input resources are read. In the case of a task being connected with a single input resource there is no pattern. In the case that the same input resource originates from two or more tasks, an 'OR' node is inserted before the task. In the case that different input resources originate from different tasks, an 'AND' node is inserted before the task denoting that both tasks from which the resources originated are required. For more than two input resources, the algorithm allows the case for both an 'AND' and an 'OR' to be inserted before a task.

5.4 Main remarks

This chapter presented a proposed representation for capturing, visualising and quantifying business process designs. The proposed representation consists of three main elements:

1. *The visual perspective*; a diagrammatic representation of business process designs using flowcharts,

2. *The quantitative perspective*; a representation based on mathematical constructs that captures the business process elements using TAM and TRM matrices, and
3. *The Process Composition Algorithm*; an algorithmic approach that provides a bridge between the two perspectives and ensures feasibility of the captured business process design.

In this chapter, business processes were specified accordingly by providing definitions based on the context of this research. The representation aimed to cover these specified elements of business processes. The aim of the representation is to make a business process design amenable to EMOAs, according to the optimisation focus of this research. The different elements of the representation target to capture and express different aspects of the business process for the different requirements posed by the EMOAs. Chapter 6 discusses how EMOAs utilise the proposed representation in order to produce alternative optimised business process designs.

The first objective of the representation was to provide visual means of communicating the business process design. This objective is covered by the visual perspective of the representation that provides a diagrammatic depiction of the business process. A simple flowchart is chosen as it is straight-forward in communicating the elements of a business process that are identified as essential in the context of this research. The second objective was to provide the capability of mathematically capturing and expressing the same elements of the design that are captured by the visual perspective. The quantitative perspective expresses all of the business process elements using mathematical constructs and introduces two matrices for capturing and evaluating the process design. TAM stores the task attributes for the tasks that participate in the business process designs thus assisting in calculating the process attributes. Having calculated the process attributes for different designs, they can be evaluated and compared. TAM is a construct that satisfies the third objective of the representation.

TRM and PCA are a proposed solution to the last objective about generating alternative designs. TRM accommodates a number of tasks and PCA checks to see whether they form a feasible design. TRM mapping reflects the relationship between a task and a resource. PCA reads TRM and attempts to compose a feasible process diagram. It is important to note that PCA does not aim to discover feasible or optimal designs from the tasks in the task library. PCA checks whether a design stored in TRM is feasible and measures its

Degree of Infeasibility. This way a design compiled in the form of the quantitative aspect of the proposed representation can be evaluated and assessed on its extend of feasibility.

5.5 Summary

This chapter introduced the proposed representation of business processes that can be part of the business process optimisation framework discussed in the next chapter. The requirements of the representation were shaped with the optimisation focus of this research in mind. As a result, the representation covers the visual aspect of a business process design, its quantitative expression and an approach for measuring its extent of feasibility. The next chapter discusses how each of the representation aspects fits with the EMOA optimisation approach of business process designs and presents the complete framework and its challenges that lie at the heart of this research.

CHAPTER 6

Business Process Optimisation Framework

This chapter introduces the proposed optimisation framework for business processes. It starts by providing the mathematical formulation of the business process optimisation problem and detailing the challenges that the proposed framework needs to address. The framework is constructed using two main components: the proposed business process representation that was detailed in the previous chapter and a series of evolutionary multi-objective optimisation algorithms. The inputs, outputs and the main operation of the framework are discussed before its main steps are elaborated. The chapter concludes with the implementation of the framework into a prototype software in order to generate experimental results.

6.1 Problem formulation

This section presents the formulation of the business process optimisation problem. The problem formulation assumes that the business process design requirements are captured based on the proposed representation in chapter 5. Table 6.1 shows the problem parameters based on the proposed business process representation.

Parameter	Description	Parameter	Description
n	Number of tasks in the library	N	Set of the n tasks
n_d	No. of tasks in the design	N_d	Set of the n_d tasks (subset of N)
n_{min}	Minimum number of tasks in the design	N_{in}	Set of library tasks to be included in the process design (subset of N)
r	No. of available resources	N_{ex}	Set of library tasks to be excluded for the process design (subset of N)
t_{in}	No. of task input resources	S_d	Set of the different process sizes
t_{out}	No. of task output resources	DoI	Degree of Infeasibility (as calculated by the PCA algorithm)
r_{in}	No. of process input resources	TAM	Matrix that stores the task attribute values for each of the n_d tasks in the process design
r_{out}	No. of process output resources	PA	Set of the p process attribute values
p	No. of task/process attributes		

Table 6.1. Parameters for business process design optimisation problem

The multi-objective problem formulation for business process optimisation is as follows:

For a business process design with a set of n_d tasks and *p* process attributes:

Minimise / maximise $(PA_1, PA_2, \dots, PA_p)^T$

Subject to:
1. DoI = 0
2. $n \geq n_d > 0$
3. $r \geq r_{in}, r_{out}, t_{in}, t_{out} > 0$
4. $p \geq 2$

{
a. $n_d \geq n_{min} > 0$
b. $n_d \in S_d$
c. $N_d \cap N_{ex} = \emptyset$
d. $N_{in} \subseteq N_d$ }

We assume that the process attributes are used as the optimisation objectives. A process attribute (PA_j) can be calculated as an aggregate of the corresponding task attributes stored in TAM for all the n_d tasks in the process design according to the following equation:

$$PA_j = \sum_{i=1}^{n_d} TAM_{ij} \qquad \text{(Equation 6.1)}$$

The problem formulation assumes that there are more than one process attributes used as optimisation objective and thus is considered as a *multi-objective optimisation* problem. The problem formulation also involves 8 constraints, 4 compulsory and 4 optional. Constraint (1) ensures that only *feasible* business process designs are evaluated. The Degree of Infeasibility (DoI), as discussed in the previous chapter, is a result of the PCA execution and measures to which degree a set of n_d tasks forms a feasible business process design. The only case that a design is feasible is when DoI equals to zero. The second constraint ensures that the available tasks in the library (*n*) are more than or at least equal to the tasks required to compose a design (n_d) and that both (*n*, n_d) are greater than zero. Constraint (3) ensures that all the resource-related parameters are greater than zero and that the available resources (*r*) are more than those required by the process and task inputs and outputs. Finally, the fourth constraint assumes that there are at least two task/process attributes and thus the problem is multi-objective or at least bi-objective.

Each of the constraints that belong to the second set (a-d) is optional and is provided in order to make the problem more flexible in terms of business process designs generated. Constraint (a) sets a lower limit (n_{min}) to the number of tasks that can formulate a design. In the case that $n_d = n_{min}$, an acceptable solution contains *exactly* n_d tasks in the design. An extension of constraint (a) is constraint (b) where it considers a design as acceptable only if its size belongs to a specified range of process sizes (S_d). Constraint (c) employs the N_{ex} set –a set of tasks in the library that should not be included to the set of N_d tasks that form the solution. This optional constraint ensures that the solution does not contain any undesired tasks from the library. Constraint (d) is the opposite of (c) as it enforces particular tasks to be included in the solution. It introduces the N_{in} set –a set of tasks in the library that are required in the solution. This optional constraint ensures that tasks that are strongly favoured appear in the solution. These last two constraints (c, d) tackle any bottlenecks that can appear in the business process design by showing preference against or towards particular tasks from the task library.

6.2 Optimisation challenges

The problem formulation described in the previous section gives rise to some challenging issues in terms of generating optimised business process designs. The main challenges for business process optimisation are:

- *Nature of the problem.* Based on the problem formulation, business process optimisation is a *discrete* problem as the main variable is a set of tasks (N_d) that form the business process design. A discrete problem is more challenging and less flexible to optimise than a continuous one, as the variables are significantly constrained in terms of different values that they can take. Also, in a discrete problem, a minor change in one of the variables can have a detrimental and uncontrollable effect on the optimisation process. The discrete nature of the problem poses a serious challenge to the effectiveness of the optimisation process.

- *Framework output.* Based on the problem formulation of the previous section, the problem is to identify a set of n_d tasks with optimal process attribute values. The proposed framework apart from optimising the process attributes of a set of n_d tasks should also compose and produce the corresponding feasible business process design.

The process of design composition needs to be embedded in the optimisation process as it is one of the framework's main outcomes.

- *Multi-objective formulation of the problem.* In addition to its discrete nature, the business process optimisation problem is formulated as multi-objective. Assuming that the participating objectives are conflicting and that each solution represents a different trade-off between the objectives, discovering the Pareto-optimal front across all of the objectives is another major challenge for the proposed optimisation framework.

- *Solution representation.* The problem formulation requires different aspects of the business process design for different stages. For example, evaluation of the objectives would require TAM, while to check the infeasibility constraint (1) would require TRM and the execution of PCA. In addition to these, the application of EMOAs to the framework requires appropriate solution representation for each of the genetic operators (selection, crossover, mutation). The framework needs to devise a strategy in order for a solution to address the different requirements that emerge during the optimisation process.

- *Constraint handling (selection of solutions).* There is a series of compulsory and optional constraints in the problem. Appendix B shows that the infeasibility constraint (1) on its own yields a significant number of infeasible solutions. The constraints need to be managed in a way that allows for feasible solutions to be generated and preserved during the optimisation process. Handling the constraints appropriately, by repairing or penalising a solution, can significantly affect the quality and diversity of the optimisation results. The framework needs to handle the constraints with flexibility in order to discover feasible solutions and evolve them towards the optimal ones.

- *Degree of Infeasibility.* The first constraint of the problem requires the execution of the PCA algorithm in order to measure the Degree of Infeasibility (DoI) of a solution. However, as discussed in chapter 5, PCA also *updates* the solution (either removing or replacing tasks in the N_d set) in order to ensure its feasibility. This is a major challenge for the optimisation framework, to handle a solution that is modified by an algorithm during the optimisation process. The framework should ensure that the different phases that a solution undergoes are consistent during all of the optimisation stages. Additionally, the DoI constraint is an equality constraint. This adds additional complexity to the problem as equality constraints are much harder to satisfy compared to inequality constraints.

- *Solution size.* Business process optimisation requires solutions of variable size. The PCA algorithm –as described in chapter 5– composes a business process design that can have a *maximum* of n_d tasks *or fewer*, provided that the design requirements are met. Having a fixed number of tasks in the design would be a major barrier towards design composition and it is directly conflicting with the aim of the framework (business process optimisation = lean business process designs). Therefore, the framework must be capable of handling solutions of variable size for the same design requirements.

- *Design evolution.* Similar to the previous challenge, the optimisation process must allow for diverse business process designs to be generated in terms of patterns. By no means should the framework restrain the design of a solution. On the contrary, the proposed framework should promote the evolution of solutions with different design patterns.

- *Open to the selected EMOAs.* The potential of business process optimisation is examined with the application of a range of different EMOAs. Therefore, the framework must be structured in a way that the problem remains independent and not tied to a particular algorithm. Each of the EMOAs should operate as a plug-in to the framework, whereas the main steps of the optimisation process should remain as generic as possible. Given the previous challenges of the problem formulation, keeping the framework open to a range of different optimisation algorithms adds an extra layer of difficulty.

The next section introduces the proposed optimisation framework that addresses the above-mentioned challenges in order to generate optimal results in the form of alternative business process designs.

6.3 Proposed optimisation framework (bpo^F)

This section provides an in-depth description of $\mathbf{bpo^F}$ – the proposed evolutionary multi-objective optimisation framework for business process designs. The main components of the proposed framework are two, as shown in figure 6.1, and are:

i. the proposed business process representation technique

ii. a series of Evolutionary Multi-objective Optimisation Algorithms (EMOAs)

The proposed business process optimisation framework (**bpo^F^**) applies a series of existing EMOAs to a business process design captured using the proposed representation. The outcome of the framework is a series of alternative optimised designs again in the form of the proposed business process representation. The challenge of the framework is to fully utilise the proposed representation technique and the capabilities of the EMOAs in order to generate alternative optimised designs.

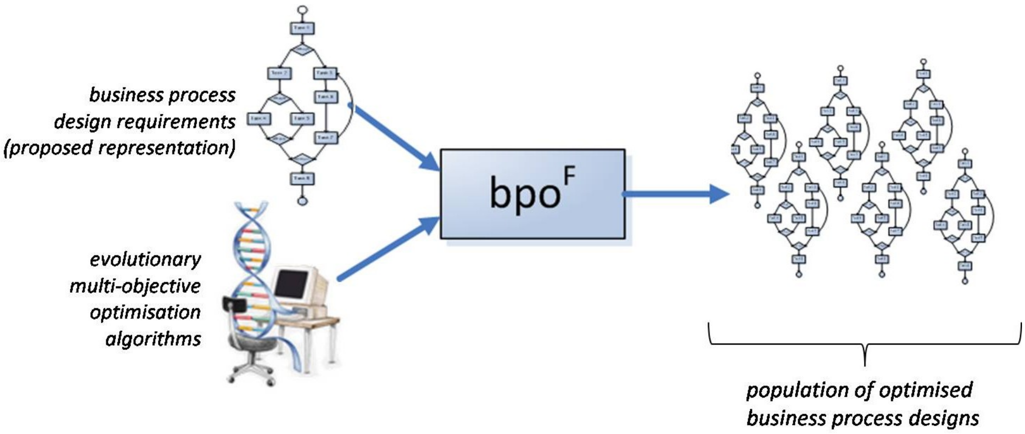


Figure 6.1. Main components of **bpo^F^** – the proposed optimisation framework

6.3.1 Main operation, inputs and outputs

The proposed optimisation framework utilises the aim of this research which is the optimisation of business processes using EC techniques.

The aim of the proposed evolutionary multi-objective optimisation framework for business process designs (**bpo^F^**) is to apply state-of-the-art EMOAs to given business process requirements in order to generate a series of alternative optimised designs.

Based on the aim, the main operation of the framework is the generation and optimisation of business process designs. To achieve this, figure 6.2 demonstrates the inputs and outputs of **bpo^F^**.

There are four inputs to the framework:

1. The *process requirements* for the design in the form of the required process inputs (R_{in}) and process outputs (R_{out}). All the generated designs must start from the same inputs and conclude to the same outputs.

2. The *process size* (n_d). The process size denotes the *maximum* number of tasks in the process designs. During the optimisation process PCA is allowed to generate designs with fewer tasks.
3. The *library of tasks* (N). This set contains all the tasks that can potentially participate in a process design. Given the process size, TRM is formed with n_d tasks from the library to create a potential feasible solution.
4. The *process attribute functions* are the formulas for each of the process attributes. The optimisation framework uses these functions as optimisation objectives. The process attributes functions are always dependent on the task attributes. Equation 6.1 provides an example of an aggregation process attribute function.

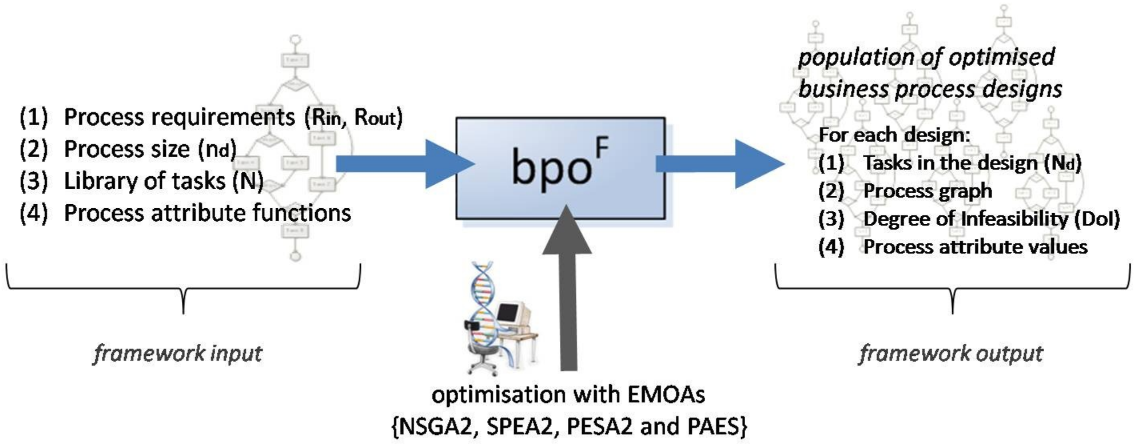


Figure 6.2. The inputs and outputs of the proposed optimisation framework

The proposed optimisation framework employs a series of Evolutionary Multi-objective Optimisation Algorithms (EMOAs) in order to optimise business process designs. The selected EMOAs are: NSGA2, SPEA2, PESA2 and PAES. All of the selected algorithms are state-of-the-art and each has distinctive features that enhance the optimisation process. Appendix D presents an overview and the main steps of the algorithms. Also section 6.3.3 discusses the key differences of the employed algorithms and the expectations of their performance as part of the proposed business process optimisation framework (also detailed in Appendix D). Employing a range of EMOAs provides the opportunity to compare their performance and determine their suitability for the problem.

The proposed framework generates a population of optimised business process designs using the inputs in conjunction with one of the evolutionary algorithms. This operation is discussed step-by-step in the following sub-section. The outcome of the framework is the population of optimised business process designs. For each design the framework produces:

1. The *tasks in the design*, stored in the N_d set.
2. The *process graph*, which is the diagrammatic representation of the design.
3. The *Degree of Infeasibility (DoI)*, which for the optimised process designs should be equal to zero (as discussed in chapter 5 – sub-section 5.3.4).
4. The *process attribute values*, which are calculated based on the input functions. These are the objective values which quantitatively show how well the design performs based on the criteria it has been assessed with.

Given the problem formulation (section 6.1), figure 6.2 shows that business process optimisation is not a typical optimisation problem in the sense of optimising a series of objective functions given the constraints. The outcome of the framework involves the generation of business process diagrams –an outcome which is not explicitly included in the original problem formulation. The only outcomes that the problem formulation requests are the process attribute values (optimisation objectives) and DoI (constraint – feasibility check). The outcomes (1), (2) and (3) of the framework are the result of the PCA execution –as discussed in the previous chapter. Based on the problem formulation, PCA is triggered to check the first constraint – zero degree of infeasibility. Therefore the proposed framework involves an anomaly in the optimisation process: *Most of the outcomes are the result of an algorithmic procedure (PCA) within a constraint of the problem.* Only the process attribute values are the products of the objective functions. This novelty of the framework is discussed in the next sub-section where the main steps of the framework and the optimisation process are demonstrated.

6.3.2 The main steps of bpo^F

The main steps and the structure of the proposed business process optimisation framework are shown in figure 6.3. Essentially, the framework employs a generic optimisation structure (blue-shaded boxes) which is handled each time by a specific EMOA. Each of these optimisation steps however, is adjusted to reflect the business process problem and ensure that the framework utilises the inputs and produces the outputs as demonstrated in figure 6.2. This sub-section describes the generic optimisation process and the business process oriented adjustments in each step, while the next sub-section discusses the details of the framework operation for the selected EMOAs. The proposed optimisation framework consists of five steps (figure 6.3):

(1) Generate random population

The first step of the optimisation process is the generation of random population. This step occurs only once in the optimisation process as then the population is evolved for a defined number of generations. The generation of random population creates a fixed number of sets of n_d tasks. The number of the sets generated equals the specified population size that the algorithm is working with. Each of the population sets contains n_d randomly allocated tasks from N – the task library. However, for each of the sets there is a constraint in the random allocation of tasks. The constraint is that *a task must appear only once in the same set*. This constraint avoids having duplicate tasks in one set –and in a potential business process design. After the random population is generated, steps 2-5 are repeated for a predefined number of generations.

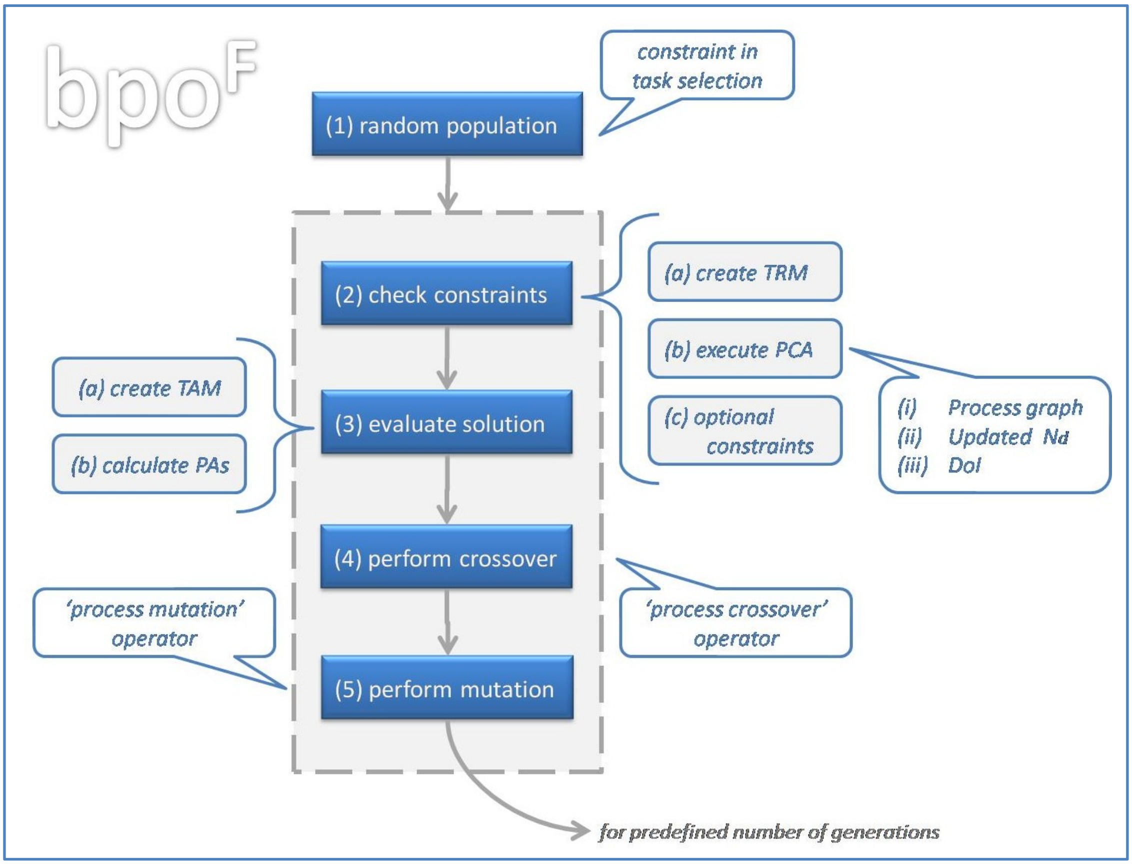


Figure 6.3. The main steps of the proposed optimisation framework

(2) Check constraints

For each solution of the population, the problem constraints are checked (see problem formulation – section 6.1). Contrary to most optimisation procedures where the solution is first evaluated, **bpo**$^{\mathbf{F}}$ checks the constraints prior to solution evaluation due to a specific reason: *the constraints modify the solution.* The first constraint measures the Degree of Infeasibility (DoI) of the solution. For this to happen, two actions are triggered on the basis of the proposed business process representation: (i) TRM is formed and (ii) PCA is executed. TRM reflects the relationships of the tasks in the N_d set with the resources of the problem and PCA uses this information to compose a business process design based on the process requirements. The outcome of the PCA is the diagrammatic version of the business process design, its DoI and the updated N_d – the updated set of tasks in the design which is necessary for two reasons:

a. A design might be composed with less than n_d tasks; therefore the remaining tasks are removed from the N_d set and
b. A design might have been repaired during composition; therefore some tasks in N_d might have been replaced.

PCA ensures that there is one-to-one relationship between the input and the output solution to ensure consistency in the optimisation process. That means that for an N_d set the *same* updated N_d will be produced each time that PCA is executed. At this stage of the optimisation process, the process design has been created, its DoI is measured and the N_d set is updated to reflect the actual tasks in the solution. These are three of the bpo^F outcomes. The last part of this step is to check the (updated) solution on whether it violates any of the problem's optional constraints –if any is included in the problem. Each of the selected EMOAs employs a different strategy in terms of constraint violation and how to handle a solution. These different strategies are discussed in more detail in the following sub-section.

(3) Evaluate solution

The solution evaluation involves two stages based on the proposed representation: (i) TAM is created and (ii) the various process attributes are calculated based on their functions. TAM is created based on updated version of the solution involving the tasks in the design and their attribute values. Based on this matrix the solution is evaluated in terms of the process attribute values. The reason for solution evaluation after the constraint checking is that only the tasks that participate in the process design are actually taken into account in the evaluation process. Each of the EMOAs employs a different strategy to evaluate, compare and select the solutions of the population that will pass through the genetic operators. These different approaches are discussed in the next sub-section.

(4) Perform crossover

After evaluation, the solutions undergo *crossover* – a genetic operator that exchanges information between two solutions. For the business process optimisation problem, crossover occurs directly in the N_d set of each solution. Figure 6.4 demonstrates how the 'process crossover' operator works for designs with $n_d = 5$ tasks. Initially, the solutions are selected for crossover based on a given crossover probability –defined separately by each of the EMOAs. The solutions that are chosen for crossover are split into pairs. For each pair a unique crossover-point is defined based on a random number (between 1 and n_d-1). Based on this crossover-point, the parent solutions exchange their tasks after this point in order to form the child solutions. At the end of the process, each of the child solutions contains tasks from both the parents. The process crossover operator does not check whether the solution is feasible; this is the concern of step 2.

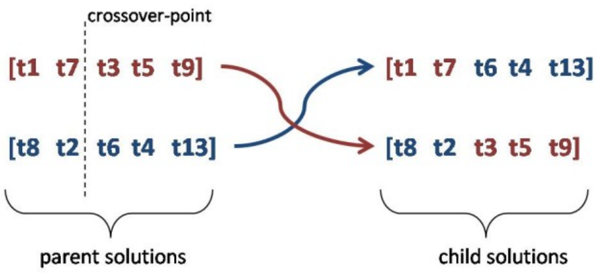


Figure 6.4. The 'process crossover' operator

(5) Perform mutation

The last operator of the optimisation process is *mutation* – a genetic operator that randomly alters information in a chosen solution. Similar to 'process crossover', the 'process mutation' operator is applied on the N_d set of tasks of a particular solution. The probability of mutation occurring is again defined by the EMOA. Figure 6.5 shows the 'process mutation' operator for a chosen task. When mutation occurs for a chosen task, the task is replaced with an *arbitrary* task from the task library.

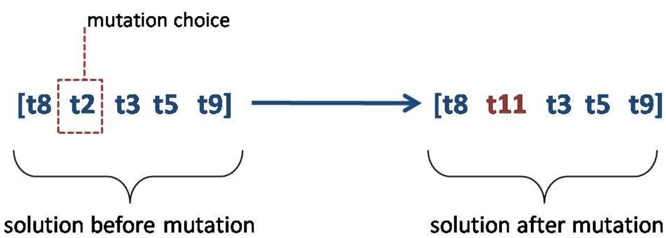


Figure 6.5. The 'process mutation' operator

6.3.3 The framework under different EMOAs

The proposed optimisation framework is structured in a way that it can operate with any of the four employed EMOAs. This section describes how each of the EMOAs manages the framework differently and what the different impact is. Each EMOA is different in three main areas:

1. The type of parameters that it is using (e.g. population size, number of generations, crossover and mutation probability),
2. The selection operator process, i.e. fitness assignment and constraint handling,
3. The genetic operators during the optimisation process (e.g. PAES does not use crossover).

Figure 6.6 shows the main steps of the framework and stresses the differences of the algorithms. The key differences lie in the creation of the empty archive set and in the selection process where each EMOA applies its own method. Appendix D discusses each algorithm in more detail. Below is a brief description of the unique characteristics of each algorithm and their impact on the proposed optimisation framework for business processes:

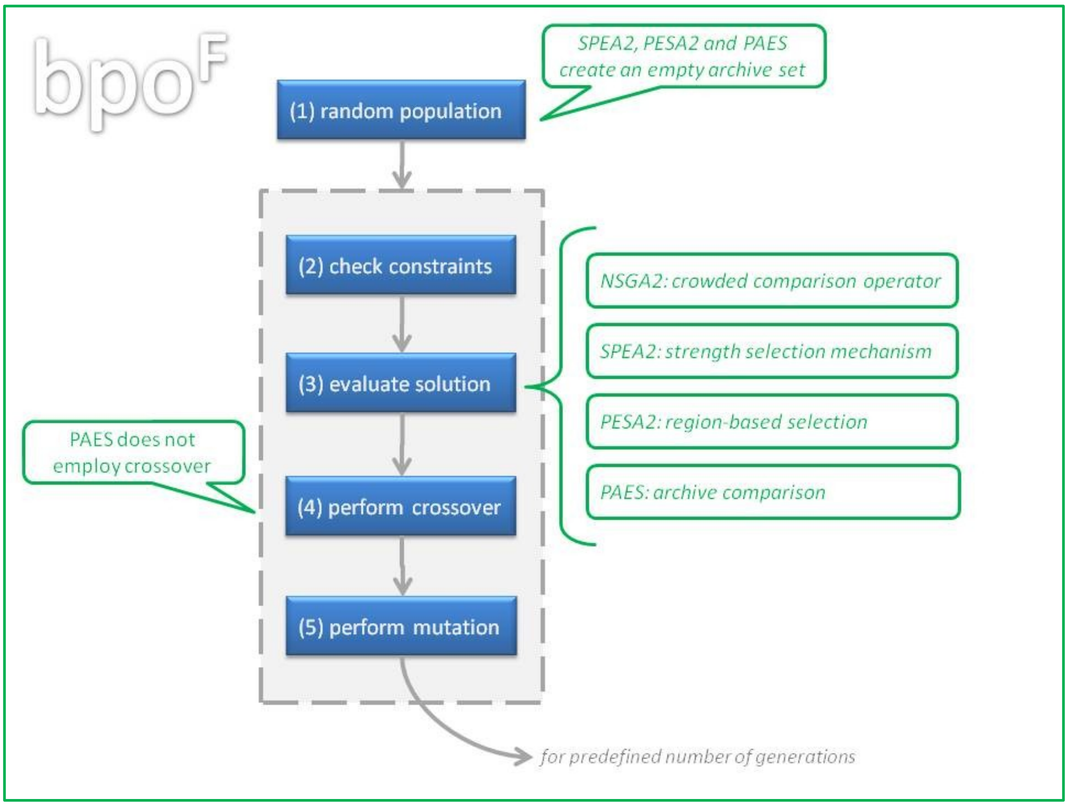


Figure 6.6. Key differences of the various EMOAs across the stages of bpo^F

(1) Non-dominated Sorting Genetic Algorithm 2 (NSGA2)

NSGA2 is considered a high-performing multi-objective optimisation algorithm. It is an elitist algorithm that uses a parent and a child population in each generation in order to maintain 'good' solutions. The diversity among non-dominated solutions is introduced in NSGA2 by using the crowded comparison operator that is used in the tournament selection during the population reproduction phase (step 3 in bpo^F). The crowded comparison operator guides the selection process at various stages of the algorithm towards a uniformly spread-out Pareto front. However, NSGA2 is known not to perform well in problems with multiple local fronts. The fitness assignment strategy of NSGA2 ceases to produce the driving force towards the global front once most of the solutions of the population share the shame non-domination level. This is further augmented due to the use of elitism and NSGA2 suffers from the tendency of getting trapped in local fronts (pre-mature convergence). The proposed framework optimises business process designs of different sizes thus creating multiple local fronts.

Utilising NSGA2 will examine its capability of discovering and optimising solutions of variables sizes in terms of business process designs. The main parameters of NSGA2 are population size, number of generations along with crossover and mutation probabilities.

(2) Strength Pareto Evolutionary Algorithm 2 (SPEA2)

SPEA2 is another elitist evolutionary algorithm. SPEA2 has been popular in the evolutionary multi-objective optimisation community and has been used in a variety of optimisation problems. SPEA2 works by maintaining an external population at every generation storing all non-dominated solutions discovered so far beginning from the initial population. This external population participates in all genetic operations and is created in the beginning as an empty set (step 1 in bpo^F).

In step 3 (evaluate solution), SPEA2 uses a novel selection strategy in which a 'strength' is associated with each member of the archive. The 'strength' of a solution is based on the number of solutions in the internal population which it dominates. Selection is biased towards minimising the strength of the solution thus preferring the exploration of less populated regions of the objective space. Because of this strength selection mechanism, it is expected that SPEA2 will demonstrate flexibility in converging to optimal solutions across the search space. The main parameters of SPEA2 are (internal) population size, archive size, number of generations and crossover/mutation probabilities.

(3) Pareto Envelope-based Selection Algorithm 2 (PESA2)

PESA2 uses an internal population and an external (or archive) population. The archive is initialised in step 1 of the framework. PESA2 uses region-based selection in step 3 in order to evaluate and select the non-dominated solutions. Region-based selection takes places on a hyper-grid division of the objective space in order to maintain diversity. PESA2 uses this crowding measure to decide what solutions to introduce into the external population (i.e. the archive of non-dominated solutions found along the evolutionary process).

The region-based selection mechanism may be a key factor for PESA2 to outperform the other three EMOAs. Dividing the objective space in hyper-boxes (for multiple objectives) or squares (for two objectives) creates what is called the 'squeeze factor'. PESA2 uses this 'squeeze factor' both in selection and in archive update of solutions. If we assume that the algorithm will accurately create at least one hyper-box for a group of business process designs with the same size, then PESA2 will be capable of locating optimal solutions across most design sizes of the search space. Apart from the standard parameters (number of generations, population/archive size), PESA2 has one parameter concerning the hyper-box.

(4) Pareto Archived Evolution Strategy (PAES)

PAES s the simplest possible non-trivial algorithm capable of generating diverse solutions in the Pareto optimal set. The algorithm is identified as being a (1+1) evolution strategy, using local search but using a reference archive of previously found solutions in order to identify the approximate dominance ranking of the current and candidate solution vectors. This makes PAES also an elitist algorithm. The archive is initialised in step 1 of the framework and serves two separate purposes. First, it stores and updates all of the non-dominated solutions (subject to diversity criteria) generated, ready for presentation at the end of a run. Second, during the run (bpo^F – step 3), it is used as an aid to the accurate selection between the current and candidate solution vectors by acting as an approximation to the current non-dominated front.

PAES can be useful when local search seems superior to or competitive with population-based methods. In the proposed framework, the search for optimised process designs of different size might be tackled better using local search. Also PAES is able to generate a diverse set of good solutions and it does so in significantly less time. Producing a series of optimised business process designs in a timely fashion could be an additional strength of the proposed optimisation framework taking into account the complexity of the problem.

As shown in figure 6.6, PAES does not use the crossover genetic operator. The algorithm is confined to local search and therefore it employs a small change (mutation) operator only to move from a current solution to a nearby neighbour. In the framework and since the solution is stored as a fixed-size array, a small change to a task can lead to process designs with different process size.

Additional remarks

The four evolutionary algorithms employ the 'process mutation' described in the previous sub-section as their mutation operator with the same probability (0.2). The 'process crossover' operator is employed but NSGA2, SPEA2 and PESA2 as the recombination operator with a probability of 0.8.

All the four evolutionary algorithms that are employed by the proposed optimisation framework are *elitist*. Elitism ensures that the search is driven towards the global Pareto front. The elitism approach of NSGA2 is through a selection operator that creates a mating pool by combining child and parent populations, and selecting the best (with respect to fitness and spread) N solutions. In SPEA2, PESA2 and PAES elitism is present through an archive of non-dominated solutions. This elitism ensures that the 'good' solutions of the population are not lost, thereby creating a selection pressure towards the global Pareto front.

6.3.4 Solution representation

Three of the optimisation challenges in section 6.2 are related with how a solution is represented, handled and updated by the optimisation framework:

- A solution should meet the different requirements of each optimisation stage (solution representation),
- A solution should be able to accommodate designs of different sizes (solution size)
- A solution should not restrain a design in terms of process patterns (design flexibility).

The framework addresses these challenges by *transforming* and *updating* the solution according to the different requirements of the optimisation process. During this process, the framework ensures the transformation of the solution is consistent and occurs through formulated processes (i.e. the PCA algorithm). The different forms that a solution can take are based on the proposed business process representation. Figure 6.7 shows the different forms of the solution across the different stages of bpo^F.

The initial form of the solution is the N_d set of participating tasks (also depicted as a single-dimensional array). The random population is created in this form. The sequence of tasks in the set does not matter since it is the PCA that attempts to arrange them in a process design. The second step ('check constraints') is the step that the solution changes the most. In the beginning, the solution takes the shape of TRM in order to reflect the relationship of the tasks in the solution with the available resources. This is necessary for the execution of PCA. PCA transforms TRM into a process graph in order to measure the DoI of the solution but also uses the graph to update the N_d set with the tasks that actually participate in the design. Therefore in step 2, the solution from TRM is transformed to a graph and then to the updated N_d set in order to check the remaining optional constraints. PCA is allowed to create a graph with fewer tasks than those in the original solution, provided the graph meets the process requirements. In the case that the updated solution contains fewer tasks, the solution set preserves its original n_d size filling the remaining of the set with the element '-1' to denote the absence of tasks. This helps in having solutions with variable size but keeping fixed the size of the set for the remaining genetic operators.

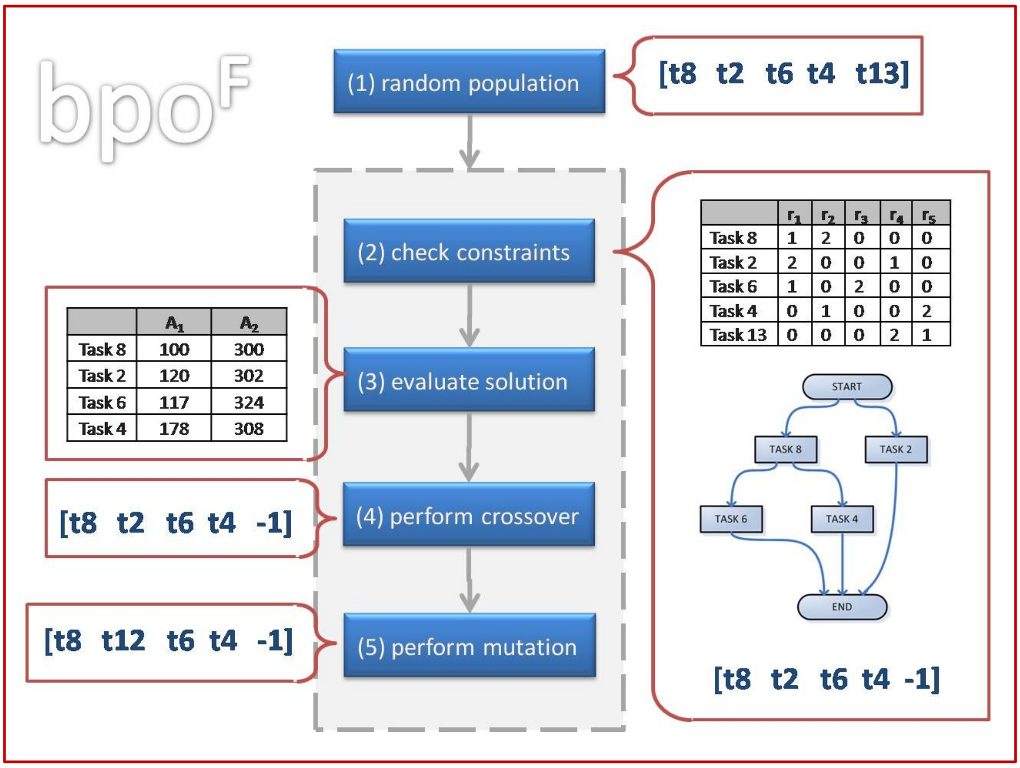


Figure 6.7. The different forms of a solution across the stages of bpo^F

The selection operator in step 3 evaluates the solution by transforming it to TAM based on the updated N_d set. TAM helps in calculating the process attributes (optimisation objectives) and thus in selecting the fittest solutions. Crossover occurs in the N_d set of the solution. The actual size of solution is not an issue since two parent solutions exchange information based on the fixed size of the set. Mutation also is performed in the N_d set and it is not affected by the solution size. It is possible during mutation for a '-1' element to be altered with an actual task from the library. This enhances the mutation operator as it not only alters a solution in terms of participating tasks but also in terms of process size.

The transformation of the solution through the different stages of the optimisation process helps in addressing the challenges stated in the beginning of this sub-section. More specifically:

- The solution is transformed to meet the requirements of each stage: TRM and process graph for the constraints, TAM for the evaluation and N_d for random population, mutation and crossover. All these transformations are based on the proposed representation for business process designs.
- The solution as a set of tasks keeps a fixed size in order to undergo all the genetic operators. However, within the set the absence of a task is denoted with '-1' and thus designs with fewer tasks can be accommodated and managed by the framework.
- Finally, the design diversity (in terms of patterns) is not constrained by the chosen representation. Since the sequence or the number of tasks in a solution are not restrained or manipulated by the representation, PCA is able to pick a solution (set of tasks) and compose a business process design by applying the pattern rules.

6.4 Framework implementation

The framework is programmed using the Java programming language. Java was selected because of its object-oriented approach and the large availability of Java libraries – collections of programs that implement various algorithms. The framework was programmed as a combination of three Java libraries. Two of the libraries (jMetal and jGraphT) were open-source and available on-line and the third (Vergidis) was developed for the purpose of the framework. Figure 6.8 shows the relationship between these three libraries based on their main packages.

jMetal library

jMetal is a Java library that implements a variety of EMOAs including those that are employed by bpo^F. jMetal takes full advantage of the capabilities that Java offers and is structured in a way that a problem can be developed as an independent class from the algorithm that solves it. Figure 6.8 shows the main packages of jMetal. As an open-source library, the user can modify or add his/her own Java classes in each of the packages. The EMOAs are under the package 'Algorithms', while the genetic operators are developed under the package 'Operators'. It is in this package that 'process crossover' and 'process mutation' (the genetic operators developed for bpo^F) are programmed and incorporated by the author. Another package is 'Type' that implements the various types of problems to be solved (e.g. Real, Binary). In this package, the class for handling an Integer problem (underlined in figure 6.8) is programmed by the author in order to handle the discrete business process optimisation problem. Finally, jMetal implements a variety of standard multi-objective optimisation problems (e.g. Kursawe, ZDT, DTLZ) but also allows for custom user-defined problems to be developed. It is in this package that a pointer towards the business process optimisation problem was developed. The pointer directs the execution to the Vergidis library where all the specific business process related components are programmed.

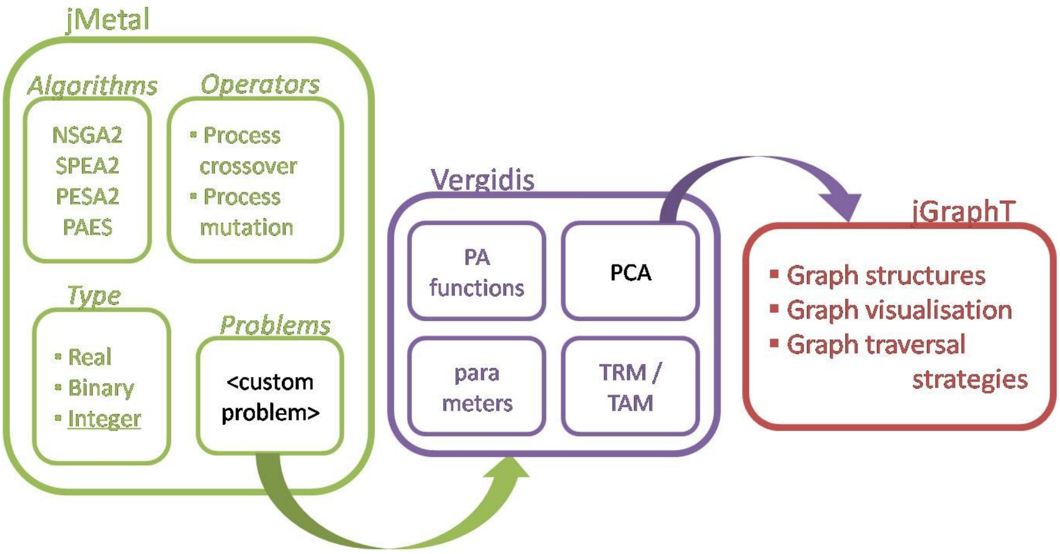


Figure 6.8. The three main Java libraries of the framework

An important reason for the selection of jMetal in the optimisation framework is the fact that the EMOAs in this library are tested for their performance with standard multi-objective optimisation problems. As Durillo *et al.* (2006) suggest, an important issue when programming various EMOAs is their performance under the specific programming language compared to the performance reported by the original EMOA author. jMetal algorithms are tested against their original implementations (e.g. in C) in order to ensure and verify the EMOA performance within the Java programming environment. Durillo *et al.* (2006) report that, unlike other libraries (e.g. PISA), jMetal produces competitive performance results.

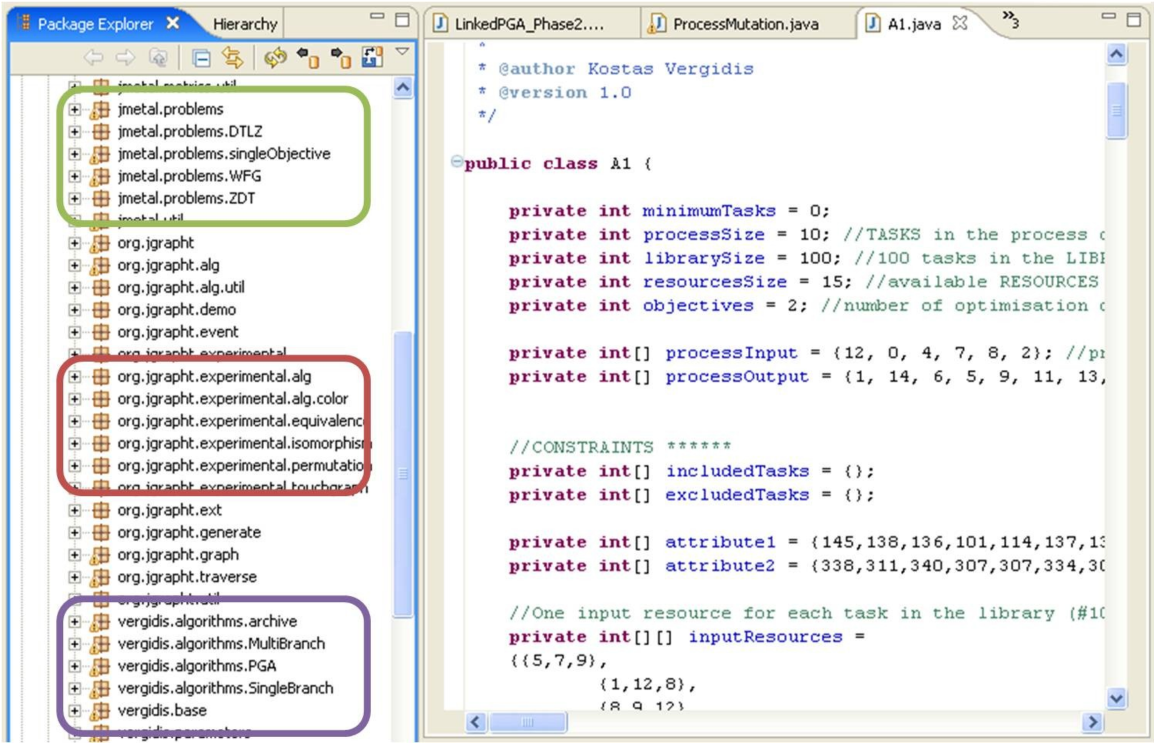


Figure 6.9. Screenshot of the Java programming environment

Vergidis library

The second library of the framework is developed exclusively for the business process optimisation problem. Figure 6.8 shows four main packages developed. The first package 'PA functions' defines the different functions for the process attributes that can be used as objective functions. In its current state the package implements only two aggregation functions of the various task attributes. However, further functions for other process attributes can be easily added and thus work as optimisation objectives making the framework truly multi-objective. The second package is called 'Parameters' and stores classes with various problem parameters (e.g. process size, library size, number of objectives, etc.) in order to test the framework for different problems. Figure 6.9 shows a screenshot of the Java programming environment that on the left-hand side has the three different Java libraries (jMetal, Vergidis, jGraphT) and on the right-hand side has a class of the package 'Parameters' with specific values. These values can be changed manually (by the user) to test the framework for different combinations. The implementation does not impose any restriction in the parameter definition other than the constraint that the parameters should be above zero.

Another package of classes converts a solution to the TRM/TAM representation. However, the central package within the Vergidis library is the one that implements the PCA algorithm. This collection of classes executes the algorithm and produces most of the outcomes of the framework. PCA converts TRM into a process graph and this is why during its execution a library that manages graph structures is necessary.

jGraphT library

jGraphT is another open-source free Java library that provides graph manipulation. There are three main reasons why the framework employs this library:

1. *Graph structures.* jGraphT provides the appropriate structures to construct (using PCA) and store a business process design as a graph. The library has graph objects that store a graph's nodes and edges. This proves essential when converting a design from TRM to a graph.
2. *Graph traversal strategies.* jGraphT also implements the various graph traversal strategies that PCA requires (depth-first, breadth-first). This capability enables the smooth traversal of the business process design in order to add/eliminate tasks or apply the pattern rules.
3. *Graph visualisation.* Finally, jGraphT provides the appropriate classes (Java applets) to visualise the graph as the outcome of the framework. As the process graph is one of the main outcomes of bpo^F it is crucial to be able to visualise the optimised business process designs at the end of the framework execution.

6.5 Main remarks

This chapter presented bpo^F – the proposed evolutionary multi-objective framework for business process optimisation. The framework employs existing EMOAs as the multi-objective optimisation techniques in conjunction with the representation technique for business process designs. The mathematical expression of the problem formulation gave rise to the optimisation challenges that the framework is required to address. This section discusses the framework's approach towards the optimisation challenges.

Table 6.2 groups the optimisation challenges based on the two main components of the framework: (i) the proposed business process representation and (ii) the various EMOAs.

It also provides the framework's approach towards these challenges. The reason for this classification is that the optimisation challenges related to the representation are concerned mostly with *adjusting* and *tuning* the representation into the framework.

Challenge type	Optimisation challenges	bpo^F approach
Business process related	*-Strategy for solution representation*	-The framework's strategy for solution representation is shown in figure 6.7. The solution is transformed and updated in a consistent way in order to address the requirements of the optimisation process.
	-Solution size	-Solution size is fixed in terms of the array that stores the tasks. However using the array element '-1' for designs with fewer tasks, the framework is able to represent solutions of varying sizes.
	-Degree of Infeasibility	-This equality constraint is checked using the execution of PCA that measures DoI for a solution. The framework adjusts to the updated solution from PCA by also updating the N_d set of tasks.
	-Design evolution	-The framework does not restrain the design evolution at any stage. Using only the set of participating tasks, PCA is able to arrange them in a process design and apply the pattern rules.
	-Framework output	-The framework output is created by executing PCA and then evaluating the generated business process design. Therefore, both the outcomes of PCA and the process attribute values are generated for each solution in the optimisation process.
EMOAs related	*-Nature of the problem*	-The performance of the EMOAs given the discrete nature of the problem needs to be checked with a series of experiments.
	-Constraint handling (Selection of solutions)	-Each of the EMOAs devises a different strategy towards assigning fitness to a solution and managing the constraint violation (e.g. crowded comparison operator, 'strength' approach, region-based selection). These different approaches will have a significant effect on the optimisation results of each EMOA.
	-Multi-objective formulation of the problem	-The problem is formulated as multi-objective with the optimisation of two or more process attributes. The capability of the EMOAs to converge towards a diverse Pareto-optimal front is the key performance criterion.
	-Open to the selected EMOAs	-The framework employed a generic optimisation process customisable towards two directions: (i) the business process problem (section 6.3.2) and (ii) the use of different EMOAs in the optimisation process (section 6.3.3). This provides the capability of testing a series of EMOAs and assessing their performance on a problem.

Table 6.2. The optimisation challenges and the bpo^F approach

However, the challenges related to the EMOAs are concerned with the *performance* of the framework in terms of employing the EMOAs to produce optimised business process designs. The framework utilises the proposed representation in order to address the optimisation challenges. In order to work with designs of different sizes, the framework keeps a fixed size set (equal to n_d) and replaces the redundant tasks with the element '-1'. It also fully utilises the PCA to produce the framework outcomes (visual design, DoI and updated N_d) executing it as part of the infeasibility constraint –and not as part of the objective functions as one would expect in the optimisation process. Another novelty of the framework is that during one generation, the solution is updated and transformed in order to address the challenges of the different optimisation stages. The solution transformation and update occur in a consistent way ensuring the correctness of the solution. Also, the optimisation process works with the participating tasks and not with how they connect to each other. This provides the flexibility to PCA to discover and compose novel process designs during the optimisation process.

The EMOAs employed by the framework also have a series of challenges to address. The discrete nature of the problem in conjunction with its multi-objective formulation can make the process of discovering feasible solutions very hard for the algorithms. Each algorithm has to unfold its own strategy along with its strengths and weaknesses in order to generate the Pareto optimal front of solutions. During the optimisation process, each EMOA has a three-fold task:

1. To identify and preserve feasible solutions
2. To converge these solutions towards optimal by optimising the objectives, and
3. To maintain diversity of the solutions across the Pareto front.

In order to assess the performance of the EMOAs it is necessary to test the framework for different business process scenarios. In order to test the framework systematically, a strategy for generating business process scenarios of varying difficulty needs to be devised. Testing the framework will reveal the capabilities and limitations of the EMOAs in dealing with business process optimisation and can demonstrate the strengths and weaknesses of the methodology followed in this research.

6.6 Summary

This chapter presented bpo^F – the proposed evolutionary multi-objective optimisation framework for business processes. The proposed optimisation approach employs existing state-of-the-art EMOAs. The selection of this optimisation technique was justified by the requirements posed by business processes. The problem formulation raised a number of challenges for the business process optimisation framework such as the solution representation and the constraint handling. The proposed framework employed the representation technique for business processes to effectively address the various challenges. For the proposed framework to be properly assessed and evaluated chapter 7 introduces a strategy for creating experimental business process scenarios and chapter 8 assesses the performance of the framework under different sets of scenarios showing the experimental results.

CHAPTER 7

Generating Experimental Business Process Scenarios

This chapter presents a strategy that generates experimental business process scenarios for the evaluation of the proposed optimisation framework. The proposed strategy will assess the capabilities and limitations of the business process optimisation framework introduced in chapter 6. This chapter starts by stating the purpose of the experimental scenarios and the methodology that is followed for devising the proposed strategy. The main steps of the strategy are described in detail and a sample scenario is generated in order to demonstrate the working of the proposed strategy.

7.1 Purpose and main steps of the proposed strategy

This chapter proposes a strategy for generating experimental business process scenarios that can be used to evaluate the performance of the proposed business process optimisation framework.

An **experimental business process scenario** is a set of parameters –based on the business process problem formulation– that is systematically generated in order to assess one or more aspects of the proposed business process optimisation framework (bpo^F)

7.1.1 Aim of the proposed strategy

The aim of devising a strategy for generating experimental business process scenarios is to evaluate the *performance* of one or more features of the proposed business process optimisation framework in a systematic way in order to specify the boundaries of the proposed research.

7.1.2 Approaches in assessing the framework performance

The performance of the optimisation capabilities of the proposed framework can be assessed through:

- The capability to discover and preserve *feasible* solutions (DoI = 0).
- The capability to identify the *optimal* feasible solutions based on the optimisation objectives (convergence of solutions).

- The capability to identify solutions on all the available –based on the problem parameters– process sizes (diversity of solutions).

7.1.3 Main steps of strategy formulation

In order to devise a strategy that generates experimental business process scenarios in order to test the proposed optimisation framework, four main steps are necessary. Each of these steps is briefly discussed below:

(1) Exploration of the search space of the problem

The first step for assessing the performance of the proposed optimisation framework is to investigate the search space of a typical business process optimisation problem and explore its basic characteristics. The search space of the problem can provide an initial guide to the challenging issues surrounding the problem and can affect the performance of the framework in terms of locating the optimal solutions.

(2) Specification of the problem features

Visualising the search space also contributes to specifying the problem features that need investigation in order to define the boundaries of the proposed optimisation approach. The problem features are characteristics of the problem that can pose challenges to the business process optimisation framework in terms of generating optimised business process designs.

(3) Identification of the corresponding parameters and their effect on the problem

The specified problem features correspond to specific problem parameters as those are expressed in the problem formulation. The effect of these parameters on the problem complexity needs to be investigated. Tuning these parameters appropriately will result in business process scenarios of varying complexity and therefore the boundaries of the proposed framework can be tested using a series of different scenarios.

(4) Introduction of control parameters

Each of the main problem parameters can be linked to a control parameter. The control parameters can measure the effect of a problem parameter on the complexity of a scenario. Classifying the control parameters according to their effect can provide a guide for creating scenarios based on the combinations of specific aspects that need investigation in relation to the proposed framework's performance.

Each of these steps is discussed in detail and elaborated in the following sections. Section 7.5 introduces the strategy to generate experimental business process scenarios for testing the proposed optimisation framework based on these steps.

7.2 Investigation of the problem search space

The first step towards a strategy for generating tuneable business process scenarios is the exploration of the problem search space. This section investigates the search space of the business process optimisation problem as formulated in chapter 6 (section 6.1). Visualising the search space is important in order to understand which parameters affect the quality and quantity of the solutions. Also, this section demonstrates an algorithm that performs a large scale search in order to identify the search space of the problem. This is necessary in order to compare the solutions generated by the different EMOAs in each experimental scenario.

7.2.1 Generic shape of the search space

The objectives of the business process optimisation problem are concerned with the minimisation or maximisation of the various process attributes (PA). The process attribute values are calculated based on the attribute values of the tasks (TA) that participate in the process design. The proposed framework supports aggregation functions of the attributes values, e.g.

$$PA_j = \sum_{i=1}^{n_d} TAM_{ij} \qquad \text{(Equation 7.1)}$$

The framework aims to discover business process designs with optimised attribute values. In order to determine the generic shape of the problem search space, two assumptions are put forward: (i) the problem formulation involves the optimisation of *two* objectives, and (ii) for each task attribute *a*, each task in the library takes specific values based on the uniform distribution given the a_{min} and a_{max} values. These two assumptions give to the problem a bi-objective focus and to each task attribute an upper and a lower boundary. Based on these assumptions and assuming that both the process attributes are calculated based on equation 7.1, figure 7.1 shows the generic shape of the search space for the business process optimisation problem (for business process designs with *n* to *n+2* participating tasks).

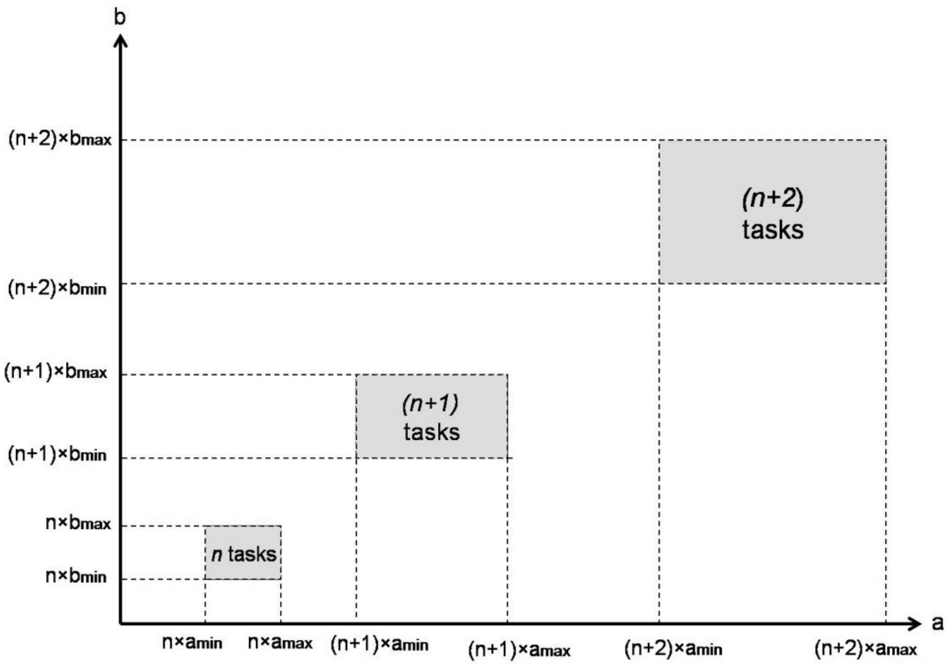


Figure 7.1. Generic shape of the search space

The search space consists of a number of 'islands' where each island accommodates business process designs with the same number of participating tasks. The number of islands depends on the minimum and maximum number of tasks that are allowed in a business process design. In the case that the problem formulation allows for a specific number of tasks in the design, the feasible region contains only one island. The boundaries of each island are determined by the minimum and maximum values of the task attributes. In figure 7.1 it is assumed that there is no overlap among the regions although it might not be the case as discussed later in this chapter. The density of solutions inside each region is determined by the feasibility probability as determined in Appendix B. Based on figure 7.1 it can be concluded that the shape of the search space is:

- *fragmented*, due to the different islands of solutions based on different number of tasks in the process design, and
- *discrete*, due to the fact that in each island the solutions are not produced by a continuous function, but represent feasible combinations of tasks generated by the PCA.

These two features of the search space introduce additional complexity to the business process optimisation problem posing the challenge of discovering the optimal solutions across all the regions. As the number of tasks in the design increases, so does the size of the corresponding island of solutions, since its boundaries become wider. It can be assumed that the number of feasible solutions also increases due to the increased number of task combinations, although that also depends on the problem parameters and constraints.

Based on the generic shape of the search space shown in figure 7.1, it is decided to focus on min-max optimisation problems; problems that aim at minimising the first attribute and maximising the second. The reason for this choice is that a min-min problem would focus on solutions in the island closer to the axes centre, and a max-max problem would focus on solutions on the uppermost island in the search space. Working with min-max problems would prose an additional challenge to business process design optimisation as it will have to locate solutions in all the available islands in the search space. Thus the optimised solutions will involve business process designs with all the available process sizes offering a variety to the business process analysts.

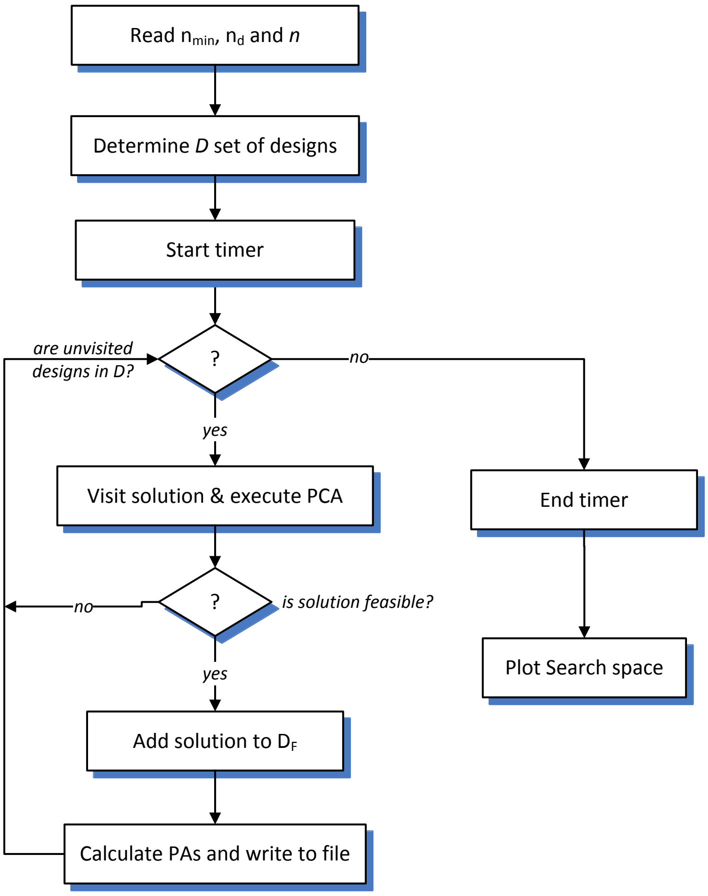


Figure 7.2. The main steps of the Large Scale Search Algorithm (LSSA)

7.2.2 The Large Scale Search Algorithm (LSSA)

The Large Scale Search Algorithm (LSSA) aims to identify the majority of the solutions for a given business process scenario and thus the shape of the problem search space. The algorithm identifies all the possible business process designs for a given scenario and checks one by one whether they correspond to a feasible solution. Figure 7.2 shows the mains steps of LSSA. Initially the algorithm reads all the parameters that are related with the process size: n_{min} – minimum size of the process design, n_d – maximum size of the process design and *n* – size of the task library. Based on these parameters, the size of the D set (exhaustive set of business process designs – see appendix B) and all the possible combinations of tasks are calculated. Determining the search space is a computationally expensive process and a timer is used in order to acquire a precise picture of the duration it takes.

LSSA visits one by one all the solutions stored in *D*. For each solution it applies the PCA to determine its feasibility. In the case that a solution is feasible, it is inserted in the D_F set (the set of feasible solutions). The process attribute values of the particular solution are calculated and stored in a separate file. After all the potential solutions in *D* are visited, the algorithm plots the attribute values of the feasible solutions in order to visualise the search space of the problem. The timer showed that on average LSSA takes 23 hours to produce the search space of a business process optimisation problem.

Parameter	Value	Description
n	100	Number of tasks in the library
n_d	10	No. of tasks in the design
n_{min}	7	Minimum number of tasks in the design
r	20	No. of available resources
t_{in} / t_{out}	3	No. of task input/output resources
r_{in} / r_{out}	5	No. of process input /output resources
p	2	No. of task/process attributes *(α, β)*
α	100 – 110	First task/process attribute ($\alpha_{min} - \alpha_{max}$)
β	200 – 220	Second task/process attribute ($\beta_{min} - \beta_{max}$)

Table 7.1. Problem parameters for the business process scenario

7.3 Main features and corresponding problem parameters

The next step towards a strategy for generating tuneable business process designs is the specification of the problem features and the corresponding problem parameters that affect these features. The identification of the problem features and main parameters is necessary as the proposed strategy will include different combinations of values for each of the parameters in order to assess the different features of the proposed optimisation framework.

The features of the problem are tightly related with the attributes of the problem search space and the optimisation goals of the EMOAs employed by the framework. In order to obtain satisfactory results, the EMOAs need to achieve two goals: (i) convergence to the Pareto-optimal front (in order to obtain optimal business process designs) and (ii) maintenance of the population diversity across the front (in order to obtain a variety of different sizes of business process designs). Additionally, sub-section 7.1.2 sets the measures of performance for the framework which, apart from diversity and convergence across the Pareto-optimal front, stress the ability of the EMOAs to identify solutions that correspond to feasible business process designs. Based on these performance measures, the features of the problem that require further investigation are three:

A. The number of feasible solutions in a given business process scenario
B. The different acceptable process sizes of a feasible business process design, and
C. The ranges of the task attribute values.

Each of these problem features is equally important as it is related with the performance goals of the framework and the EMOAs. The number of solutions in a business process scenario (feature A) can affect the ability of the EMOAs to discover feasible solutions and converge to the optimal. In scenarios with scarce feasible solutions the task of discovering these solutions might prove challenging. Therefore it is essential to test whether the framework can deal with problems in which there are not many alternative designs. Also, in the opposite case of abundant designs, the framework needs to demonstrate its ability to discover the fittest solutions based on specific objectives.

The framework also needs to demonstrate that it is able to discover optimal designs with different process sizes (i.e. different number of tasks in the design). A business process scenario might test the framework for optimised business processes with large number of tasks in the design. The second feature of the problem is important in order to assess the framework's capability of locating optimal solutions across all the process sizes of a given business process scenario. This feature puts to test both the convergence and diversity capabilities of the EMOAs. The optimisation algorithms need not only to discover feasible solutions across all the islands of the feasible region, but also to converge to the optimal.

This feature will determine the framework's limits in terms of dealing with ranges of process sizes.

The last feature –which is related with the shape of the search space– tests the framework's ability to maintain the diversity of optimal solutions across the different islands. The shape of the entire search space is determined by the size, shape and distance of the different islands. These characteristics of the islands are directly related with the boundaries of the task attribute values. Large, overlapping or distant islands can pose a significant challenge to the framework's attempt to maintain a diverse Pareto-optimal front. Testing for this feature will determine the ranges of task attribute values that the framework can effectively deal with and locate the Pareto-optimal front.

No	Parameters	Relate with	Effect on search space	Hinders	Description / Example
1	n_{min}, n_d	feature B	Number of neighbouring islands	Diversity (Convergence)	n_{min} = 3 and n_d = 10, result in 8 feasible islands in the search space.
2	S_d	feature B	Number and continuity of islands	Diversity (Convergence)	S_d = {3, 5, 8} results in 3 discontinuous islands in the search space.
3	TA_i	feature C	Size / Distance / Shape of islands	Diversity (Convergence)	Various min and max boundaries of the task attributes affect the size, distance and shape of the islands in the search space.
4	n	feature A	Density of solutions per island	Diversity & Convergence	The size of the task library affects the density of solutions per island.

Table 7.2. Main parameters of the business process optimisation problem

The problem features can significantly affect the performance of the proposed business process optimisation framework. The problem features are related with specific problem parameters. Table 7.2 presents an overview of the main parameters that are identified as challenges for the performance of the proposed business process optimisation framework. The problem parameters form the basis for the construction of tuneable business process scenarios in order to assess the features of the proposed optimisation framework and investigate its boundaries. For each parameter, its effect on the search space is identified along with the potential challenges that can arise. The next section examines the problem parameters and introduces a control parameter for each.

7.4 Introduction of control parameters

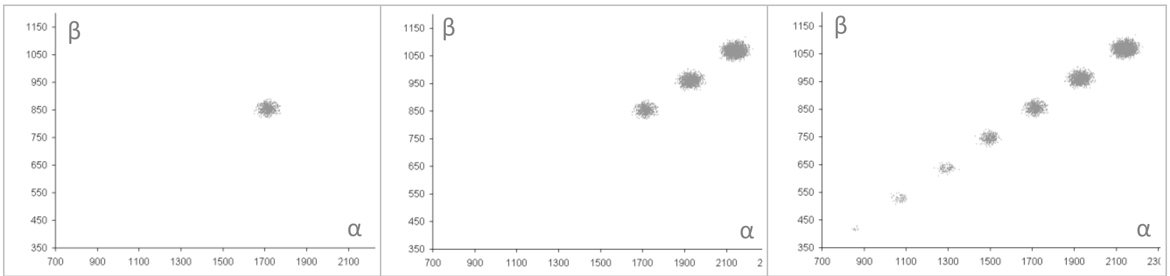
This section links each of the problem parameters with a control parameter. The control parameters are introduced so that they can measure the effect of a problem parameter on an experimental business process scenario. This section concludes with a summary of the control parameters introduced and an overview of their various classifications. The classifications of the control parameters were determined based on initial experiments with the proposed optimisation framework as detailed in Appendix C. The classification of the control parameters is an essential part of the proposed strategy for generating tuneable experimental scenarios and testing the proposed optimisation framework.

7.4.1 Number of neighbouring islands

The number of tasks in the business process design is defined by the n_{min} – minimum tasks in the design and n_d – maximum tasks in the design. Based on these two parameters, L the number of *neighbouring islands* in the search space can be defined.. Based on that, L equals:

$L = n_d - n_{min} + 1$ (Equation 7.2)

Figure 7.3 shows three examples of the search space as a result of different n_{min} and n_d values. In the first instance, $n_{min} = n_d = 8$ tasks. This results to a single feasible island as the problem's search space (L = 1). In the second case, $n_{min} = 8$ and $n_d = 10$ tasks, which provides L = 3 neighbouring islands for business process designs with 8, 9 and/or 10 tasks. Finally, figure 7.3 (c) shows 7 neighbouring islands as the result of $n_{min} = 4$ and n_d = 10 tasks which create a search space that can accommodate business process designs with L = 7 different process sizes.



(a) single island (L = 1) | (b) 3 neighbouring islands (L = 3) | (c) 7 neighbouring islands (L = 7)

Figure 7.3. Search space with varying numbers of neighbouring islands (L)

The number of neighbouring islands can hinder the convergence and diversity of the population with respect to the Pareto-optimal front. Business process scenarios that include a significant number of islands (large L values) are expected to pose a greater challenge to the EMOAs in locating the optimal solutions. The optimisation framework maintains in each generation solutions that correspond to all the different process sizes and thus reside in all the islands in the search space. In the case of significant number of islands this means fewer solutions per island are evolved in each generation. Thus the convergence towards the optimal solutions might prove challenging especially in islands with sparse population of solutions.

The aim is to provide a guide on what could be characterised as 'small' or 'large' number of neighbouring islands in the context of the proposed optimisation framework. A series of initial experiments with the optimisation framework was used as a guide for classifying the different levels that L can take (see Appendix C). These experiments provided a strong indication of the impact of the number of neighbouring islands on the framework performance. Based on these, L can be classified as:

- *low*, for 0-4 neighbouring islands,
- *moderate*, for 5-9 neighbouring islands,
- *large*, for 10-14 neighbouring islands, and
- *very large*, for more than 15 neighbouring islands

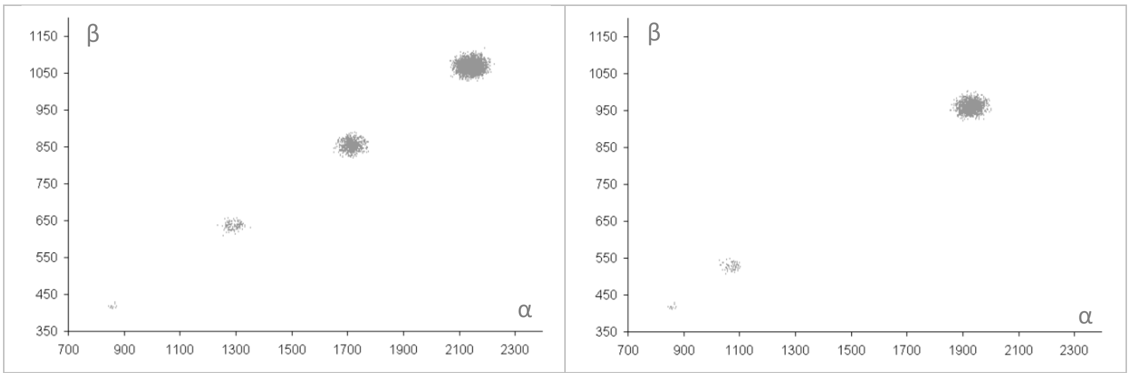
7.4.2 Number and Continuity of islands

One of the optional constraints in the problem formulation defines S_d as a set that contains different acceptable process sizes. In the case that this constraint is included in the problem, a process design is considered feasible only if its process size belongs to S_d. The process sizes in S_d can be equal to or less than the maximum allowable process size n_d. The cardinality of S_d defines the number of islands in the search space. The different process sizes stored in S_d can result in a search space that has *discontinuous* or *non-neighbouring* islands. These islands have more than ± 1 task distance from each other without any neighbouring islands in between. Figure 7.4 provides two examples of discontinuous islands in the search space.

We define D as the *average distance of the islands* included in S_d. This is useful measure in order to assess the discontinuity of the process sizes in a given S_d set. The calculation of D involves the following stages:

1. the different process sizes in S_d are placed in ascending order,
2. the distances in between the islands are calculated and aggregated,

3. the sum of the distances is divided by the number of distances (s_d -1) between the various islands.



(a) 4 discontinuous islands (D = 2)

(b) 2 continuous and 1 discontinuous island (D = 3)

Figure 7.4. Search space with discontinuous islands

The calculation of D is demonstrated using two examples. Figure 7.4(a) shows the search space for $S_d = \{4, 6, 8, 10\}$ (step 1 – S_d process sizes in ascending order). The search space consists of 4 discontinuous islands each having distance of 2 tasks from the next which provides an aggregate of 6 (step 2 – calculation and aggregation of the in-between island distances). Based on that, the islands have average distance, $D = \frac{6}{4-1} = \frac{6}{3} = 2$.

Figure 7.4(b) shows three islands with larger distances. The search space is a result of $S_d = \{4, 5, 9\}$ that creates two neighbouring islands for process designs with 4 and 5 tasks and a non-neighbouring island for designs with 9 tasks. The sum of distances is 1 + 4 = 5 and (s_d -1) = (3 – 1) = 2. Thus, the three islands have an average distance of $D = \frac{5}{2} = 2.5$. In the case that S_d contains only neighbouring islands, D equals to 1. D is an indication of the discontinuity between the feasible islands but should always be considered in conjunction with the number of islands in the search space.

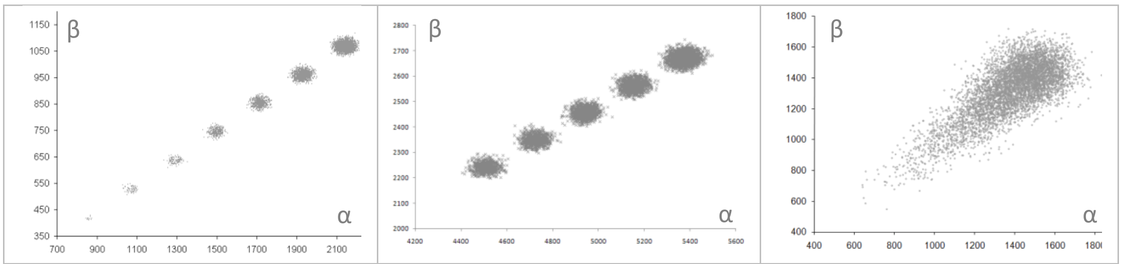
The number and continuity of feasible islands can affect the capability of maintaining diversity of the optimisation algorithms especially in the case of distant islands. During the optimisation process, the framework can discover solutions of different sizes (that belong to different islands) using operators such as process crossover and process mutation. Solutions that belong to neighbouring islands are easier to discover because of their ±1 task difference. In the case that there is a discontinuous island, the framework might not be able to modify the solution size drastically enough to discover solutions in that island and thus may limit itself to neighbouring islands. This constraint can also affect the convergence of solutions towards optimal in the case of significant number of discontinuous islands. The proposed optimisation framework might discover feasible solutions in most of the islands of the feasible region but might not be able to push towards the optimal in all of them.

The aim is to investigate the framework's performance for average distance of islands larger than 1. This would provide a guide to the tolerance of the framework for a fragmented search space where the islands are scattered across. The result of initial experiments (see Appendix C) provided a classification that characterises the different values that D can take. Based on these, the average distance between the islands in the search space (D) can be classified into three categories:

- *short*, for average distance between 1 and 1.5,
- *moderate*, for average distance between 1.5 and 3, and,
- *distant*, for average distance above 3.

7.4.3 Size / Distance / Shape of islands

The task attribute values have a significant effect on the *size*, *shape* and *distance* of the islands that constitute the search space. As *size* of the island we define the size of the area in the search space which accommodates solutions with a particular process size. *Shape* of the island is the form that an island can take based on the proportion of its two attribute values. Finally, *distance* between two neighbouring islands is the distance of their two areas based on their attribute values. The size and the distance of the islands are closely interrelated as larger islands will have shorter distance from each other or might even overlap. The shape of the islands is based on the ratio between the two attribute values.



(a) small-sized islands | (b) medium-sized islands | (c) full overlap

Figure 7.5. Search space different size and distance among the islands

As mentioned earlier, each task receives a value for each of its attributes based on a uniform distribution and a given set of minimum and maximum boundaries for each attribute. Business process designs that have the same number of tasks (i.e. process size)

belong to the same island (see figure 7.1). The dimensions of the island grow as the process size (n_d) increases but it also depends on the difference between the min-max values of the task attributes. Attributes with larger boundaries result in larger islands where the solutions are more spread. In the case of small difference between the min-max boundaries, the feasible solutions in an island are restrained to a small area (figure 7.5.a).
For a task attribute α we define:

- α_{min} – the minimum value of the attribute
- α_{max} –the maximum value of the attribute, and
- $\alpha_d = \alpha_{max} - \alpha_{min}$

Based on these, for two process attributes (α, β), the size of an island (I_{size}) of process size n_d equals with the area of the rectangle that is created (see figure 7.1):

$$I_{size} = \alpha_d \cdot \beta_d \cdot n_d^2 \qquad \text{(Equation 7.3)}$$

Smaller islands pose a challenge for convergence and diversity as the optimal solutions are difficult to locate and maintain. The size of the islands as defined by equation 7.3 also affects the distance between two neighbouring islands. The distance ($I_{distance}$) between two neighbouring islands (based on attribute α) with (n_d-1) and n_d tasks respectively equals with the distance between two rectangle areas (see figure 7.1)

$$I_{(n_d-1),n_d} = n_d \cdot \alpha_{min} - (n_d - 1) \cdot \alpha_{max} \qquad \text{(Equation 7.4)}$$

Two islands *overlap* when at least for one attribute $I_{distance} < 0$. In the case that all the attributes for all the islands in the search space overlap, there is a *full overlap* and the search space consists of a unified area instead of distinct islands (figure 7.5.c). Overlapping islands affect the shape of the Pareto-optimal front and can pose a challenge for the diversity of solutions. The reason for that is that in overlapping areas particular process sizes might dominate over others and thus the framework might ignore the latter. This would result in a Pareto-front that lacks variety and does not adequately reflect the design possibilities in terms of different process sizes. Measuring the potential overlap of the islands in the search space can help define the complexity imposed on the problem based on the size and the distance of the islands. Therefore, the *degree of overlap* (λ) measures the percentage of the overlapping islands against the maximum overlapping regions and equals:

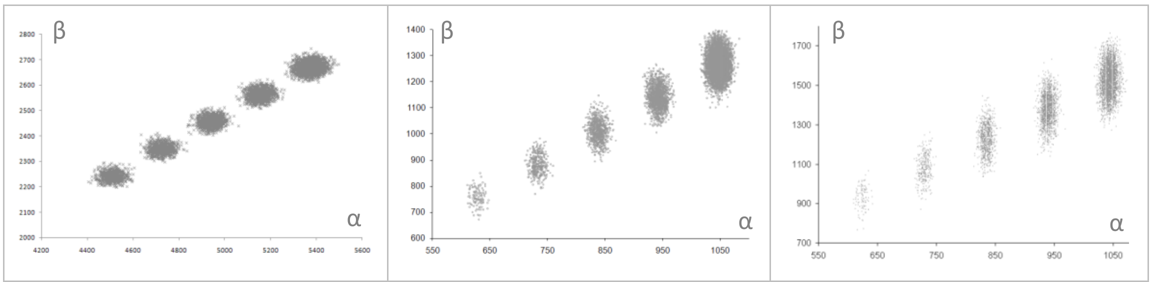
$$\lambda = \frac{L^*}{L-1}\% \qquad \text{(Equation 7.5)}$$

L -1 are the maximum overlapping regions between L neighbouring islands. L^* is a measure of the overlapping regions, regions for which $I_{distance} < 0$. Based on initial experiments cited in Appendix C, λ can be classified as:

- *no overlap*, for values between 0 and 0.2,
- *medium overlap*, for values between 0.3 and 0.5,
- *dominant overlap*, for values between 0.6 and 0.9 and,
- *full overlap*, for values equal to 1.

Finally, the ratio of the two optimisation attributes affects the shape of the islands and the Pareto-optimal front. We specify the ratio of the two attribute values as μ equal:

$$\mu = \frac{\alpha_d}{\beta_d} \qquad \text{(Equation 7.6)}$$



(a) round-shaped islands	(b) oval-shaped islands	(c) bar-shaped islands

Figure 7.6. Search space for different size, distance and shape of the islands

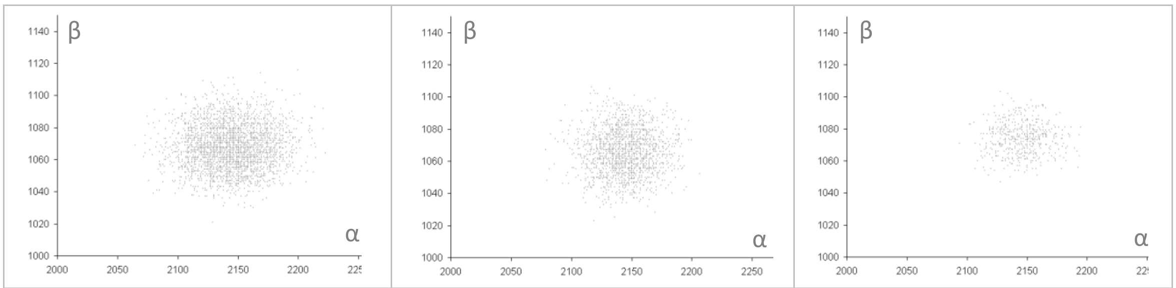
Figure 7.6 shows how the shape of the islands can be affected by changing the ratio of the two process attributes. The shape of the islands can affect the convergence and diversity of the solutions with respect to the optimal front. Initial experiments (in Appendix C) assessed the problem for different values of μ in order to classify the ranges that the parameter can take based on its effect to the search space. Based on these, μ can be classified as:

- *normal*, for values between 1 and 3,
- *challenging*, for values between 3 and 6,
- *hard*, for values between 6 and 10, and,
- *extreme*, for values above 10.

7.4.4 Density of solutions per island

The density of solutions corresponds to the number of feasible business process designs per island. It is mainly affected by two parameters: (i) the library size and (ii) the feasibility constraints (process requirements, etc.). Appendix B demonstrates that even for large library sizes, the feasible designs are only a small percentage of the total number of potential designs due to strict feasibility constraints. This calculation is based on the probability of a design being feasible multiplied by the number of potential designs (based on the library size). Assuming that the probability of feasibility is constant, the effect of the library size on the search space is further investigated. The library size is a central concept of the proposed business process representation and optimisation framework. Defining an acceptable and/or desirable library size for the optimisation framework to unfold its full potential is considered as an important aspect of this research and its orientation towards business processes with real-life elements where large libraries of tasks might not be at hand. In order to further investigate the relationship between *n* and n_d, we introduce γ – the ratio between tasks in the library and tasks in the process design.

$$\gamma = \frac{n}{n_d} \qquad \text{(Equation 7.7)}$$



(a) high density island (n = 100, γ = 10)	(b) medium density island (n = 50, γ = 5)	(c) low density island (n = 30, γ = 3)

Figure 7.7. Single-island search space with varying density of solutions

Figure 7.7 shows an island of feasible process designs with n_d = 10 participating tasks and varying library sizes (*n*), everything else kept same. As the ratio of tasks in the library to tasks in the design decreases (γ), the number of feasible solutions is reduced drastically. The density of solutions per island affects both convergence and diversity of the optimisation results. A low density search space challenges the optimisation algorithms in terms of first discovering feasible solutions and then converging to optimal. Also, having scarce number of solutions can hinder diversity as the framework might trap itself on a local optimum.

The aim is to come up with a classification of the different γ values in order to characterise an experimental business process scenario based on its density of solutions. Below is a classification of the values that γ can take to affect the density of solutions in the search space (based on initial experiments – refer to Appendix C):

- *abundant*, for γ equal to or above 10,
- *satisfactory*, for γ between 5 and 10, and,
- *scarce*, for γ less than 5.

7.4.5 Overview of the classification of control parameters

The previous sub-sections identified the main parameters of the business process optimisation problem and introduced a series of control parameters. This sub-section summarises the classification of the control parameters based on the effect they have on the problem search space. The proposed classification of the control parameters will be used as an integral part of the proposed strategy for generating tuneable business process scenarios. The proposed control parameters are summarised in table 7.3. For each of the control parameters a brief description is provided along with its link to the actual problem parameter(s). The control parameters can be used to create business process scenarios of varying complexity and help assess the framework's optimisation performance under different conditions.

No	*Control Parameter*	Description	*Problem parameter(s)*	Relates with	Affects
1	*L*	number of neighbouring islands	n_{min}, n_d	feature B	Diversity (Convergence)
2	*D*	average distance of islands	S_d	feature B	Diversity (Convergence)
3	*λ*	degree of island overlap	TA_i	feature C	Diversity (Convergence)
	μ	ratio of task attributes			
4	*γ*	ratio of tasks in the library vs. design	N	feature A	Diversity & Convergence

Table 7.3. Control parameters of the business process optimisation problem

Table 7.4 shows the different levels that each control parameter can take based on specific value ranges (taken from Appendix C). This table will constitute an important part of the proposed strategy for generating experimental business process scenarios that is described in the next section. Using this table, one can select different levels of complexity for each of the basic parameters and thus create an experimental business process scenario that focuses on specific aspects of the problem (e.g. low density of solutions). Then, the performance of the proposed business process optimisation framework can be tested and evaluated based on the specific scenario.

L	D	λ	μ	γ
low (0-4)	*short* (1-1.5)	*no overlap* (0-0.2)	*normal* (1-3)	*abundant* (>= 10)
moderate (5-9)	*moderate* (1.5-3)	*medium overlap* (0.3-0.5)	*challenging* (3-6)	*satisfactory* (5-10)
large (10-14)	*distant* (>3)	*dominant overlap* (0.6-0.9)	*hard* (6-10)	*scarce* (<5)
very large (>15)		*full overlap* (>1)	*extreme* (>10)	

Table 7.4. Summary of the classification of control parameters

Apart from generating experimental scenarios, the classification in table 7.4 can be used to assess the complexity of an existing (real-life) business process scenario. This will be useful in chapter 9 where the proposed optimisation framework will be tested with business process scenarios with real-life elements. For each of these scenarios the proposed classification can point their complexity on specific aspects (e.g. small library size) and thus help in defining the expectations regarding the performance of the proposed optimisation framework providing a more accurate explanation based on the generated optimisation results.

7.5 Framework Evaluation Strategy (FES)

This section presents FES – the Framework Evaluation Strategy for generating the experimental tuneable business process scenarios and testing the performance of the proposed optimisation framework for business processes. FES consists of 9 main steps as demonstrated in figure 7.8. These steps are classified in two phases. Phase I involves the scenario formulation where the scenario goal, the problem features and parameters are specified. Phase I is completed with the full experimental scenario formulation. The second phase involves the testing and evaluation of the optimisation framework. In this phase the search space of the problem is generated, the proposed optimisation framework is executed and the EMOA results are evaluated. Section 7.6 follows the proposed strategy and demonstrates the generation of a sample experimental business process scenario. Chapter 8 focuses on the testing phase of FES and demonstrates the framework evaluation results on a series of experimental scenarios. The remaining of this section details the main steps of FES.

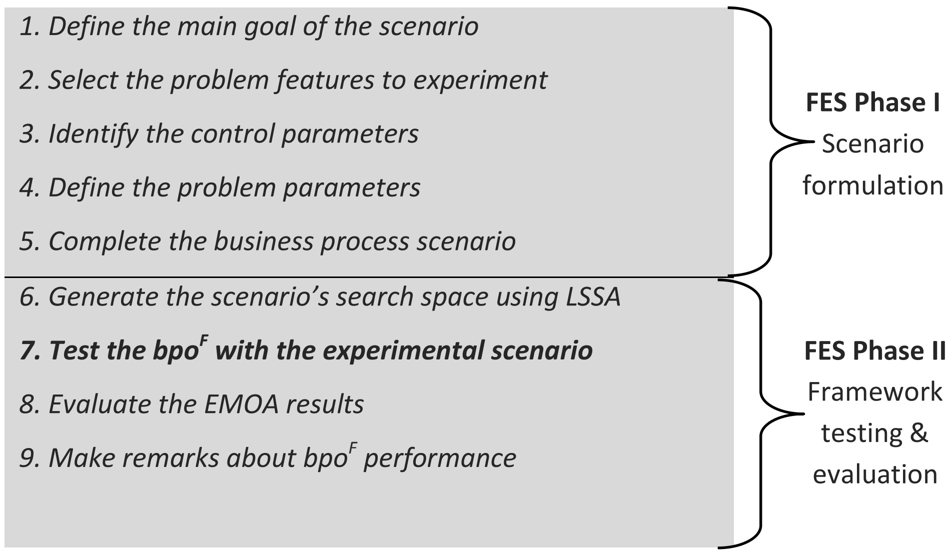


Figure 7.8. The main steps and two phases of FES

(1) Define the main goal of the scenario

The first step of FES is to define the main goal of the experimental scenario –the purpose of creating and testing a particular scenario. The goal is usually linked to one or more of the problem features and can be either generic or specific. A generic goal can be framed as *'to investigate the effect of the island shape on the framework's performance'*. A specific goal can be expressed as *'to investigate the effect of medium overlap degree (λ = 0.4) on the EMOAs convergence capability'*. Based on the nature of the goal, FES might generate different versions of a scenario (also called sub-scenarios) in order to properly assess the goal. The last step of FES is closely linked with the defined goal of the scenario, as the remarks about the framework's performance are drawn based on the extent that the goal is satisfied as demonstrated by the generated results.

(2) Select the problem features to experiment

The second step locates the problem features that the scenario testing focuses on. These features are derived from the scenario goal. A scenario might seek to investigate more than one feature in a combined way. Since each of the identified features affects the framework's performance, one of the main outcomes of the scenario testing is to investigate the effect –and the degree of this effect– of each feature on the problem in order to draw appropriate conclusions regarding the framework's capabilities and limitations.

(3) Identify the control parameters

For each of the problem features and their corresponding parameters, section 7.4 introduced a series of control parameters and presented a classification of different levels for each of these parameters based on a typical business process scenario. This step of the proposed evaluation strategy identifies which control parameters will be tuned in the particular scenario and defines the level(s) –or combination of levels if more than one– of the parameters that need investigation. As a result of this step, different versions of the scenario might be created in order to investigate the framework's performance under different sets of parameters.

(4) Define the problem parameters

The control parameters are related with particular problem parameters. Having defined the levels of the control parameters, specific values can be assigned to the problem parameters. These values will vary across the different sub-scenarios in order to assess the problem in a systematic way.

(5) Complete the business process scenario(s)

The last step of the problem formulation phase (FES Phase I) is to define the remaining parameters for each sub-scenario. The remaining parameters can be defined using the typical business process scenario (in Appendix C) as a guide. The end result is the complete set of sub-scenarios with all their parameters specified. These sub-scenarios form the complete scenario for investigating and testing bpo^F.

(6) Generate the scenario's search space

The first step of the scenario testing and evaluation phase (FES Phase II) is to generate the search space for the scenario. The search space is generated using the LSSA algorithm (sub-section 7.2.2). The generation of the search space provides a clear picture of the problem's boundaries.

(7) Test the bpo^F with the scenario

The next step involves testing the bpo^F with the scenario. This is the central step of the proposed strategy where the formulated scenario is incorporated in the proposed optimisation framework. The framework applies the four employed EMOAs –NSGA2, SPEA2, PESA2 and PAES– and generates optimisation results for each of the algorithms.

(8) Evaluate the EMOA results

The generated optimisation results are then plotted along with the problem search space and evaluated using appropriate metrics (see chapter 8). The metrics assess the performance of each of the EMOAs in generating optimal business process designs according to the scenario requirements.

(9) Make remarks about bpo^F performance

The last step of FES assesses the performance of the framework in comparison to the initial goal. If the goal was a hypothesis, the remarks either confirm or disregard it based on the framework results. The importance of this step is that it defines the framework's capabilities and limitations based on its optimisation performance. It is essential to state a clear goal and generate tuneable scenarios that address the issues raised by the goal in order to reach to conclusions about the framework. The remarks drawn in this stage might call for another set of experimental scenarios or further investigation. They might also point towards the fittest algorithm for a particular business process optimisation problem.

7.6 Generating an experimental scenario using FES

This section creates a sample experimental scenario in order to demonstrate Phase I of FES. It shows how the problem features are elicited based on the scenario goal, how the control parameters are identified and how the problem parameters are calculated. The example concludes with the formulation of the experimental scenario which is compromised of three sub-scenarios. This example aims to demonstrate the process of scenario generation as part of the proposed strategy for evaluating the business process optimisation framework. Chapter 8 will thoroughly demonstrate the second phase of the strategy which is testing a series of generated experimental scenarios in order to systematically evaluate the optimisation capabilities of the proposed framework. The sample scenario follows the first five steps of FES, starting with the scenario goal definition:

(1) Scenario goal

The sample scenario goal is defined as: *'Investigate business process optimisation problems with limited feasible designs and different design sizes'*.

(2) Locate the problem features

The scenario goal focuses on two features of the problem: the number of feasible solutions (feature A) and the acceptable process sizes of a feasible business process design (feature B). Feature A is restrained to *'limited* feasible designs', while feature B is concerned with 'different design sizes' –without the extent of difference of the designs being explicitly specified in terms of process size. Therefore, this scenario provides the opportunity of examining how the framework performs on a limited number of feasible designs scattered across non-neighbouring islands.

(3) Identify the control parameters

The control parameter related with feature A is γ –the ratio of tasks in the library vs. tasks in the process design. It has been previously identified that the number of feasible solutions (or the density of solutions per island) is dependent on the library size. Since the scenario goal is referring to *'limited feasible designs'*, γ –based on the classification table 7.4 will be in the region of 'scarce' ($\gamma < 5$). The control parameter related to feature B is D – the average distance of (non-neighbouring) islands. Since the level of this parameter is not explicitly defined in the scenario, we will investigate the effect across all the levels of D (short / moderate / distant) in order to systematically investigate the effect of low density islands.

	Sub-scenario A	**Sub-scenario B**	**Sub-scenario C**
γ = 4 (*scarce*)	D = 1.5 (*short*)	D = 2 (*moderate*)	D = 3 (*distant*)

Table 7.5. Sub-scenarios based on control parameter classification

At this stage, with the problem features identified and the levels of the control parameters defined, the scenario goal can also be expressed as: '*Investigate the effect of low density islands across scenarios with varying distance of non-neighbouring islands'*. The sub-scenarios created initially are three, each at a level of the D parameter, as shown in table 7.5.

(4) Calculate the problem parameters

The control parameters are related with particular problem parameters. Thus having defined the levels of the control parameters, specific values can be assigned to the problem parameters. Assuming that the maximum process size in all sub-scenarios n_d = 10, n becomes equal to 40. Also the S_d set can be defined based on D for all sub-scenarios. Table 7.6 shows the calculated problem parameters based on the control parameters.

	Sub-scenario A	Sub-scenario B	Sub-scenario C
n_d =10 n = 40	S_d = {4, 5, 6, 8, 10}	S_d = {5, 7, 9}	S_d = {3, 6, 9}

Table 7.6. The main problem parameters for the sub-scenarios

(5) Complete the business process scenario(s)

The remaining parameters can be defined on the basis of the typical business process scenario. Table 7.7 demonstrates the complete set of sub-scenarios with all their parameters specified. These sub-scenarios form the complete scenario to be used fpr investigating and testing bpo^F. In bold are the problem parameters defined as a result of the control parameters. This step completes the first phase of FES, which is the scenario formulation.

Parameter	Sub-scenario A	Sub-scenario B	Sub-scenario C
n	**40**	**40**	**40**
$\boldsymbol{n_d}$	**10**	**10**	**10**
$\boldsymbol{S_d}$	**{4, 5, 6, 8, 10}**	**{5, 7, 9}**	**{3, 6, 9}**
r	20	20	20
t_{in} / t_{out}	3	3	3
r_{in} / r_{out}	5	5	5
p	2	2	2
α	100 – 110	100 – 110	100 – 110
β	200 – 220	200 – 220	200 – 220

Table 7.7. Problem parameters for the sample scenario

The second phase is the focus of chapter 8 where the proposed optimisation framework is tested with a series of experimental scenarios generated using FES. A series of scenarios such as the sample scenario above will be used to assess the capabilities and limitations of the proposed framework in generating optimised business process designs.

7.7 Main Remarks / Summary

This section summarises the chapter and highlights the main remarks. The chapter introduced FES – a strategy to assess the framework that chapter 6 detailed. The strategy largely encompasses the creation of business process scenarios in order to experiment with specific aspects of the framework. These remarks summarise the main contribution of this chapter:

- The search space of the business process optimisation problem advocates min-max problems in order to acquire optimal designs across the available process sizes.
- LSSA is a computationally expensive algorithm that generates the search space for a given business process optimisation problem. It provides a basis for comparing the quality of the bpo^F generated results.
- The business process optimisation problem depends on three main features: (i) the number of feasible solutions, (ii) the available process sizes and (iii) the ranges of task attribute values.
- For these three features, the main problem parameters are identified and a corresponding set of control parameters is introduced.
- FES - The proposed strategy for generating experimental business process scenarios and evaluating the proposed framework is largely dependent on the classification of the control parameters (see table 7.5).
- The classification of the control parameters can play a dual role: (i) as part of the proposed FES for experimental scenario generation and (ii) as part of assessing the complexity of a given (real-life) business process scenario.
- The Framework Evaluation Strategy (FES) consists of two phases: phase I which involves the formulation of the experimental scenario (as demonstrated by the generation of a sample scenario in section 7.6) and phase II which involves the testing and evaluation of the framework utilising the generated experimental scenario.

In order to draw conclusions about the business process optimisation framework that this research is proposing, it is essential to have a systematic testing procedure in place and this is where the main contribution of chapter 7 lies. Chapter 8 will utilise FES generating specific business process scenarios and evaluating the basic features of the proposed optimisation framework.

CHAPTER 8

Performance Evaluation of the Proposed Framework

This chapter evaluates the performance of the proposed optimisation framework that was introduced in chapter 6. The performance evaluation occurs through three experimental scenarios. These scenarios are formulated based on the experimental strategy that was detailed in chapter 7. Each experimental scenario focuses on a particular aspect of the framework in order to investigate its boundaries for the optimisation of business process designs. The performance of the optimisation algorithms is also assessed in order to determine the fittest for a particular problem.

8.1 Purpose of performance evaluation

This chapter investigates the performance of the framework using a series of different experimental scenarios. These scenarios are generated based on the proposed strategy that was introduced in the previous chapter. The aim of the performance evaluation is two-fold:

- To investigate the boundaries of the framework in optimising business process designs, and,
- To assess and compare the performance of the optimisation algorithms and determine the most suitable for the business process optimisation problem.

The overall assessment of the framework's performance will be based on these two evaluation aspects. The first aspect will determine the flexibility of the framework in relation to specific parameters of the business process representation (e.g. process size). This will determine the boundaries of specific business process design parameters that the framework can work with and produce optimised results in an effective way. The selection of the specific parameters to be investigated will be based on the main problem features as identified in chapter 7.

The second aspect will evaluate the performance of the optimisation algorithms in terms of generating optimised business process designs. Each of the algorithms will be tested for all the experimental scenarios and the optimisation results will be compared and evaluated using appropriate metrics. The outcome of this evaluation will determine the suitability of one or more EMOAs for specific cases and overall for the business process optimisation problem.

8.2 Focus of performance evaluation

This chapter tests the framework using three experimental scenarios (A, B and C). The experimental scenarios are generated based on specific problem features. Chapter 7 introduced three basic features of the problem, the main parameters and a series of control parameters. To evaluate the performance of the proposed framework in a systematic way, there is a need to focus on specific aspects of the problem by asking the right questions. These questions are important in determining the limits of the proposed optimisation framework. As a result, a series of experimental scenarios can be generated in a systematic way. These questions are:

? What is the minimum *library size* that the framework can operate with? (scenario A)

? What is the maximum *size of a business process design* that can be optimised? (scenario B)

? What is maximum number of *islands* that the framework can handle? (scenario C)

The rationale behind these questions is as follows: The proposed business process representation introduced the *task library* that allows the composition of equivalent alternative business process designs and thus enables optimisation based on evolutionary algorithms. Scenario A is pre-occupied with the task library as it is important to determine the minimum available number of tasks in the library that the framework can work effectively with. The second scenario is based on the question that seeks to define the maximum size of a process design that can be optimised. This will assess whether the framework can work with large business process designs. Finally, the third scenario investigates the maximum number of islands that the framework can handle. This comes as a result of the framework's capability of generating business process designs of varying sizes; it seeks to determine the boundaries in terms of simultaneous capturing of different process sizes (islands) for specific process requirements.

The three experimental scenarios stem from the questions stated above and are oriented towards evaluating the capabilities of the framework in relation to optimising business process designs. The range of the attribute values of a business process design does not vary in the proposed experimental scenarios and therefore its effects on the search space (e.g. island overlap) and their relation to the performance of the framework are not examined by this research but are left as future work. However, it is also important to evaluate the performance of the optimisation algorithms employed by the framework. Chapter 6 discusses some of the optimisation challenges for the algorithms (table 6.2). Based on these challenges, the results of the experiments in this chapter will seek to answer the following questions:

- Is an evolutionary multi-objective optimisation approach effective for the discrete nature of the business process optimisation problem?
- Is the performance of the EMOAs satisfactory given the multi-objective and highly constrained formulation of the problem?
- From the selected EMOAs, is one or more significantly better in the business process context? Do some EMOAs perform better under specific conditions (e.g. for large process size)?

The first two questions seek to verify whether the employment of EMOAs is appropriate for the business process problem and how effective they are for the particular context. The effectiveness of the EMOAs will be assessed based on specific metrics as discussed in the next section. The last question seeks to identify whether there is a single EMOA that performs better in comparison to the others, overall and/or in particular problems. In such a case, an additional contribution of this research would be the nomination of a particular EMOA as the fittest for business process optimisation.

The questions related to EMOA performance will be also investigated by the three experimental business process scenarios. Each of these scenarios is formulated based on the first set of questions and discussed in a separate section in this chapter (scenario A in section 8.4, scenario B in section 8.5 and scenario C in section 8.6). The optimisation results of each scenario are assessed based on the second set of questions regarding the performance of the algorithms.

8.3 EMOA parameters and performance metrics

This section shows the parameters for each EMOA as used for the experimental scenarios. It also describes the metrics that will be used for the performance evaluation of the framework's optimisation capabilities and the performance of the EMOAs in each of the experimental scenarios.

8.3.1 Parameter specification

Table 8.1 shows the parameters for each of the EMOAs employed by the proposed optimisation framework. The EMOAs are described in detail in Appendix D. The parameters for each algorithm were tuned based on initial experimentation with particular emphasis on the number of generations and the population size for each algorithm. Although 25,000 evaluations might seem excessively high; for most algorithms it helped produce better quality results in comparison with lower numbers (e.g. 10,000, 1,000 and 500).

Parameter	NSGA2	PAES	PESA2	SPEA2
Population	500	-	500	500
Archive	-	1000	500	500
Bisections	-	5	5	-
Generations	25,000	25,000	25,000	25,000
Crossover prob.	0.8	-	0.8	0.8
Mutation prob.	0.2	0.2	0.2	0.2

Table 8.1. Parameter specification for the EMOAs employed in bpo^F

For each experiment, each algorithm is executed for 30 independent runs. The results from one of the typical runs are shown in a diagram and are appropriately evaluated using the metrics discussed in the following sub-section.

8.3.2 Evaluation metrics

To evaluate the performance of Evolutionary Multi-objective optimisation algorithms, a series of metrics have been proposed in relevant literature (Deb, 2001). The majority of the proposed metrics evaluate the optimisation algorithms based on two attributes: (i) the *convergence* and (ii) *diversity* of the optimised solutions on the Pareto-optimal front. The convergence refers to the capability of the optimisation algorithm to discover non-dominated solutions, and the diversity to discovering solutions across the Pareto-optimal front.

A *non-dominated* solution is defined here as one that has better attribute values for all the optimisation attributes compared to the solutions generated as part of the large scale search.

Using the concept of non-domination, the performance of the EMOAs can evaluated based on the number of non-dominated solutions and the time it takes to generate them. Chapter

7 (section 7.2.1) investigated the search space of the business process optimisation problem and showed that it consists of various islands with each representing business process designs of different size. Given the context of the problem (business processes) and the shape of the search space, the evaluation criteria of the optimisation algorithms lie in three categories:

- The *time* it takes to generate the optimisation results,
- The *number of non-dominated solutions* generated by the algorithm compared to the large scale search, and,
- The *number of islands* discovered.

The time element can compare the efficiency of the algorithms in generating optimal solutions. However, it can be used only in conjunction with the quality of the generated solutions in order to provide an accurate indicator of the algorithms' performance. The second criterion, the number of non-dominated solutions, evaluates the convergence capability of the optimisation algorithms in discovering optimal solutions. Calculating the ratio of non-dominated solutions against all the generated solutions of a particular algorithm can provide a good indication of its convergence capability. The third evaluation criterion is focused on the diversity of solutions. In the business process context, the diversity of solutions is assessed on whether the optimisation algorithm is capable of locating non-dominated solutions across all the islands in the search space and thus generating business process designs of all the available sizes. Based on these evaluation criteria, three metrics are put forward:

$M_1 = \frac{non - dominated\ solutions}{unique\ generated\ solutions}\%$	*Success ratio*

This metric is based on Deb's (2001) error ratio. M_1, or *success ratio*, measures the convergence capability of the optimisation algorithms. It measures the non-dominated solutions generated by an evolutionary algorithm (in relation to the solutions generated by the large scale search) against the unique generated solutions that the algorithm produces. The number of unique generated solutions of an evolutionary algorithm varies and can be equal to or less than its population size. This is because in the final population some solutions maybe copied several times. The *success ratio* provides the percentage of an algorithm's solutions that are non-dominated thus providing a measure for its convergence performance in a given problem.

$$M_2 = number\ of\ \ islands\ without\ non-domimated\ solutions$$

M_2 measures the diversity capability of the optimisation algorithms. In the context of the business process optimisation problem, diversity is defined as the capability to discover solutions across all the available islands in the search space. Counting the islands without non-dominated solutions (in relation to the solutions generated by the large scale search) is a straight-forward way of comparing the diversity capability of the algorithms and also assessing the performance of the framework in generating optimised solutions for all the process sizes.

$M_3 = \frac{time\ to\ execute\ the\ algorithm}{non-dominated\ solutions}$	*Time per non-dominated solution*

M_3, or time per non-dominated solution, measures the time taken in minutes to produce a single non-dominated solution. This helps to normalise the time comparison between the different optimisation algorithms in the case they produce different number of non-dominated solutions.

8.4 Experimental scenario A

The first experimental scenario is related with the first question about the task library (see section 8.2) as it investigates the capability of the proposed optimisation framework to generate optimised business process designs for problems with limited library sizes. It is important to investigate the flexibility of the framework in relation to the task library as it is an element that it is not expected to be in real-life business processes. The outcome of this experiment is expected to define the lowest size of the library for which the framework provides satisfactory results. In terms of EMOA performance, limiting the library size can affect the EMOA's capability to converge to optimal solutions. For experimental scenario A, the steps in generating the scenario and testing the framework – as detailed in chapter 7- are elaborated and detailed below:

(1) Scenario goal

The goal of the experimental scenario A is defined as: *'Determine the minimum library size for a business process design such that the framework can generate satisfactory optimisation results'*. Based on this goal, the outcome of this experimental scenario would be the size of the task library (related to the process size) that allows the framework to generate satisfactory optimisation results for a particular scenario. 'Satisfactory' optimisation results are considered those that offer a selection of at least a handful alternative optimised business process designs.

(2) Problem features

The goal of the experimental scenario is related with one of the three problem features as identified in chapter 7. This feature is feature A – the number of feasible solutions of a business process scenario. As shown in the previous chapter, the library size affects the number of feasible solutions of an experimental scenario. By experimenting with different library sizes, essentially we experiment with the number of feasible solutions that the framework can work with.

(3) Control parameters

The control parameter related with feature A is γ – the ratio of tasks in the library vs. tasks in the process design. Essentially, the scenario goal seeks to define the minimum number of tasks in the library for a given number of tasks in the process design. Therefore, this experimental scenario investigates the framework performance for different γ values. By defining the minimum acceptable γ value, we can define the minimum task library size for a given business process design. Table 8.2 shows the γ values for 4 sub-scenarios that will be tested within the framework as parts of experimental scenario A. The γ values for these sub-scenarios are selected based on the classification of the control parameters as defined in table 7.5 (refer to chapter 7 – classification of control parameters). It is expected that the optimisation results both in terms of diversity and convergence will deteriorate as γ decreases.

Sub-scenario A.1	Sub-scenario A.2	Sub-scenario A.3	Sub-scenario A.4
$\gamma = 10$ (abundant)	$\gamma = 5$ (satisfactory)	$\gamma = 3$ (scarce)	$\gamma = 2$ (scarce)

Table 8.2. γ values for each of the four sub-scenarios

(4) Corresponding problem parameters

Having defined the various γ values for each sub-scenario, the corresponding problem parameters can be calculated. As the scenario aims to test the framework for a variety of library sizes, we assume that n_d is constant and equals to 10. Table 8.3 calculates the library sizes for each sub-scenario based on the formula $n = \gamma \cdot n_d$.

Sub-scenario A.1	Sub-scenario A.2	Sub-scenario A.3	Sub-scenario A.4
n = 100	*n* = 50	*n* = 30	*n* = 20

Table 8.3. Task library size (*n*) for each of the four sub-scenarios

(5) Complete the experimental scenario parameters

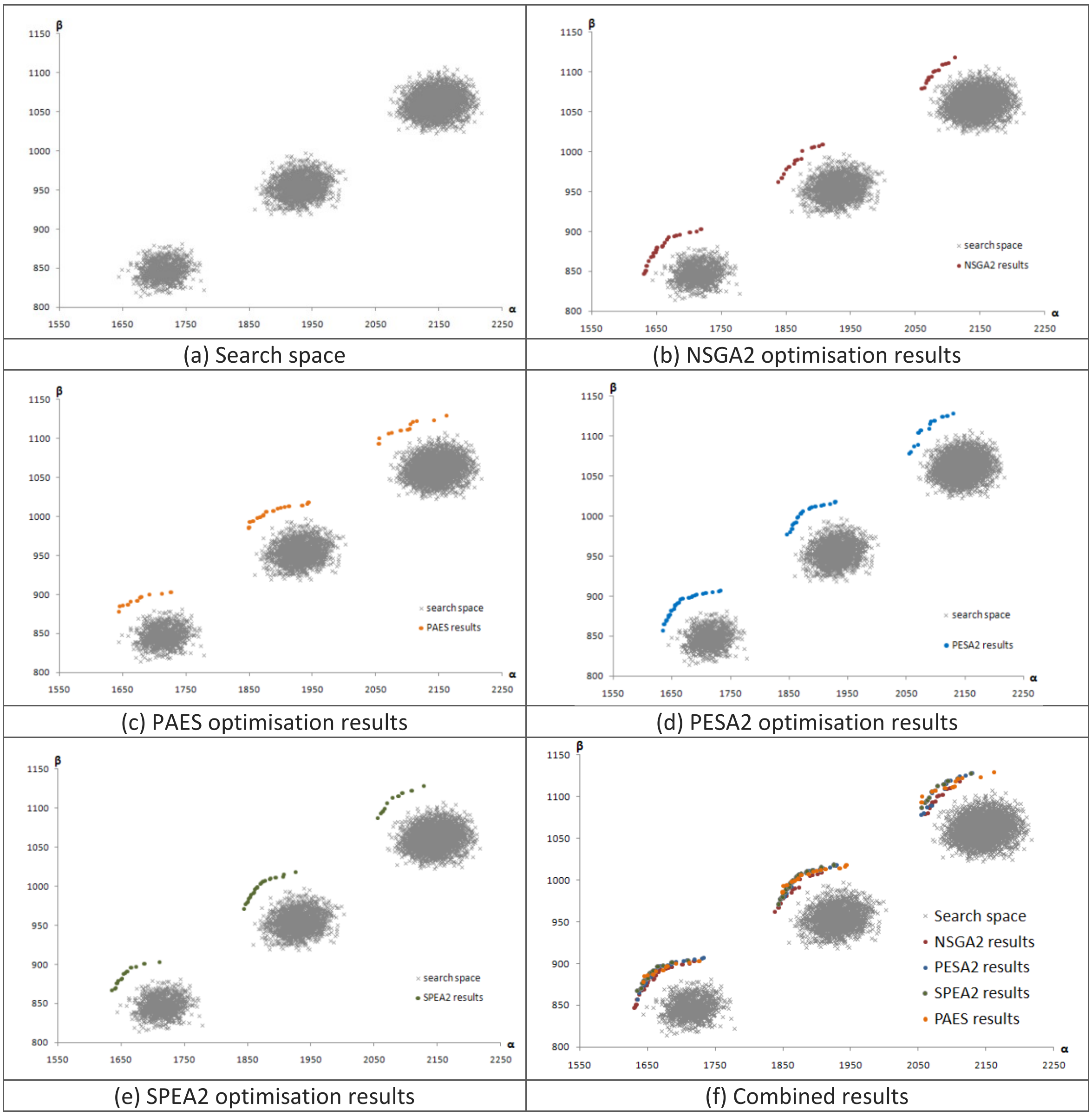
The next step is to define the remaining parameters which are common across the sub-scenarios. Table 8.4 shows these parameters based on the typical business process scenario as introduced in chapter 7. Keeping all the parameters constant for the sub-scenarios allows for the performance of the EMOAs to be dependent only on the size of the task library. Thus conclusions about the framework's performance with varying library sizes can be safely drawn. This completes phase I of the strategy for scenario formulation. Phase II involves the framework testing and it entails the following three steps: (6) generation of the search space, (7) testing the framework and (8) evaluation of the EMOA results. These steps are shown below for each sub-scenario while the remarks about the scenario (step 9) are documented in the end for the complete scenario. These remarks evaluate the performance of the EMOAs based on their unique features and assess the optimisation capabilities of the framework in the context of the specific experimental scenario.

Parameter	Scenario A
n	***defined by each sub-scenario***
n_d	10
n_{min}	8
r	20
t_{in} / t_{out}	3
r_{in} / r_{out}	5
p	2
α	100 -115
β	200 -230

Table 8.4. Remaining problem parameters for the sub-scenarios A.1 - A.4

Sub-scenario A.1

Figure 8.1 shows the search space (a) and the optimisation results (b, c, d, e and f) generated by bpo^F for each of the optimisation algorithms. The search space consists of 3 neighbouring islands with the density of solutions being similar across the islands. The approximation of search space is generated based on the LSSA algorithm. All the EMOAs identified non-dominated solution in all the neighbouring regions of this sub-scenario.



(a) Search space

(b) NSGA2 optimisation results

(c) PAES optimisation results

(d) PESA2 optimisation results

(e) SPEA2 optimisation results

(f) Combined results

Figure 8.1. Optimisation results for sub-scenario A.1

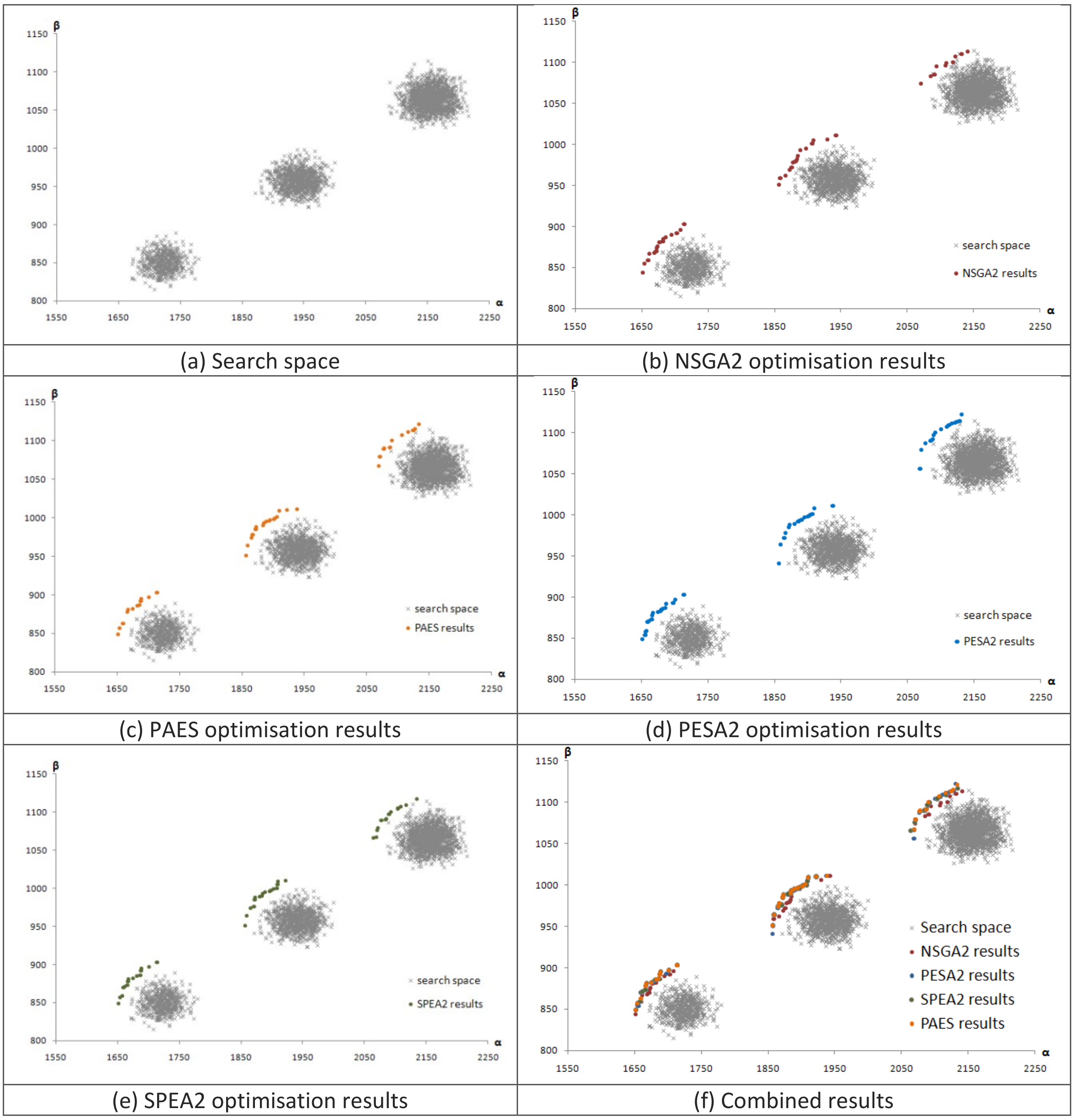
	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	1,147.2	5,021	-	-	-	-
NSGA2	24.6	45	45	100%	0	0.55
PAES	5.2	36	36	100%	0	0.14
PESA2	11.7	54	54	100%	0	0.22
SPEA2	16.7	45	45	100%	0	0.37

Table 8.5. Optimisation data for sub-scenario A.1

Table 8.5 shows the data for the optimisation results. NSGA2 took the most time to execute (0.55 minutes per non-dominated solution) whilst PESA2 produced the non-dominated solutions in less than half the time. All the algorithms achieved 100% success in generating non-dominated solutions in relation to the large scale search. PAES has fastest execution time with 0.14 minutes per non-dominated solution.

Sub-scenario A.2

The results of the second sub-scenario are shown in figure 8.2. The 50% reduction in the library size results in smaller and less dense islands in the search space. The capability of convergence of the optimisation algorithms is more restrained in this sub-scenario.



(a) Search space | (b) NSGA2 optimisation results

(c) PAES optimisation results | (d) PESA2 optimisation results

(e) SPEA2 optimisation results | (f) Combined results

Figure 8.2. Optimisation results for sub-scenario A.2

Table 8.6 shows the optimisation data for the sub-scenario A.2. The first observation is the significant drop in the number of solutions identified in the search space by the large scale search (2,758) which is almost 50% compared to sub-scenario A.1 (5,021). However, the algorithms generate similar number of non-dominated solutions. The success ratio is maintained in 100% despite the increase complexity of the sub-scenario. PESA2 generates the highest number of non-dominated solutions, achieves a 98% success ratio and is the second fastest algorithm. PAES is once again the fastest to generate optimised results.

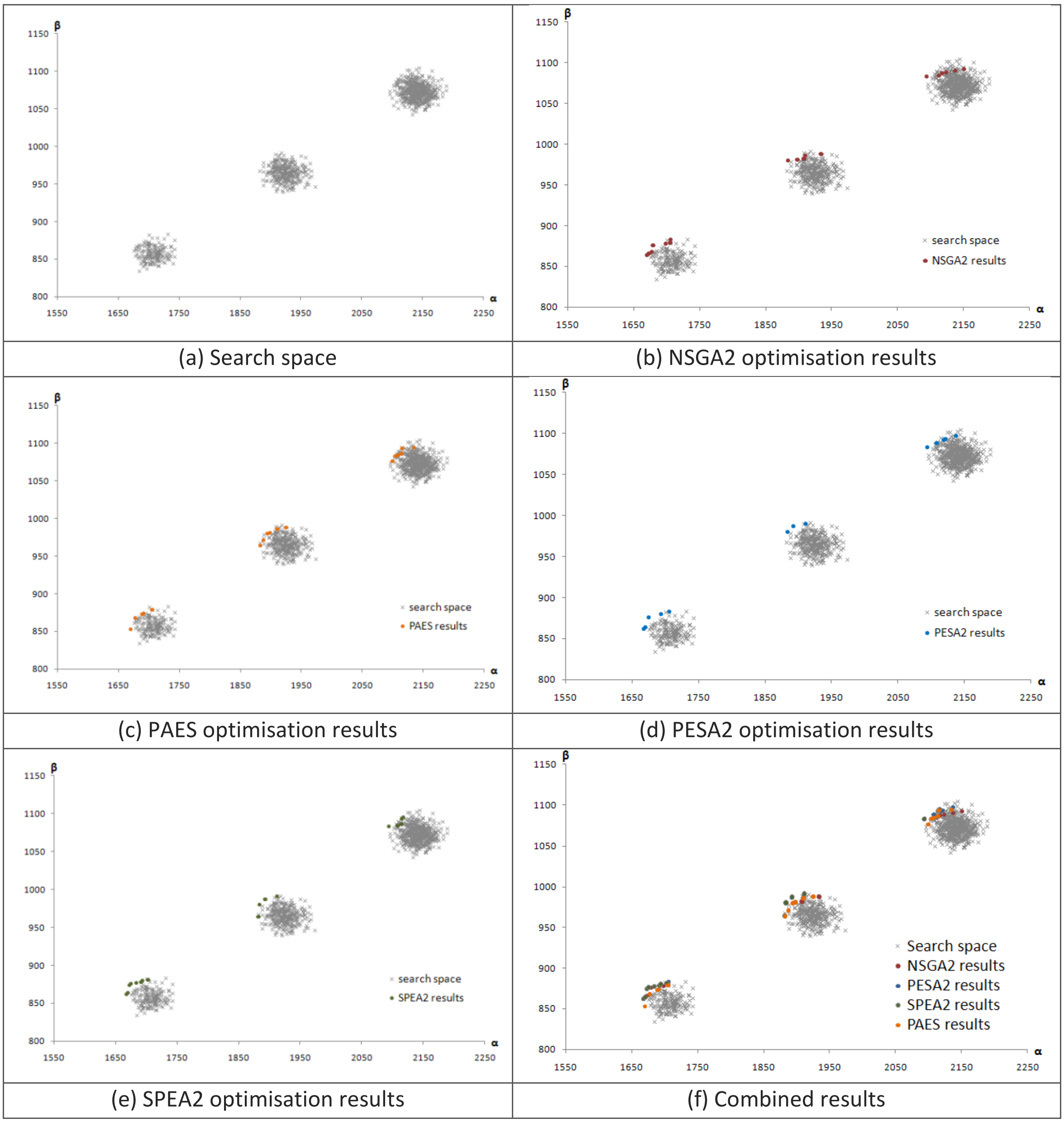
	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	958.7	2,758	-	-	-	-
NSGA2	23.7	40	40	100%	0	0.59
PAES	6.5	38	38	100%	0	0.17
PESA2	17.7	50	50	100%	0	0.35
SPEA2	16.7	30	30	100%	0	0.56

Table 8.6. Optimisation data for sub-scenario A.2

Sub-scenario A.3

Figure 8.3 shows the results for sub-scenario A.3 where $\gamma = 3$ (scarce). The search space islands are significantly shrunk compared to the previous sub-scenarios and the algorithms discover fewer non-dominated solutions. The performance of the algorithms drops sharply in the uppermost island where the EMOAs struggle to locate non-dominated solutions.

Table 8.7 provides the data of the optimisation results for sub-scenario A.3. The solutions identified in the search space by the large scale search are a fraction of the previous sub-scenarios as they are less than a thousand (994). This significantly affects the number of unique solutions generated by each of the optimisation algorithms. However, the effect of the drastic reduction of task library is the number of non-dominated solutions generated by the algorithms. NSGA2 and PAES perform poorly having a success ratio well below 50% resulting in 2-3 non-dominated solutions per island. NSGA2 also takes double the time to produce a non-dominated solution. Although PESA2 and SPEA2 generate similar number of solutions, they identify more non-dominated solutions and thus perform comparatively better.



(a) Search space | (b) NSGA2 optimisation results | (c) PAES optimisation results | (d) PESA2 optimisation results | (e) SPEA2 optimisation results | (f) Combined results

Figure 8.3. Optimisation results for sub-scenario A.3

	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	1,002.1	994	-	-	-	-
NSGA2	24.2	18	8	44.4%	0	3.03
PAES	5.5	17	6	35.3%	0	0.92
PESA2	11.2	15	11	73.3%	0	1.02
SPEA2	18.7	17	13	76.5%	0	1.44

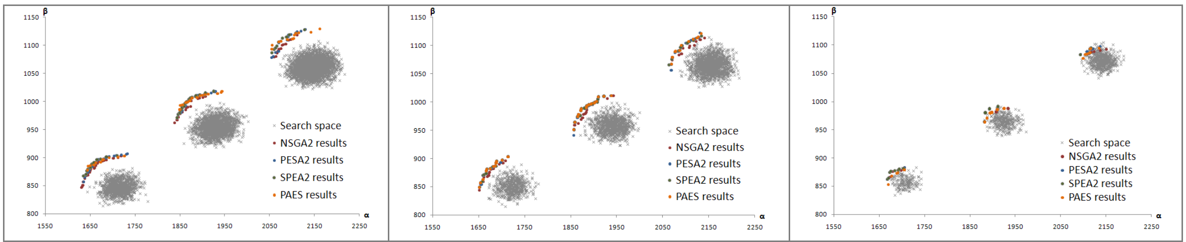
Table 8.7. Optimisation data for sub-scenario A.3

Sub-scenario A.4

The last sub-scenario of this experiment involves the execution and testing of the framework for $\gamma = 2$. For this sub-scenario the LSSA algorithm and the EMOAs could not locate a solution that corresponds to a feasible business process design. Following the search space shrinkage through the previous sub-scenarios, we can assume that for $\gamma = 2$ the framework cannot compose a feasible design and thus optimisation cannot happen.

(9) Remarks about bpo^F performance

The experimental scenario A involved four sub-scenarios with which the framework was tested. Figure 8.4 summarises the optimisation results and demonstrates the key differences across the three sub-scenarios.



(a) sub-scenario A.1	(b) sub-scenario A.2	(c) sub-scenario A.3

Figure 8.4. Summary of optimisation results for experimental scenario A

In more detail:

- Reducing the library size affects the search space significantly as shown by the results of the large scale search. The number of solutions per island are drastically reduced which consequently reduces the density of the islands and affects the convergence of solutions.
- The island density directly affects the capability of the EMOAs to generate non-dominated solutions. For *abundant* (figure 8.4.a) and *satisfactory* library size (figure 8.4.b) all the algorithms perform well whereas for *scarce* library size (figure 8.4.c) all the algorithms generate poor results.
- Chapter 7 discussed that the library size affects not only the convergence but also the diversity capability of the EMOAs. The results from experimental scenario A verify this hypothesis: As the task library size is reduced, fewer non-dominated solutions (convergence) and fewer solutions per island (diversity) are discovered.

Figure 8.5 shows the average time of execution for the EMOAs in all sub-scenarios. PAES is clearly the fastest with average time of execution 5.7 minutes and NSGA2 is the slowest with 24.2 minutes. The speed of PAES can be justified by the fact that it does not implement the crossover operator and thus it omits a step that the other EMOAs include in the optimisation process. The lengthy execution time of NSGA2 is triggered by the large population size (500). Experiments with lower population sizes improved its speed but resulted in less competitive optimisation results.

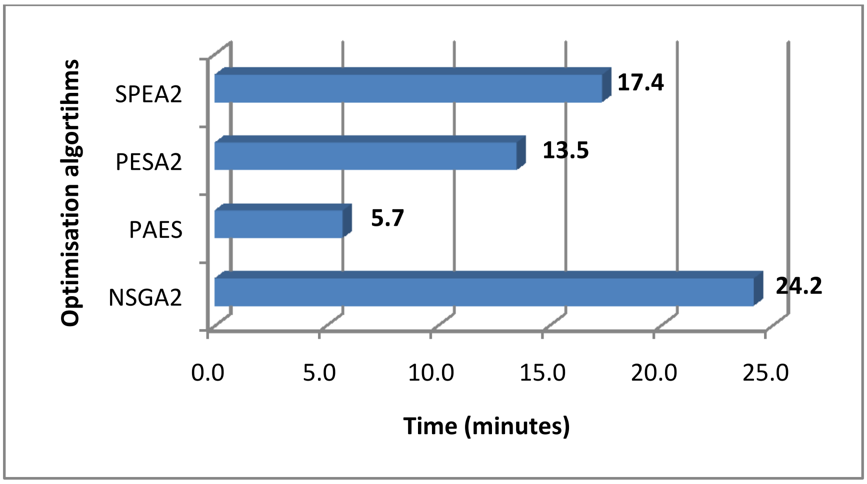


Figure 8.5. Average execution times of the EMOAs for experimental scenario A

To assess the effect on the convergence capability of the algorithms and the framework's optimisation performance, figure 8.6 provides an overview of the non-dominated solutions as generated by each EMOA in each sub-scenario. PESA2 has generated the most non-dominated solutions overall thus providing a more dense Pareto-optimal front and more alternatives in terms of optimised business process designs. This can be attributed to its region-based selection process that seems to be more efficient in the particular context. NSGA2 and SPEA2 provided around 80 non-dominated solutions across the sub-scenarios. However, SPEA2 discovered 50% more solutions in the challenging third sub-scenario thus making it preferable over NSGA2. The SPEA2 'strength' selection mechanism performs better than the crowded comparison operator of NSGA2. PAES, despite being the fastest, provided the fewest non-dominated solutions and performed poorly in the last sub-scenario. The simple optimisation process and the local search that this algorithm is based on proved ineffective for small library sizes.

Figure 8.6 also shows how the number of non-dominated solutions reduces from sub-scenario A.1 to A.3 as a consequence of the library size reduction. The goal of the scenario was to discover the minimum library size (in relation to process size) for which the framework can provide *satisfactory* results. After generating the results for the different γ

values, the question is: *Are the optimisation results of the third sub-scenario considered as satisfactory?*

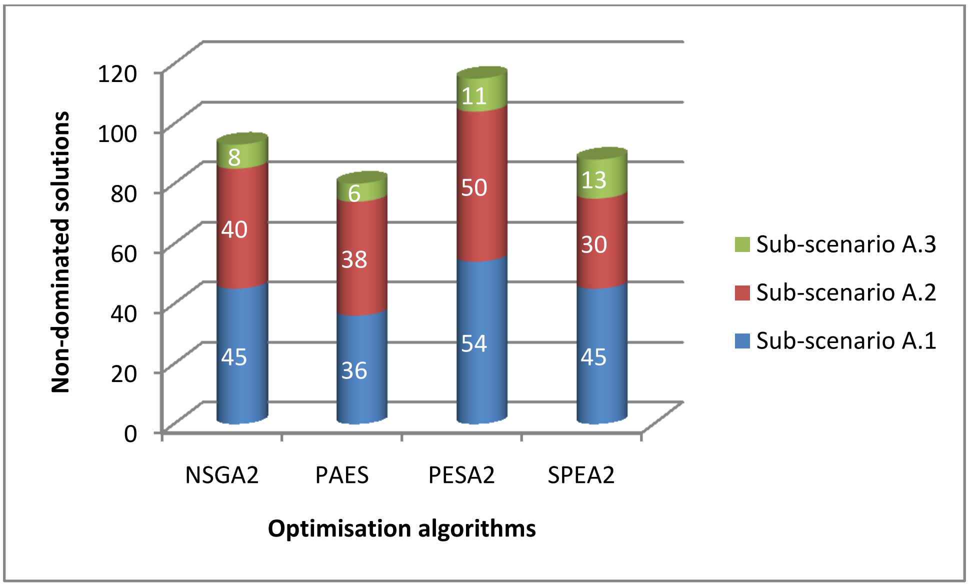


Figure 8.6. Non-dominated solutions for the experimental scenario A

On average, the algorithms provided 10 unique non-dominated solutions each. In the context of business processes, this means 10 different business process designs with optimised attribute values and three different process sizes. This result is considered as satisfactory because it provides the business analyst with a number of optimised alternative designs for a business process.

As a result, **a ratio of 1:3** and below between tasks in the process design and tasks in the library (γ >= 3) is considered as acceptable for the framework to generate satisfactory optimisation results.

In terms of EMOA performance, all the algorithms performed well without any major differences in terms of quality of results. PESA2 provided the most non-dominated solutions and in a good time comparative to the other EMOAs. This algorithm because of its region-based selection provided a relatively dense front with the most solutions per island in all three sub-scenarios.

8.5 Experimental scenario B

The second experimental scenario is focused on the second question of section 8.2. It seeks to investigate the maximum size of business process designs that the framework can optimise. The previous experiment defined the minimum acceptable ratio between library size and process size. Using this information, this experiment seeks to investigate the maximum size of process design for limited library sizes that the framework can operate with. The outcome of this experiment is expected to suggest the maximum size of business process designs for which the framework provides satisfactory results. This is essential in assessing the boundaries of the framework in terms of business processes it can optimise as for large process sizes it is expected that the performance of the algorithms will drop both in terms of convergence and diversity. The steps of generating the experimental scenario B and testing the framework are described below:

(1) Scenario goal

The goal of the experimental scenario B is defined as: *'Given $\gamma = 3$, determine the maximum process size for a business process such for which the framework can generate satisfactory optimisation results'*. The goal statement uses the finding from the first experiment in order to further investigate the framework limits. Based on the scenario goal, the outcome of this experimental scenario will be a suggested process size that the framework can operate with.

(2) Problem features

The goal of this experimental scenario is related with the second feature of the problem. Feature B is related with the different process sizes of a feasible business process design. This has to do both with the number of islands (which is discussed in the final experimental scenario) and with the maximum process size which is the focus of this particular experiment. As shown in the previous chapter, feature B can hinder convergence of solutions towards the optimal and it also affects diversity. This experiment will test the framework for different process sizes to investigate its flexibility in generating optimal solutions.

(3) Control parameters

The control parameters related with feature B are L and D none of which is directly related to controlling the maximum size of a process design (n_d). Therefore in this experiment we will define the various sub-scenarios by directly defining a series of values for the main problem parameters.

(4) Corresponding problem parameters

Table 8.8 shows four sub-scenarios, each with different process sizes. The scenario goal defines the ratio γ between library and process size, thus table 8.8 also calculates the corresponding library size for each sub-scenario. As a result of these sub-scenarios, the framework will be tested for four different maximum process sizes.

	Sub-scenario B.1	**Sub-scenario B.2**	**Sub-scenario B.3**	**Sub-scenario B.4**
	$n_d = 10$	$n_d = 20$	$n_d = 25$	$n_d = 30$
γ = 3	*n* = 30	*n* = 60	*n* = 75	*n* = 90

Table 8.8. Process size (n_d) and library size (*n*) for each of the four sub-scenarios

(5) Complete the experimental scenario parameters

The next step is to define the remaining parameters for the sub-scenarios. Table 8.9 shows these parameters based on the typical business process scenario (see Appendix C). The parameters that vary across sub-scenarios are highlighted in bold. Library (*n*) and process size (n_d) vary based on the design of experiments. Also, for each sub-scenario, there are $L = 5$ neighbouring islands and the n_{min} (minimum size of process design) is calculated accordingly ($n_{min} = L - n_d + 1$). The reason is that we need to capture the maximum process size so in each sub-scenario we test for a significant range of process sizes in order to get a more accurate picture of the framework's boundaries. Another set of parameters that varies across sub-scenarios are the process requirements as those are expressed by the process inputs and process outputs. The reason is that as the process size elaborates, the process requirements need to increase in order to acquire the desired process size. For example, a business process with 5 output resources cannot justify 30 tasks in the design to produce such a low number of process outputs.

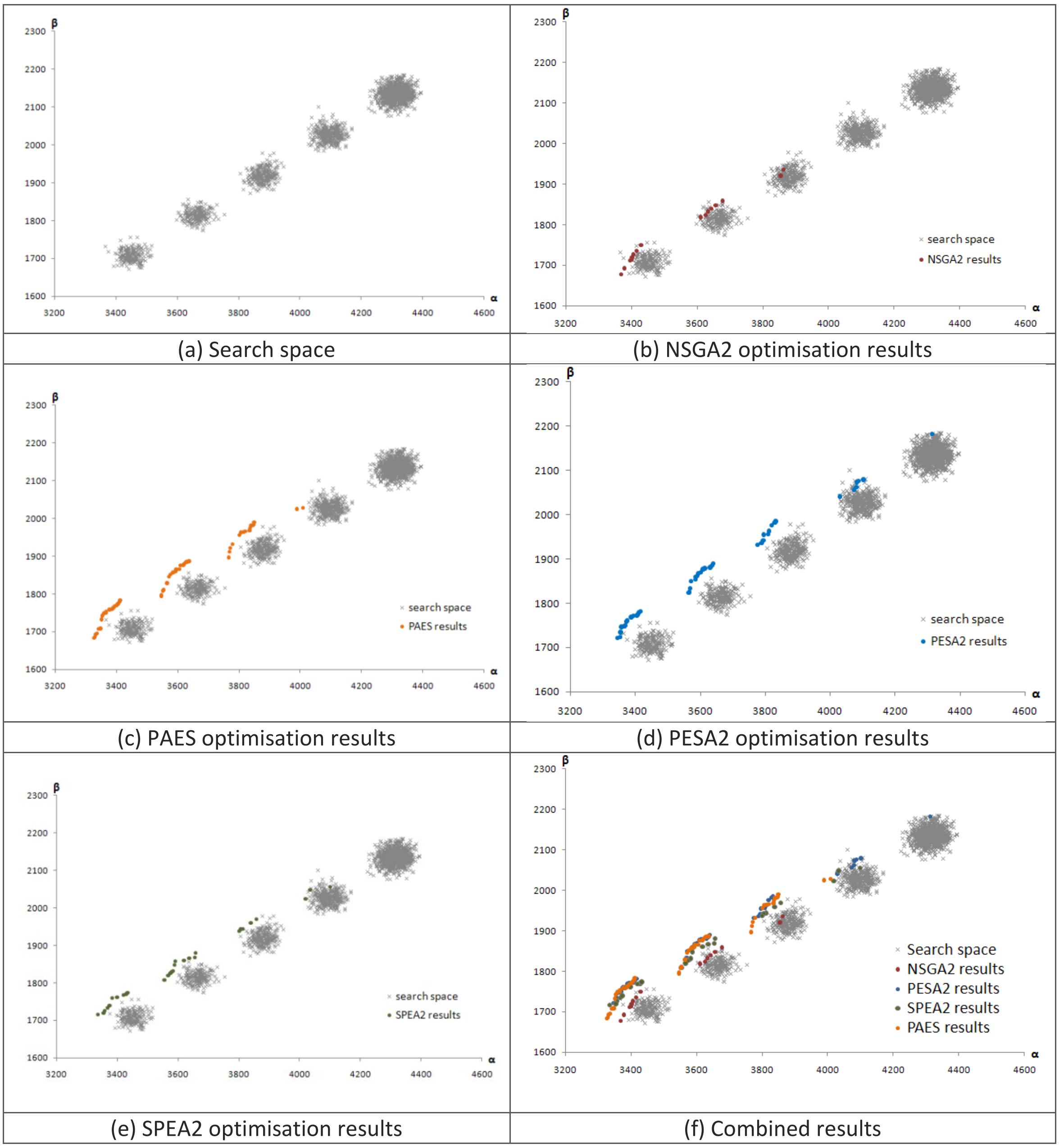
Parameter	Sub-scenario B.1	Sub-scenario B.2	Sub-scenario B.3	Sub-scenario B.4
n	***30***	***60***	***75***	***90***
n_d	**10**	**20**	**25**	**30**
n_{min}	**6**	**16**	**21**	**26**
r	20	20	20	20
t_{in} / t_{out}	3	3	3	3
r_{in} / r_{out}	**5 / 5**	**5 / 10**	**5 / 10**	**10 / 10**
p	2	2	2	2
α	100 -115	100 -115	100 -115	100 -115
β	200 -230	200 -230	200 -230	200 -230

Table 8.9. Problem parameters for the sub-scenarios B.1 - B.4

This completes the scenario formulation. The steps for scenario testing are described below for each sub-scenario.

Sub-scenario B.1

The parameters defined for sub-scenario B.1 are identical to sub-scenario A.3 apart from the number of neighbouring islands. However, as the results of A.3 demonstrate the three uppermost islands of the search space (process designs with 8, 9 and10 tasks) are used here as the results of sub-scenario B.1. As previously discussed, the EMOAs manage to identify non-dominated solutions although for n_d =10 they struggle to identify more than 10.



(a) Search space | (b) NSGA2 optimisation results

(c) PAES optimisation results | (d) PESA2 optimisation results

(e) SPEA2 optimisation results | (f) Combined results

Figure 8.7. Optimisation results for sub-scenario B.2

Sub-scenario B.2

Figure 8.7 shows the optimisation results for scenario B.2. It is evident that the EMOAs struggle to identify non-dominated solutions in the two uppermost islands that correspond to process sizes of 19 and 20 tasks. This inadequacy of the algorithms is attributed to the limited library size (60 tasks) which makes it challenging to compose a business process design with large number of tasks (20).

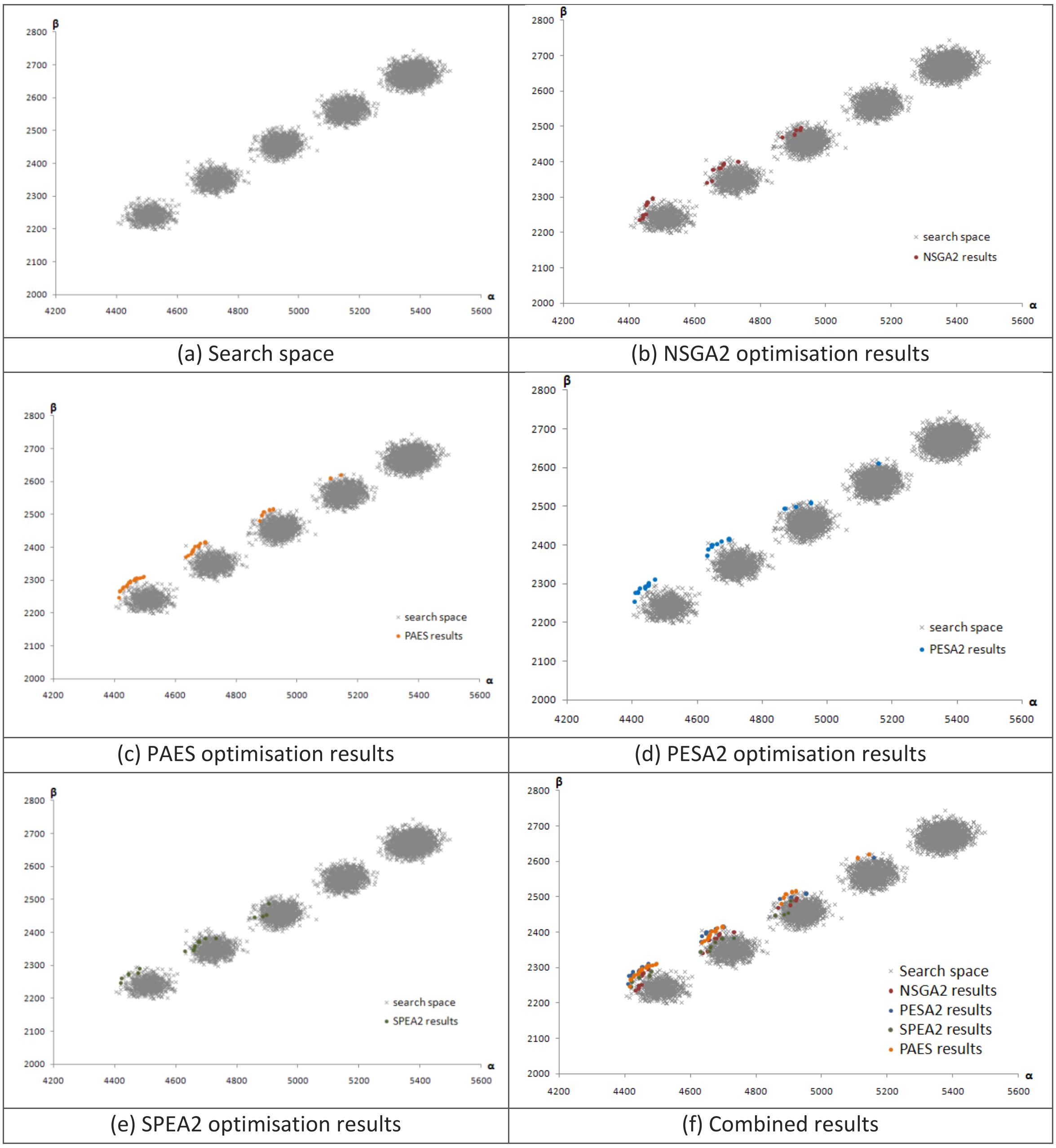
Table 8.10 shows the optimisation data for sub-scenario B.2. NSGA2 performs poorly by identifying only 3 non-dominated solutions and taking the most time to generate the results with 6.9 minutes per non-dominated solution which is 10 times more than the second SPEA2 (0.69 minutes). PESA2 and SPEA2 perform similarly in the number of non-dominated solutions but PESA2 generates the results in 50% less time. PESA2 is also the only EMOA that discovers a non-dominated solution in the uppermost island (20-task designs). The algorithm, however, with very good performance in this particular sub-scenario is PAES which generates by far the most non-dominated solutions (50) in the shortest time (0.21 minutes per solution) achieving a 100% success ratio. PAES identifies two non-dominated solutions for n_d = 19 tasks.

	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	2,435	2,094	-		-	-
NSGA2	20.7	16	3	18.8%	3	6.90
PAES	10.7	50	50	100%	1	0.21
PESA2	13.5	48	33	68.8%	0	0.41
SPEA2	19.2	31	28	90.3%	1	0.69

Table 8.10. Optimisation data for sub-scenario B.2

Sub-scenario B.3

The results of the third sub-scenario are shown in figure 8.8. In this sub-scenario the library is increased to 75 tasks and the maximum process size investigated is 25 tasks. Similar to the previous experiment, the EMOAs have a difficulty in locating non-dominated solutions for the uppermost islands.



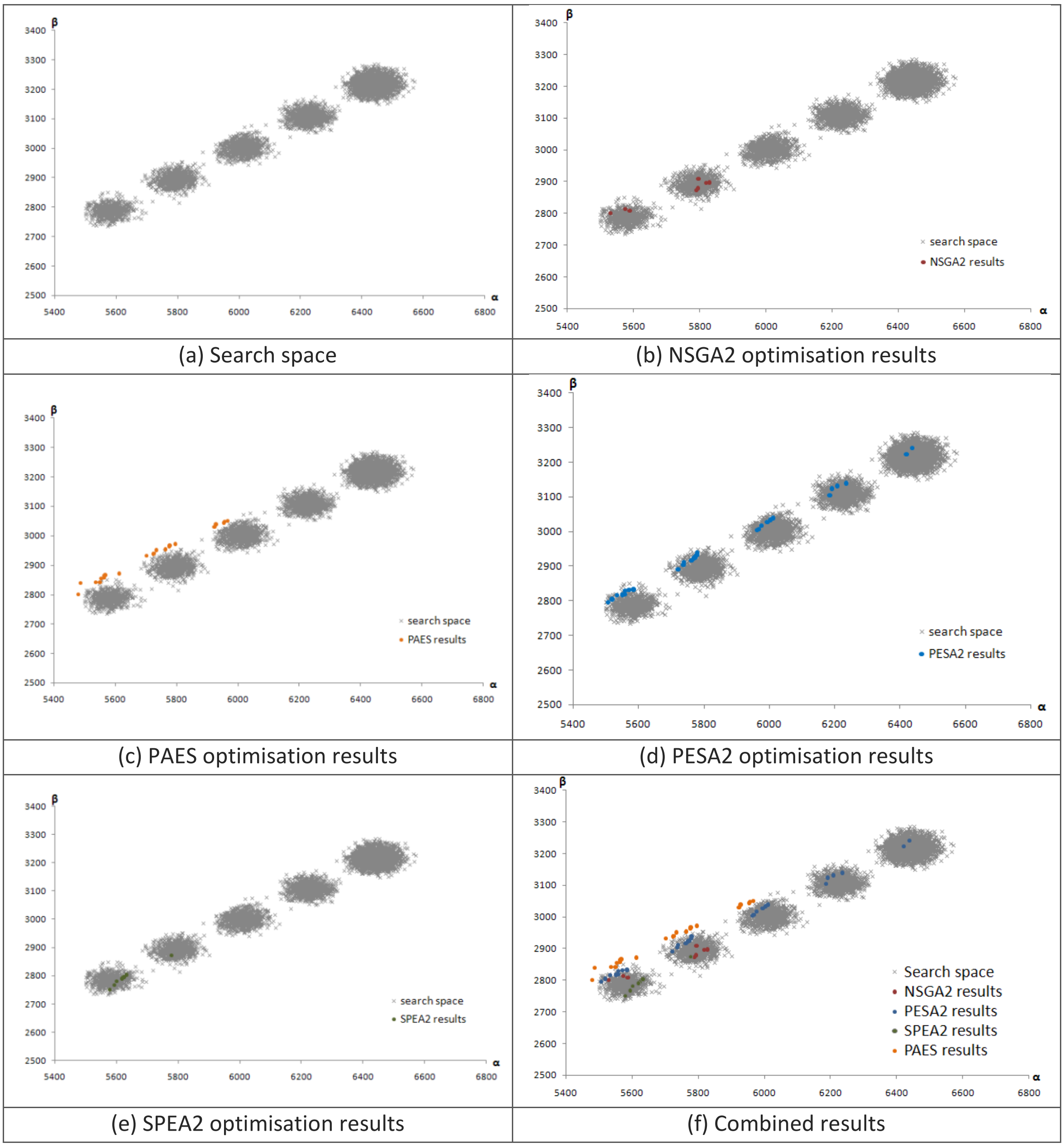
(a) Search space
(b) NSGA2 optimisation results
(c) PAES optimisation results
(d) PESA2 optimisation results
(e) SPEA2 optimisation results
(f) Combined results

Figure 8.8. Optimisation results for sub-scenario B.3

	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	2,778	6,411	-		-	-
NSGA2	17.6	20	2	10.0%	3	8.80
PAES	13.5	36	32	88.9%	1	0.42
PESA2	15.9	20	14	70.0%	2	1.14
SPEA2	31.8	16	4	25.0%	3	7.95

Table 8.11. Optimisation data for sub-scenario B.3

In this sub-scenario both NSGA2 and SPEA2 perform poorly both in terms of time and number of non-dominated solutions as table 8.11 shows. As shown also in the combined results (figure 8.8.f) PAES provides the best fronts for the three first islands and locates two non-dominated solutions in the fourth island. It generates in total 32 non-dominated solutions at 0.42 minutes each and maintains a success ratio close to 90% while PESA2 comes second with 70% and 14 non-dominated solutions



(a) Search space
(b) NSGA2 optimisation results
(c) PAES optimisation results
(d) PESA2 optimisation results
(e) SPEA2 optimisation results
(f) Combined results

Figure 8.9. Optimisation results for sub-scenario B.4

Sub-scenario B.4

The results of the last sub-scenario are shown in figure 8.9. In this sub-scenario the proposed optimisation framework is tested for 30-task business process designs with a library of 90 tasks. NSGA2, SPEA2 and PESA2 discover zero non-dominated solutions while PAES demonstrates a -comparatively- good performance.

Table 8.12 shows the results for the fourth sub-scenario. NSGA2 and SPEA2 perform poorly identifying few feasible solutions none of which are non-dominated. PESA2, although fails to converge to non-dominated solutions, it identifies solutions in all the islands. This shows the hindrance to the convergence capability that the large process sizes cause. PAES identifies 30 non-dominated solutions achieving a 100% success ratio. The algorithm is the only one that identifies non-dominated solutions in the first three islands having the fastest execution time once again.

	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	4,231	6,821	-		-	-
NSGA2	19.7	8	0	0%	5	-
PAES	15.6	30	30	100%	2	0.52
PESA2	18.9	28	0	0%	5	-
SPEA2	23.1	7	0	0%	5	-

Table 8.12. Optimisation data for sub-scenario B.4

(9) Remarks about bpo^F performance

The experimental scenario B involved four sub-scenarios with which the proposed optimisation framework was tested. Figure 8.10 summarises the combined optimisation results for each of the sub-scenarios in order to draw some remarks about the performance of the framework for dealing with varying process sizes.

In more detail:

- Increasing the process size while keeping the γ ratio constant affects significantly the performance of the optimisation algorithms. With the exception of PAES, the performance of the remaining algorithms deteriorates as the maximum process sizes increase.
- The diversity of solutions suffers throughout the experiment. The algorithms do not locate non-dominated solutions in the uppermost island in sub-scenarios B.2,

B.3 and B.4. Also in the fourth island there are only scarce non-dominated solutions in sub-scenarios B.2 and B.3 and none in B.4.

- The capability of the EMOAs to converge towards optimal solutions decreases as the process size increases. A good example is PESA2 which in the last sub-scenario locates fronts across all the five islands but cannot push these solutions towards the Pareto-optimal front.
- PAES performs consistently well in this experimental scenario. It discovers the most non-dominated solutions and it is the fastest in all the sub-scenarios. Its performance shines in the last sub-scenario where 3 EMOAs fail to discover even a single non-dominated solution and PAES generates 30 achieving a success ratio of 100%.

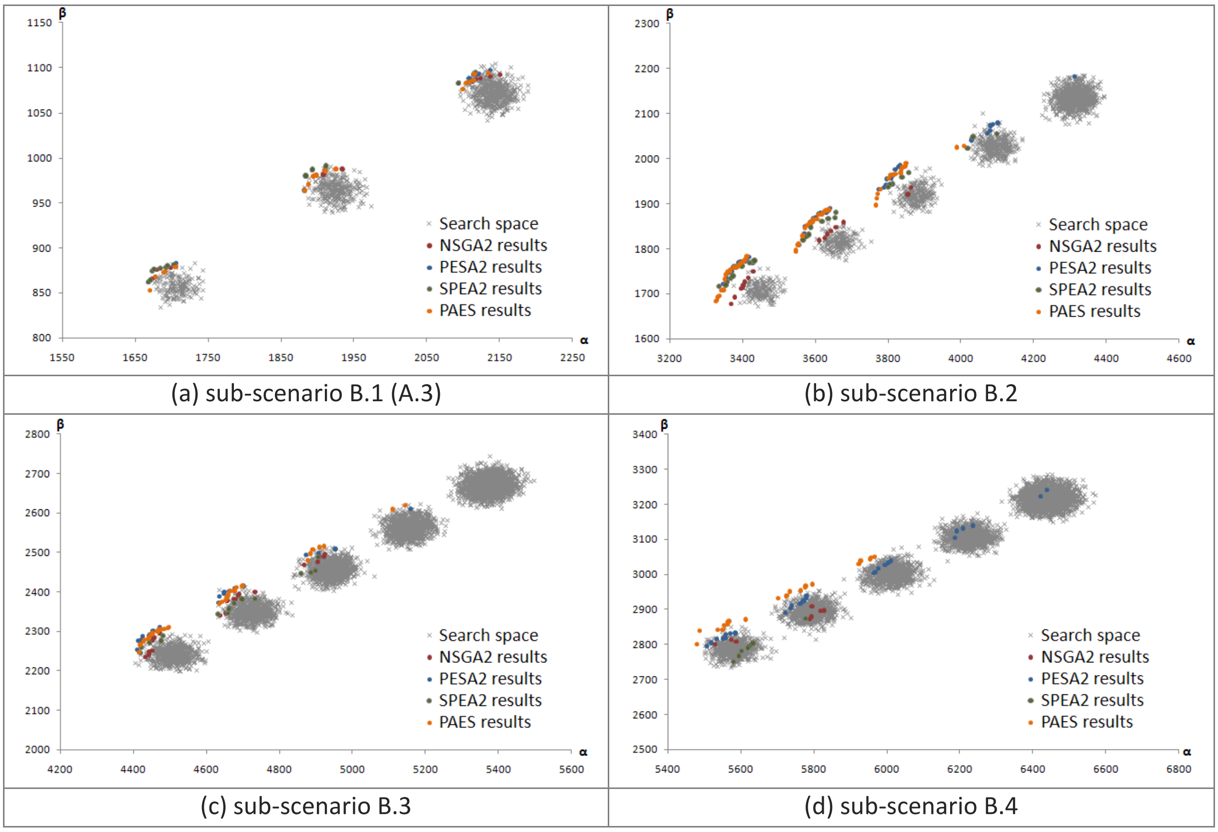


Figure 8.10. Summary of optimisation results for experimental scenario B

To demonstrate better the EMOA performance, figure 8.11 shows the percentage of the non-dominated solutions per algorithm generated in all sub-scenarios of experimental scenario B. PAES generated 50% of the non-dominated solutions overshadowing the other three EMOAs. Its execution time per solution was 0.5 minutes. PAES local optimisation approach and simplicity in the optimisation process helps to efficiently discover non-dominated solutions with large number of tasks whereas more complex optimisation approaches cannot discover and push towards optimal solutions. On the other side of the spectrum NSGA2 generated only 6% of the non-dominated solutions requiring 6.2 minutes per solution. The performance of SPEA2 is also considered poor since it produced 4 solutions in sub-scenario B.3 and zero in the last one. The performance of PESA2 is considered satisfactory. It contributed 25% of the non-dominated solutions; it required 0.9 minutes per solution and failed only in the last sub-scenario where it did not converge. Also in this experimental scenario is demonstrated that the region-based selection approach that PESA2 is employing is more effective for the business process problem compared to NSGA2 and SPEA2 selection mechanisms

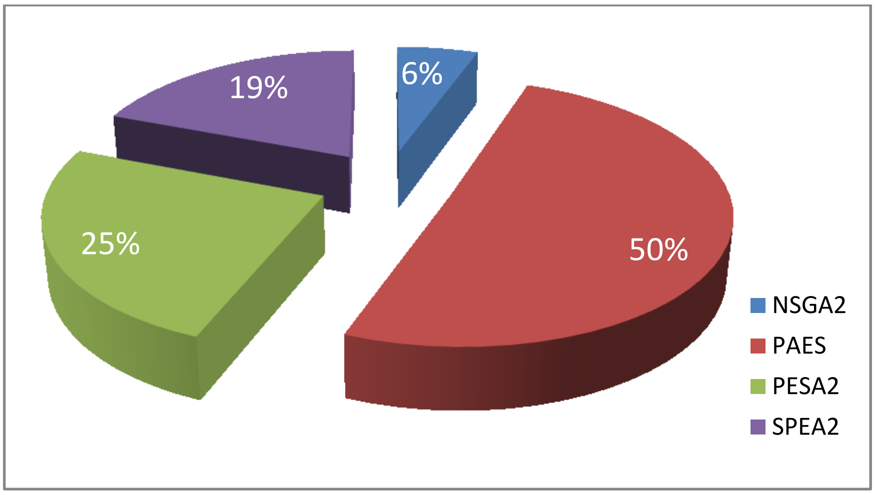


Figure 8.11. Percentage of non-dominated solutions generated in scenario B

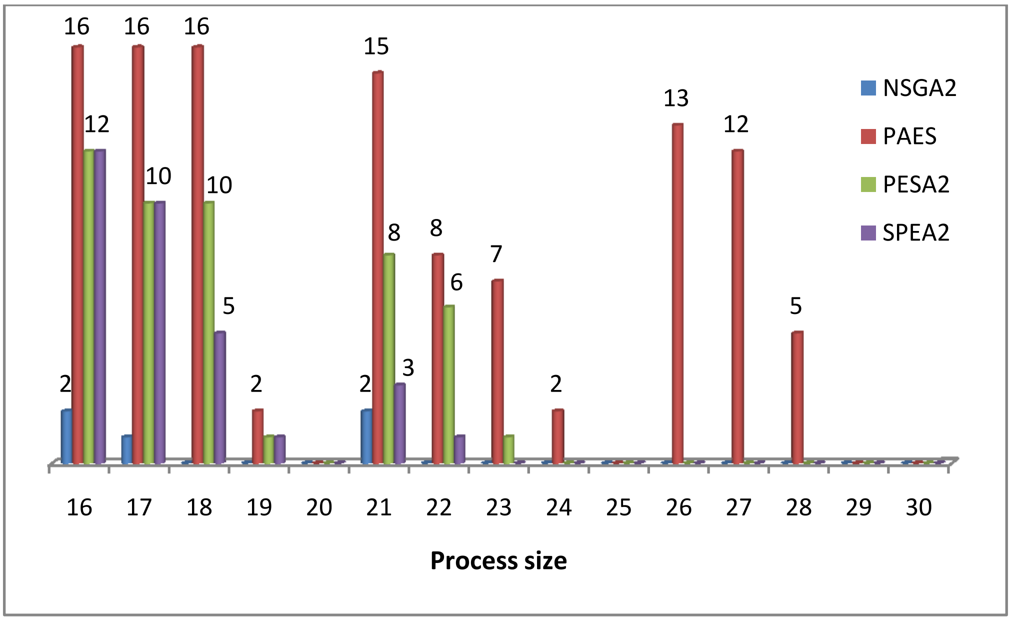


Figure 8.12. Non-dominated solutions generated in scenario B per EMOA

Figure 8.12 shows the non-dominated solutions generated by the EMOAs sorted by process size for the last three sub-scenarios (B.2, B.3 and B.4). This figure shows that for each sub-scenario, the algorithms have a challenge to identify business process designs with the maximum number of tasks (20, 25 and 30). This is due to the fact that the library is restrained accordingly in each sub-scenario making it hard for the framework to compose a feasible design with the maximum number of tasks and thus the EMOAs show preference to lower process sizes.

Figure 8.12 will help answer the quest for determining a satisfactory process size for the framework. The goal of the scenario was to discover the maximum process size ($\gamma = 3$) for which the framework can provide *satisfactory* results. Figure 8.12 shows that for the sub-scenarios B.3 and B.4 the optimisation algorithms consistently struggled to generate optimised results with the exception of PAES. However, for process sizes in the range of 26-28 tasks, PAES is still capable of generating a satisfactory number of alternative optimised business process designs. Therefore the answer to the scenario goal is two-fold:

- Overall, using any of the EMOAs employed by the framework, a business process design with size in the range of **16-20 tasks** can be optimised with confidence by PESA2, SPEA2 and PAES.
- In the case of using PAES, however, the proposed optimisation framework can optimise business process designs in the range of **26-28 tasks** and provide a satisfactory range of alternatives.

8.6 Experimental scenario C

One of the novelties of the proposed optimisation framework is the capability to generate optimised business process designs of varying sizes for the same business process requirements. This capability allows for the generation of business process designs with different number of tasks. It also poses a challenge for the optimisation algorithms as they have to maintain and simultaneously optimise feasible business process designs from all the different acceptable process sizes. The final experimental scenario is investigating the number of different process sizes that the framework can effectively handle during the optimisation process. Having defined the minimum library size and the maximum process design size for generating satisfactory results, this experiment seeks to identify the number of islands in the search space that the framework can manage. The steps of generating the experimental scenario C and testing the framework are described below:

(1) Scenario goal

The formulation of the scenario goal is largely based on the third question of section 8.2 related to the number of neighbouring islands, it also uses the findings of the previous two experiments. The goal of the experimental scenario C is defined as: *'Based on the findings of the previous two experiments, determine the maximum number of neighbouring islands for which the framework can produce optimised business process designs'*. The final experimental scenario tests the boundaries of the framework further and this is why it is based on the findings of the previous experiments. Assuming that for large libraries the framework can manage large number of neighbouring islands, this experiment will preserve the 1:3 ratio between the tasks in the design and the tasks in the library as dictated from scenario A. Also, it will utilise the main finding of the previous scenario using process designs with the maximum number of tasks (20) that the framework can handle effectively. Therefore, it will challenge the problem further adding to scenarios with restrained libraries and maximum process size the additional parameter of the number of neighbouring islands.

(2) Problem features

Similar to experimental scenario B, this scenario is also related with the second feature of the problem. Feature B is related with the different sizes of a feasible business process design. As shown in the previous chapter and in the previous experimental scenario, feature B can hinder convergence and diversity of solutions. However, this experiment will focus more on the diversity capability of the framework –its capability to discover non-dominated solutions across all the available process sizes. Therefore, this experimental scenario will test the framework for different ranges of neighbouring islands to investigate its flexibility in generating optimal solutions across all the acceptable process sizes.

(3) Control parameters

The control parameters related with feature B are L and D. In this experiment we will focus on L – the number of continuous neighbouring islands. It is important to first investigate the number of neighbouring islands (L) before experimentation takes place with the distance (D). Essentially the scenario seeks to test the framework for different ranges of L. Defining the maximum L value that the framework can operate will provide an answer to the scenario goal. Table 8.13 shows the L values for 3 different sub-scenarios. These values are selected based in the classification of the control parameters as defined in table 7.5 (chapter 7).

Sub-scenario C.1	Sub-scenario C.2	Sub-scenario C.3
L = 5 (moderate)	L = 10 (large)	L = 20 (very large)

Table 8.13. *L* values for each of the three sub-scenarios

(4) Corresponding problem parameters

Having defined the various *L* values for each sub-scenario, the corresponding problem parameters can be calculated. The experimental scenario aims to test the framework for a range of neighbouring islands based on the maximum process size defined by the previous scenario. Therefore, for scenario C we assume that n_d is constant and equals to 20. Table 8.14 calculates the minimum process size (n_{min}) for each sub-scenario based on the formula $n_{min} = n_d - L + 1$.

	Sub-scenario C.1	Sub-scenario C.2	Sub-scenario C.3
	L = 5	L = 10	L = 20
$n_d = 20$	$n_{min} = 16$	$n_{min} = 6$	$n_{min} = 1$

Table 8.14. Minimum process size (n_{min}) for each of the three sub-scenarios

(5) Complete the experimental scenario parameters

The next step is to define the remaining parameters across the sub-scenarios. Table 8.15 shows these parameters based on the typical business process scenario with the exception of $n = 60$ that is defined based on the experimental scenario A and $n_d = 20$ which is defined based on the experimental scenario B.

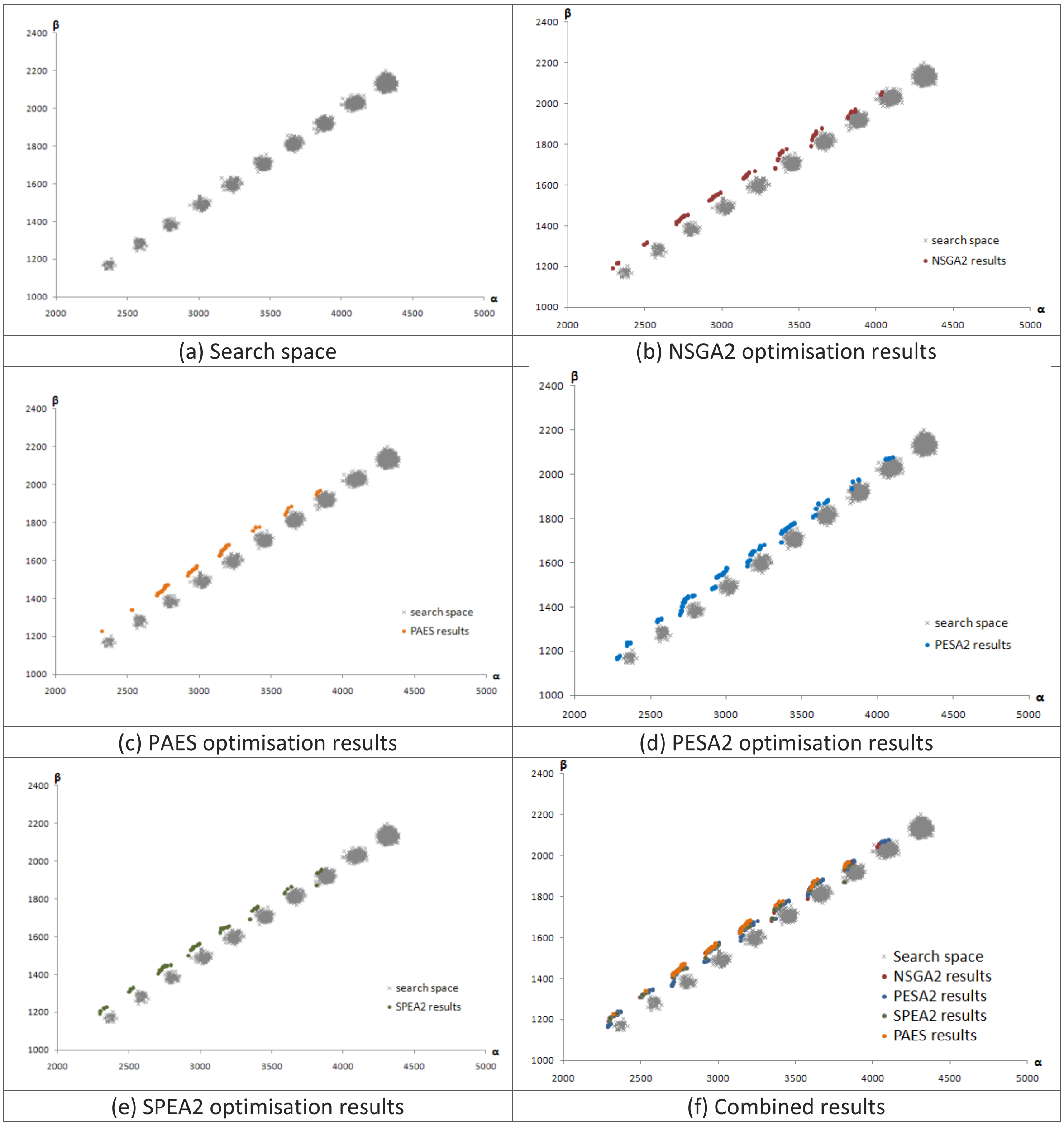
Parameter	Sub-scenario C.1	Sub-scenario C.2	Sub-scenario C.3
n	*60*	*60*	*60*
n_d	*20*	*20*	*20*
n_{min}	**16**	**6**	**1**
r	20	20	20
t_{in} / t_{out}	3	3	3
r_{in} / r_{out}	5 / 10	5 / 10	5 / 10
p	2	2	2
α	100 -115	100 -115	100 -115
β	200 -230	200 -230	200 -230

Table 8.15. Problem parameters for the sub-scenarios C.1, C.2 and C3

This completes the scenario formulation. The steps for scenario testing are described below for each sub-scenario and step 9 (main remarks) is detailed at the end of the experimental scenario C.

Sub-scenario C.1

The parameters defined for sub-scenario C.1 are identical to sub-scenario B.2, therefore the results of sub-scenario C.1 are demonstrated in figure 8.7. This particular sub-scenario involves 5 different islands and the EMOAs struggle to identify non-dominated solutions in the two uppermost islands that correspond to process sizes of 19 and 20 tasks.



(a) Search space | (b) NSGA2 optimisation results
(c) PAES optimisation results | (d) PESA2 optimisation results
(e) SPEA2 optimisation results | (f) Combined results

Figure 8.13. Optimisation results for sub-scenario C.2

Sub-scenario C.2

Figure 8.13 shows the optimisation results for sub-scenario C.2. The EMOAs manages to identify non-dominated solutions in the majority of the 10 available islands. However, similar to the previous sub-scenario, they struggle to identify solutions for process designs with 19 and 20 tasks.

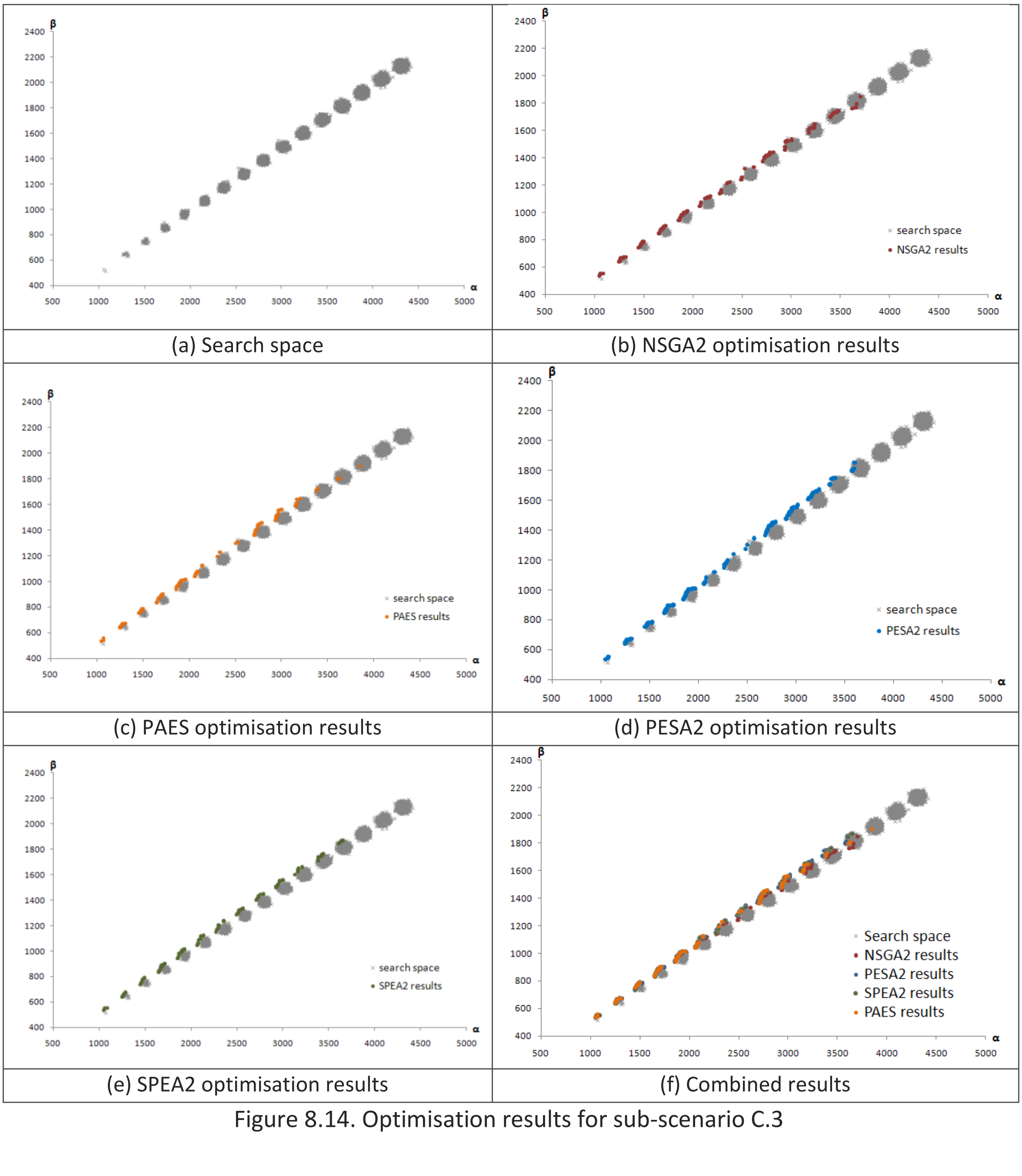
	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	4,449	2,575	-		-	-
NSGA2	21.7	62	48	77.4%	2	0.45
PAES	10.2	50	47	94.0%	2	0.22
PESA2	17.6	82	71	86.6%	1	0.25
SPEA2	18.8	53	48	90.6%	2	0.39

Table 8.16. Optimisation data for sub-scenario C.2

Table 8.16 shows the optimisation data for sub-scenario C.2. NSGA2, PAES and SPEA2 identified around 48 non-dominated solutions whereas PESA2 provides more dense fronts with 30% more solutions. For the total generated solutions per algorithm, PAES has a success ratio of 94% and once more it is the fastest algorithm. In terms of the various islands –which is the focus of the current experiment– none of the EMOAs discovered a (non-dominated) solution for $n_d = 20$. SPEA2 and PAES did not discover solutions also for $n_d = 19$ while PAES discovered a single solution in $n_d = 6$ and $n_d = 7$ which shows poor diversity in the specific islands. NSGA2 discovered 2 dominated solutions in $n_d = 19$ and PESA2 3 non-dominated solutions making it the best performing algorithm in this sub-scenario in terms of diversity and number of solutions.

Sub-scenario C.3

The third sub-scenario of this experiment tests the framework for a range of 20 different process sizes, from 1 to 20 tasks. However the framework discovers feasible process designs with 5 tasks and above, resulting in 16 different islands. 16 are still classified as *very large* number of neighbouring islands (according to the control parameters classification) and therefore the sub-scenario parameters remain unchanged. The performance of the algorithms is considered satisfactory as they locate non-dominated solutions for the majority of the islands with the exception of the three uppermost.



(a) Search space | (b) NSGA2 optimisation results | (c) PAES optimisation results | (d) PESA2 optimisation results | (e) SPEA2 optimisation results | (f) Combined results

Figure 8.14. Optimisation results for sub-scenario C.3

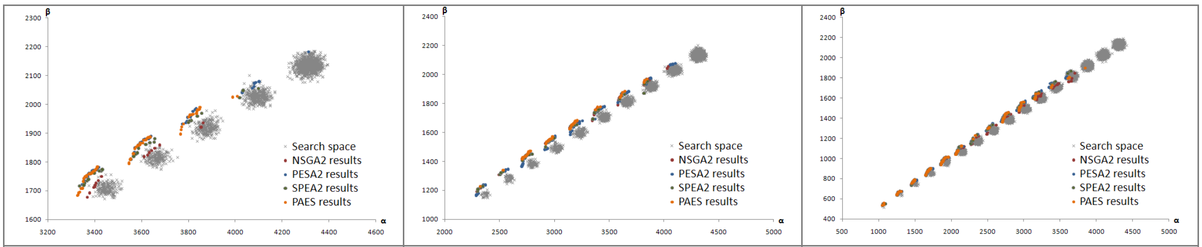
	Time (mins.)	Unique Generated Solutions	Non-dominated solutions	Success ratio (M_1)	Islands without n-d solutions (M_2)	Time per n-d solution (M_3)
Search space	7,109	11,866	-	-	-	-
NSGA2	19.2	105	71	67.6%	6	0.27
PAES	9.4	91	85	93.4%	5	0.11
PESA2	18.6	122	114	93.4%	4	0.16
SPEA2	19.7	88	78	88.6%	3	0.25

Table 8.17. Optimisation data for sub-scenario C.3

Table 8.17 shows the optimisation data for each of the EMOAs. Due to the large number of islands all the algorithms generate a considerable number of non-dominated solutions with PESA2 reaching 114, 30% more than the others. PESA2 and PAES have an equal success ratio of 93.4% while PAES is the fastest with 9.4 minutes execution time compared to the other three that took on average 19 minutes to generate the results. In terms of diversity, NSGA2 discovers non-dominated solutions in 10 of the islands, PAES in 11, PESA2 in 12 and SPEA2 in 13. Also, NSGA2, PAES and PESA2 discover very few non-dominated solutions for n_d = 11 and 12. Overall in this sub-scenario, PESA2 discovers the most non-dominated solutions but SPEA2 discovers solutions in the most islands and PAES is the fastest utilising 50% less time to generate results.

(9) Remarks about bpo^F performance

The experimental scenario C involved the testing of the proposed optimisation framework with three sub-scenarios. Figure 8.15 summarises the combined results of the sub-scenarios.



(a) sub-scenario C.1 (B.2)	(b) sub-scenario C.2	(c) sub-scenario C.3

Figure 8.15. Summary of optimisation results for experimental scenario C

From these results we can observe that:

- Despite the gradual and significant increase in the range of acceptable process sizes –which results in increased number of neighbouring islands in the search space– the optimisation algorithms are capable of locating non-dominated solutions across the majority of these islands.
- As expected, the performance in terms of diversity deteriorates as the number of islands increase but it is the case only for the uppermost islands in each sub-scenario.
- The convergence capability of the algorithms remains high through all the sub-scenarios as shown by the success ratio which –with the exception of NSGA2– remains at 80%.

- There is no clear winner in terms of algorithm in this scenario. PESA2 generates significantly more non-dominated solutions but SPEA2 is slightly better in discovering solutions in more islands.

The final experimental scenario focused more on the diversity capability of the proposed optimisation framework. To acquire a better overview of the performance of the optimisation algorithms, figure 8.16 shows the number of islands without non-dominated solutions per algorithm per sub-scenario. The most consistent algorithm is SPEA2, while NSGA2 demonstrates poor performance in terms of diversity. PESA2 discovers solutions across all islands for the low range sub-scenario (C.1), but omits 4 islands in the high range sub-scenario (C.2). Finally, PAES performs well C.1 and C.2, but omits five islands in C.3.

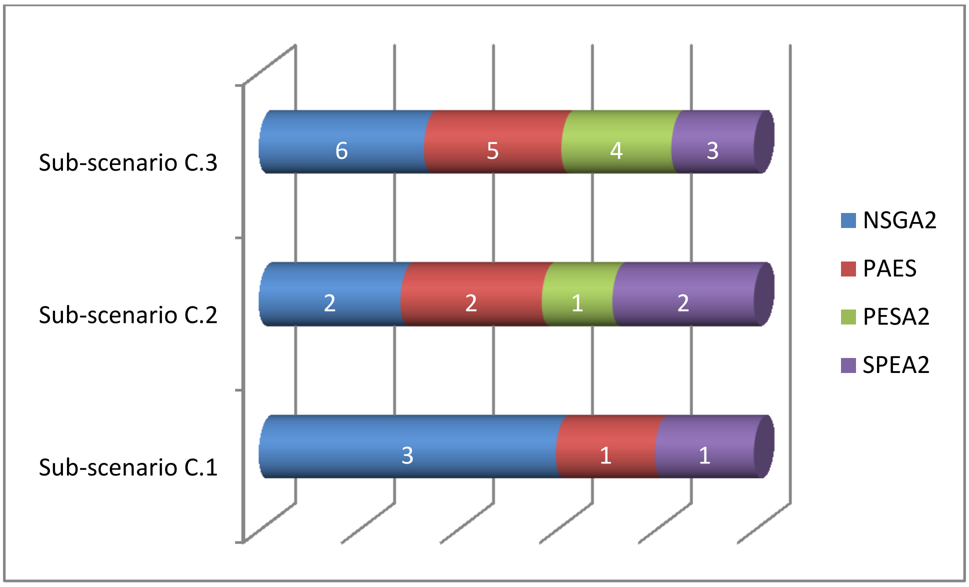


Figure 8.16. Number of islands without non-dominated solutions

In terms of the framework's performance, the goal was to discover the maximum number of neighbouring islands for which the framework can generate optimised results. Figure 8.17 shows the percentage of islands without non-dominated solutions in each sub-scenario. The percentage is calculated based on the average performance of all the algorithms per sub-scenario. In sub-scenario C.1 ($L = 5$) and sub-scenario C.3 ($L = 20$) the framework fails to discover non-dominated solutions in a quarter of the islands. However, in sub-scenario C.2 only 18% of the islands are left without non-dominated solutions. Although the percentage seems high for C.1 it corresponds to 1.3 islands, for C.2 it corresponds to 1.8 islands and for C.3 to 4.5 islands.

The answer to this scenario goal is not straight-forward. Despite the restrained library size and large process size, the framework discovers non-dominated solutions for most of the islands even with 16 islands in the search space. We can assume that for larger library sizes the quality of results improves as the number of feasible process designs increases. All the algorithms dealt with the challenge of more islands by increasing the number of non-dominated solutions. Unlike scenario B there is no clear winner in performance. PESA2 generated the most non-dominated solutions but SPEA2 performed more consistently in the diversity aspect. PAES was fast to generate results but did not perform as good as in the previous scenario.

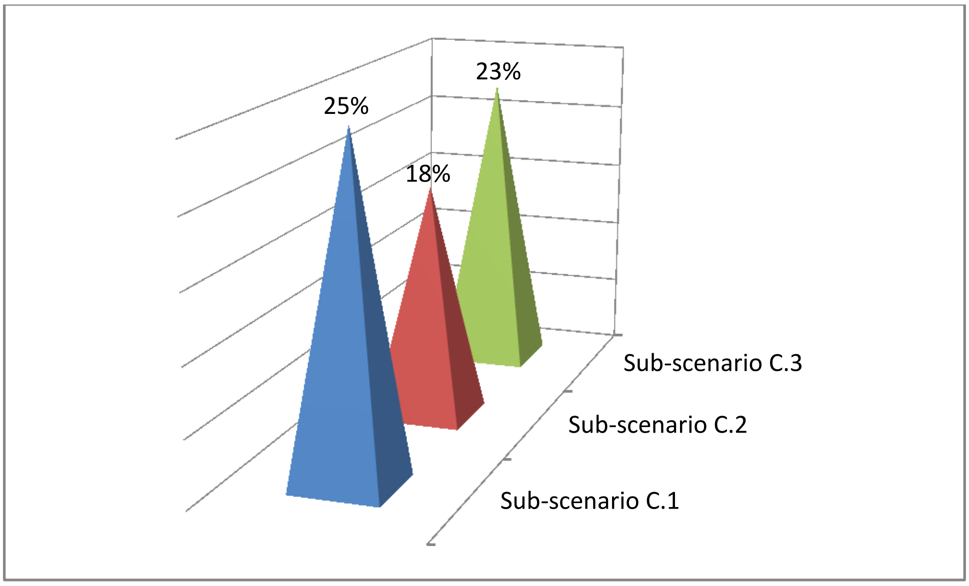


Figure 8.17. Percentage of islands without non-dominated solutions in each sub-scenario

8.7 Main remarks

This section summarises the chapter and highlights the main findings of the experiments. The proposed business process optimisation framework was tested with three experimental scenarios. These scenarios were generated based on the proposed strategy for scenario generation and framework testing (presented in chapter 7). The focus of this performance evaluation was two-fold: (i) to assess the capability of the EMOAs in optimising effectively business process designs and (ii) to investigate the potential of the framework in generating feasible business processes designs with optimal values based on the proposed business process representation and problem formulation.

In terms of the EMOA performance the following remarks can be provided as the answers to the relevant questions (see section 8.2):

? *Is an evolutionary multi-objective optimisation approach effective given the discrete nature of the business process optimisation problem?*

? *Is the performance of the EMOAs satisfactory given the multi-objective and highly constrained formulation of the problem?*

Employing EMOAs as the optimisation technique for the proposed framework proved an effective choice. This is demonstrated by the results of the experiments. In the majority of cases the algorithms identified a number of non-dominated solutions across various processes sizes despite the highly constrained problem formulation. Putting this in the business process context means that each solution corresponds to an alternative business process design with optimised attribute values which shows the advantage of using a multi-objective optimisation technique. Therefore, the variety of results that these algorithms offered under varying parameters makes them an attractive optimisation technique for business processes.

? *From the selected EMOAs, is one or more significantly better overall in the business process context? Do some EMOAs perform better under specific conditions (e.g. large process size)?*

⊖ NSGA2 shows average/poor results throughout the experiments. It generates less non-dominated solutions and it is much slower that the other EMOAs. However, NSGA2 is known not to perform well in problems with multiple local fronts (Tiwari, 2001) and business process optimisation in one of them. The fitness assignment strategy of NSGA2 ceases to produce the driving force towards the global front once most of the solutions of the population share the shame non-domination level. This is further augmented due to the use of elitism and NSGA2 suffers from the tendency of getting trapped in local fronts (pre-mature convergence). Further testing and parameter tuning might help improve the quality of results. Based on the experiments shown here, NSGA2 is considered unfit for business process optimisation.

⊖ PAES has been consistently the fastest of the algorithms generating competing results as it reported by its creators (Knowles and Corne, 1999). In scenario B the algorithm demonstrated much better performance than the other algorithms. PAES is strongest in cases when local search seems superior to or competitive with population-based methods. Scenario B accommodated large solution sizes which the EMOAs found hard to cope with. PAES using the simple (1+1) evolution strategy managed to be more effective in discovering optimised business process designs.

- PESA2 demonstrated a consistently good performance throughout the experiments. It generated the most non-dominated solutions and showed strong diversity capabilities. It outperformed the other EMOAs in scenario A and in some sub-scenarios of C. If one algorithm was recommended for the business process optimisation problem it would be PESA2. This is attributed to its sophisticated selection strategy. Using region-based selection and breaking the search space into hyper-boxes proved suitable and effective for the business process optimisation problem. The search space consists of separate areas (islands) and PESA2 was capable on working on multiple fronts using the 'squeeze factor' in all the islands simultaneously thus locating more optimal solutions in more islands that its counterparts.

- SPEA2 showed a good/ average performance in the experiments. It generated a good number of non-dominated solutions but showed limited diversity capabilities. It is also the second slowest algorithm. The 'strength' selection technique that SPEA2 is using did not prove as effective in comparison to PAES and PESA2 but it demonstrated better results than NSGA2.

In terms of the framework's business process optimisation capability the following remarks can be provided as the answers to the relevant questions (see section 8.2):

? *What is the minimum library size that the framework can operate with?*
As experimental scenario A showed, the minimum library size that a framework can operate with is 3 times the size of the process design. With such a ratio the framework can produce a satisfactory number of business process designs.

? *What is the maximum size of a business process design that can be optimised?*
The maximum size of a business process design was the focus of experimental scenario B. The framework can optimise business processes with 20 tasks given the ratio with the library is 1:3 or smaller. In the case of larger library the framework can work with larger designs as the possible combinations of tasks are increased exponentially.

? *What is maximum number of neighbouring islands that the framework can handle?*

The framework can handle a number of neighbouring islands in the search space as scenario C demonstrated. However, for limited libraries and large process designs it is advisable not to exceed a range of 8-10 neighbouring islands.

8.8 Summary

This chapter evaluated the performance of the proposed optimisation framework for business processes using three experimental scenarios. The design of the experimental scenarios helped test the framework and provide answers about the basic parameters of the proposed business process representation and problem formulation such as library size and process design size. The evaluation both of the EMOAs and the framework optimisation capability is considered as satisfactory and interesting for further investigation and research. NSGA2 proved unfit for business process optimisation whilst PESA2 showed the best results due to its sophisticated region-based selection technique. Chapter 9 moves from experimental scenarios to real-life business process scenarios testing the framework with three business processes with real-life elements.

CHAPTER 9

Validation using Real-life Business Process Scenarios

This chapter validates the capability of the proposed optimisation framework in dealing with business process designs that encompass real-life elements. To achieve this, three different scenarios –that are current practice in the service industry– are selected and tested by the framework. The procedure of tuning a real-life scenario within the proposed optimisation framework is described in detail along with the justification of selection of the particular scenarios. Also, this chapter presents the results of a series of workshops that aimed to compare the framework output with the current practice in business process design composition and optimisation.

9.1 Purpose of real-life business process scenarios

This section justifies the reason for testing the proposed business process optimisation framework with real-life scenarios and lists the main steps. Below is a definition of what is perceived by this research as a real-life business process scenario:

A **real-life business process scenario** is a set of parameters –based on the business process problem formulation– that are extracted from a real-world business process model reported in literature or captured from industry practice.

9.1.1 Aim of real-life scenarios

Testing the proposed optimisation framework with real-life scenarios aims to validate the framework's capability in capturing, composing and optimising designs of business processes that are current practice in real-life situations.

9.1.2 Main steps for testing bpo^F with real-life scenarios

The testing of the framework with real-life scenarios encompasses three main steps; these are:

1. Specification of the context in which the *selection* of the real-life business process scenarios will occur,
2. Definition of the steps for *tuning* the scenarios and *testing* them within bpo^F, and,

3. Capturing the current practice about the scenarios to compare it with the framework's optimisation approach.

The first step is to specify the context of business processes. Although the framework has been developed for generic business processes, real-life business processes might lack some of the framework's presumed elements, such as the task library. Therefore, the context will be oriented to existing business processes that encompass most of the elements assumed by the proposed optimisation framework. The context selection and specification is detailed and justified in section 9.2. Based on the context, three real-life scenarios are selected. The scenarios are taken from relevant literature. The selection of the particular scenarios is largely based on the degree they meet the requirements posed by the framework.

Once the scenarios are selected, the second step involves the tuning of the scenarios in order to use them with the proposed optimisation framework. Tuning the real-life scenarios essentially means following a set of steps that bring the scenarios in such a form that they can work with the framework. The strategy for tuning the scenarios is described in section 9.3. When the scenarios are tuned for the framework, they are executed in order to obtain the optimisation results. In the case of real-life scenarios, the framework is not only evaluated based on the performance of the algorithms but also whether it generates meaningful business process designs that can efficiently replace the one currently in practice. Each of the real-life scenarios is presented in sections 9.4 (scenario A), 9.5 (scenario B) and 9.6 (scenario C).

Finally, the third step involves capturing the current practice of business process composition and optimisation to effectively evaluate the performance of the proposed optimisation framework. For this reason, a series of workshops with business process experts took place as the concluding part of this research. The outcome of the workshops – discussed in section 9.7– helped to get the opinion of the experts on the output of the proposed research, compare it with the current practice and raise some issues about the strengths, weaknesses and future orientation.

9.2 Selection of real-life scenarios

This section discusses the selection of the real-life business process scenarios that the proposed optimisation framework is validated with. The context of the real-life scenarios is first presented and the resulting features of the scenarios are then discussed. The section concludes with a brief description of the three real-life scenarios and a justification for their selection.

9.2.1 Context of real-life scenarios

This sub-section sets the context of selecting the real-life scenarios by presenting two perspectives of business processes: (i) business process automation and (ii) business processes as a Service System. These two perspectives are selected by the researcher as they provide the real-life features of business process scenarios that the framework can be validated with.

Business process as a Service System

A *Service System* is a configuration of technology and organisational networks designed to deliver services that satisfy the customer (Reijers, 2002). Since the context of this research is business processes in the service industry (see chapter 3), business processes can be perceived as a generic type of a Service system. The issues of designing a Service system have given rise to SOA (Service-Oriented Architecture) which is an architectural style for creating and using business processes packaged as services throughout their life-cycle. SOA regards the tasks in a business process as separate units which can be distributed over a network and can be combined and reused to create business applications. These units are called *web services*. A web service is a discretely defined set of contiguous and autonomous business or technical functionality implemented over a network.

Essentially, perceiving a business process as a Service system allows for the tasks in the business process design to be implemented by web services.

Another advantage of adopting the Service system perspective towards business processes is the service specification attributes that can be used as process attributes and consequently as optimisation objectives by the proposed framework. According to Lakin *et al.*, (1996), any service can be completely, consistently and clearly specified by means of the following 12 *service specification attributes*:

1. Service Consumer Benefit(s)
2. Service-specific Functional Parameter(s)
3. Service Delivery Point
4. Service Consumer Count

5. Service Readiness Time(s)
6. Service Support Time(s)
7. Service Support Language(s)
8. Service Fulfilment Target (SFT)
9. Maximum Impairment Duration per Incident
10. Service Delivering Duration
11. Service Delivery Unit
12. Service Delivering Price (SDP)

Business Process Automation (BPA)

Business process automation (BPA) is the replacement of manual business processes with automated ones using advanced technologies (Grigori *et al.*, 2004). The benefits of automation are that processes can be executed faster, with lower costs (due to the reduced human involvement), and in a controlled way, since the enactment system can detect exceptions or delays in process executions and react to them in the way specified by the process designer. As more and more processes become automated, the focus of both industry and academia shifts from deployment to process monitoring, analysis, and optimisation (Grigori *et al.*, 2001). Business process automation assumes that the business processes are correctly designed, their execution is supported by a system that can meet the workload requirements, and the (human or automated) process resources are able to perform their work items in a timely fashion (Castellanos *et al.*, 2004). According to these authors, a business process can be classified according to its automation level, as

- *Manual*, with little or no application support for the business process operation,
- *Semi-automated*, with several (non co-ordinated) software systems to perform the process, and,
- *Automated*, which assumes an end-to-end fully automated process from input to final output (integrated to or orchestrated by a single application).

9.2.2 Features of real-life business processes

The context of the real-life scenarios, as discussed in the previous section, helps in aligning the real-life scenarios with business process elements suggested by the proposed representation and optimisation framework. The specific features of the real-life scenarios and their match with the elements proposed by this research are shown below:

Real-life scenarios :: Different levels of Business Process Automation (BPA)

The classification based on BPA levels provides a good starting point for validating the framework. Each of the three real-life scenarios belongs to one of the three levels of business process automation. This can help in determining whether the proposed framework can handle business processes with different automation levels. It is expected that the automated business process scenario will already have a design and the benefit from the framework application will lie more on the optimisation side, whereas the manual business process will benefit both from the automation of the process composition activity and also of the optimisation activity.

Task library :: (On-line) libraries of web services

A web service is a software system that performs a task and is designed to interact over a network. There is a trend to compose business processes (or mash-ups) with web services as the participating elements (tasks). Taking into account that there are libraries of web services available through the Internet and that the proposed framework is largely based on a library of tasks, adopting the 'web service' perspective can be an opportunity to (a) demonstrate the automation and optimisation capabilities of the framework using real examples of business processes created by available web services and (b) to stress the importance of embracing web services as a crucial element for the future of business process automation and improvement. Therefore, the real-life business process scenarios will be composed of web services that are stored in relevant on-line libraries.

Optimisation objectives :: Service specification attributes

The proposed optimisation framework is oriented towards bi-objective optimisation. It is also focused on min-max problems, i.e. the minimisation of the first objective and the maximisation of the second. Based on the service specification attributes, *Service Delivery Price (SDP)* is selected as the first objective. SDP specifies the amount of money the service customer has to pay for the consumption of distinct service volumes, i.e. the cost to use the service. The second objective is the maximisation of *Service Fulfilment Target (SFT)*. SFT specifies the service provider's promise of effective and seamless delivery of the defined benefits to any authorised service consumer requesting the service within the defined service times. It is expressed as the promised maximum number of successful individual service deliveries with respect to the total counts of individual service deliveries. SFT can be measured and calculated per service consumer or per consumer group and may be referred to different time periods. These two services attributes will be used as the optimisation objectives for the real-life scenarios.

9.2.3 Real-life scenarios

This chapter validates the proposed business process optimisation framework with three real-life business process scenarios. These scenarios are selected based on the business process automation classification. A scenario from each category is tested with the proposed optimisation framework. The aim is to show the versatility and capability of the framework to automate and optimise business processes for each level. All three scenarios are taken from literature and are tuned appropriately based on the proposed business process representation (see chapter 5) in order for them to be tested with the framework for the generation of optimised business process designs.

For the real-life scenario selection the criteria were three: (i) to be reported as current practice in business process literature, (ii) to be able to tune with the proposed framework (i.e. encompass all or most of the required elements) and (iii) to be able to demonstrate that adopting web services as the participating tasks in the process design progresses business processes closer to SOA. The selected scenarios are:

- Scenario A, which describes an *automated* business process (**On-line order placement**) and it is discussed in detail in section 9.4,
- Scenario B a *semi-automated* business process about **Sales forecasting** discussed in section 9.5, and,
- Scenario C a *manual* process about **Fraud investigation** detailed in section 9.6.

Each of the three real-life scenarios belongs to a different classification in terms of automation. Scenario A is about an On-line order placement business process. The process is fully automated, enacted by one or more software applications and published over the Internet. This scenario takes advantage of the proposed representation of business processes in the context of web services. The proposed framework is expected to demonstrate its capability to generate a number of alternative optimised designs given the structured nature of the process.

The second scenario discusses the semi-automated process of Sales forecasting. This process involves both manual and automated elements. The challenge for the proposed framework is not only to optimise, but also to model effectively the business process and automate it. Although the optimised alternatives are not expected to be large in number, the main benefit of adopting the approach suggested from this research will lie in the automation of the design composition activity.

Finally, the third scenario involves the manual business process of Fraud investigation. This business process is selected due to its unstructured and loose nature. Fraud investigation might involve a number of different activities at different times. In this case the composition capability of the framework is expected to demonstrate flexibility in composing a large variety of designs. These three scenarios were selected in order to demonstrate and emphasise on the composition and optimisation capabilities of the proposed framework. The methodology for tuning and testing each of these scenarios is presented in detail in the following section.

9.3 Tuning and testing the real-life scenarios

This section presents the main steps for tuning the real-life scenarios in order to test them with the proposed optimisation framework. Figure 9.1 shows these steps which are classified in two phases: Phase I that involves the tuning steps and Phase II that involves the steps required for testing the scenario and obtaining the results. Each of these steps is briefly discussed below:

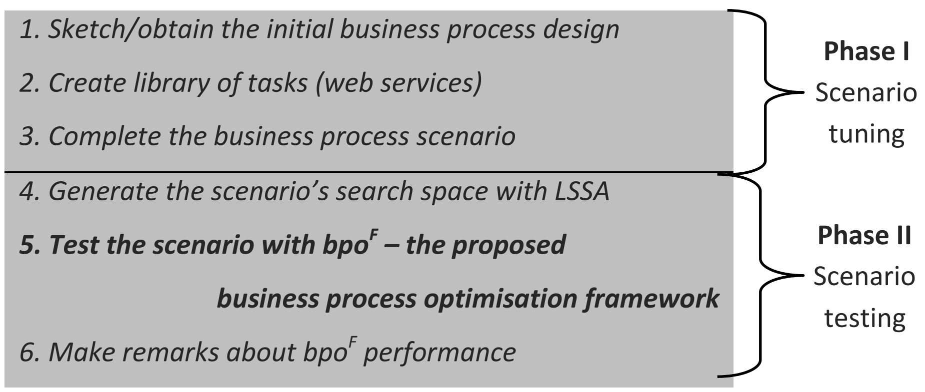


Figure 9.1. Main steps of tuning and testing a real-life business process scenario

(1) Sketch/obtain the initial business process design

The first step towards testing a real-life business process scenario is to sketch the initial business process design or to obtain it (if it is provided by the source). This is essential in order to capture and understand the operation and the flow of the business process. The elements that are captured are: (i) the main steps (tasks) of the process and (ii) the process inputs and outputs. Based on these, an initial sketch of the business process design can be created. This helps in acquiring a first idea of the solutions that the framework is expected to generate and also for compiling the task library in the following step.

(2) Create library of tasks (web services)

In the second step, the initial business process design is used as a guide for locating relevant web services. These web services will constitute the task library for the particular scenario. The main on-line sources of web services that were used in the context of this research are listed in appendix E. These libraries provide relevant web services in order to compile the task library for each of the real-life business process scenarios. As the web services are obtained from different sources, their input and output resource names were modified (for those web services that perform the same operation and have similar inputs and outputs).This facilitates the framework's operation as it creates a degree of similarity among the web services in the library and thus provides more alternatives to the framework.

(3) Complete the business process scenario

The last step of tuning the scenario is to complete the scenario by specifying the remaining parameters such as the task attribute values. As soon as the scenario is completed, it is encoded in a dedicated Java class and implemented within the framework

(4) Generate the scenario's search space

The first step of the scenario testing phase (Phase II) is to approximate the scenario's search space using large scale search. The search space is generated using the LSSA algorithm (see chapter 7). The search space provides an outlook on all the possible solutions of the problem.

(5) Test the scenario with bpo^F

The next step involves testing the scenario with the bpo^F. This is the central step where the formulated scenario is incorporated in the proposed optimisation framework. The framework applies the four employed EMOAs –NSGA2, SPEA2, PESA2 and PAES– and generates optimisation results for each of the algorithms. Also, a set of representative optimised business process designs is demonstrated for each of the examples. It is important in the case of real-life scenarios to demonstrate how an optimised solution corresponds to an actual business process design. The generated optimisation results are plotted along with the problem's search space and evaluated using the appropriate metrics (see chapter 8). Also, the generated designs can be compared with the initial design in order to identify framework's strengths and/or weaknesses in terms of design composition and optimisation.

(6) Make remarks about bpo^F performance

The last step assesses the performance of the framework in relation to the real-life scenario. It draws remarks about the framework optimisation performance and its capability of generating alternative optimised business process designs. These remarks define the capabilities of the framework in dealing with real-life scenarios, its limitations and the potential for expansion in order to fully address real-life business processes.

9.4 Scenario A: On-line order placement

The first scenario discusses the business process of placing an order in an on-line store (Havey, 2005). This process is considered as *automated* as the online store already has an end-to-end integrated application for successfully receiving customer orders. The aim of this scenario is to show the optimisation potential of the framework for an automated business process. In automated process like the one described here, the process steps are clearly defined and the process is end-to-end integrated. It is expected that the framework will showcase a range of alternative equivalent process designs that perform the same operation with optimised attributes. To get the framework results, the steps introduced in the previous section for tuning and testing a real-life scenario are applied below to the on-line order placement scenario.

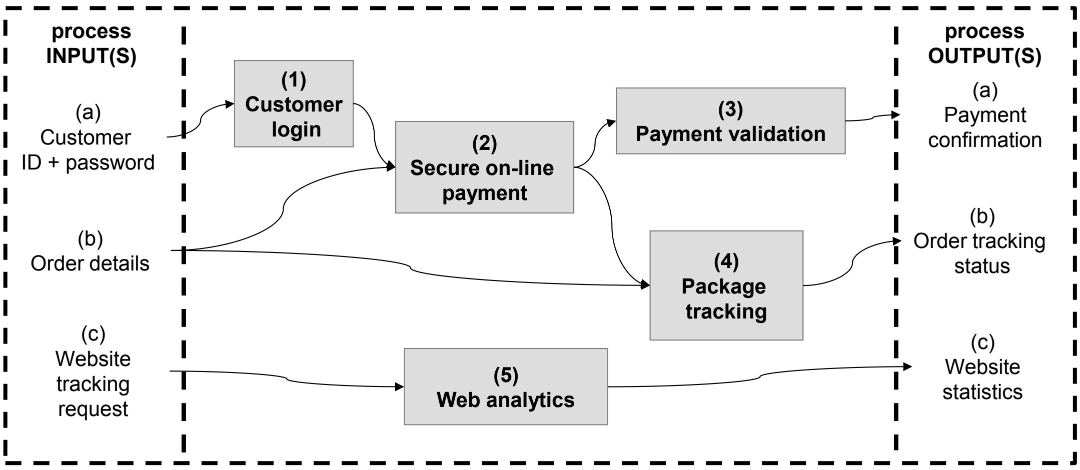


Figure 9.2. Initial business process design of on-line order placement (scenario A)

(1) Initial business process design

Figure 9.2 shows the initial business process design for scenario A. The design involves the main process steps, how they are interconnected and the identification of the process inputs and outputs. It is important to note that the process design is sketched from the business analyst's perspective not the customer's. Therefore, the inputs of the business process design are what the business analyst considers as process requirements even if they do not directly impact the customer's interaction with the process (e.g. website tracking request). The same applies to the process outputs that might not affect the customer (e.g. website statistics) but have a crucial informative role for the company. Scenario A starts with three process inputs: (a) Customer ID & password, (b) Order details and (c) Website tracking request. The basic steps of the process are five. The customer credentials are necessary to access the on-line store (step 1) together with the order details to place the order and pay for it (step 2). Paying for the order invokes the payment validation (step 3) and the monitoring of the order progress (step 4). Also, the web analytics (step 5) track the customer's behaviour in the website. The three outputs of the process are: (a) payment confirmation, confirms that the payment processing is successful, (b) order tracking status returns the order status in terms of delivery to the customer and (c) website statistics record the customer's behaviour in the website and influence the store's marketing strategy in terms of customer's individual needs.

No.	Resource name
0	Customer account credentials
1	Customer account details
2	Order details
3	Payment details
4	Payment confirmation
5	Order tracking status
6	Website tracking request
7	Website statistics

Table 9.1. Available resources (R) for scenario A

(2) Library of tasks (web services)

Having sketched the initial business process design, we can compile the library of alternative web services based on the main steps of the generic process design. Relevant research on the selected on-line libraries of web services (see appendix E) resulted in a selection of 29 web services from different providers that can potentially implement scenario A. Appendix E provides the list of web services along with the source from which they were obtained, a description of their operation (as provided by their source) and most importantly the identified inputs and outputs. The inputs and outputs of each web service were either provided by their source or derived from the web service description. In both cases, the resource names have been modified (i.e. similar resources among different web services are given the same name) in order for these services to be considered as alternatives by the proposed optimisation framework.

No.	Task Name	Input(s)	Output(s)	SDP	SFT
0	Achworks Soap (T$$ - Rico Pamplona)	1, 2	3	208	113
1	BAX Global Tracking Service	2, 3	5	219	109
2	CDYNE Death Index	1, 3	4	229	115
3	Credit Card Processor	1, 2	3,4	202	109
4	D&B Business Credit Quick Check	1, 3	4	203	108
5	Drupal authentication	0	1	200	103
6	ecommStats Web Analytics	6	7	218	112
7	Entrust login	0	1	206	103
8	FedEx Tracker	2, 3	5	211	109
9	FedEx / UPS Package Tracking	2, 3	5	224	103
10	FraudLabs Credit Card Fraud Detection	1, 3	4	220	113
11	Google Analytics	6	7	218	107
12	Google Checkout	1, 2	3	206	105
13	GUID Generator	0	1	203	110
14	Internet Payment Systems	1, 2	3	226	105
15	LID login	0	1	222	114
16	OpenID login	0	1	228	100
17	Paypal online payment	1, 2	3	215	102
18	Real Time Check Verification (T$$ - Rico Pamplona)	3	4	229	108
19	Rich Payments NET	2	3, 4	208	105
20	SAINTlogin users validation	0	1	219	105
21	Servicetrack	6	7	212	113
22	SmartPayments Payment	2	3	214	105
23	Smartpayments CardValidator	3	4	206	107
24	StrikeIron Global Address Verification	1, 3	4	201	105
25	SXIP login	0	1	203	105
26	Typekey authentication service	0	1	224	114
27	UPS Tracking	2, 3	5	225	109
28	VeriSign Payment	1, 2	3	230	103

Table 9.2. Task library for scenario A

(3) Complete business process scenario

Having gathered all the necessary information (library of web services and input/output resources for each), the problem parameters can be defined based on the business process problem formulation. Table 9.1 shows the 8 different resources among the web services in the library (counting starts from zero due to Java programming constraints). The process input and output resources can be also identified in this table. Table 9.2 summarises the task library of web services as selected from the on-line libraries. The library tasks (web services) are laid in alphabetical order and since they are 29, each gets a unique number from 0-28. For each task, its input and output resources are in the adjacent columns. The proposed optimisation framework is tested for two objectives, SDP and SFT. As the service providers do not provide detailed information on the performance of each web service, the values allocated are based on the uniform distribution of a small range of values on each web service. This makes the available web services competitive as they have little difference from each other. Based on these, table 9.3 summarises the parameter values for scenario A. This makes a complete scenario ready to be tested within the proposed optimisation framework.

Parameter	Value
n	29
n_d	5
r	8
t_{in} / t_{out}	1-2
r_{in} / r_{out}	3
p	2
SDP	200 – 230
SFT	100 – 115

Table 9.3. Parameter values for Scenario A

(4) Generate the scenario's search space

The first step towards obtaining the results is to generate the scenario's search space. The LSSA algorithm is executed for n_d = 8 and n_{min} = 4 tasks. The initial business process design in figure 9.2 involves 5 main steps. A design with less that 5 tasks shows that there is a web service that consolidates two or more tasks. Additionally, a design with more tasks shows that one step requires two or more web services to be implemented. The search space for this scenario is shown on figure 9.3(a). The search space consists of five different regions, each corresponding to a group of designs with same number of tasks (4, 5, 6, 7 or 8).

(5) Test the scenario within bpo^F

The challenge for the EMOAs in the framework is to identify non-dominated (optimised) solutions in each of the regions in the search space. Figure 9.3(b, c, d and e) shows the results for each of the optimisation algorithms and figure 9.3(f) shows the combined results. All the algorithms identify the same Pareto-front in terms of optimum solutions. This is a strong indicator of the performance of the algorithms and the confidence in the generated designs being optimal.

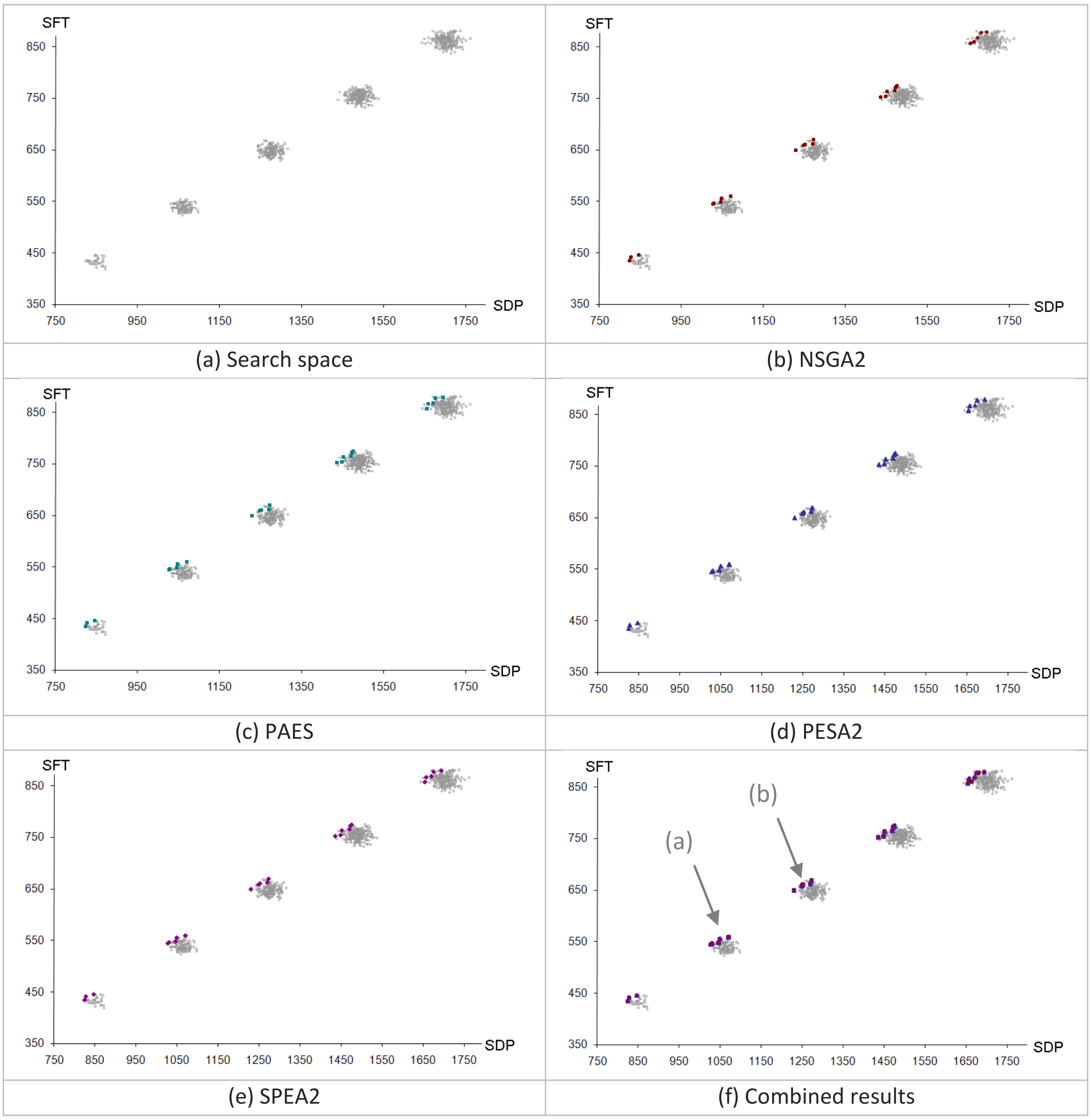
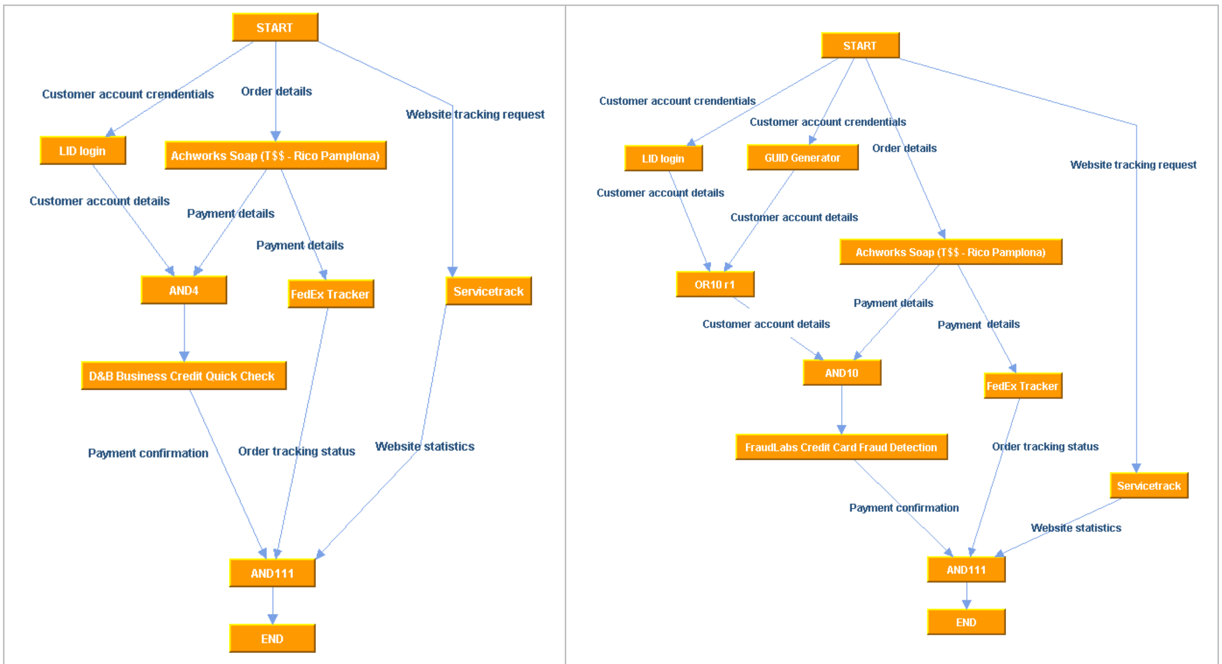


Figure 9.3. Search space and EMOA results for scenario A

Figure 9.4 demonstrates two optimised business process designs (a) and (b), one with 5 tasks and one with 6 tasks. Each of these designs belongs to a different island based on its solution size. The arrows in figure 9.3(f) indicate the island where each design in figure 9.4 originates. The design in figure 9.4(a) is following the initial design of figure 9.2. It implements a web service for each of the steps and demonstrates the AND pattern which is required in two cases: (i) for the payment validation and (ii) to ensure that all of the process outputs are produced in the end. For each of the generic steps, figure 9.4(a) has selected a specific web service implementation (e.g. 'FedEx tracker' for tracking the order) optimising the already automated process by selecting the web services with better combination of attribute values.

Figure 9.4(b) shows a 6-task process with different implementations for the majority of tasks compared to (a). It provides two alternatives for the login process making the process more reliable and thus achieving bigger SFT (Service Fulfilment Target) values compared to (a). These two examples select different service implementations in login, payment and validation steps as for each design the particular combination provides optimal attribute values. Both examples implement the 'FedEx tracker' and 'Service track' web services for tracking the order and monitoring the website statistics respectively. The design in figure 9.4(a) has lower price (low SDP) but it is also less reliable (low SFT). On the contrary, design (b) has increased price but also it is more reliable. For example it provides two alternative web services for login.



(a) 5-task process design	(b) 6-task process design

Figure 9.4. Optimised business process designs for scenario A

Apparently these two services have a clear advantage in terms of attribute values compared to their alternatives. The two demonstrated business process designs and the complete set of generated designs are all optimal and equivalent in terms of trade-offs between their two objectives. Selecting one depends on higher level preference criteria which may involve: (i) specific number of tasks in the design, (ii) emphasis on one of the two objectives or (iii) other parameters.

(6) Remarks for scenario A

This scenario demonstrated how an automated process can be benefited from the proposed optimisation framework:

- For each of the main process steps, a series of alternatives was identified. This was relatively straight-forward as the process is clearly defined.
- The optimisation framework managed to identify optimal designs based on the two objectives for all the available process sizes.
- The generated designs select and incorporate different web services arranged with the appropriate process patterns so that (i) the process input and output requirements are satisfied and (ii) the attribute values are optimised.

9.5 Scenario B: Sales forecasting

The second scenario describes the business process of sales forecasting (Grigori *et al.*, 2004). This process is considered as *semi-automated* as it involves the interaction of some applications but it is not streamlined and still requires human involvement in the process of generating and visualising the requested forecasts. The framework is expected to fully automate the process by selecting and implementing relevant web services and propose a set of optimised designs that fulfil the process requirements having optimal attribute values.

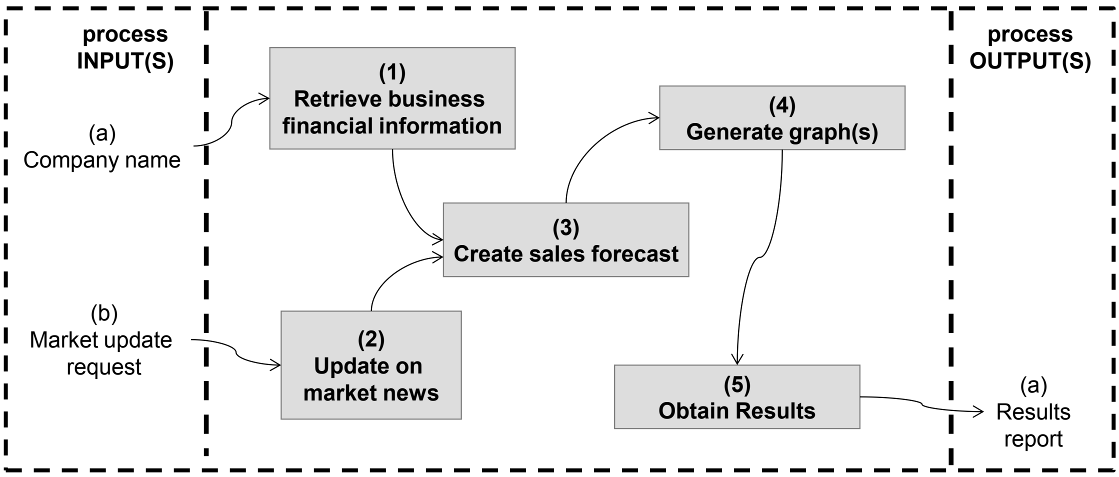


Figure 9.5. Initial business process design of Sales forecasting (scenario B)

(1) Initial business process design

Figure 9.5 shows the generic business process design for scenario B. It involves two input resources: (a) company name and (b) market update request. The first resource is necessary for the web service to extract relevant data for the specified company. The second resource is a request for a market update that needs to be considered for the sales forecast. The initial process design consists of five steps: the first step is to retrieve the relevant financial business information for the relevant forecast. Parallel with that, the latest market levels (e.g. stock level) need to be updated as they are taken into account for the relevant forecast (step 2). The outcome of these two is then fed to a Monte-Carlo simulation that generates the sales forecast (step 3). This forecast is then plotted in a graph (step 4) and then communicated back to the person requesting it (step 5). The outcome of this process is a report containing the forecast results.

(2) Library of tasks (web services)

The second step is the compilation of the web services library based on the main steps of the process. Relevant research on the selected on-line libraries of web services resulted in a selection of 20 web services from different providers. Appendix E provides this list of web services for scenario B along with the source from which they were obtained, a description of their operation (as provided by their source) and the identified inputs and outputs. The inputs and outputs of each web service were either provided by their source or derived from their description. In both cases, the resource names have been modified in order for these services to be considered alternatives by the proposed framework.

No.	Resource name
0	Business details
1	Business query
2	Chart / graph
3	Company name
4	Fax (on-line)
5	Financial data
6	Market update request
7	Recent market trends
8	Time-series forecast

Table 9.4. Available resources (R) for scenario B

(3) Complete business process scenario

Having gathered all the necessary information (library of web services and input/output resources for each), the problem parameters can be defined. Table 9.4 shows the 9 different resources among the web services in the library. Table 9.5 summarises the task library of 20 web services as selected from the on-line libraries. For each task, its input and output resources as well as the SDP and SFT values are in the adjacent columns. Based on these, table 9.6 summarises the parameter values for scenario B.

No.	Task Name	Input(s)	Output(s)	SDP	SFT
0	D&B Business Verification	3,1	0,5	206	103
1	Fax.com	8,2	4	220	103
2	Gale Group Business Information	3,0	5	223	106
3	Gale Group Business Intelligence	3,1	0,5	229	113
4	GraphMagic's Graph & Chart Web Service API	5,8	2	203	107
5	interfax.net	8,2	4	222	113
6	Lokad Business time-series forecasting and analysis	5,7	8	228	110
7	Midnight Trader Financial News	6,3	7	217	101
8	StrikeIron Company Search	3	3,0	230	114
9	StrikeIron Get Business Prospect	3,1	0,5	205	110
10	StrikeIron Lookup Business	3	3,0	201	110
11	Wall Street Horizon Real-Time Company Earnings	3,1	5	210	105
12	Xignite Get Balance Sheet	3,1	5	216	112
13	Xignite Get Chart Url	8	2	228	110
14	Xignite Get Chart Url Preset	8	2	228	101
15	Xignite Get Growth Probability	5,7	8	215	109
16	Xignite Get Market News Headlines	6	7	221	114
17	Xignite Get Market Summary	6	7	203	112
18	Xignite Get Topic Chart	3,5,7	8	218	112
19	Xignite Get Topic Data	3,5,7	8,2	222	109

Table 9.5. Task library for scenario B

Parameter	Value
n	20
n_d	5
r	9
t_{in} / t_{out}	1-3
r_{in} / r_{out}	2 / 1
p	2
SDP	200 – 230
SFT	100 – 115

Table 9.6. Parameter values for Scenario B

(4) Generate the scenario's search space

The first step towards obtaining the results is to generate the scenario's search space. The generic business process design contains 5 basic steps. In this scenario we will push a bit further and see if the process steps can be consolidated to 3 steps and therefore we will test the framework for designs that contain 3-6 services. The search space for this scenario is shown on figure 9.6(a) The search space consists of four different regions, each corresponding to a group of designs with same number of tasks (3, 4, 5 or 6), it is scarce with only 4 solutions in the lowest region (designs with 3 tasks).

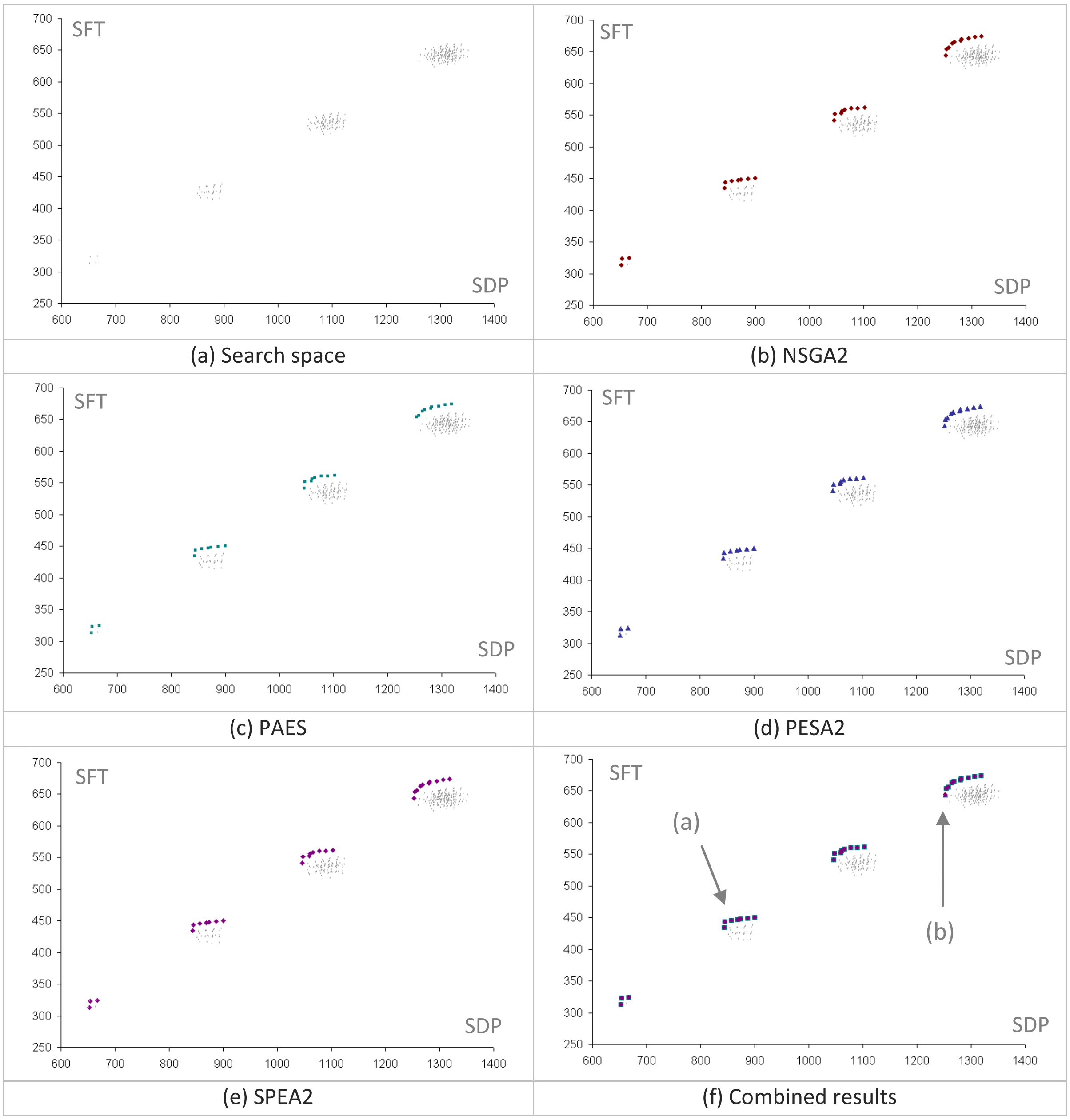
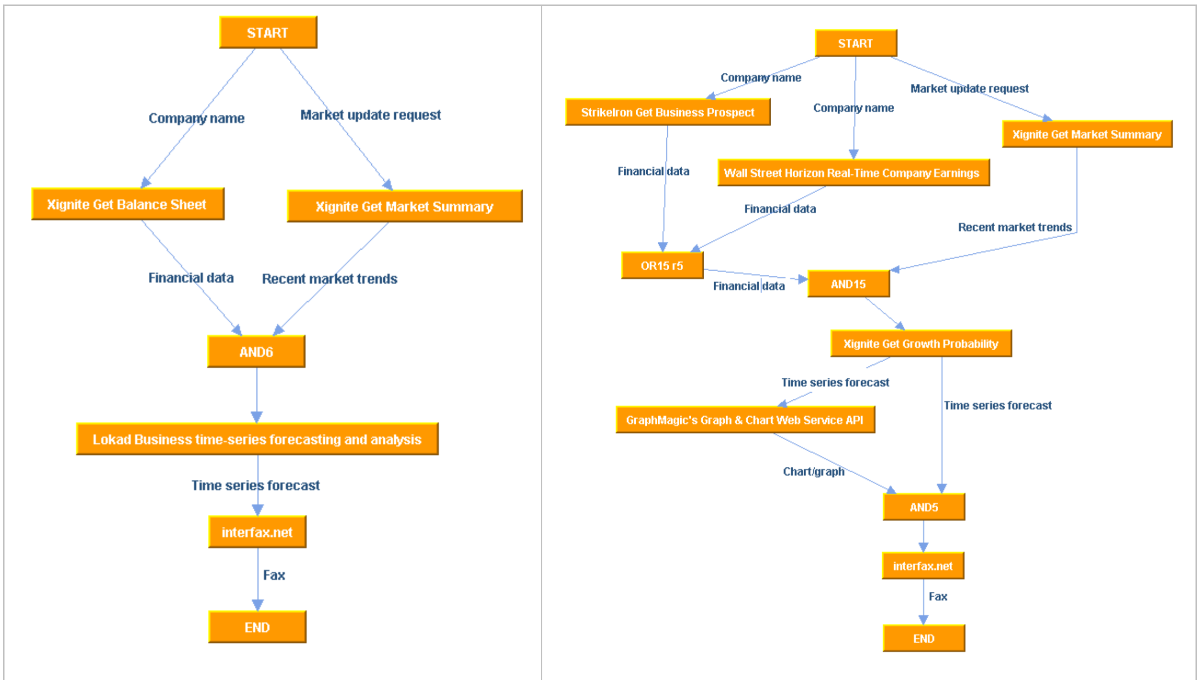


Figure 9.6. Search space and EMOA results for scenario B

(5) Test the scenario within bpo[F]

The challenge for the EMOAs in the framework is to identify non-dominated (optimised) solutions in each of the regions. Figure 9.6(b, c, d and e) shows the results for each of the optimisation algorithms and figure 9.6(f) shows the combined results. All the EMOAs identify non-dominated solutions in all the regions and they all shape the same Pareto-front. This demonstrates the working of the framework and the performance of all the algorithms. Also, it shows good quality of the results since the generated designs are optimal in terms of their attribute values.

Figure 9.7 demonstrates two optimised business process designs (a) and (b), one with 4 and one with 6 tasks. Each of these designs belongs to a different island based on its solution size. The arrows in figure 9.6(f) indicate the island where each design in figure 9.7 originates. Figure 9.7(a) shows a business process design with one of the generic steps missing. The forecasting results are not plotted into a graph but they are just faxed back to the requestor. The proposed framework reduces cost in this instance. Therefore, in a semi-automated process the framework can take 'initiative' and alter the generic design provided that the process input and output requirements are still satisfied. Figure 9.7(b) is composed of 6 services and involves two tasks for obtaining the company's financial data either from selecting one or both (OR is not exclusive choice). This provides better confidence in terms of accuracy of the data obtained and more reliability to the process execution itself.



(a) 4-task process design | (b) 6-task process design

Figure 9.7. Optimised business process designs for scenario B

(6) Remarks for scenario B

In addition to the remarks based on scenario A, creating and testing a semi-automated process (scenario B) within the proposed framework raises additional benefits:

- For each step in the generic design, more flexible alternatives can be identified.
- The framework can modify the initial design by adding or removing steps provided that the process requirements are satisfied, thus enhancing the decision making capabilities of the process analyst.
- The number of tasks in the design directly affects the optimisation objectives either by reducing cost (less tasks) or by increasing the process reliability (more tasks).

9.6 Scenario C: Fraud investigation

The third scenario describes the business process of fraud investigation (Havey, 2005) which takes places when there is a suspicion of customer identity fraud and consequent loss by misusing company goods or services. This process is considered as *manual* as there is no standard procedure to be followed in the investigation; there is no complete software application that can track, identify or prevent fraud. As a result, fraud investigation involves manual investigation of data that the company maintains. The benefits of adopting a web services approach and implementing the proposed optimisation framework would be (a) standardising the process, (b) making it more reliable, (c) automating it and (d) optimising it. The sections below detail the necessary steps towards that direction.

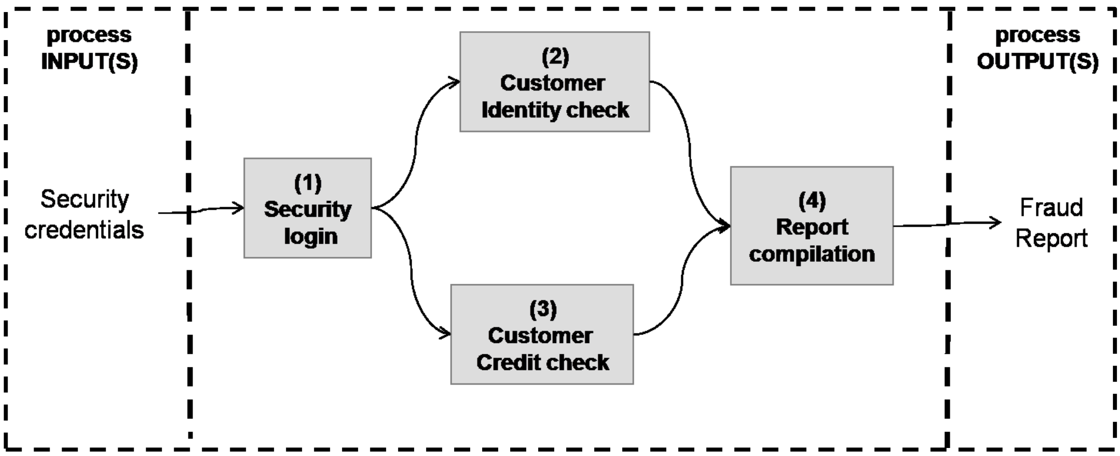


Figure 9.8. Initial business process design of fraud investigation (scenario C)

(1) Initial business process design

Figure 9.8 shows the initial business process design for scenario C. It consists of four steps, two of which are specific and two generic due to the manual nature of the process.

The process starts by requesting the security login credentials in order to access the data (step 1 – specific). Then two parallel checks occur: one related to the customer's identity check (e.g. address verification) and one related to credit check. Both these steps (2 and 3) are considered generic as there is no standard procedure for identifying a possible fraud. After the checks are completed, the outcomes are compiled into a report which is the single outcome of this scenario. Based on the fraud investigation report, the company can then take further action.

No.	Task Name	Input(s)	Output(s)	SDP	SFT
0	Address Doctor Global Address Verification	2	3	209	106
1	cbarron bankValidate	1	0	212	114
2	CDYNE Death Index	2, 1	3, 0	227	102
3	CDYNE Email Verifier	2	3	210	105
4	CDYNE Phone Verifier	2	3	212	109
5	D&B Business Credit Quick Check	1	0	201	101
6	D&B Business Verification	2, 1	3, 0	207	110
7	Dimple Email Address Validator	2	3	208	112
8	Drupal authentication	5	2, 1	205	103
9	Dun & Bradstreet Business Credit Quick Check	1	0	215	109
10	Dun & Bradstreet Business Verification	2	3	228	108
11	Entrust login	5	2, 1	228	104
12	FraudLabs Credit Card Fraud Detection	2, 1	0	213	115
13	Google Docs	3, 0	4	216	106
14	GUID Generator	5	2,1	219	109
15	LID login	5	2,1	224	104
16	OpenID login	5	2, 1	225	113
17	Real Time Check Verification (T$$ - Rico Pamplona)	1	0	211	109
18	SAINTlogin users validation	5	2, 1	211	103
19	Smartpayments CardValidator	1	0	224	110
20	StrikeIron 24-hour Accurate Residential Lookup	2	3	204	112
21	StrikeIron 24-hour Accurate Reverse Phone Lookup	2	3	211	111
22	StrikeIron Email Verification	2	3	215	113
23	StrikeIron Gender Determination	2	3	215	102
24	StrikeIron Global Address Verification	2	3	211	111
25	StrikeIron Reverse Phone Residential Intel	2	3	203	111
26	StrikeIron Reverse Residential Lookup	2	3	222	107
27	SXIP login	5	2, 1	216	110
28	Typekey authentication service	5	2, 1	206	114
29	Web Services Security Monitor	5	2, 1	227	110
30	webba E-Mail validator	2	3	202	112

Table 9.7. Task library for scenario C

(2) Library of tasks (web services)

The second step is the compilation of the web services library based on the main steps of the process. In this scenario however, steps 2 and 3 are not specific enough to be implemented by standard web services. This is due to the fact that the scenario deals with a manual process for which there is no existing application integration or orchestration approach. Relevant research on the selected on-line libraries of web services resulted in a selection of 31 web services from different providers that implement different types of checks (ID and/or credit). Appendix E provides this list of web services along with the source from which they were obtained, a description of their operation (as provided by their source) and the identified inputs and outputs. The available resources in this scenario are only 6 as the inputs and outputs of steps 2 and 3 are also kept generic

No.	Resource name
0	Credit assessment
1	Customer Credit details
2	Customer ID details
3	ID verification outcome
4	Risk Assessment Report
5	Security login credentials

Table 9.8. Available resources (R) for scenario C

Parameter	Value
n	31
n_d	4
r	6
t_{in} / t_{out}	1-2
r_{in} / r_{out}	1 / 1
p	2
SDP	200 – 230
SFT	100 – 115

Table 9.9. Parameter values for Scenario B

(3) Complete business process scenario

Table 9.7 summarises the task library of the 31 web services as selected from the on-line libraries. Table 9.8 shows the 6 different resources among the web services in the library. For each task, its input and output resources as well as the SDP and SFT values are in the adjacent columns. Based on these, table 9.9 summarises the parameter values for scenario C.

(4) Generate the scenario's search space

The generic business process design contains 4 basic steps. However, two of them are not clearly specified; they are implemented by a series of different web services. Therefore we will test the framework for process designs that span across 4 to 15 tasks. A process design with many tasks will implement more checks and therefore perform a more strenuous investigation for high profile cases. A process design with fewer tasks will perform a cost-efficient less-extensive investigation suitable for low-risk cases. The search space for this scenario is shown on figure 9.9(a). The search space consists of twelve different regions, each corresponding to a small dense group of designs with same number of tasks.

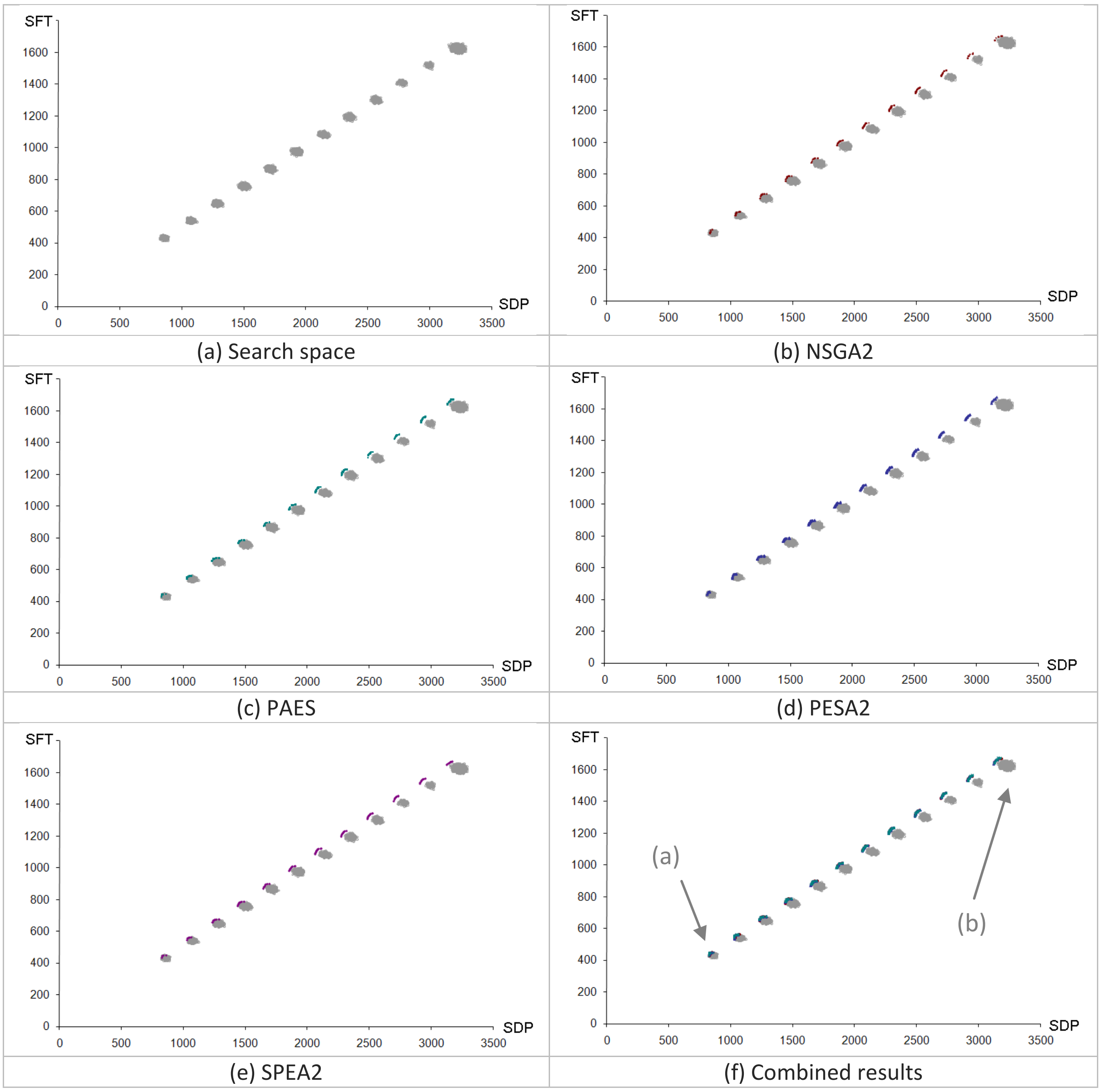
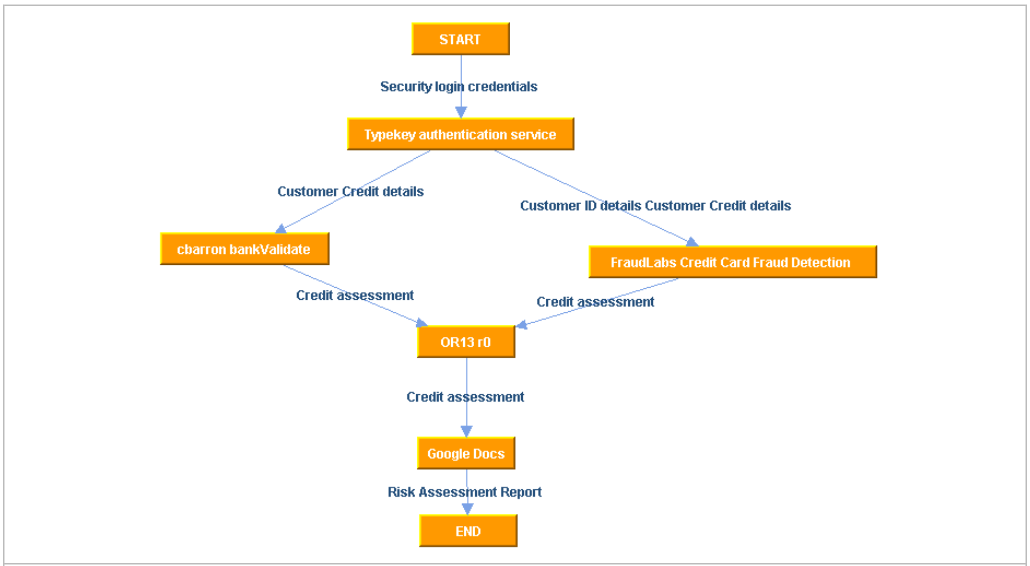


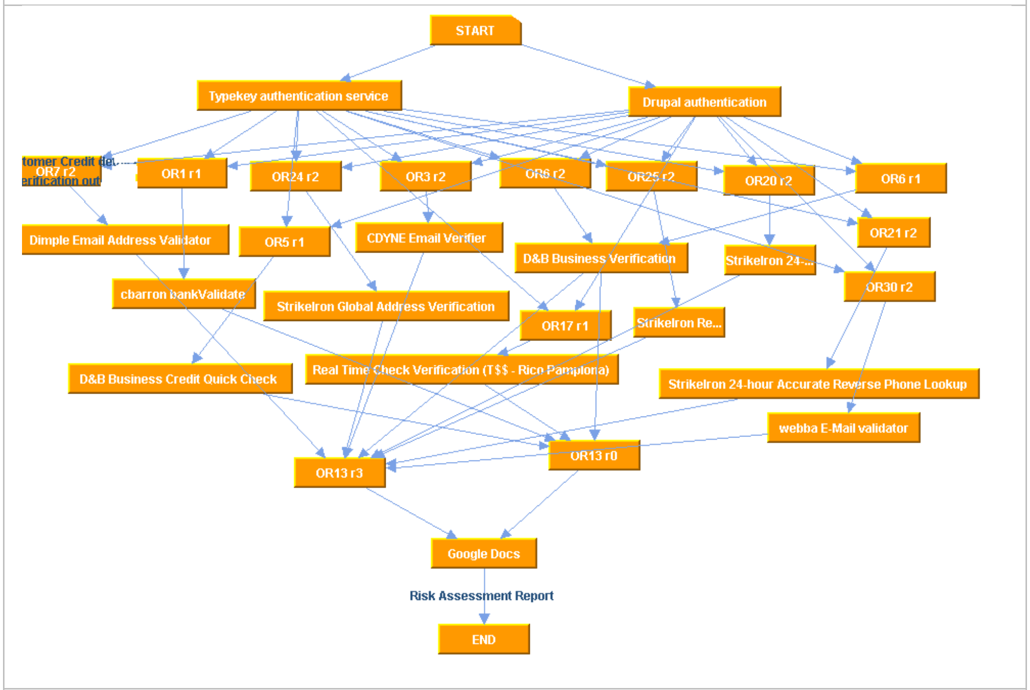
Figure 9.9. Search space and EMOA results for scenario C

*(5) Test the scenario within bpo^F^*

The challenge for the EMOAs is to identify non-dominated (optimised) solutions in each of the regions. Figure 9.9(b, c, d and e) shows the results for each of the optimisation algorithms and figure 9.9(f) shows the combined results. All the EMOAs identify non-dominated solutions in all the regions and they all shape the same Pareto-front. This demonstrates the working of the framework and the performance of all the algorithms. Also, it shows good quality of the results since the generated designs are optimal in terms of their attribute values.



(a) 4-task process design



(b) 15-task process design

Figure 9.10. Optimised business process designs for scenario C

Figure 9.10 demonstrates two optimised business process designs (a) and (b), one with 4 tasks and one with the maximum 15 tasks. Each of these designs belongs to a different island based on its solution size. The arrows in figure 9.9(f) indicate the island where each design in figure 9.10 originates. Figure 9.10(a) demonstrates a business process design where the participating web services implement the same number of steps as in the initial design in figure 9.8. There is one service that performs the credit check and one service that performs the identity check. Figure 9.10(b) shows a business process design with the maximum number of tasks (15). The implementation of this design offers a wide combination of fraud checks for customer ID and customer credit. This design also offers two alternative security login services. Note that the resource labels are removed from the figure for clarity purposes.

(6) Remarks for scenario C

Automating and optimising a manual process demonstrates the potential of the proposed framework provided that there is a library of web services of appropriate size. Since the process is not defined in terms of specific steps, the framework has the capability of generating optimal designs with bigger variation in terms of process sizes. The generated designs can emphasise not only the preference to one or other objective but also shape the process to make it more (or less) reliable, efficient and strenuous. In the particular scenario, the number of tasks involved in the design, affect the process itself in terms of the quality of the operation itself (fraud investigation). The proposed framework can capture, automate and optimise a manual process with benefit not only to the optimisation attributes (goals) but also to the operation itself by streamlining it and providing alternatives with regard to efficiency and reliability.

9.7 Comparison with current practice

This section presents the outcome of a workshop with business process experts. The workshop attempts to demonstrate, compare and validate the proposed research as a significant shift from the current practice in business process composition and optimisation. This section describes the workshop details and records the responses, comments and issues raised by the participants.

9.7.1 Workshop details

In order to compare the proposed approach of this research with the current practice in business process composition and optimisation, a series of workshops were organised based on one of the real-life scenarios previously discussed (scenario B – Sales forecasting). This sub-section discusses the details of the workshop.

(a) Purpose of the workshop

The purpose of the workshop was to demonstrate the current manual practice in business process composition and optimisation and compare it with the approach proposed in this research.

(b) Workshop structure and material

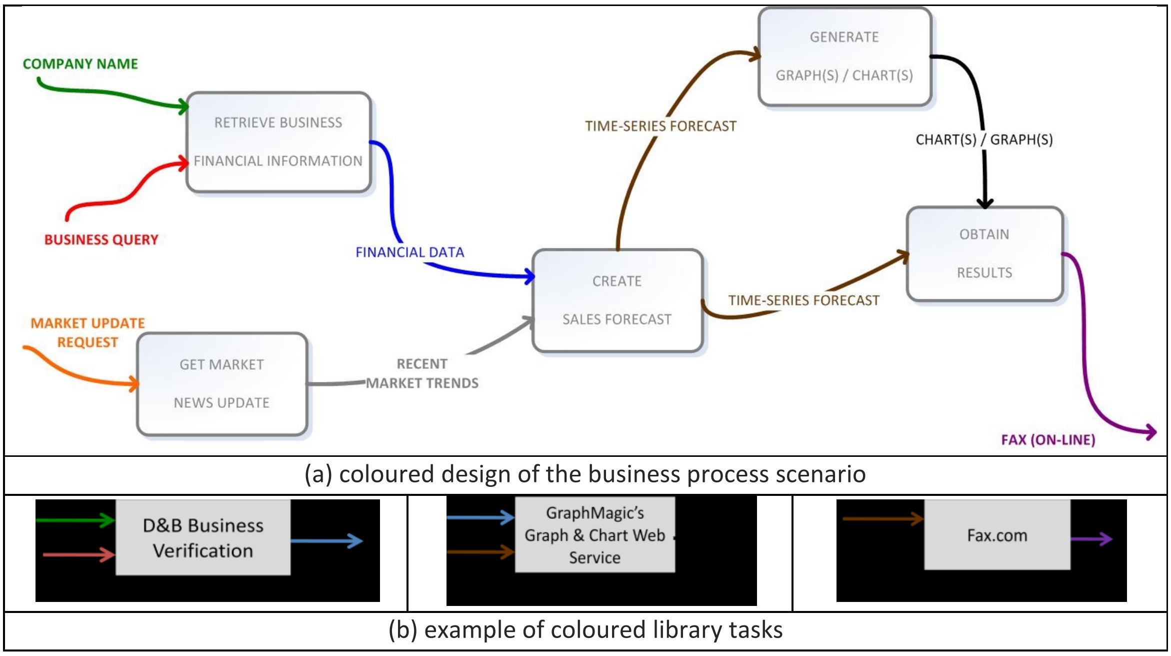
The workshop was structured in the following way:

a. Presentation of the main concepts involved in this research, the current practice and the proposed approach.
b. Exercises and questions related to the current practice of *business process composition* followed by demonstration of the proposed framework's composition capabilities.
c. Exercises and questions related to the current practice of *business process optimisation* followed by demonstration of the proposed framework's optimisation capabilities.
d. General questions, remarks and discussion.

Also, each participant in the workshop was handed the following material:

1. Hand-outs of the workshop presentation,
2. Exercise material, which involved the initial design of a real-life scenario with the library of tasks, and,
3. A questionnaire including the workshop exercises and questions (attached in Appendix E). The participants filled in the questionnaire and returned it at the end of the workshop.

The scenario that was selected for the workshop is scenario B (sales forecasting). The main reason for selecting the particular scenario was the small size of task library compared to the other two scenarios. The exercise material was coloured in order to facilitate the workshop exercises. In particular, each resource was assigned a particular colour. This helped the participants to identify the task(s) that linked to a particular resource. Figure 9.11. shows the exercise material that was handed to the workshop participants. The coloured business process design was printed in A3 paper and each library task was provided as a separate piece. This facilitated the composition and optimisation exercises as the participants had to place the appropriate library task in the grey boxes of the initial design. The coloured version of the exercises resulted after piloting the workshop.



(a) coloured design of the business process scenario

(b) example of coloured library tasks

Figure 9.11. Exercise material for the workshop

Workshop	Participants	Location	Date
a	2	BT, Ipswich	18 April 2008
b	1	Barclaycard, Northampton	28 April 2008
c	5	Cranfield University	6 May 2008

Table 9.10. Details of the workshops

(c) Workshop dates and participants

In total, three workshops took place with 8 experts in business processes. Table 9.10 provides the details of these workshops. The following sub-sections summarise the main findings of the workshops.

9.7.2 Business process composition

The first section of the workshop was related with business process composition. As chapter 5 described, PCA is capable of composing feasible business process designs based on the given requirements. The current practice is manual business process composition given a library of web services. This section of the workshop compared these two practices in order to highlight the contribution of PCA as part of the proposed framework. The questionnaire involved two exercises and three questions related to business process composition.

The first exercise asked the participants to compose a business process design with five tasks, given the initial design and the library of alternatives (figure 9.11). Since the initial design also has five tasks, the participants had to replicate the initial design finding specific web services from the library. The second exercise was of increased difficulty and asked for a design with four tasks, requesting essentially from the participants either to omit a task from the design or to locate a web service that consolidates two or more tasks. For both exercises, each participant was requested to mark the duration of the exercise (in minutes) and to rate the difficulty of the exercise on a scale 1 (easy) to 5 (very difficult).

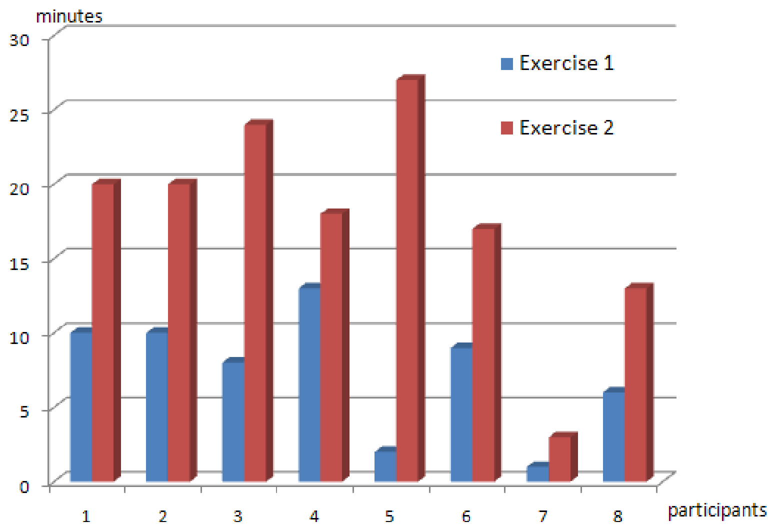


Figure 9.12. Time to complete the business process composition exercises

Figure 9.12 shows the duration for each exercise per participant. The first composition exercise took on average 7.3 minutes to complete and was rated as 'average' (2.5) in terms of difficulty. All the participants were able to compose a correct business process design with the exercise material. The second composition exercise took significantly longer to complete (average of 17.7 minutes) and was rated as 'difficult' (4.25). The majority of participants had a hard time composing a business process design with fewer tasks than the original. Some of the participants came up with designs that were not correct and had to try again within the time allotted. The aim of these exercises was to demonstrate that even with the initial design at hand, it is a challenging activity to identify the relevant web services and compose a new design. In particular, in the case of producing a modified design (with less steps), the composition process proved complicated and time consuming.

After the composition exercises, the composition capability of the proposed framework was demonstrated to the participants. PCA produces results of varying process sizes in less than a minute. After the demonstration, the participants were asked to rate the proposed composition algorithm based on the time it takes to compose business process designs. Six participants rated the algorithm as 'very efficient/fast' and the remaining two rated the algorithm as 'satisfactory/average' (this might be due to fact that there was some waiting time and the results are not generated immediately). The second question was related with the variety of results. All the workshop participants agreed that the framework is capable of producing abundant business process designs, the majority of which would possibly be overlooked by a human designer. Finally, the third question of this section was seeking to identify the advantages of the proposed composition algorithm (PCA) that the participants have identified. The participants consider the proposed algorithm as a significant shift from the current practice as:

- it automates the process composition activity,
- it is time efficient compared to current practice,
- it is capable of generating a range of alternative designs, and,
- it composes end-to-end business process designs.

There were also some comments and remarks for improvement. These are discussed in sub-section 9.7.4.

9.7.3 Business process optimisation

The second section of the workshop was related with business process optimisation. Chapter 6 presented in detail the proposed optimisation framework for business process designs employing a series of EMOAs. Unlike business process composition, there is no established practice for business process optimisation or improvement as chapter 2 has discussed. The proposed framework composes feasible business process designs and then evaluates them based on the process attributes. In the real-life scenarios demonstrated in the previous sections, the process attributes are SDP and SFT. This section of the workshop aimed to demonstrate that following the current manual practice of business process composition, any attempt to manually compose an optimal design would be time consuming and not robust.

The relevant questionnaire section had one exercise (exercise 3) that was an extension of exercise 1. The exercise asked from the participants to compose a design with 5 tasks but with *minimum* SDP value (single objective optimisation). For this exercise, each participant was requested to mark the duration of the exercise (in minutes) and to rate the difficulty of the exercise on a scale 1 (easy) to 5 (very difficult). Also the participants were asked to rate the difficult of the exercise if it also involved SFT maximisation (bi-objective optimisation).

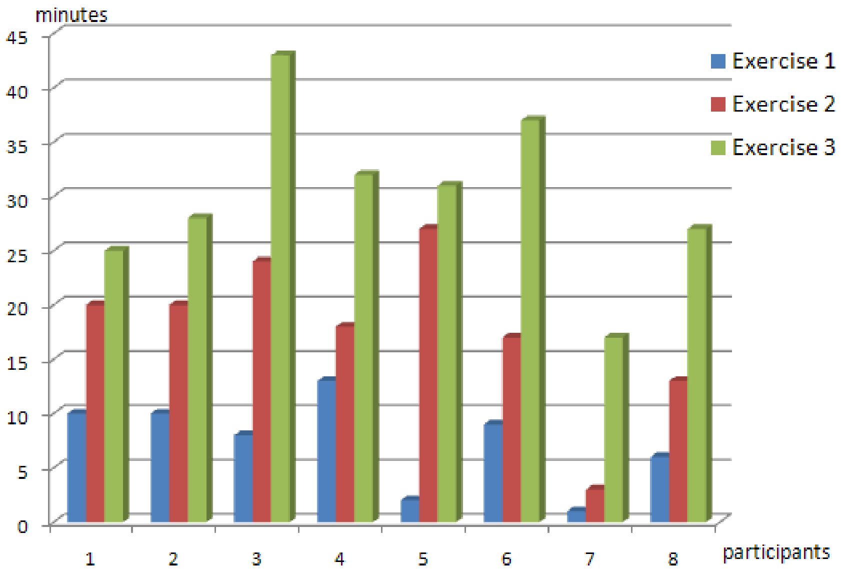


Figure 9.13. Time to complete the business process composition/optimisation exercises

Figure 9.13 shows the duration of the optimisation exercise (exercise 3) per participant in comparison to the first two composition exercises. The average duration for completion of the optimisation exercise was 30 minutes. In addition, only five of the participants composed the design with the actual minimum SDP value, while the remaining three were not able to discover the combination of web services that delivers the minimum SPD but composed a design they perceived as optimal. On average, the participants rated the optimisation exercise as 'difficult' (3.7) while in the case that the exercise involved both objectives it was rated as 'very difficult' (4.5). After the completion of the optimisation exercise, the proposed optimisation framework was demonstrated to the participants. In the question of assessing the capability of the framework to produce alternative optimised results, six participants answered 'abundant number of optimised designs'. The remaining two rated the number of generated optimised designs as 'satisfactory'. One respondent added that a business analyst would actually require *less* optimised alternatives and suggested adding high-level preference criteria would reduce the optimised designs between 5 and 10.

The next question was focused on comparing the proposed optimisation approach with a manual practice. All the participants agreed that only an algorithmic optimisation approach –as the one proposed by this research– can ensure the generation of optimal business process designs in a timely fashion compared to manual approach, especially for large library sizes. The final question asked whether the proposed research can have any benefit for formal business process improvement initiatives. The vast majority of respondents (7) noted that an algorithmic approach towards business process optimisation is a much sought after capability. One respondent stated that based on the proposed research, its benefits towards business process optimisation are largely based on the nature and the frequency that the process occurs (and thus the need to be optimised). As a concluding remark, all the respondents agreed that it is almost impossible for a human designer to deal with large number of task alternatives, multiple objectives and any number of tasks in the final design, thus the proposed optimisation framework for business process designs constitutes a significant shift in business process improvement initiatives. The next section discussed any further remarks and comments that were made during the workshop.

9.7.4 General questions and remarks

The final section of the workshop included two questions and requested from the participants to provide any final remarks for discussion. The first question asked whether web services are current practice within the respondents' organisations. One respondent answered positively to this question, one responded said that he/she has not come across the concept before and the remaining six said that they are familiar with the concept although it is not current practice. Further to that question, the second question asked whether the context of the proposed approach (business process composition and optimisation employing web services as tasks) is considered practical. All the participants agreed that web services are the future for designing business processes and one of the participants revealed that his/her organisation is moving to a new business process environment exclusively based on web services. However, the design and employment of web services-based business processes will be entirely manual; thus the proposed research constitutes a potential shift for automating the design, composition and optimisation of business processes.

The participants also raised a number of issues regarding the proposed framework and the particular context for validation. The following points summarise the main areas of concern of the participants and the responses provided by the researcher:

- *How do you ensure the relevance of web services in the task library? How did you specify your main resources (web service libraries)?*
  Each web service comes with a brief description of its main operation (see web services in appendix E). The web services for each scenario were selected primarily based on the provided description. The resources for web services were also selected based on their relevance to the particular scenarios and on the number of web services on offer.

- *The workshop assumes no previous knowledge for the business process while it is not usually the case. The business process analyst has some knowledge and/or experience about the particular domain that enhances the manual activity of (re)designing a business process.*
  This is a valid point. In manual business process composition, the business process analyst is equipped with knowledge and experience about the business process he/she is working with. In an attempt to balance this factor, prior to the exercises there was a brief presentation that described the scenario. Also the workshop exercises facilitated appropriate colouring in order to reduce the complexity of the composition activity.

- *In real life 50 optimised alternatives of a business process design are considered unnecessary and impractical. Does the proposed framework employ a method of reducing the alternatives applying high-level preference criteria or a rating between the objectives?*
  The aim of the proposed optimisation framework for business processes is to generate a set of optimised designs. A possible extension of the framework could have a filter that reduces the generated results based on high level preference criteria or a weight attached to each of the objectives. However, what needs to be pointed out is that the proposed approach is an aid for the process analyst and not a replacement. The proposed approach will help the analyst in the exploration of new and optimal business process designs, which currently is time-consuming.

The participants were also asked about any future directions that this research could be oriented to. Their remarks are summarised in the following two points:

- The proposed approach could be more interactive in the future. The business process analyst could have the capability of altering the scenario while it is composed based on his/her priorities. Making the framework more interactive could benefit the proposed approach by combining the analyst's experience and knowledge about the particular domain with the automated composition and optimisation of designs.

- The framework could have an add-in that compares the different generated designs. The results of this comparison would demonstrate the different flow, drivers and characteristics between the optimised designs and thus enhance the selection process.

The overall feeling from the experts was that the proposed framework is a significant shift from the current manual practice and provides a basis for automated business process composition and optimisation. The participants also stated that the proposed approach is particularly beneficial in the case of expensive processes.

9.8 Main remarks

This section discusses the main observations from the application of the proposed framework to the three real-life scenarios and the main outcomes of the expert validation through the workshops. Each real-life scenario was selected based on the classification of business processes into: automated, semi-automated and manual automation levels. The aim was to demonstrate the versatility of the proposed framework and to also highlight any strengths or weaknesses when applied to business process examples that encompass real-life elements. All the three scenarios are about business processes that are reported as current practice. In all the scenarios, the library of tasks consisted of publicly available web services. And in all scenarios, optimal designs were generated as a result of the framework application and were demonstrated as an example of the automation and optimisation capabilities of the proposed research.

All the scenarios were provided with an initial design that sketched the main steps. Although this is not a prerequisite for the framework, it is usually the case in real world where the business analyst usually has a clear idea of how the process design will flow. In automated processes such as scenario A, the process flow is much more rigid in terms of the steps that need to be performed. The framework demonstrated its ability to generate different designs using different sets of alternative services for the main process steps that perform the same operation with optimal attribute values. In semi-automated processes, such as scenario B, the process flow is largely defined but there is still room for automation, design variations and optimisation of the operation. In this case the framework generated not only optimised designs, but also designs that 'eliminated (based on the process requirements) one of the initial process steps. For example, the framework produced designs that did not plot the results but compiled the results report only with the time-series forecast results. The framework considered this option as an equivalent process alternative with lower costs since in the process outputs there was no explicit reference to a graph or a chart but to a report. Therefore, while attempting to optimise the process, the framework 'decided' to omit a step since there was no impact on the process (output) requirements.

The framework 'initiatives' are more apparent in scenario C, the manual process. Scenario C came with a very generic design as it is a manual process largely dependent on the person who enacts it. The flexibility of the process design and the generic nature of the process requirements allowed the compilation of a library of diverse web services that were still performing a similar operation. Unlike the previous scenarios, there was no prerequisite to find clearly defined web services such as 'generate a chart'. This flexibility is also apparent in the generation of alternative designs. The framework was allowed to add as many tasks in order to make the investigation process more rigorous. Business process designs with significant difference in the number of tasks demonstrated better the trade-offs between the two objectives. A process with fewer tasks can be employed in a low-cost low-risk investigation whereas a process with larger number of task can be used in a high-risk high-profile case. In all these cases, the framework demonstrated its capability to offer automated and optimal alternatives of the process. In summary, the proposed approach towards automated composition and optimisation of business process designs offers:

- The capability to employ real web services and compile them appropriately to formulate a business process design.
- The *automated composition* of alternative business process designs based on the process design input and output requirements.
- The proposed framework's non-reliance on any existing or initial business process design that gives to the framework the flexibility to 'decide' whether to omit or include any steps in the generated business process designs.
- The capability of identifying designs that satisfy the process requirements and have *optimal* attribute values for all the available process sizes.
- The benefit of utilising a range of state-of-the-art evolutionary algorithms to deliver optimal results.

In all three scenarios, all the EMOAs delivered the same Pareto-front. This shows that unlike the experimental scenarios that were artificially complex, in scenarios with real-life elements (e.g. small libraries of alternatives) these state-of-the-art algorithms can be used with confidence and they provide optimal results. Finally, the workshop exposed the proposed framework to a number of business process experts to get their feedback on the issues of business process composition and optimisation compared to the current manual practice. The focus of the real-life scenarios was on web services as the business process as is the context that offers most of the elements that the framework presumes for business process optimisation. The main outcome of the workshops was that the proposed research constitutes a shift in the area of automated business process improvement. However, there is plenty of room for further research and additional functionalities regarding optimisation and selection of optimal business process designs.

9.9 Summary

This chapter presented three real-life scenarios of business process designs. As a context for scenario selection two perspectives were adopted in relation to business processes: (i) business processes as a Service system, and (ii) business process automation. Both these perspectives point to the composition of business processes using web services. The three real-life scenarios demonstrated that the framework can automate the process composition and also identify business process designs with optimised attribute values. Scenarios that involved already automated processes stressed on the optimisation capability, whereas the manual process scenario demonstrated the framework's ability to enhance the process operation itself. The framework's contribution was also validated by business process experts in a series of workshops. The next chapter provides an overview of the research and a critical discussion on the contribution, the limitations and the potential for further research.

CHAPTER 10

Discussion & Conclusions

This chapter concludes the thesis with a discussion on the findings of this research. The discussion involves the key observations throughout the course of this research and identifies the main contributions. This chapter also underlines the limitations of the proposed approach towards business process optimisation and highlights the corresponding future research activities that can push forward the research in the area.

10.1 Key observations

This section summarises the key observations of this research as made through each of its main stages.

10.1.1 Literature survey

The literature survey in this research focused on business process *definition*, *modelling*, *analysis* and *optimisation*. The review of modelling, analysis and optimisation approaches was based on a proposed classification based on the types of business process models (diagrammatic, mathematical and business process languages). The proposed classification resulted in visually highlighting a number of interesting observations and especially the lack of optimisation approaches for business processes.

Business process optimisation has not received as much attention as compared to business process modelling and analysis techniques. Business process modelling has attracted the attention of researches from a variety of fields which resulted in a variety of modelling approaches that are used for business processes. Each of these modelling approaches has distinctive advantages but still what is missing is a holistic approach that involves both visual and mathematical constructs in capturing a business process design. There is a need for defining operational and reusable business process models within different types of enterprises, in different contexts and at the required level of detail. Therefore, there is an increasing need for formal methods and techniques to support both the modelling and the analysis of business processes.

For most of the business process models there is no structured and repeatable improvement technique reported. However, despite the existence of many formal process modelling notations, the majority of the business process community still uses simple diagrammatic modelling techniques that have no potential for performance analysis and optimisation.

Performance analysis can be directly used for decision-support and further improvement of the process. Performance indicators are critical for the control and monitoring of a business process. However, there are very few attempts reported in literature to combine performance evaluation and process optimisation. Regarding the latter, there are some successful attempts reported, but they are highly complicated and yet address only simple sequential business processes. The lack of optimisation approaches can be attributed to the static and complex models and to the unwillingness of business analysts towards 'black box' process improvement.

10.1.2 Service industry survey

In order to ground the research within the service industry context, a survey with 25 participants was carried out as part of this research. The service industry survey enabled the identification of the current practice related to business processes. The main observations revealed the *industrial context* of the research.

Both the academic researchers and service industry practitioners feel more confident in dealing with structured and defined business processes. This is justified by the fact that a structured process with expected (or predefined) inputs and outputs can be subject to quantification and measurable evaluation. Investigation of business process modelling both in literature and in service industry proved that simple diagrammatic techniques such as flowcharts still dominate the area. This reflects the need for a simple, communicative and effective illustration of business processes. However, in literature there are also plenty of advanced modelling techniques and methods for business processes. These advanced – and perhaps more complex– modelling methods do not guarantee a more formal and structured approach towards business processes; they might even discourage the industry practitioners.

Process analysis is still largely perceived as the manual inspection of diagrams. Due to the qualitative nature of business process modelling, quantitative analysis and process evaluation are hard to apply. Manual or qualitative analysis approaches, such as diagram inspection, are used for performance analysis to aid process improvement initiatives. Quantitative process evaluation currently takes place only in a small number of the participating organisations. The survey participants are not using a structured methodology for improving their business processes; thus there is a large gap and a potential for a methodology for automated improvement (process optimisation) based on a standard process model.

10.1.3 Business process representation

As a first step towards a comprehensive optimisation framework for business processes, this research proposed a representation for *capturing*, *visualising* and *quantifying* business process designs. The representation covers a range of specified business process elements such as tasks, resources, patterns and attributes. The aim of the representation is to make a business process design amenable to a variety of EMOAs, according to the optimisation focus of this research. The different elements of the representation target to capture and express different aspects of the business process for the different requirements posed by the EMOAs.

The first objective of the representation is to provide visual means of communicating the business process design. This objective is covered by the visual perspective of the representation that provides a diagrammatic depiction of the business process. A simple flowchart is chosen as it is straight-forward in communicating the elements of a business process (tasks, resources, attributes, patterns) that are identified as essential in the context of this research. The second objective is to provide the capability of mathematically capturing and expressing the same elements of the design that are captured by the visual perspective. The quantitative perspective expresses all of the business process elements using mathematical constructs and introduces matrices for capturing (TRM) and evaluating (TAM) the process design.

TAM stores the task attributes for the tasks that participate in the business process designs thus assisting in calculating the process attributes. Having calculated the process attributes for different designs, they can be evaluated and compared. TAM is a construct that satisfies the third objective of the representation. The proposed representation also presented an algorithm (PCA) as a proposed solution to the last (fourth) representation objective about generating and evaluating alternative designs. Using this algorithm, a business process design in the form of quantitative perspective of the proposed representation can be: (i) transformed to the equivalent visual perspective and (ii) assessed based on whether it corresponds to a *feasible* design based on its degree of feasibility (DoI).

10.1.4 Business process optimisation framework - bpo^F

The central part of this research is bpo^F – the proposed evolutionary multi-objective framework for business process optimisation. The framework incorporates two main components: (i) the proposed business process representation and (ii) existing state-of-the-art EMOAs as the multi-objective optimisation technique. The aim of the framework is to address a series of optimisation challenges regarding the business processes representation and the optimisation algorithms. The challenges related to the representation were concerned mostly with *adjusting* and *tuning* the representation into the optimisation framework. The challenges related to the optimisation algorithms are concerned with the *performance* of the framework in terms of employing the EMOAs to produce optimised business process designs.

The framework utilises the proposed representation in order to address the optimisation challenges. In order to work with process designs of varying sizes, the framework keeps a fixed size set for the solution representation and replaces the redundant tasks with the element '-1'. It also fully utilises the PCA to produce the framework outcomes. It executes the PCA as part of the infeasibility constraint (DoI = 0) and not as part of the objective functions –as it would be normally expected. Another novelty of the framework is that during one generation, the solution is updated and transformed in order to address the representation challenges of the different optimisation stages. The solution transformation and update occur in a consistent way ensuring the correctness of the solution. Also, the optimisation process works with the participating tasks and not with how they connect to each other. This provides the flexibility to PCA to discover and compose novel process designs during the optimisation process.

The EMOAs employed by the framework also have a series of challenges. The discrete nature of the problem in conjunction with its multi-objective formulation can make the process of discovering feasible solutions challenging for the algorithms. Each algorithm has to unfold its own strategy along with its strengths and weaknesses in order to generate the Pareto optimal front of optimised solutions. In order to acquire a clear picture about the performance of the optimisation algorithms, a thorough experimental strategy of the proposed framework is required.

10.1.5 Generation of experimental business process scenarios

In order to assess the performance of the EMOAs it is necessary to test the framework for different business process scenarios. In order to test the framework systematically, a strategy of generating business process scenarios of varying complexity is devised. The strategy largely encompasses the creation of business process scenarios in order to experiment with specific aspects of the framework.

The business process optimisation problem has three main features: (i) the number of feasible solutions, (ii) the available process sizes and (iii) the ranges of task attribute values. For these three features, the main problem parameters are identified and a corresponding set of control parameters is introduced. The proposed strategy for generating and testing experimental business process scenarios is largely dependent on the classification of the control parameters. The proposed strategy consists of two phases: phase I which involves the formulation of the experimental scenario and phase II which involves the testing the optimisation framework with the particular scenario and drawing appropriate remarks.

10.1.6 Performance evaluation of bpo^F

Utilising the strategy for generating experimental scenarios, the next step was to evaluate the proposed optimisation framework. The focus of the framework performance evaluation was two-fold: (i) to assess the capability of the EMOAs in optimising *effectively* business process designs and (ii) to investigate the potential of the framework in generating *feasible* business processes designs with *optimal* values using the proposed business process representation and problem formulation.

Employing EMOAs as the optimisation technique in the proposed framework proved an effective choice as it is demonstrated by the results of the experiments. In the majority of cases the algorithms identified a significant number of non-dominated solutions across various process sizes despite the highly constrained problem formulation. The variety of results that these algorithms offered under challenging parameters makes them an *attractive optimisation technique* for business processes. In terms of the framework's business process optimisation capability, the design of the experimental scenarios helped test the framework and provide performance boundaries around the basic parameters of the proposed business process representation and problem formulation such as library size, design size and different process sizes.

10.1.7 Validation using real-life business process scenarios

The proposed framework was tested with three real-life scenarios. The three real-life scenarios were respectively: *automated*, *semi-automated* and *manual* in nature. The aim was to demonstrate the versatility of the proposed framework and to also highlight any strengths or weaknesses when applied to business process examples that encompass real-life elements. All the three scenarios are about business processes that are reported as current practice. In all the scenarios, the library of tasks consisted of real and publicly available web services. And in all scenarios, optimal designs were generated as a result of the framework application and were demonstrated as examples of the automation and optimisation capabilities of the proposed research.

The performance of the EMOAs employed within the framework is considered as very good. In all three scenarios, all the algorithms delivered the same Pareto-front. This shows that unlike the experimental scenarios that were artificially complex, in scenarios with real-life elements (e.g. small process designs) these state-of-the-art algorithms can be used with confidence and they provide optimal results.

Finally, the workshop exposed the proposed framework to a number of business process experts in order to get their feedback on the issues of business process composition and optimisation compared to the current (manual) practice. The main outcome of the workshops was that the proposed research constitutes a shift in the area of automated business process improvement. However, there is plenty of room for further research and additional functionalities regarding business process optimisation and selection of a small range of optimal designs.

10.2 Main research contributions

The overall contribution of this research is a framework that generates optimised business process designs. This section presents in detail the main contributions of this research. The aim of this research was to develop and propose a new framework for business process optimisation capable of: (i) representing business processes in a quantitative way, (ii) algorithmically composing business process designs based on specific requirements and (iii) identifying the optimal processes utilising the state-of-the art evolutionary multi-objective optimisation algorithms.

The research has provided an understanding about the current state of business processes within literature and service industry and highlighted the lack of a comprehensive optimisation approach. The representation for business processes that was put forward was capable of capturing both the visual and the quantitative elements of a business process design. PCA was the composition algorithm that can generate alternative business process designs based on specific process requirements. This research proposed an evolutionary multi-objective optimisation framework (bpo^F) that encompassed state-of-the-art optimisation algorithms.

In addition, the research devised a strategy for generating experimental problems (business process scenarios) and thus assessing the performance of the proposed optimisation framework. A series of real-life scenarios were also tested within the proposed optimisation framework to demonstrate its capability in optimising business process designs with real-life elements. The following remarks describe in detail how the research aim was achieved as well as the contributions to knowledge that emerged from this work:

- *Current state of business process optimisation in literature and service industry*

  The literature survey carried out as part of this research in conjunction with the industry survey in the service sector defined the current state of business processes in the areas of definition, modelling, analysis and optimisation. Based on the comparison and analysis between the literature and service sector surveys, this research highlighted the similarities and gaps along with the lack of a comprehensive optimisation approach towards business processes.

- *Business process modelling for optimisation*

  This research introduced a representation technique that models the visual and quantitative elements of a business process design. Part of the proposed representation is a clear and accurate definition of business processes along with the identification of the main elements that make a business process. The proposed representation satisfied the identified needs for a diagrammatic depiction of the business process designs and a formal (mathematical) background that will facilitate the design optimisation.

- *Process composition algorithm*

  As part of the proposed representation and as stated in the research aim, a composition algorithm was put forward. The Process Composition Algorithm (PCA) is a central part of the representation which elaborates (constructs) a business process design and also checks whether it is feasible based on the design requirements of the business process. As a result, PCA can generate alternative designs based on the same process requirements.

⊖ *Multi-objective business process optimisation approach*

The central contribution of this research is the proposed optimisation framework. The framework operates for multiple objectives offering the advantage of multi-objective optimisation. It employs state-of-the-art evolutionary algorithms which are known to operate efficiently in multi-criteria optimisation problems. Also, the proposed framework utilised the representation technique in order to compose, assess and optimise business process designs. The performance evaluation of the framework revealed that the framework is capable of optimising a series of challenging problems and producing satisfactory results in terms of alternative optimised business process designs.

⊖ *New application domain for Evolutionary Multi-objective Optimisation Algorithms (EMOAs)*

The results of the performance evaluation also showed that the employment of EMOAs as a business process optimisation technique is effective. These optimisation algorithms are capable of tackling the discrete and highly fragmented search space of the problem and discover a series of diverse optimal solutions. Therefore, an outcome of this research is a new application domain for EMOAs –that of business processes.

⊖ *Development of a strategy for generating experimental business process scenarios*

The performance evaluation of the framework occurred through the introduction of a strategy for generating experimental problems (business process scenarios). As it is difficult to find real-life business processes that extensively test all the aspects of the framework, the proposed strategy was able to generate problems with specific orientation towards one or more problem features. In this way, controlled testing of the framework was made possible across all the problem features and their corresponding control parameters.

⊖ *Applicability to real-life business processes*

This research demonstrated the capability of the proposed framework to capture, represent and optimise business process designs with real-life elements. Adopting the classification of business process automation (automated, semi-automated and manual) the research utilised the concept of web services as process tasks and was able to create a library with alternatives. The outcome of the framework demonstrated a series of alternative optimised business process designs composed of web services. This contribution not only shows the validity of the optimisation framework but also is in accordance with the latest trend of implementing business processes with web services.

The results of the performance analysis and the real-life scenarios validate the hypothesis in chapter 3 that: *Business process optimisation using an evolutionary multi-objective framework can produce a number of alternative optimised business processes for a range of experimental and real-life business process scenarios.*

10.3 Business impact analysis

The outcome of this research can be potentially used by the service industry with significant benefits in terms of business process re-design through optimisation. That is one of the reasons why the service industry survey and the workshop with the business process experts were key parts of this research.

The service industry survey helped in understanding the current practice in issues related to business process and optimisation. The industrial context of this research (as discussed in chapter 4) summarised the main observations and shaped the orientation of this research in response to the needs of the service industry. Based on these, this research proposed the following:

1. A business process representation with a visual perspective that communicates the business process design as a diagram.
2. An optimisation framework that can generate alternative designs with optimised attribute values based on specific and measurable objectives.
3. A business process automation approach in which business processes are composed of web services and implemented over a network.

The representation of the business process designs communicates with a process diagram the participating tasks, the flow of the process (patterns) and the resources that are involved. This answers the need posed by the service industry practitioners for simple and communicative technique for visualising the key elements of a business process. The optimisation approach provides a tool that can compose alternative designs with optimal attribute values based on specific requirements. The framework is inherently multi-objective and thus closer to real-life business process problems that usually have multiple objectives. It is also able to generate equivalent designs of different sizes. The objectives are measurable which facilitates the comparison and evaluation of generated designs. The proposed optimisation framework is a complete approach that can result in improvement and/or re-design of a business process using evolutionary multi-objective optimisation techniques.

The optimisation capabilities of the framework were demonstrated by generating optimised designs that are composed of web services. The service industry sees in web services the future of business processes. This research showed how a library of web services can be incorporated in the framework, leading to a series of alternative designs. However, there can be issues related to scaling the framework to a real business process scenario. As the experimental results showed, the library size has a lower limit in relation to the tasks in the design. Therefore, for a real scenario an abundance of web services is necessary. Also, the number of tasks in the design cannot be increased without taking into account the library size. Finally, it is expected that the framework will generate satisfactory results for more than two objectives although further testing is required.

The workshop exposed the outcomes of this research to business process experts and verified that the proposed research constitutes a shift in the area of automated business process improvement compared to the current practice in the service industry. However, for the previously stated benefits of the proposed research to be fully realised by the service industry, the framework and the prototype tool need to evolve to a fully functional software tool. Also, an organisation that would like to adopt the proposed approach would need to follow these steps:

- Create a pool of web services and/or a library of tasks that can potentially participate in a business process design.
- Identify the input and output resources of each task/web service and of the process. The process input and output resources will be used as the requirements for the generation of new and optimised designs.
- Develop a technique for incorporating high level preferences. This technique (e.g. weighted sum) will facilitate the selection of a single optimised business process design using as input the optimised population of bpo^F.

10.4 Limitations of research

An attempt has been made in this research to keep it as general as possible regarding the business processes and the optimisation algorithms. However, as with any other research, this work has also some limitations. The following sections group and identify some of the limitations of this research.

10.4.1 Limitations of the Research Methodology

The following are the main limitations of the methodology used in this research:

- The literature and industry surveys identified a number of issues around business process definition, modelling, analysis and optimisation. However, this research is primarily focused on business process optimisation and based on this orientation it suggests a specification and modelling approach. Therefore, the research methodology focuses on developing an optimisation framework for business processes.
- In this research, the service sector survey was facilitated with the use of semi-structured and structured questionnaires. Although the use of questionnaires is useful, it also has its limitations as it captures mainly the *perception* of the interviewee and not the actual situation (which can be captured more accurately through observation for example). The organisations did not cover a full spectrum of the Service sector but there was more emphasis on IT, Telecommunications, Banking, Finance and Consultancy
- The multi-objective optimisation approach utilised within the proposed framework is based on Evolutionary Multi-objective Optimisation Algorithms (EMOAs) and did not explore or consider any other optimisation techniques. This is justified by the multi-objective formulation of the business process optimisation problem and the necessity to 'evolve' during the optimisation process infeasible business process designs towards feasible ones.
- Although the strategy for experimental business process scenarios suggested a number of control parameters to test the proposed optimisation framework (chapter 7), the performance analysis focused only on discovering the flexibility of the framework towards its basic parameters (chapter 8) and did not explore more complex combinations. This helped in evaluating the optimisation algorithms and assessing its optimisation capabilities. It also allows for future testing on more advance problems and real-life cases.

- The real-life case studies that are used in this research are borrowed from literature. This has provided limited insight on how to model an actual process from the service industry and incorporate it with the proposed optimisation framework.

10.4.2 Limitations of the proposed representation technique

The following are the main limitations of the proposed representation technique:

- The elements that were selected in the proposed business process specification (chapter 5) were a subset of the proposed business process schema (chapter 2). The actors were not involved in the proposed representation as this research is oriented towards design optimisation and not business process enactment.
- The patterns that were selected were a subset of the proposed patterns for business processes. Exclusive choice (XOR) and discriminator are among the main patterns that were not taken into consideration.
- The proposed representation assumes the concept of task library, a repository of tasks that can be potentially used to create a business process design. Although this element is in accordance with the evolutionary optimisation orientation of the research, it may not be readily available for real-life business processes.
- The values of the task / process attributes are static (constant values) as opposed to dynamic (based on a distribution of values). Although this facilitates the calculation of objectives and the evaluation of process designs, the real-life process/task attribute values might change dynamically during execution.

10.4.3 Limitations of the proposed optimisation framework

The following are the main limitations of the proposed optimisation framework:

- The proposed framework is closely bound with the representation. This means that only a business process expressed in the proposed representation can be subjected to evolutionary multi-objective optimisation as suggested by this research.
- The proposed optimisation framework is capable of multi-objective optimisation of business process designs. However, in all the experiments presented in this thesis the framework is tested for two objectives only. This is common practice in assessing multi-objective optimisation problems and algorithms. All the EMOAs used by the framework have been tested with standard multi-objective optimisation problems (e.g. ZDT, DTLZ and WFG) and are known to perform well for 4-5 objectives. We can

assume that the framework would also perform in a satisfactory way although thorough testing of the framework for more objectives is left as future research.

- The optimisation approach employed by the proposed framework is based exclusively on evolutionary algorithms and thus the framework structure is influenced by this orientation.

10.5 Future research

Despite the potential of a multi-objective optimisation approach for business processes and the recognition of its benefits by the service industry, it was observed from the service industry survey that none of the participating organisations uses such an approach. In order to address this situation, additional research activities are required that push the proposed research by addressing its limitations.

The proposed research suggests a generic framework for multi-objective optimisation of business process designs using evolutionary algorithms. As part of the framework, this research proposed a specification and representation of business process designs. There is a need to extend this representation and to enhance the framework in order to fully address the needs of real-life business process optimisation.

The research activities for future elaboration of the proposed business process *specification* can be summarised as follows:

- To include more elements in the specification and more specifically:
    - To involve the *actors* responsible for the enactment of the task/process. This would help in task/process ownership and reliable execution of the business process within the organisation.
    - To acknowledge *sub-processes* as separate entities in a business process design. This frequently occurs in high-level abstract business processes (e.g. at the strategic level).
- To include more business patterns such as exclusive choice (XOR), multi-choice, discriminator, cancel task, kill process, etc.

The research activities for future elaboration of the proposed business process *representation* can be summarised as follows:

- The task attributes stored in TAM can be *dynamic* (distribution-based) instead of static (constant). They can also be *actual* values (from past executions) instead of estimated.

- *Qualitative* task/process attributes can also be introduced for a more accurate depiction of the process characteristics.
- The process attribute functions can be more complex taking into account the process patterns. This is essential for the calculation of attributes such as the process duration.

The research activities for future elaboration of the proposed *business process optimisation framework* can be summarised as follows:

- To test the framework for more that two objectives.
- The influence of the various process patterns can be measured in the calculation of the attributes and taken into account in design optimisation.
- The framework could incorporate process execution feedback (historical data) in order to refine/optimise a business process design.
- The research in business process optimisation can also move towards the direction of *execution flow optimisation* and *automatic process modification.* Execution flow optimisation is the notion of deciding the optimum path for a business process during its enactment (execution). Automatic process modification is the *real-time* composition of a business process design according to specific needs. The optimum process design is created based on the selection and combination of different alternative web services scattered over a network.

Business process optimisation has a potential growth with direct benefit to the business process community and there are still a number of research activities to be addressed.

10.6 Conclusions

This final section of the thesis compares the achievements of this research with the objectives stated in chapter 3. The following discussion analyses each research objective (in italics) and compares it with what is achieved in this research.

- *Investigate and establish the state-of-the-art regarding business process modelling, analysis and optimisation.*

  This research carried out a literature survey about business processes regarding modelling, analysis and optimisation (chapter 2). To facilitate the survey, the research proposed a classification of existing modelling approaches based on their visual, mathematical and language capabilities. This classification helped in assessing existing business process modelling approaches along with their analysis and optimisation capabilities. The result of this classification was a comprehensive overview of the state of business process modelling, analysis and optimisation in literature. What emerged was the lack of any systematic optimisation approach for business processes.

- *Explore the industrial context of this research through a survey that identifies the main issues regarding business processes in the service industry.*
  This research also carried out an industry survey within the service sector (chapter 4). This survey helped in contrasting and comparing the business process practices in literature with these in real-life. The service industry survey identified that business process experts are more comfortable in dealing with simple diagrams of business processes. As a result, the analysis takes place as a simple inspection of the process diagram with little or no room for quantitative performance analysis. The survey highlighted the scarce improvement initiatives in terms of business processes and the lack of any systematic optimisation approaches. However, the majority of the participants underlined that a business process optimisation framework would have a significant impact and benefit in the organisation.

- *Provide a formal specification and a representation technique for modelling business processes quantitatively so that they can be subjected to evolutionary optimisation techniques.*
  Chapter 5 proposed a specification and a representation technique for business process designs. Both were based on the issues identified from the literature and industry surveys. The specification included a definition for business processes and a selection of the main elements (tasks, resources, attributes, patterns) to be included in the representation technique. The representation technique encompassed two perspectives of modelling a business process design: the *visual* perspective – a diagrammatic representation of the business process and the *quantitative* perspective – a formal representation based on mathematical parameters.

- *Develop an algorithmic technique that composes new business process models based on specific requirements.*
  The Process Composition Algorithm (PCA) is proposed by this research (in chapter 5) as an algorithmic approach that can compose a business process design based on the representation technique and measure its degree of infeasibility (DoI). PCA composes the visual perspective of a design given its quantitative form and thus provides a bridge between the two aspects of the proposed representation. Moreover, during the process composition, the algorithm checks on whether the captured design

corresponds to a feasible one. PCA composes designs based on specific predefined requirements and given the task library it can modify an infeasible design to a feasible.

- *Construct an evolutionary multi-objective optimisation framework for business processes.*
  This research proposed an evolutionary multi-objective optimisation framework for business processes – bpo^F (chapter 6). The framework incorporated the proposed representation technique –and PCA as part of it– along with state-of-the-art Evolutionary Multi-objective Optimisation Algorithms (EMOAs). The proposed optimisation framework is capable of working with a population of designs with the same process requirements and optimising them based on multiple objectives. The outcome of the framework is a series of alternative feasible business process designs with optimal attribute values.

- *Identify the basic features of the problem and suggest a strategy for generating tuneable business process scenarios in order to systematically evaluate the performance of the optimisation framework.*
  Chapter 7 identified the basic features of the problem and suggested a strategy for generating tuneable business process scenarios. The main features that require investigation are the number of feasible solutions in a given problem, the different sizes that an optimised design can have and the ranges of the attribute values. Based on these features the problem parameters were identified and a series of control parameters was introduced. This research proposed a strategy for generating tuneable business process scenarios in order to systematically simulate the features of the business process optimisation problems. Chapter 8 generated three experimental scenarios based on the proposed strategy and examined the flexibility of the framework and the performance of the optimisation algorithms. The framework was able to optimise challenging scenarios generating satisfactory results in terms of optimised business process designs.

- *Validate the business process representation technique, composition algorithm and optimisation framework using a set of business process scenarios with real-life elements.*
  Chapter 9 introduced three real-life scenarios of business process designs and validated the framework with them. The framework was able to generate optimised results and demonstrate the alternative business process designs. Finally, a series of business process experts assessed the framework as a significant contribution towards business process optimisation.

The aim of this research was to develop and propose a new framework for business process optimisation capable of: (i) representing business processes in a quantitative way, (ii) algorithmically composing business process designs based on specific requirements and (iii) identifying the optimal processes utilising the state-of-the art evolutionary multi-objective optimisation algorithms. The following achievements of this research summarise the step towards achieving the research aim:

1. Critical analysis of business process definitions, modelling, analysis and optimisation,
2. Business process specification and representation oriented towards optimisation,
3. Composition algorithm for generation of business process designs,
4. Business process optimisation framework based on evolutionary multi-objective algorithms,
5. Strategy for generation of experimental business process scenarios, and,
6. Steps for transforming and optimising business processes with real-life elements.

In this way, this research has proposed a fully tested and validated framework for representing and optimising business process designs.

References

Abate, A.F., Esposito, A., Grieco, N. and Nota, G. (2002), 'Workflow Performance Evaluation Through WPQL', in *Proceeding of the 14th International Conference on Software Engineering and Knowledge Engineering Vol. 27*, ACM Press, New York, pp. 489-495.

Agerfalk, P., Goldkuhl, G. and Cronholm, S. (1999), 'Information Systems Actability Engineering - Integrating Analysis of Business Process and Usability Requirements', in *Proceedings of the 4th International Workshop on the Language Action Perspective on Communication Modelling* Copenhagen, Denmark, pp. 1245-1252.

Aguilar-Saven, R.S. (2004), 'Business Process Modelling: Review and Framework', *International Journal of Production Economics*, Vol. 90, pp. 129-149.

Aldowaisan, T.A. and Gaafar, L.K. (1999), 'Business Process Reengineering: an Approach for Process Mapping', *Omega, International Journal of Management Science*, Vol. 27, pp. 515-524.

Back, T., Hammel, U. and Schwefel, H.P. (1997), 'Evolutionary Computation: Comments on the History and Current State', *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, pp. 3-17.

Badica, C., Badica, A. and Litoiu, V. (2003a), 'A New Formal IDEF-Based Modelling of Business Processes', in Manolopoulos Y. and Spirakis, P. (eds.), *Proceedings of the First Balkan Conference in Informatics (BCI):* Thessaloniki, Greece, pp. 535-549.

Badica, C., Badica, A. and Litoiu, V. (2003b), 'Role Activity Diagrams As Finite State Processes', in *Second International Symposium on Parallel and Distributed Computing*, London, pp. 15-22.

Bal, J. (1998), 'Process Analysis Tools for Process Improvement', *The TQM Magazine*, Vol. 10, No. 5, pp. 342-354.

Basten, T. and van der Aalst, W.M.P. (1999), 'Inheritance of Behaviour', *Computing Science Report 99/17*, Eindhoven University of Technology, Eindhoven.

Basu, A. and Blanning, R.W. (2000), 'A Formal Approach to Workflow Analysis', *Information Systems Research*, Vol. 11, No. 1, pp. 17-36.

Bellabdaoui A. and Teghem J. (2006), 'A Mixed-Integer Linear Programming Model for the Continuous Casting Planning', *International Journal of Production Economics*, Vol. 104, No. 2, pp. 260-270.

Biazzo, S. (2000), 'Approaches to Business Process Analysis: a Review', *Business Process Management Journal*, Vol. 6, No. 2, pp. 99-112.

Biazzo, S. (2002), 'Process Mapping Techniques and Organisational Analysis', *Business Process*

Management Journal, Vol. 8, No. 1, pp. 42-52.

Boekhoudt, P., Jonkers, H. and Rougoor, M. (2000), 'Graph-Based Analysis of Business Process Models', in Mastorikis, N.ed. *Proceedings of the WSES/MIUE/HNA International Conference, Montego Bay, Jamaica, December 2000:* pp. 227-235.

Bosilj-Vuksic, V., Giaglis, G.M. and Hlupic, V. (2000), 'IDEF Diagrams and Petri Nets for Business Process Modeling: Suitability, Efficacy and Complementary Use', in *Proceedings of International Conference on Enterprise Information Systems* Stafford, UK, pp. 242-247.

Burns, R. (2000), *Introduction to Research Methods* (International edition), Sage Publications Limited, London.

Castellanos, M., Casati, F., Umeshwar, D. and Ming-Chien, S. (2004), 'A Comprehensive and Automated Approach to Intelligent Business Processes Execution Analysis', *Distributed and Parallel Databases*, Vol. 16, pp. 1-35.

Chapin, N. (1971), *Flowcharts*, Auerbach Publishers, Princeton.

Chapin, N. (1974), 'New Format for Flowcharts', *Software - Practice and Experience*, Vol. 4, pp. 341-357.

Chen, Y., Li, Q. and Zhang, F. (2001), *Business Process Re-Engineering and Systems Integration*, Tsinghua University Press, Beijing.

Coello Coello, C.A. (2000), 'Handling Preferences in Evolutionary Multi-objective Optimization: A Survey', in *Proceedings of the Congress on Evolutionary Computation (CEC'2000)* San Diego, USA, pp. 30-37.

Coello Coello, C.A. (2005), 'Recent Trends in Evolutionary Multi-objective Optimization.', in Abraham, A., Jain, L. and Goldberg, R. (editors), *Evolutionary Multi-objective Optimization: Theoretical Advances and Applications*, Springer-Verlag, London, pp. 7-53.

Corne, D.W., Jerram, N.R., Knowles, J.D. and Oates, M.J. (2001), 'PESA-II: Region-Based Selection in Evolutionary Multi-objective Optimization', in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001)* San Francisco, California, pp. 283-290.

Davenport, T.H. (1993), *Process Innovation: Reengineering Work Through Information Technology*, Harvard Business School Press , Boston.

Davenport, T.H. and Short, J.E. (1990), 'The New Industrial Engineering: Information Technology and Business Process Redesign', *Sloan Management Review*, pp. 11-27.

Deb, K. (2001), *Multi-objective Optimisation Using Evolutionary Algorithms*, John Wiley & Sons, New York.

Dewan, R., Seidmann, A. and Walter, Z. (1998), 'Workflow Optimisation Through Task Redesign in Business Information Processes', in *Proceedings of the Hawaii International Conference on*

System Sciences Vol. 1, pp. 240-252.

Donatelli, S., Ribaudo, M. and Hillston, J. (1995), 'A Comparison of Performance Evaluation Process Algebra and Generalised Stochastic Petri Nets', *Sixth International Workshop on Petri Nets and Performance Models - PNPM '95*, pp. 158.

Durillo, J.J., Nebro, A.J., Luna, F., Dorronsoro B. and Alba E. (2006), 'JMetal: a Java Framework for Developing Multi-Objective Optimisation Metaheuristics', *ETS Ingenieria Informatica TECH-REPORT: ITI-2006-10*, pp. available at: http://mallba10.lcc.uma.es/wiki/pdfs/TechRep.ITI-2006.10.pdf (accessed: December 2007).

Ernst, A.T., Jiang, H., Krishnamoorthy, M. and Sier, D. (2004), 'Staff Scheduling and Rostering: A Review of Applications, Methods and Models', *European Journal of Operational Research*, Vol. 153, pp. 3-27.

Fan, Y.S. (2001), *Fundamental of Workflow Management Technology*, Springer-Verlag, New York.

Fathee, M.M., Redd, R., Gorgas, D. and Modarres, B. (1998), 'The Effects of Complexity on Business Process Reengineering: Values and Limitations of Modelling and Simulation Techniques', in Medeiros D.J., Watson E.F., Carson J.S. and Manivannan M.S. (eds.), *Proceedings of the 1998 Winter Simulation Conference:* pp. 1339-1345.

Feldman, C.G. (1998), *The Practical Guide to Business Process Reengineering Using IDEF0*, Dorset House, New York.

Ferscha, A. (1998), 'Optimistic Distributed Execution of Business Process Models', in *Proceedings of the 31st Annual Hawaii International Conference on System Sciences Vol. 7*, Kohala Coast, HI, pp. 723-732.

Floudas, C.A. and Lin, X. (2005), 'Mixed Integer Linear Programming in Process Scheduling: Modelling, Algorithms and Applications', *Annals of Operations Research*, Vol. 139, pp. 131-162.

Gao, M., Zhou, M., Huang, X. and Wu, Z. (2003), 'Fuzzy Reasoning Petri Nets', *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, Vol. 33, No. 3, pp. 314-324.

Greasley, A. (2003), 'Using Business-Process Simulation Within a Business-Process Reengineering Approach', *Business Process Management Journal* , Vol. 9, No. 3, pp. 408-420.

Greenfield, T. (2002), *Research Methods for Postgraduates* (2nd edition), Arnold, London.

Grigori, D., Casati, F., Castellanos, M., Dayal, U., Sayal, M. and Shan, M.-C. (2004), 'Business Process Intelligence', *Computers in Industry*, Vol. 53, pp. 321-343.

Grigori, D., Casati, F., Dayal, U. and Shan, M.-C. (2001), 'Improving Business Process Quality through Exception Understanding, Prediction and Prevention', in *Proceedings of the 27th Very Large Data Base (VLDB) Conference*, Los Angeles, pp. 872-888.

Guha, S., Kettinger, W.J. and Teng, J.T.C. (1993), 'Business Process Reengineering: Building a Comprehensive Methodology', *Information Systems Management*, Vol. 10, pp. 13-22.

Gummesson, E. (1991), *Qualitative Methods in Management Research*, Sage Publications Limited, London.

Gunasekaran, A. and Kobu, B. (2002), 'Modelling and Analysis of Business Process Reengineering', *International Journal of Production Research*, Vol. 40, No. 11, pp. 2521-2546.

Gutjahr, W.J., Strauss, C. and Wagner, E. (2000), 'A Stochastic Branch-and-Bound Approach to Activity Crashing in Project Management', *INFORMS Journal on Computing*, Vol. 12, No. 2, pp. 125-135.

Hammer, M. and Champy, J. (1993), *Reengineering the Corporation: a Manifesto for Business Revolution*, N. Brealey, London.

Han, H.-S. (2003), 'Business Process Change Design From Decision Model Perspective', *Omega, International Journal of Management Science*, Vol. 9, No. 2, pp. 21-45.

Hart, E., Ross, P. and Corne, D. (2005), 'Evolutionary Scheduling: A Review', *Genetic Programming and Evolvable Machines*, Vol. 6, pp. 191-220.

Havey, M. (2005), *Essential Business Process Modelling*, O'Reilly, U.S.A.

Hermanns, H., Herzog, U. and Katoen, J.-P. (2002), 'Process Algebra for Performance Evaluation', *Theoretical Computer Science*, Vol. 274, pp. 43-87.

Hofacker, I. and Vetschera, R. (2001), 'Algorithmical Approaches to Business Process Design', *Computers & Operations Research*, Vol. 28, pp. 1253-1275.

Holt, A.P. (2000), 'Management-Oriented Models of Business Processes', *Lecture Notes in Computer Science*, Vol. 1806, pp. 99-109.

Irani, Z., Hlupic, V. and Giaglis, G.M. (2002), 'Business Process Reengineering: An Analysis Perspective', *International Journal of Flexible Manufacturing Systems*, Vol. 14, pp. 5-10.

Jaeger, T., Prakash, A. and Ishikawa, M. (1995), 'A Framework for Automatic Improvement of Workflows to Meet Performance Goals'*, in Proceedings of the International Conference on Tools with Artificial Intelligence*, pp. 640-646.

Johanson, H.J., McHugh, P., Pendlebury, A.J. and Wheeler, W.A.I. (1993), *Business Process Reengineering - Breakpoint Strategies for Market Dominance*, Wiley, Chichester.

Kallrath J. (2002), 'Planning and Scheduling in the Process Industry', *OR Spectrum*, Vol. 24, pp. 219-250.

Kettinger, W.J., Teng, J.T.C. and Guha, S. (1997), 'Business Process Change: A Study of Methodologies, Techniques and Tools', *MIS Quarterly*, Vol. 21, No. 1, pp. 55-80.

Kiepuszewski, B., ter Hofstede, A.H.M. and van der Aalst, W.M.P. (2003), 'Fundamentals of Control Flow in Workflows', *Acta Informatica*, Vol. 39, No. 3, pp. 143-209.

Kim, C.H., Weston, R.H., Hodgson, A. and Lee, K.H. (2003), 'The Complementary Use of IDEF and UML Modelling Approaches', *Computers in Industry*, Vol. 50, pp. 35-56.

Knowles, J. and Corne, D. (1999), 'The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Multi-objective Optimization', in *Proceedings of the 1999 Congress on Evolutionary Computation (CEC'1999)* IEEE Press, Piscataway, NJ, pp. 98-105.

Knuth, D.E. (1963), 'Computer-Drawn Flowcharts', *ACM Communications*, Vol. 6, No. 9, pp. 555-563.

Koenig, J. (2004), *JBoss JBPM White Paper, Version 2004*, available at: http://www.jboss.com/pdf/jbpm_whitepaper.pdf (accessed 2006).

Koubarakis, M. and Plexousakis, D. (2002), 'A Formal Framework for Business Process Modelling and Design', *Information Systems*, Vol. 27, pp. 299-319.

Kusiak, A. and Zakarian, A. (1996a), 'Reliability Evaluation of Process Models', *IEEE Transactions on Components, Packaging and Manufacturing Technology - Part A*, Vol. 19, No. 2, pp. 268-275.

Kusiak, A. and Zakarian, A. (1996b), 'Risk Assessment of Process Models', *Computers and Industrial Engineering*, Vol. 30, No. 4, pp. 599-610 .

Kusiak, A., Larson, N.T. and Wang, J. (1994), 'Reengineering of Design and Manufacturing Processes', *Computers and Industrial Engineering*, Vol. 26, No. 3, pp. 521-536.

Lakin, R., Capon, N. and Botten, N. (1996), 'BPR Enabling Software for the Financial Services Industry', *Management Services*, Vol. 40, pp. 18-20.

Lee, D.-H., Kiritsis, D. and Xirouchakis, P. (2001), 'Branch and Fathoming Algorithms for Operation Sequencing in Process Planning', *International Journal of Production Research*, Vol. 39, No. 8, pp. 1649-1669.

Lee, Y.H., Min, K.G., Han, C., Chang, K.S. and Choi, T.H. (2004), 'Process Improvement Methodology Based on Multivariate Statistical Analysis Methods', *Control Engineering Practice*, Vol. 12, pp. 945-961.

Levas, A., Boyd, S., Jain, P. and Tulskie, W.A. (1995), 'Panel Discussion on the Role of Modelling and Simulation in Business Process Reengineering', in *Winter Simulation Conference Proceedings* Piscataway, NJ, USA, IEEE, pp. 1341-1346.

Lewis, C. (1993), 'A Source of Competitive Advantage', *Management Accounting (CIMA)*, Vol. 71, No. 1, pp. 44-46.

Li, H., Yang, Y. and Chen, T.Y. (2004a), 'Resource Constraints Analysis of Workflow Specifications', *Systems and Software*, Vol. 73, pp. 271-285.

Li, J., Fan, Y. and Zhou, M. (2004b), 'Performance Modelling and Analysis of Workflow', *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, Vol. 34, No. 2, pp. 229-242.

Li, J., Maguire, B. and Yao, Y. (2003), 'A Business Process Centred Software Analysis Method', *International Journal of Software Engineering* , Vol. 13, No. 2, pp. 153-168.

Lin, H., Zhap, Z., Li, H. and Chen, Z. (2002), 'A Novel Graph Reduction Algorithm to Identify Structural Conflicts', in *Proceedings of the 35th Hawaii International Conference on System Sciences*

Lindsay, A., Downs, D. and Lunn, K. (2003), 'Business Processes - Attempts to Find a Definition', *Information and Software Technology*, Vol. 45, pp. 1015-1019.

Mayer, R.J., Paintec, M.K. and Dewitte, P.S. (1994), 'IDEF Family of Methods for Concurrent Engineering and Business Reengineering Applications', *Technical Report*, Knowledge Based Systems Inc., Texas.

Melao, N. and Pidd, M. (2000), 'A Conceptual Framework for Understanding Business Process Modelling', *Information Systems*, Vol. 10, pp. 105-129.

Mendez, C.A., Cerda, J., Grossmann, I.E., Harjunkoski, I. and Fahl, M. (2006), 'State-of-the-Art Review of Optimisation Methods for Short-Term Scheduling of Batch Processes', *Computers and Chemical Engineering*, Vol. 30, No. 6-7, pp. 913-946.

Menzel, C. and Mayer, R. (1998), 'The IDEF Family of Languages', *in Bernus, P., Mertins K. and Schmidt, G. (eds.), Handbook on Architectures for Information Systems*, Springer-Verlag, New York.

Moon, C. and Seo, Y. (2005), 'Evolutionary Algorithm for Advanced Process Planning and Scheduling in a Multi-Plant', *Computers and Industrial Engineering*, Vol. 48, No. 2, pp. 311-325.

Oppenheim, A.N. (1992), *Questionnaire Design, Interviewing and Attitude Measurement*, Pinter, London.

Orman, L.V. (1995), 'A Model Management Approach to Business Process Reengineering', in *Proceedings of the 1995 American Conference on Information Systems* Pittsburgh, PA, USA, pp. 238-247.

Ould, M.A. (1995), *Business Processes: Modelling and Analysis for Re-Engineering and Improvement*, John Wiley & Sons, Chichester.

Pall, G.A. (1987), *Quality Process Management*, Prentice-Hall, Englewood Cliffs, NJ, USA .

Peters, L. and Peters, J. (1997), 'Using IDEF0 for Dynamic Process Analysis', in *Proceedings of the 1997 IEEE International Conference on Robotics and Automation* Albuquerque, New Mexico, pp. 3203-3208.

Phalp, K. and Shepperd, M. (2000), 'Quantitative Analysis of Static Models of Processes', *Systems and Software*, Vol. 52, pp. 105-112.

Pinto, J.M. and Grossmann, I.E. (1998), 'Assignment and Sequencing Models for the Scheduling of Process Systems', *Annals of Operations Research*, Vol. 81, pp. 433-466.

Powell, S.G., Schwaninger, M. and Trimble, C. (2001), 'Measurement and Control of Business Processes', *System Dynamics Review*, Vol. 17, No. 1, pp. 63-91.

Quatrani, T. (2001), 'Introduction to Unified Modelling Language (UML)', *Rational Developer Network*.

Raposo, A.B., Magalhaes, L.P. and Ricarte, I.L.M. (2000), 'Petri Nets Based Coordination Mechanisms for Multi-Flow Environments', *International Journal of Computer Systems Science and Engineering*, Vol. 15, No. 5, pp. 315-326.

Reijers, H.A. (2002), 'Product-Based Design of Business Processes Applied Within the Financial Services', *Journal of Research and Practice in Information Technology*, Vol. 34, No. 2, pp. 110-122 .

Reijers, H.A. and Liman-Mansar, S. (2005), 'Best Practices in Business Process Redesign: an Overview and Qualitative Evaluation of Successful Redesign Heuristics', *Omega, International Journal of Management Science*, Vol. 33, pp. 283-306.

Reimer, U., Margelisch, A. and Staudt, M. (2000), 'EULE: A Knowledge-Based System to Support Business Processes', *Knowledge-Based Systems*, Vol. 13, pp. 261-269.

Reyneri, C. (1999), 'Operational Building Blocks for Business Process Modelling', *Computers in Industry*, Vol. 40, pp. 115-123.

Riehle, D. and Zullinghoven, H. (1996), 'Understanding and Using Patterns in Software Development', *Theory and Practice of Object Systems*, Vol. 2, No. 1, pp. 3-13.

Robson, C. (2002), *Real World Research*, Blackwell, Oxford.

Rommelfanger, H.J. (2004), 'The Advantages of Fuzzy Optimization Models in Practical Use', *Fuzzy Optimization and Decision Making*, Vol. 3, pp. 295-309.

Rummel, J.L., Walter, Z., Dewan, R. and Seidmann, A. (2005), 'Activity Consolidation to Improve Responsiveness', *European Journal of Operational Research*, Vol. 161, pp. 683-703.

Sadiq, W. and Orlowska, M. (2000), 'Analyzing Process Models Using Graph Reduction Techniques', *Information Systems*, Vol. 25, No. 2, pp. 117-134 .

Salimifard, K. and Wright, M. (2001), 'Petri Net-Based Modelling of Workflow Systems: An

Overview', *European Journal of Operational Research*, Vol. 134, pp. 664-676.

Scheer, A.W. (1994), *Business Process Reengineering, Reference Models for Industrial Enterprises*, Springer, Berlin.

Shah, N. (1998), 'Single- and Multi-site Planning and Scheduling: Current Status and Future Challenges', in Pekny J.F. and Blay G.E. (eds.), *Proceedings of the Third International Conference on Foundations of Computer-Aided Process Operations:* Snownbird, Utah, USA, pp. 75-90.

Shen, H., Wall, B., Zaremba, M., Chen, Y. and Browne, J. (2004), 'Integration of Business Modelling Methods for Enterprise Information System Analysis and User Requirements Gathering', *Computers in Industry*, Vol. 54, pp. 307-323.

Shimizu, Y. and Sahara, Y. (2000), 'A Supporting System for Evaluation and Review of Business Process through Activity-Based Approach', *Computers & Chemical Engineering*, Vol. 24, pp. 997-1003.

Smith, H. (2003), 'Business Process Management - the Third Wave: Business Process Modelling Language and Its Pi-Calculus Format', *Information and Software Technology*, Vol. 45, pp. 1065-1069.

Soliman, F. (1998), 'Optimum Level of Process Mapping and Least Cost Business Process Re-Engineering', *International Journal of Operations and Production Management*, Vol. 18, No. 9/10, pp. 810-816.

Stock, J.R. and Lambert, D.M. (2001), *Strategic Logics Management*, McGraw-Hill, Irwin.

Stohr, E.A. and Zhao, J.L. (2001), 'Workflow Automation: Overview and Research Issues', *Information Systems Frontiers*, Vol. 3, No. 3, pp. 281-296.

Tinnila, M. (1995), 'Strategic Perspective to Business Process Redesign', *Business Process Re-Engineering and Management Journal*, Vol. 1, No. 1, pp. 44-59.

Tiwari, A. (2001), 'Evolutionary computing techniques for handling variables interaction in engineering design optimisation'. PhD Thesis, SIMS, Cranfield University, Cranfield, U.K.

Tiwari, A., Vergidis, K. and Majeed, B. (2006), 'Evolutionary Multi-Objective Optimisation of Business Processes', in *Proceedings of IEEE Congress on Evolutionary Computation 2006* Vancouver, Canada, pp. 3091-3097.

Valiris, G. and Glykas, M. (1999), 'Critical Review of Existing BPR Methodologies, The Need for a Holistic Approach', *Business Process Management Journal*, Vol. 5, No. 1, pp. 65-86.

Valiris, G. and Glykas, M. (2004), 'Business Analysis Metrics for Business Process Redesign', *Business Process Management Journal*, Vol. 10, No. 4, pp. 445-480.

van der Aalst, W.M.P. (1995), 'A Class of Petri Nets for Modelling and Analyzing Business Processes', *Computing Science Reports 95/26*, Eindhoven University of Technology, Eindhoven.

van der Aalst, W.M.P. (1996), 'Petri Net Based Scheduling', *OR Spectrum*, Vol. 18, pp. 219-229.

van der Aalst, W.M.P. (1998a), 'The Application of Petri-Nets to Workflow Management', *Journal of Circuits, Systems and Computers*, Vol. 8, No. 1, pp. 21-66.

van der Aalst, W.M.P. (1998b), 'Formalisation and Verification of Event-Driven Process Chains', *Computing Science Reports 98/01*, Eindhoven University of Technology, Eindhoven.

van der Aalst, W.M.P. (2001), 'Re-Engineering Knock-Out Processes', *Decision Support Systems*, Vol. 30, pp. 451-468.

van der Aalst, W.M.P. (2003), 'Challenges in Business Process Management: Verification of Business Processes Using Petri Nets', *Bulletin of the EATCS*, Vol. 80, pp. 174-198.

van der Aalst, W.M.P. (2004), 'Business Process Management: A Personal View', *Business Process Management Journal*, Vol. 10, No. 2, pp. 135-139.

van der Aalst, W.M.P. and ter Hofstede, A.H.M. (2002), 'Workflow Patterns: On the Expressive Power of (Petri-Net-Based) Workflow Languages', *in Proceedings of the Fourth Workshop on the Practical Use of Coloured Petri Nets and CPN Tools (CPN 2002),* Vol. 560, University of Aarhus, Aarhus, Denmark, pp. 1-20.

van der Aalst, W.M.P. and ter Hofstede, A.H.M. (2003), 'YAWL: Yet Another Workflow Language (Revised Version)', *QUT Technical Report, FIT-TR-2003-04*, Qeensland University of Technology, Brisbane.

van der Aalst, W.M.P. and van Hee, K.M. (1996), 'Business Process Redesign: A Petri-Net-Based Approach', *Computers in Industry*, Vol. 29, pp. 15-26.

van der Aalst, W.M.P., Hirnschall, A. and Verbeek, H.M.W. (2002), 'An Alternative Way to Analyse Workflow Graphs', *Lecture Notes in Computer Science*, Vol. 2348, pp. 535.

van der Aalst, W.M.P., ter Hofstede, A.H.M. and Weske, M. (2003), 'Business Process Management: A Survey', *Lecture Notes in Computer Science*, Vol. 2678, pp. 1-12.

van der Aalst, W.M.P., van Hee, K.M. and Houben, G.J. (1994), 'Modelling and Analysing Workflow Using a Petri-Net Based Approach', in *Proceedings of 2nd Workshop on Computer-Supported Cooperative Work, Petri Nets and Related Formalisms,* pp. 31-50.

Vergidis, K., Tiwari, A. and Majeed, B. (2006), 'Business Process Improvement Using Multi-Objective Optimisation', *BT Technology Journal*, Vol. 24, No. 2, pp. 229-235.

Volkner, P. and Werners, B. (2000), 'A Decision Support System for Business Process Planning', *European Journal of Operational Research*, Vol. 125, pp. 633-647.

Wamelink, J.W.F., Stoffele, M. and van der Aalst, W.M.P. (2002), 'Workflow Management in Construction: Opportunities for the Future', in Agger K., Christiansson P. and Howard R. *Proceedings of CIB W78 Conference: Distributing Knowledge in Building:* Aarhus, Denmark, pp. 115-122.

Wang, K., Salhi, A. and Fraga, E.S. (2004), 'Process Design Optimisation Using Embedded Hybrid Visualisation and Data Analysis Techniques Within a Genetic Algorithm Optimisation Framework', *Chemical Engineering and Processing*, Vol. 43, pp. 663-675.

Wang, M. and Wang, H. (2005), 'Intelligent Agent Supported Business Process Management', in *Proceedings of the 38th Hawaii International Conference on System Sciences*, Hawaii, pp. 567-577.

WfMC (1995), *The Workflow Reference Model, WFMC-TC-1003, 19-Jan-95, 1.1*, available at: http://www.wfmc.org (accessed 2006).

Wohed, P., van der Aalst, W.M.P., Dumas, M. and ter Hofstede, A.H.M. (2002), 'Pattern-Based Analysis of BPEL4WS', *QUT Technical Report, FIT-TR-2002-04*, Queensland University of Technology, Brisbane.

Wohed, P., van der Aalst, W.M.P., Dumas, M., ter Hofstede, A.H.M. and Russell, N. (2004), 'Pattern-Based Analysis of UML Activity Diagrams', *BETA Working Paper Series, WP 129*, Eindhoven University of Technology, Eindhoven.

Zakarian, A. (2001), 'Analysis of Process Models: A Fuzzy Logic Approach', *International Journal of Advanced Manufacturing Technology*, Vol. 17, pp. 444-452.

Zakarian, A. and Kusiak, A. (2000), 'Analysis of Process Models', *IEEE Transactions on Electronics Packaging Manufacturing*, Vol. 23, No. 2, pp. 137-147.

Zakarian, A. and Kusiak, A. (2001), 'Process Analysis and Reengineering', *Computers & Industrial Engineering*, Vol. 41, pp. 135-150.

Zapf, M. and Heinzl, A. (2000), 'Evaluation of Generic Process Design Patterns: An Experimental Study', *Lecture Notes in Computer Science*, Vol. 1806, pp. 83-88.

Zhou, Y. and Chen, Y. (2002), 'Business Process Assignment Optimisation', in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics Vol. 3*, pp. 540-545.

Zhou, Y. and Chen, Y. (2003a), 'The Methodology for Business Process Optimised Design', in *IECON Proceedings (Industrial Electronics Conference) Vol. 2*, pp. 1819-1824.

Zhou, Y. and Chen, Y. (2003b), 'Project-Oriented Business Process Performance Optimization', in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics Vol. 5*, pp. 4079-4084.

Zitzler, E., Laumanns, M. and Thiele, L. (2001), 'SPEA2: Improving the Strength Pareto Evolutionary Algorithm', *TIK Report Nr. 103*, Computer Engineering and Networks Lab (TIK), Swiss Federal Institute of Technology (ETH) Zurich.

APPENDIX A

Questionnaires & Survey Participants

This Appendix provides the details of the interviewees, the participants and their organisations that provided the responses to the service industry survey (A.1). It also details the questions that were asked in the industry visits (A.2) and the on-line survey (A.3) by providing the relevant questionnaires. The service industry survey was performed by both the researchers of the 'Intelli-Process' project. The questionnaire sections that do not involve questions relevant to this research have been omitted.

A.1 Service industry participants

Table A.1 lists the five service industry survey participants. For each interviewee, the date of the interview, the organisation and the job title are also provided. The interviewees were asked to provide their experience in relation to business processes in order to provide an indicator of their expertise.

No.	Date of interview	Job Title	Organisation	Experience (years)
1	5 April 2006	Project Manager	BT	10
2	6 April 2006	Credit Risk Analyst	Barclaycard	2
3	6 April 2006	Project Manager	Barclaycard	2
4	18 May 2006	Logistics Manager	HTC	5
5	24 May 2006	Senior Researcher	BT	15

Table A.1. List of interviewees from the service industry visits

Table A.2 provides the list of the 20 on-line survey participants with the same details as the previous table. The on-line feature of the survey provided the capability to reach participants across the world.

No.	Job Title	Organisation	Experience (years)
1	Director of Business Process Improvement	Rockland Trust Company	11
2	Business Performance Improvement Manager	Heller College of Business Administration	15
3	Assistant Registrar	University of Chester	12
4	Principal Consultant	Colin Brook & Associates	18
5	Director	Wilde FEA Ltd	25

6	Project Manager	The Success Institute Inc.	26
7	Head of Process Design	Lloyds TSB	25
8	Director of Corporate Information Systems	Liverpool John Moores University	12
9	Professor / Consultant	Costa Rica Institute of Technology	18
10	Director of Process Management	State of Rhode Island	15
11	BPR Lecturer / Consultant	University of Maribor	10
12	Business Development and IT	DaimlerChrysler Services	12
13	Process Change Manager	British Nuclear Group	10
14	Global Director of Manufacturing Strategy Development	Sealed Air Corporation	20
15	Project Manager	Leeds City Council	1
16	Data Management Officer	Macquarie University	10
17	Assistant Professor	Universidade Catolica Portuguesa	10
18	Group Account Manager	Euro RSCG	1
19	Logistics and Distribution Manager	Schering Hellas S.A.	3
20	Credit Risk Analyst	Barclays PLC	1

Table A.2. List of participants of the on-line survey

A.2 Questionnaire for industry visits

The purpose of this questionnaire is to identify the procedures and/or practices in industry for capturing / modelling the business processes and optimising the process design. It also aims to capture the industrial requirements for an intelligent tool regarding business processes.

Section A: Business Process Management & Modelling

This section of the questionnaire tries to understand how your company perceives business process. It also deals with business process management and modelling issues.

Q1. How does your company perceive/understand the concept of **business process**?

Q2. Is there a process-centred (cross-functional) focus or a classical functional orientation?

Q3. Who holds a holistic view/knowledge about each business process?

Q4. What is the most common approach when it comes to modelling/representing a business process? Is there a particular methodology used (e.g. IDEF)?

Q5. Which would you name as the most common patterns (e.g. parallel flow, feedback loops, and decision points) of a business process?

Q6. What are the stages for introducing a new process within your organisation? Do these stages differ according to the type of process being introduced?

Q7. Are there regular reviews of business processes within your organisation?

Q8. Do you use a particular methodology for the review?

Q9. What are the main steps of the review?

<u>Section B: Business Process Mining</u>

(section that belongs to the second researcher of the 'Intelli-Process' project)

<u>Section C: Business Process Comparison</u>

(section that belongs to the second researcher of the 'Intelli-Process' project)

<u>Section D: Business Process Analysis and Optimisation</u>

The aim of this section is to investigate the current business process analysis techniques that are being used within the company. It also attempts to explore any structured improvement approaches for business processes.

Q17. Are there specific KPIs (Key Performance Indicators) to evaluate the business process performance? (Please specify and describe them)

Q18. Is business process simulation or statistical analysis used to analyse a business process? Are any other methods/approaches used for quantitative analysis of business processes?

(If yes to the last question) – Q19. Which software tools are used and why?

Q20. Do you apply any structured methodology for improving business processes manually and/or automatically?

Q21. Is this approach software assisted? How is the process represented (modelled)?

Q22. Is the business process improved according to one or more objectives simultaneously?

Q23. Does the approach take into account all aspects of the process (e.g. actors, resource constraints etc.) or focuses on particular aspects?

<u>Section E: Industry Requirements for a Software Tool</u>

This section of the questionnaire records your company's requirements for a process mining, optimisation and conformance tool

Q24. If a software tool was to be implemented, how would you describe it process modelling capabilities (i.e. visual representation of a process)?

Q25. How would you describe its main functionalities to assist business process analysis and optimisation?

Section F: Additional Multiple-choice Questions

Q26. Please select the phrase that reflects most appropriately the current practice within your company:
[ ] The various departments work independently.
[ ] Although there is a departmental segmentation, there is informal co-operation for certain processes (i.e. cross-functional teams)
[ ] The company is organised around business processes
[ ] Other, please specify: ………………………………………………………………

Q27. Which of the definitions below is closest to your understanding of the term *'business process'*?
[ ] A dynamic ordering of work activities across time and place, with a beginning, an end and clearly identified process inputs and outputs
[ ] A set of logically related tasks, performed by specific actors to achieve a defined business outcome
[ ] A construct with complex sociotechnical interrelations
[ ] Other, please specify:………………………………………………………………

Q28. For the business processes that currently exist in your organisation:
[ ] No one has explicit knowledge about the complete process flow
[ ] There is a specific process owner who is responsible for each particular business process
[ ] The process knowledge is shared among the main actors of the process
[ ] Other, please specify: …………………………………………………………………

Q29. What is the most common modelling/representation approach of business processes used in your organisation?
[ ] Simple flowcharts with no predefined notation
[ ] IDEF0 / IDEF3
[ ] Petri-nets
[ ] Other software-assisted representation method
[ ] Process documentation (no visualisation of business processes)
[ ] Other, please specify:…………………………………………………………………

Q30. Please tick the main business process patterns that you recognise in the organisation's business processes:
[ ] Sequential flow
[ ] Parallel flow (AND)
[ ] Decision points (OR / XOR)
[ ] Feedback loops (GOTO)
[ ] Process documentation (no visualisation of business processes)
[ ] Other pattern, please specify:………………………………………………………….

Q31. Do you use business process management software?

[ ] No
[ ] Yes
    [ ] ARIS
    [ ] SAP
    [ ] Tibco software
    [ ] Other software, please specify:………………………………………………………

Q32. Which methods/approaches are used for quantitative analysis of business processes and why?
[ ] Statistical analysis based on KPIs (Key Performance Indicators) [ ] Simulation [ ] Other approach, please specify:……………………………………………………………..

Please use the space below to add any comments you might have, or further explain any business process related issues that were not covered by the questions above:

A.3 On-line questionnaire

The on-line version of the questionnaire is developed based on the questionnaire developed for the industry visits. It has the same structure in terms of sections, but most of its questions are closed or require a short answer.

__Section A: Business process perception and modelling__

Q1. Please select the phrase that reflects most appropriately the current practice within your company:
[ ] The various departments work independently. [ ] Although there is a departmental segmentation, there is informal co-operation for certain processes (i.e. cross-functional teams) [ ] The company is organised around business processes [ ] Other, please specify: ………………………………………………………………

Q2. Which of the definitions below is closest to your understanding of the term *'business process'*?
[ ] A dynamic ordering of work activities across time and place, with a beginning, an end and clearly identified process inputs and outputs [ ] A set of logically related tasks, performed by specific actors to achieve a defined business outcome [ ] A construct with complex sociotechnical interrelations [ ] Other, please specify: ………………………………………………………………

Q3. For the business processes that currently exist in your organisation:
[ ] No one has explicit knowledge about the complete process flow
[ ] There is a specific process owner who is responsible for each particular business process
[ ] The process knowledge is shared among the main actors of the process
[ ] Other, please specify: …………………………………………………………………

Q4. What is the most common modelling/representation approach of business processes used in your organisation?
[ ] Simple flowcharts with no predefined notation
[ ] IDEF0 / IDEF3
[ ] Petri-nets
[ ] Other software-assisted representation method
[ ] Process documentation (no visualisation of business processes)
[ ] Other, please specify: …………………………………………………………………

Q5. Please tick the main business process patterns that you recognise in the organisation's business processes:
[ ] Sequential flow
[ ] Parallel flow (AND)
[ ] Decision points (OR / XOR)
[ ] Feedback loops (GOTO)
[ ] Process documentation (no visualisation of business processes)
[ ] Other pattern, please specify:……………………………………………………………………………………….

Q6. Do you use business process management software?
[ ] No
[ ] Yes
[ ] ARIS
[ ] SAP
[ ] Tibco software
[ ] Other software, please specify:………………………………………………………

<u>Section B: Business process mining</u>

(section that belongs to the second researcher of the 'Intelli-Process' project)

<u>Section C: Business process comparison</u>

(section that belongs to the second researcher of the 'Intelli-Process' project)

<u>Section D: Business process analysis & optimisation</u>

Q17. Are there specific KPIs to evaluate the business process performance? Please specify and describe them.

Q18. Which methods/approaches are used for quantitative analysis of business processes and why?
[ ] Statistical analysis based on KPIs (Key Performance Indicators) [ ] Simulation [ ] Other approach, please specify:…………………………………………………………………………………….

Q19. Are business processes simulated using a software environment before actual implementation? Which simulation software is used and why?

Q20. Do you apply any structured methodology for improving business processes manually and/or automatically?
[ ] No [ ] Yes If yes, which tools are being used and why?

Q21. Is the process improved according to one or more objectives? Does the approach take into account all the aspects of the process (e.g. actors, resources, constraints) or focuses on particular ones?

<u>Section E: Industry requirements for a software tool</u>

Q22. If a business process software tool was to be implemented, how would you describe its main functionalities to assist business process analysis and optimisation?

Please use the space below to add any comments you might have, or further explain any business process related issues that were not covered by the questions above:

APPENDIX B

Feasibility of business process designs

This Appendix details the cases of a design being feasible based on the proposed business process representation (see chapter 5). For each of the three identified feasibility cases a probability of occurrence is calculated. The total probability of design feasibility is then calculated as a combination of these three probabilities. Also, the extensive and the feasible sets of business process designs are identified and discussed in relation to the feasibility probability. The Appendix demonstrates the probability of a business process design being feasible with a numerical example.

B.1 Cases of design feasibility

Chapter 5 discusses the proposed representation for business processes and a proposed algorithmic approach towards composing business process designs (i.e. the PCA algorithm). A business process design is considered as *feasible* when:

4. *All* the process input resources are utilised by one or more tasks that participate in the process design,

5. *All* the process output resources are produced by one or more tasks that participate in the process design, and

6. Each task in the design is connected either with the process inputs, the process outputs or another task in the design.

Parameter	Description	Parameter	Description
n	Number of tasks in the library	n_d	No. of tasks in the design
t_{in}	No. of task input resources	t_{out}	No. of task output resources
r_{in}	No. of process input resources	r_{out}	No. of process output resources
r	No. of available resources		

Table B.1. Parameters for checking the feasibility requirements

This Appendix calculates a probability for all of these cases to occur in order to acquire a feasible business process design. For each of the cases above, a probability of occurrence is calculated; the combination of these probabilities provides the probability of a process design being feasible. To show better these probabilities the last section of this Appendix demonstrates a numerical example. The main problem parameters are in explained in table B.1.

B.2 Probability of task/process connectivity

The first probability that is calculated is the probability of *task connectivity* in the process design. During the elaboration of the business process design, a broken link occurs in the case that no task can be connected. The broken link occurs as a lack of *task connectivity* capability. *Task connectivity is the ability of the available tasks to be connected based on their common input and output resources* and it is measured by the probability of any two tasks being able to connect with each other. According to the problem statement any two tasks can connect to each other in the process design when at least one input resource of one task is common with one output resource of the other.

We assume that the probability of having a broken link in the design occurs when there is no task connectivity capability. Based on table B.1, each task has a fixed number of t_{in} input and t_{out} output resources allocated from the set of *r* available resources. Initially we examine the probability of two tasks, T1 and T2 being able to connect. Equation B.1 provides all the possible combinations for the t_{out} output resources of task T1 based on the *r* available resources.

$$\binom{r}{t_{out}} = \frac{r!}{t_{out}!\,(r - t_{out})!} \qquad r \geq t_{out} \geq 0 \qquad \text{(Equation B.1)}$$

If we exclude the t_{out} resources allocated to T1, then for a task T2 the combinations of its input resources can be calculated based on the remaining $(r-t_{out})$ resources arranged to sets of t_{in}:

$$\binom{r - t_{out}}{t_{in}} = \frac{(r - t_{out})!}{t_{in}!\,(r - t_{in} - t_{out})!} \qquad r \geq t_{in} + t_{out} \geq 0 \qquad \text{(Equation B.2)}$$

Based on equations B.1 and B.2, the probability of T2 input resources *not* containing any of T1 output resources is given by the probability P_0:

$$P_0 = \frac{\binom{r-t_{out}}{t_{in}}}{\binom{r}{t_{out}}} \qquad r \geq t_{in} + t_{out} \geq 0 \qquad \text{(Equation B.3)}$$

Equation B.3 provides the probability (P_0) of the two tasks T1 and T2 not sharing a common resource and thus not being able to connect. Based on this probability, we can define the *task connectivity probability* (P_{TC}) which is the probability of any two tasks sharing *at least* one common resource and thus being able to connect. This probability equals to:

$$P_{TC} = 1 - P_0 = 1 - \frac{\binom{r-t_{out}}{t_{in}}}{\binom{r}{t_{out}}} \qquad r \geq t_{in} + t_{out} \geq 0 \qquad \text{(Equation B.4)}$$

Based on the task connectivity probability (for any two tasks), we can define the *process connectivity probability* (P_{PC}) for a process design with n_d tasks as:

$$P_{PC} = P_{TC}^{\frac{n_d-1}{2}} = \left(1 - \frac{\binom{r-t_{out}}{t_{in}}}{\binom{r}{t_{out}}}\right)^{\frac{n_d-1}{2}} \% \qquad r \geq t_{in} + t_{out} \geq 0 \qquad \text{(Equation B.5)}$$

In a process design with n_d tasks, there are a minimum of (n_d–1) connections between the tasks if all placed in sequence. However, we assume that a process design might be formed with half of these (n_d–1) connections. This can occur either due to patterns formation or due to a feasible process design with fewer tasks. The process connectivity probability shows the probability (expressed in percentage) that at least half of the (n_d–1) tasks can be connected during the composition of a feasible business process design.

B.3 Process input and output requirements

The other two probabilities that need to be calculated are in relation to the process input and output requirements. The process input requirements is the R_{in} set that needs to be utilised in the beginning of the process and the process output requirements is the R_{out} set that needs to be produced at the end of the process design. What makes these two requirements challenging is that *all* the resources in R_{in} set need to be utilised and *all* the resources in R_{out} need to be produced for the process design to be feasible.

For all the r_{in} process inputs to be utilised, they need to exist in the input resources of the n_d tasks in the design so that the tasks can link with process input resources in the beginning of the business process design composition. If *at least one* of these resources does not exist in the input resources of the tasks, then the design is infeasible. Therefore, we need to calculate the probability of *all the process input resources existing in the set of the participating tasks' input resources*. Each of the n_d participating tasks has t_{in} input resources so the set of task input resources has size of $n_d \times t_{in}$. The probability P_{IN} of *all* the r_{in} resources to appear in the set of task input resources is calculated as:

$$P_{IN} = \frac{\binom{n_d . t_{in}}{r_{in}}}{\binom{r}{r_{in}}} \% \qquad \text{(Equation B.6)}$$

The probability is calculated by calculating all the possible combinations of r_{in} resources in the task input resources set against all combinations of the r_{in} resources in the set of r available resources. Similarly, for the process output requirements the probability P_{OUT} equals with *all of the r_{out} output resources being produced by the* n_d *participating tasks* in the process design.

$$P_{OUT} = \frac{\binom{n_d . t_{out}}{r_{out}}}{\binom{r}{r_{out}}} \% \qquad \text{(Equation B.7)}$$

Note that equations B.6 and B.7 become equivalent for the same number of input and output resources per task ($t_{in} = t_{out}$) and the same number of process input and output resources ($r_{in} = r_{out}$).

B.4 Extensive set of business process designs

The number of n_d tasks in the process design and the number of *n* candidate tasks in the library define the *extensive* set *D* of business process designs. The extensive set involves every possible combination of the *n* tasks arranged in process designs with n_d tasks and is different to the set of *feasible* designs as it does not take into account any of the process feasibility requirements. Therefore, the set D_F of feasible process designs is a subset of the extensive set with business process designs that satisfy the criteria of feasibility. The extensive set involves groups of n_d different tasks where no task is repeated more than once in each set. The order of the tasks in the set does not matter and therefore in order to calculate the cardinality of the extensive set we calculate the combinations (not permutations) of the tasks. Two designs are considered different when they contain at least a different task. The number of possible process designs for a design with n_d tasks equals the combinations of the *n* library tasks into n_d-task process designs, i.e. the binomial coefficient:

$$D = \binom{n}{n_d} = \frac{n!}{n_d!\,(n-n_d)!} \qquad n \geq n_d \geq 0 \qquad \text{(Equation B.8)}$$

We assume that calculating the extensive set of business process designs D and multiplying it with the feasibility probabilities calculated previously (P_{PC}, P_{IN} and P_{OUT}) we can acquire *an estimation* of the population of feasible business process designs (D_F) for given parameters.

B.5 Estimating the number of feasible designs

Based on the previous sections, the probability (P_F) for a business process design to be feasible equals with:

$$P_F = P_{PC} \times P_{IN} \times P_{OUT} \qquad \text{(Equation B.9)}$$

Equation B.9 multiplies all the previously calculated probabilities in order to acquire the overall probability of a design being feasible. The reason for multiplication –and not aggregation– is that the probabilities are independent to each other and we want all of them to occur for a business process design to be feasible. Having calculated the probability of feasible designs (P_F) and the extensive set of business process designs (D), we can acquire *an estimation* of the number of feasible business process (D_F) with the following formula:

$$D_F = P_F \times D \qquad \text{(Equation B.10)}$$

To demonstrate the probability of feasible designs, table B.2 has a numerical paradigm for the problem parameters.

Parameter	Value	Parameter	Values
n	20	n_d	5
t_{in}	3	t_{out}	3
r_{in}	5	r_{out}	5
r	20		

Table B.2. Parameters and corresponding values

Based on table B.2 the probability of process connectivity equals to:

$$P_{PC} = \left(1 - \frac{\binom{20-3}{3}}{\binom{20}{3}}\right)^{\frac{5-1}{2}} \% = \left(1 - \frac{\binom{17}{3}}{\binom{20}{3}}\right)^{2} \% = 16\%$$

Because, $t_{in} = t_{out}$ and $r_{in} = r_{out}$, the probability for the process input requirements equals to the probability of the process output requirements:

$$P_{IN} = P_{OUT} = \frac{\binom{5\cdot 3}{5}}{\binom{20}{5}} \% = \frac{\binom{15}{5}}{\binom{20}{5}} \% = 19\%$$

Having calculated these three probabilities, the probability of feasibility (P_F) can be calculated as:

$P_F = P_{PC} \times P_{IN} \times P_{OUT} = 16\% \times 19\% \times 19\% = 0.6\%$

The extensive set of business process designs has a size of:

$D = \binom{20}{5} = 15{,}504$ designs.

The set of feasible business process designs is:

$D_F = P_F \times D = 15{,}504 \times 0.6\% = 91$ business process designs.

It is evident that given a set of n_d tasks, the probability for a design to be feasible is extremely low, which makes the problem of design composition very challenging. The proposed Process Composition Algorithm (PCA) will have to manage the initial set of tasks accordingly in order to try and reduce the infeasibility occurrences.

B.6 Summary

This Appendix demonstrated the high probability of a business process design being feasible. For a design to be considered a feasible, three conditions must be satisfied. The probability of each condition is calculated and the combined feasibility probability (P_F) is formed. This Appendix also calculated the extensive set of business process designs ($\boldsymbol{D}$) given the task library size and the process size. For average values of the problem parameters, the probability of a design being feasible is extremely low. This proves that the business process composition problem is challenging and calls for a manipulation strategy of the large number of infeasible business process designs.

APPENDIX C

Initial Experiments

This Appendix presents a classification of the control parameters based on initial experiments with the proposed optimisation framework. The proposed classification of the control parameters is used as an integral part of the proposed strategy for generating tuneable business process scenarios. Specifying the level of each control parameter will 'tune' an experimental scenario towards the desired level of complexity. This Appendix presents the results of the experiments that helped classify the control parameters and also a typical business process scenario that is used as a guide to these experiments.

C.1 Overview of the control parameters

The proposed control parameters are summarised in table C.1. For each of the control parameters a brief description is provided along with its link to the actual problem parameters. The control parameters can be used to create business process scenarios of varying complexity and help assess the framework's optimisation performance under different conditions. However, in order to assess the framework in a systematic way, the effect of the control parameters to the problem needs to be further investigated. The following section presents a typical business process optimisation problem that will be used as a starting point for determining the effect of the control parameters to the performance of the proposed business process optimisation framework.

No	*Control Parameter*	Description	*Problem parameter(s)*	Relates with	Affects
1	L	number of neighbouring islands	n_{min}, n_d	feature B	Convergence
2	D	average distance of islands	S_d	feature B	Diversity (Convergence)
3	λ	degree of island overlap	TA_i	feature C	Diversity & Convergence
	μ	ratio of task attributes			
4	γ	ratio of tasks in the library vs. design	N	feature A	Diversity & Convergence

Table C.1. Control parameters of the business process optimisation problem

C.2 A typical business process scenario

This section describes a 'typical' business process scenario. Such a scenario is a straight-forward problem of low complexity. The main use of this scenario is to be used as a guide towards experimentation with the control parameters. Using this scenario as a basis, each control parameter will modify one problem aspect at a time (e.g. library size) in order to assess the complexity effect on the problem. Based on the impact to the problem, the control parameter will be assigned a particular classification (e.g. small / medium / large). This classification will assist in formulating a strategy for generating tuneable business process scenarios on the basis of systematically investigating the effect of multiple control parameters combined at different levels of impact. Table C.2 shows the parameters of the typical business process scenario.

Parameter	Value	Description
n	100	Number of tasks in the library
n_d	10	No. of tasks in the design
n_{min}	7	Minimum number of tasks in the design
r	20	No. of available resources
t_{in} / t_{out}	3	No. of task input/output resources
r_{in} / r_{out}	5	No. of process input /output resources
p	2	No. of task/process attributes *(α, β)*
α	100 – 110	First task/process attribute ($\alpha_{min} - \alpha_{max}$)
β	200 – 220	Second task/process attribute ($\beta_{min} - \beta_{max}$)

Table C.2. Problem parameters for the business process scenario

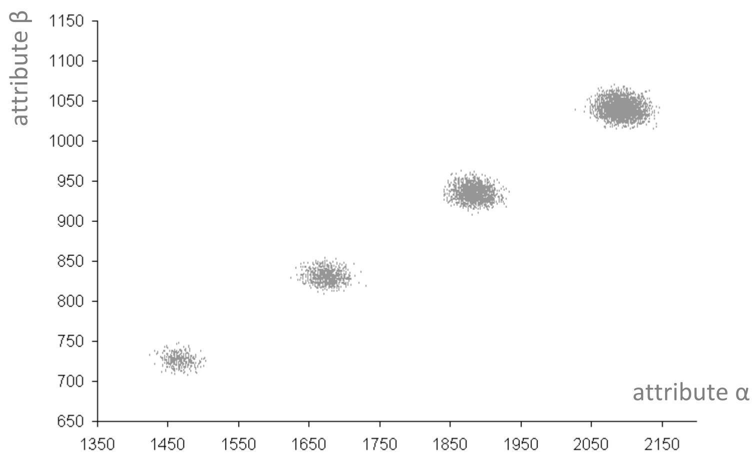


Figure C.1. Search space of the typical business process scenario

The values in table C.2 are selected based on initial experiments in order to determine a low complexity scenario. Figure C.1 shows the search space of the typical business process scenario as it was generated with the LSSA algorithm. The results shown in the following sections aim to be a guide for classifying the control parameters and therefore the performance of NSGA2 will not be scrutinised.

C.3 Experiments with number of neighbouring islands (L)

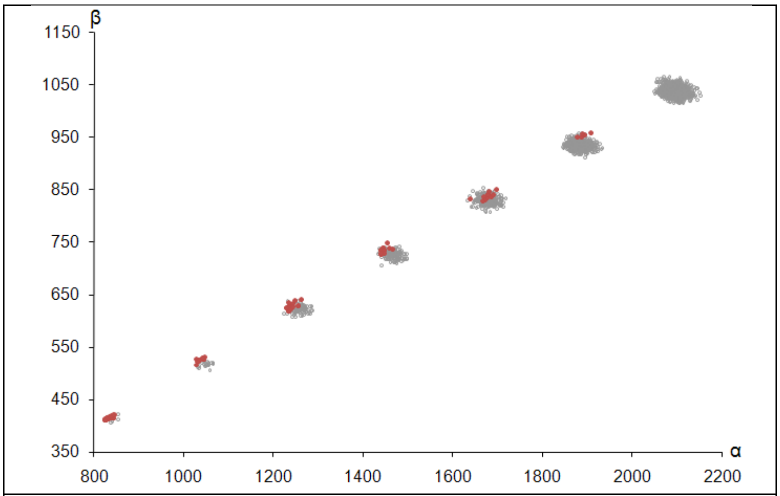
The number of neighbouring islands can potentially hinder the convergence and diversity capability of the optimisation algorithm used by the proposed framework. The aim of this section is to investigate the behaviour of the algorithm for increasing values of L and provide a guide on what could be characterised as 'small' or 'large' number of neighbouring islands in the context of the proposed optimisation framework.

Figure C.2 shows three experiments with different values for L (10, 15 and 20). The first experiment in figure C.2(a) shows the search space for $n_{min} = 1$ and $n_d = 10$. The feasible designs start with 4 or more tasks. NSGA2 is able to discover non-dominated solutions in the first four islands, few dominated solutions in the following two and zero solutions in the last island (10-task process designs). In the next experiment (L=15), the algorithm's performance deteriorates as for most islands NSGA2 discovers dominated solutions with the exception of the last two islands where no solutions are identified. Finally in the last instance (L=20), NSGA2 performance drops sharply as there are no solutions identified for the five uppermost islands.

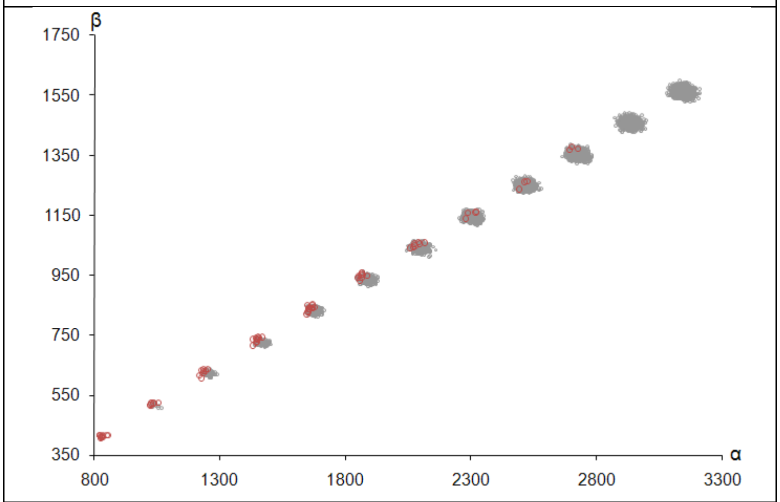
As expected the algorithm's convergence capability is significantly affected by increasing number of neighbouring islands. Also, the diversity of solutions is limited as for large numbers of L NSGA2 cannot identify solutions in all the available process sizes. This series of experiments is used as a guide for classifying the different levels that L can take and do not necessarily signify the limits of the framework in terms of neighbouring islands. The reason is that in these experiments all the other parameters remained constant whereas the framework might have performed better having larger task library size in the experiments with large neighbouring islands. However, the experiments provide a strong indication of the impact of the number of neighbouring islands in the framework performance. Based on the results shown in figure C.2, L can be classified as:

- *low*, for 0-4 neighbouring islands,
- *moderate*, for 5-9 neighbouring islands,

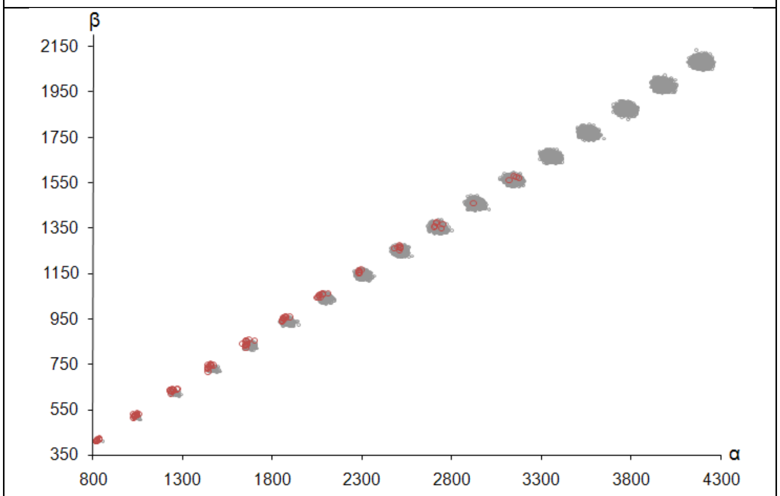
- ⊖ *large*, for 10-14 neighbouring islands, and
- ⊖ *very large*, for more than 15 neighbouring islands



(a) *L* = 10



(b) *L* = 15

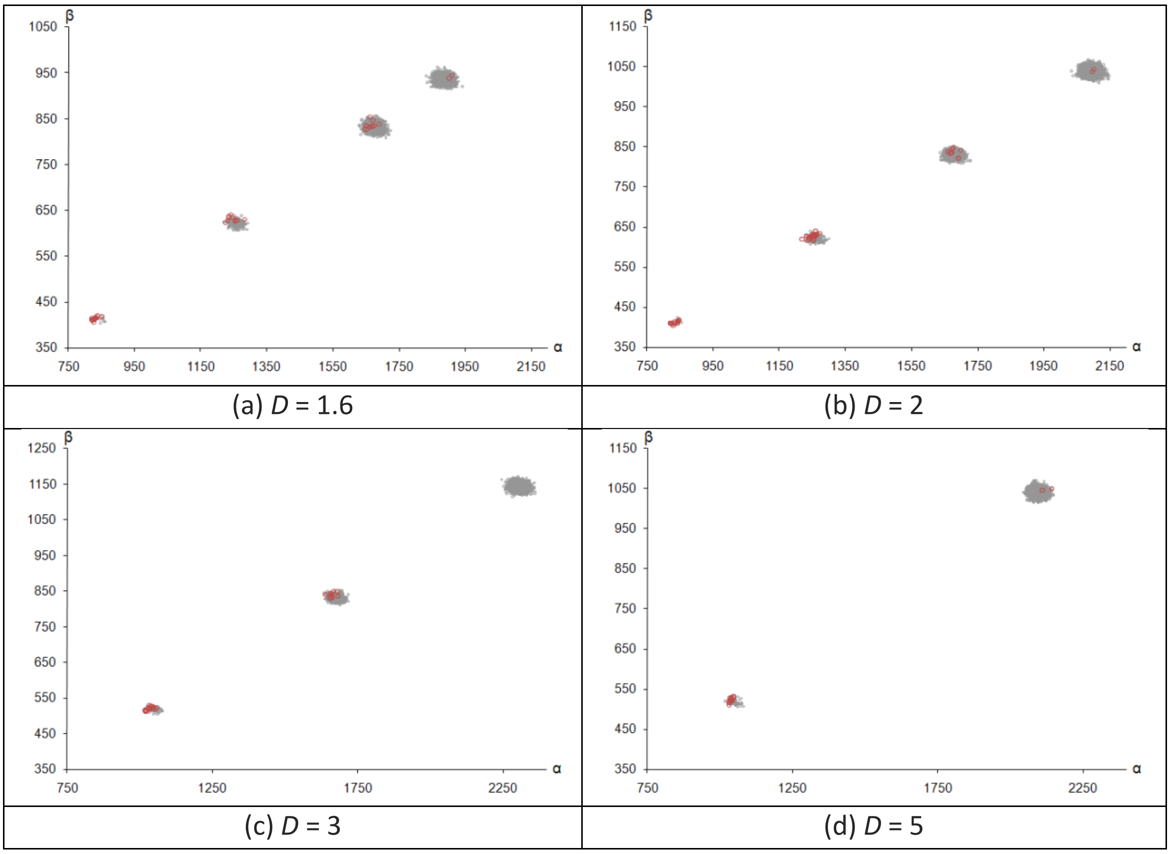


(c) *L* = 20

Figure C.2. Experiments with different L values

C.4 Experiments with number and continuity of islands (D)

The number and continuity of islands in the search space can affect mainly the diversity but also the convergence capability of the optimisation algorithms. The typical business process scenario involved neighbouring islands where $D = 1$. The aim of this section is to investigate the framework's performance for average distance of islands larger than 1. This would provide a guide on the tolerance of the framework for a fragmented search space where the islands are scattered across. The result of the experiments in this section will be a classification that characterises the different values that D takes.

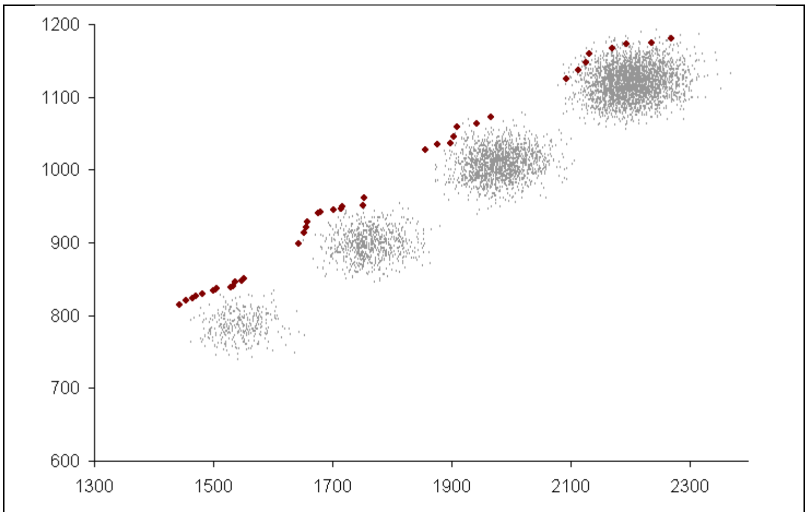


(a) D = 1.6 | (b) D = 2 | (c) D = 3 | (d) D = 5

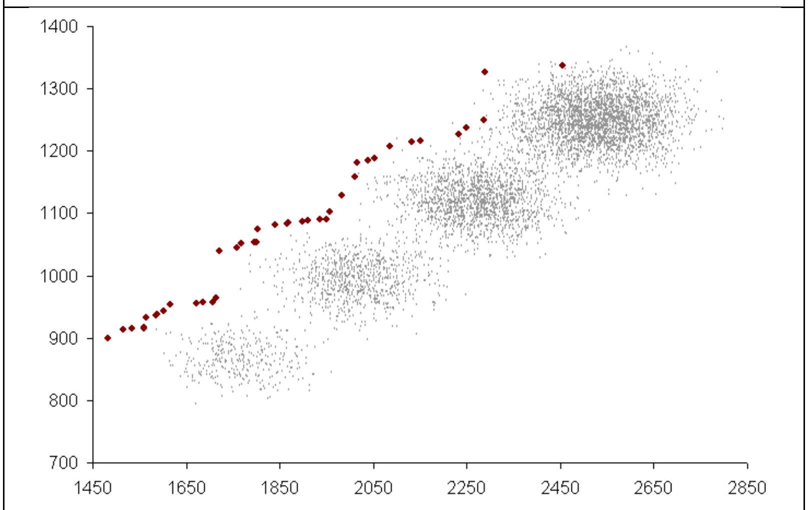
Figure C.3. Experiments with different D values

Figure C.3 shows the NSGA2 results of four experiments with different D values. The first experiment (a) is for $S_d = \{4, 6, 8, 9\}$, four islands with average distance 1.6. Although NSGA2 discovers solutions in all the four islands, only in the first two the solutions appear non-dominated. Similar results are generated in the second experiment (b) where $S_d = \{4, 6, 8, 10\}$ and D equals 2. NSGA2 locates only one solution in the uppermost island. The performance of the algorithm deteriorates in the third experiment (c) where $S_d = \{5, 8, 11\}$ and D=3. In this experiment and the next, the number of islands is reduced to show that the complexity introduced by the distance cannot be softened by reducing the islands in the search space. In the case of the results shown in figure C.3(c), NSGA2 locates non-dominated solutions in the first island, dominated in the second and none in the third. In the last experiment there are two islands with S_d = {5, 10} and *D*=5. NSGA2 locates a non-dominated front in the first island but shows a difficulty converging to the second island where only two solutions are identified. Based on the results shown in figure C.3, the average distance between the islands in the search space (*D*) can be classified into three categories:

- *short*, for average distance between 1 and 1.5,
- *moderate*, for average distance between 1.5 and 3, and,
- *distant*, for average distance above 3.



(a) λ = 0.6



(b) λ = 1

Figure C.4. Experiments with different degrees of overlap

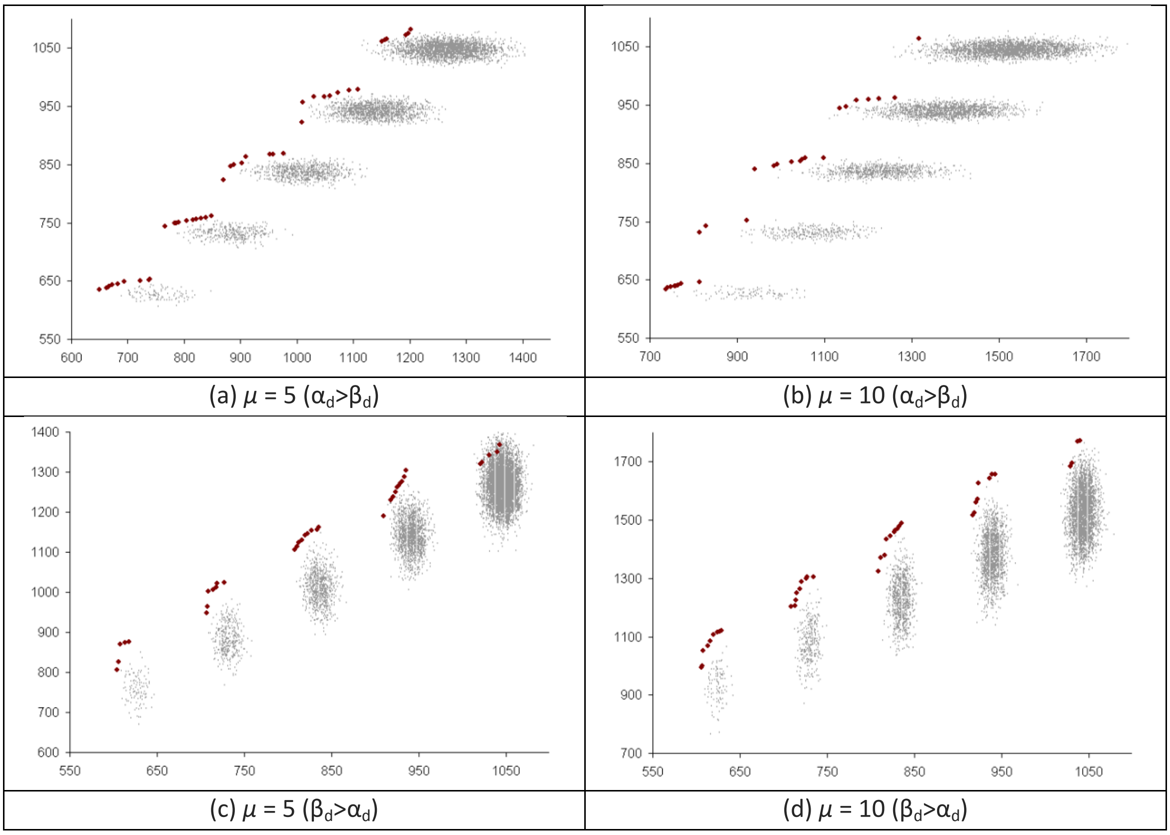
C.5 Experiments with overlap and shape of islands

This section examines the effect of island overlap and island shape with two different sets of experiments. The first set focuses on the island overlap; the effect it has on the search space and the effect it can potentially have on the framework's optimisation performance. The overlap of islands in the search space is measured by the parameter λ. In the typical business process scenario there is no overlap among any of the islands that constitute the search space ($\lambda = 0$). The aim of these experiments is to classify the different values that λ can take based on the effect of overlap to the framework's optimisation capability.

Figure C.4 shows the results of two experiments with different degrees of overlap. In the first experiment (a), the three uppermost islands are overlapping thus giving a degree of overlap equal to 0.6 (two out of three regions overlap). The results in the two uppermost islands indicate that NSGA2 cannot locate non-dominated solutions. In the second experiment (b), all the islands in the search space overlap and thus $\lambda = 1$. In this case the fronts in each island do not distinguish in a clear way, rather the result seems more like a continuous Pareto-optimal front across the islands. The front is less dense than in the typical business process scenario and the first experiment of this series. In particular, in the two uppermost islands the solutions are scarce and mostly dominated. Based on these, λ can be classified as:

- *no overlap*, for values between 0 and 0.2,
- *medium overlap*, for values between 0.3 and 0.5,
- *dominant overlap*, for values between 0.6 and 0.9 and,
- *full overlap*, for values equal to 1.

The shape of the islands can affect the convergence of the solutions towards the optimal front. As the previous section described, the shape of the islands is characterised by the ratio of the difference between the maximum and minimum values of the two attributes. The ratio is calculated based on which attribute has larger difference (it goes into the nominator of the ratio). In the typical business process scenario, $\alpha_d = 10$ and $\beta_d = 20$ which means that $\mu = \beta_d/\alpha_d = 2$. As the results in the typical scenario were satisfactory, this series of experiments will assess the problem for different values of μ in order to classify the ranges that the parameter can take based on its effect to the search space.



(a) $\mu = 5$ ($\alpha_d > \beta_d$) | (b) $\mu = 10$ ($\alpha_d > \beta_d$) | (c) $\mu = 5$ ($\beta_d > \alpha_d$) | (d) $\mu = 10$ ($\beta_d > \alpha_d$)

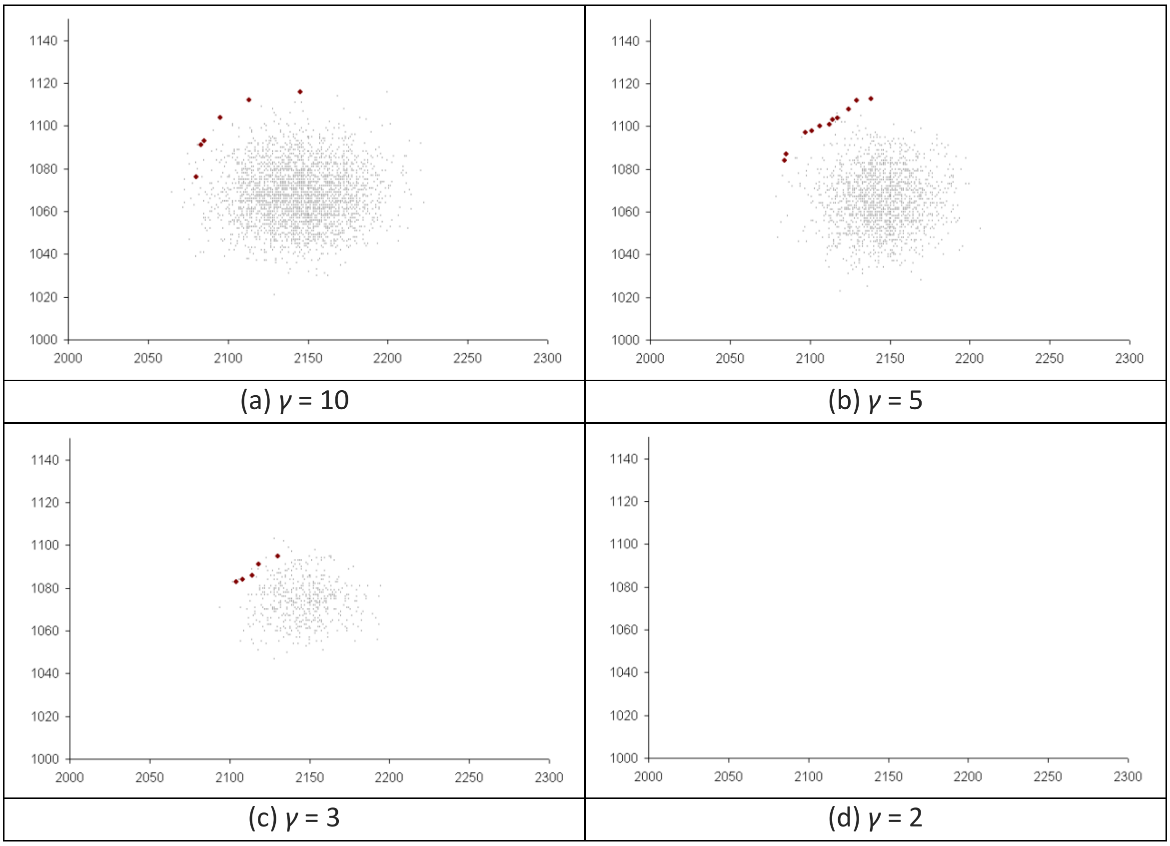
Figure C.5. Experiments with different μ values

Figure C.5 shows the results of the experiments based on the μ parameter. In the first two examples, $\alpha_d > \beta_d$, while in the last two, $\beta_d > \alpha_d$. For $\mu = 5$ NSGA2 performs relatively well in both experiments –(a) and (c)– discovering dense fronts of non-dominated solutions in almost all the islands of the search space. In both cases though, the algorithm has a difficulty in converging in the uppermost island. In experiments (b) and (d) where $\mu = 10$, the performance of the algorithm deteriorates significantly. The fronts are not as dense and in the uppermost island only a scarce number of solutions are identified. Based on these remarks, μ can be classified as:

- *normal*, for values between 1 and 3,
- *challenging*, for values between 3 and 6,
- *hard*, for values between 6 and 10, and,
- *extreme*, for values above 10.

C.6 Experiments with density of solutions per island

The density of solutions per island can affect both diversity and convergence towards the optimal solutions. In order to investigate the framework's reaction to different library sizes, γ measures the ratio between tasks in the library vs. tasks in the design. The typical business process scenario involved an abundant number of tasks in the library with a ratio 10:1 compared to the tasks in the design ($\gamma = 10$). This section investigates the island of 10-task process designs for different sizes of the task library. On this particular series of experiments the search space is reduced to a single island in order to demonstrate better the varying density of solutions. The section concludes with a classification based on the different γ values in order to characterise a business process scenario based on its projected density of solutions.



(a) $\gamma = 10$ (b) $\gamma = 5$ (c) $\gamma = 3$ (d) $\gamma = 2$

Figure C.6. Experiments with different γ values

The results of the experiments for different γ values are shown in figure C.6. It is apparent that the density of solutions becomes sparser as the values of γ decrease. Since n_d is constant to 10 tasks, assigning a value to γ calculates the value for $n = \gamma \cdot n_d$. For a ratio of 10:1 tasks (a) the solutions in the island are abundant while for a ratio of 2:1 (d) there is not a single feasible solution in the search space. This affects the performance of the optimisation framework. A large number of available solutions help the algorithm to discover feasible solutions and converge towards the optimal. A limited number of solutions make even solution discovery a challenging task. Below is a classification of the values that γ can take in effect of the projected density of solutions in the search space:

- *abundant*, for γ equal or above 10,
- *satisfactory*, for γ between 5 and 10, and,
- *scarce*, for γ less than 5.

C.7 Summary of classification of the control parameters

This section summarises the classification of each of the control parameters. Table C.3 shows the different levels that each control parameter can take based on specific value ranges. Using this table, one can select different levels of complexity for each of the basic parameters and thus create a scenario that focuses on specific aspects of the problem (e.g. low density of solutions). Then, the performance of the proposed business process optimisation framework can be assessed based on the specific scenario.

L	***D***	***λ***	***μ***	***γ***
low (0-4)	*short* (1-1.5)	*no overlap* (0-0.2)	*normal* (1-3)	*abundant* (>= 10)
moderate (5-9)	*moderate* (1.5-3)	*medium overlap* (0.3-0.5)	*challenging* (3-6)	*satisfactory* (5-10)
large (10-14)	*distant* (>3)	*dominant overlap* (0.6-0.9)	*hard* (6-10)	*scarce* (<5)
very large (>15)		*full overlap* (>1)	*extreme* (>10)	

Table C.3. Summary of classification of the control parameters

Apart from generating experimental scenarios, the proposed classification in table C.3 can be used to assess the complexity of an existing business process scenario. For each of these scenarios the proposed classification can point their complexity on specific aspects (e.g. small library size) and thus help in defining the expectations regarding the performance of the real-life scenario and providing a more accurate explanation based on the generated optimisation results.

APPENDIX D

Evolutionary Multi-objective Optimisation Algorithms

This Appendix presents the four algorithms that are incorporated in the proposed optimisation framework (bpo^F), namely NSGA2, SPEA2, PESA2 and PAES. For each algorithm its operation and the main steps are demonstrated. The Appendix briefly discusses the basics of evolutionary multi-objective optimisation and concludes with the short discussion highlighting the key differences and the expectations from each of the evolutionary algorithms.

D.1 Evolutionary multi-objective optimisation

According to Deb (2001), optimisation refers to finding one or more feasible solutions which correspond to extreme values of one or more objectives. The proposed optimisation approach regarding business process designs is built on the basis of an evolutionary multi-objective optimisation approach.

D.1.1 Basics of multi-objective optimisation

An optimisation problem that involves the task of finding the optimal solution having one objective function is called *single-objective*. In the case that an optimisation problem involves more than one objective function, it is called *multi-objective*. Multi-objective optimisation problems and algorithms have received wide attention during the last two decades due to the fact that most real-world problems naturally involve multiple objectives. A multi-objective optimisation problem is formulated as:

Minimise/maximise $(f_1(x), f_2(x), \dots, f_M(x))^T$ Equation 6.1

In this formulation, there are M objectives to be either minimised or maximised. A solution x is a vector of a decision variables. A multi-objective optimisation problem can also be subject to a number of constraints that limit the problem boundaries.

D.1.2 The concept of optimality

In multi-objective optimisation problems (with conflicting objectives) there is no single optimum solution but a series of equally optimal. A solution to such problems assumes different trade-offs among the different (conflicting) objectives. Without any further information, no solution from the set of optimal can be said to be better than any other. Unlike single-objective optimisation where finding the lone optimum solution is important, in multi-objective optimisation it is important to discover and investigate all the optimal solutions that arise from the trade-offs between the objectives. The optimal solutions are called non-dominated or Pareto-optimal solutions. The boundary of the feasible region on which these solutions are located is called the Pareto front. The two primary goals of global multi-objective optimisation are:

1. To guide the search towards the global Pareto-optimal region (convergence), and
2. To maintain population diversity in the Pareto front.

The proposed optimisation framework attempts to achieve these goals by employing evolutionary optimisation algorithms. These techniques are further discussed in the following section.

D.2 Non-dominated Sorting Genetic Algorithm 2

This section describes the Non-dominated Sorting Genetic Algorithm 2 (or NSGA2). NSGA2 is one of the four Evolutionary Algorithms incorporated in the proposed optimisation framework. The section starts with citing the main reasons for selecting this algorithm and provides an overview of its main steps.

D.2.1 Overview of the algorithm

NSGA2 is considered a high-performing multi-objective optimisation algorithm. It was developed by Deb *et al.* (2001) as an answer to the criticisms of the original NSGA. The main improvements of the second version involved:

1. A fast non-dominated sorting approach,
2. A selection operator for elitism preservation, and
3. The specification of a niching operator to ensure diversity in a population.

NSGA2 is a computationally fast evolutionary algorithm proven to maintain a better spread of solutions and convergence in difficult test problems (Deb, 2001). NSGA2 is perhaps the most popular EMOA and has been applied to many problems on a number of research areas. Also, it has provided satisfactory results in real world applications.

D.2.2 Main steps of the algorithm

The main steps of the algorithm are described below. For a more in-depth description see Deb *et al.* (2001).

1. Create a random parent population of size N.
2. Sort the population based on non-domination, and to each solution assign a fitness value equal to its non-domination level.
3. Create a child population of size N using binary tournament selection, crossover and mutation operators.
4. Combine the parent and child populations to create a global population of size 2N.
5. Sort the global population based on non-domination.
6. Create a new parent population by selecting solutions in order of their fronts until the number of selected solutions exceeds N.
7. Sort the solutions of the last accepted front using niched comparison operator.
8. Using this sorting, select solutions from the last front until the size of new parent population becomes N.
9. If the number of generations has exceeded a pre-determined value (e.g. 100), stop the process, else go to step 3.
10. Display the final solutions.

D.3 Strength Pareto Evolutionary Algorithm 2

This section describes the Strength Pareto Evolutionary Algorithm 2 (or SPEA2). SPEA2 is also incorporated in the proposed optimisation framework. The section starts with citing the main reasons for selecting this algorithm and provides an overview of its main steps.

D.3.1 Overview of the algorithm

SPEA2 is another elitist evolutionary algorithm and was evolved as an improved version of SPEA from the same group of researchers (Zitzler *et al.*, 2001). SPEA2 came after NSGA2 was implemented. SPEA2 incorporates, in addition to its original version, a fine grained fitness assignment strategy, a density estimation technique and an enhanced archive truncation method. SPEA2 has been popular in the evolutionary multi-objective optimisation community and has been used in a variety of optimisation problems. SPEA2 and NSGA2 are the most prominent EMOAs used when comparing a newly designed EMOA (Coello Coello, 2005). The main differences of SPEA2 in comparison to SPEA are:

- An improved fitness assignment scheme is used, which takes for each individual into account how many individuals it dominates and it is dominated by.
- A nearest neighbour density estimation technique is incorporated which allows amore precise guidance of the search process.
- A new archive truncation method guarantees the preservation of boundary solutions.

D.3.2 Main steps of the algorithm

SPEA2 works by maintaining an external population at every generation storing all non-dominated solutions discovered so far beginning from the initial population. This external population participates in all genetic operations. At each generation, a combined population with the external and the current population is constructed. All non-dominated solutions in the combined population are assigned a fitness based on the number of solutions they dominate and dominated solutions are assigned fitness worse than the worse fitness of any non-dominated solution. This assignment of fitness makes sure that the search is directed towards the non-dominated solutions. A deterministic clustering technique is used to ensure diversity among non-dominated solutions.

The main steps of SPEA2 are briefly described below. For a more in-depth description see Zitzler *et al.* (Zitzler *et al.*, 2001).

1. The first step is to generate an initial population of size N. Also, the external or archive set is created. This set that will contain the non-dominated solutions.
2. Next, each individual is assigned a fitness value. The fitness assignment incorporates density information into its calculation. SPEA2 uses a truncation method that preserves boundary points. In the event that individuals have the same fitness values a density estimation technique is used.
3. The archive is updated by copying all the non-dominated individuals into the archive set.
4. The termination condition check is next. If the maximum number of generations or some other stopping criteria is satisfied then the algorithm stops.
5. If the stopping condition is not met then mating selection is performed using binary tournament selection
6. Finally crossover and mutation operators are applied to the mating pool and to the resulting population. The generation counter increments. Go to Step 2.

D.4 Pareto Envelope-based Selection Algorithm 2

This section describes the Pareto Enveloped-base Selection Algorithm (or PESA2). PESA2 is an elitist algorithms that was proposed by Corne *et al.* (Corne *et al.*, 2001). It is the third algorithm to be incorporated in the proposed optimisation framework. This section provides an overview of the algorithm and discusses its main steps.

D.4.1 Overview of the algorithm

PESA2 is a revised version of the original PESA algorithm. PESA2 is identical to PESA, except for the fact that it employs region-based selection. PESA2 uses an internal population and an external (or secondary) population. It also uses a hyper-grid division of the objective space to maintain diversity (through the creation of hyper-boxes). Its selection mechanism is based on the crowding measure used by the hyper-grid. This same crowding measure is also used to decide what solutions to introduce into the external population (i.e. the archive of non-dominated solutions found along the evolutionary process). Therefore, in PESA2, the archive plays a crucial role in the algorithm since it determines not only the diversity scheme, but also the selection performed by the method.

In region-based selection, the unit of selection is a hyper-box rather than an individual. The procedure consists of: (i) selecting (using any of the traditional selection techniques) a hyper-box and then (ii) randomly selecting an individual within the hyper-box. The main motivation of this approach is to reduce the computational costs associated with traditional EMOAs (i.e., those based on Pareto ranking). Again, the role of the external memory in this case is crucial to the performance of the algorithm. Apart from the standard parameters such as crossover and mutation rates, PESA2 has two parameters concerning the population size (size of the main population set and size of the archive set) and one parameter concerning the hyper-box.

D.4.2 Main steps of the algorithm

The main steps of PESA2 are briefly discussed below:

1. Generate and evaluate each on an initial 'internal' population of solutions and initialise the 'external' (archive) population to the empty set.
2. Incorporate the non-dominated member of the internal population to the archive.

3. If the termination criterion has been reached, then stop returning the external set of solution as the result. Otherwise, delete the current contents of the internal set and generate a new population of candidate solutions
4. Return to step 2.

In the archive incorporation step (step 2) the current set of new candidate solutions is incorporated into the archive one by one. A candidate may enter the archive if it is non-dominated by any current member of the archive. Once a candidate has entered the archive, members of the archive which it dominated (if any) will be removed.

D.5 Pareto Archived Evolution Strategy

This section describes the Pareto Archived Evolution Strategy (or PAES) algorithm. PAES is the last algorithm to be incorporated in the proposed optimisation framework. The section starts with and overview of this algorithm and discusses its main steps. PAES developed in 1999 and is the youngest of the selected evolutionary algorithms.

D.5.1 Overview of the algorithm

According to its creators, PAES may represent the simplest possible non-trivial algorithm capable of generating diverse solutions in the Pareto optimal set (Knowles and Corne, 1999). The algorithm is identified as being a (1+1) evolution strategy, using local search but using a reference archive of previously found solutions in order to identify the approximate dominance ranking of the current and candidate solution vectors.

The PAES algorithm was developed with two main objectives in mind. The first of these was that the algorithm should be strictly confined to local search i.e. it should use a small change (mutation) operator only, and move from a current solution to a nearby neighbour. The second objective was that the algorithm should be a true Pareto optimiser, treating all non-dominated solutions as having equal value. However there are cases that in a pair of solutions neither one will dominate the other. This problem is overcome in PAES by maintaining an archive of previously found non-dominated solutions. The archive is used as a means of estimating the true dominance ranking of a pair of solutions. This makes PAES also an elitist algorithm. PAES consists of three main parts: (i) the candidate solution generator, (ii) the candidate solution acceptance function, and (iii) the Non-dominated-Solutions (NDS) archive. Viewed in this way, PAES represents the simplest non-trivial approach to a multi-objective local search procedure.

D.5.2 Main steps of the algorithm

PAES begins with the initialisation of a single chromosome (the current solution) which is then evaluated using the multi-objective function. A copy is made and a mutation operator is applied to the copy. This mutated copy is evaluated and forms the new candidate solution. The current and candidate solutions must then be compared. Acceptance is simple if one solution dominates the other but in the case where neither solution dominates, the new candidate solution is compared with a reference population of previously archived non-dominated solutions. If comparison to the population in the archive fails to favour one solution over the other, the tie is split to favour the solution which resides in the least crowded region of the space.

The archive serves two separate purposes. First, it stores and updates all of the non-dominated solutions (subject to diversity criteria) generated, ready for presentation at the end of a run. Second, during the run, it is used as an aid to the accurate selection between the current and candidate solution vectors by acting as an approximation to the current non-dominated front. The latter is what provides the selection pressure, always pushing the process to find better solutions. Without this process, the algorithm is unable to differentiate between good and bad solutions with the result that it wanders rather aimlessly about the search space. The archive has a maximum size which is set by the user to reflect the required number of final solutions desired. Each candidate solution generated which is not dominated by its parent (the current solution) is compared with each member of the comparison set. Candidates which dominate the comparison set are always accepted and archived. Candidates which are dominated by the comparison set are always rejected, while those which are non-dominated are accepted and/or archived based on the degree of crowding in their grid location.

To keep track of the degree of crowding in different regions of the solution space, a d-dimensional grid is used, where d is the number objectives in the problem. When each solution is generated its grid location is found using recursive subdivision and noted using a tree encoding. A map of the grid is also maintained, indicating for each grid location how many and which solutions in the archive currently reside there. When a candidate solution is in a position to join a full archive, it replaces one of the archived solutions with the highest grid-location count, so long as its own grid-location count is lower. This system is also used to select between the current and candidate solutions when the candidate is not dominated nor dominates any member in the archive. In this case the solution with the lower grid count is selected.

D.6 Expectations from the selected EMOAs

The previous sections presented the four evolutionary algorithms that are employed by the proposed optimisation framework for business processes. All the four algorithms are considered 'basic' EMOAs in the sense that their flow of control is essentially a pure evolutionary algorithm framework, while the differences between them amount to explorations of various different ways to do selection and populations maintenance in multi-objective spaces (Corne et al., 2001).

PAES represents the simplest possible, non-trivial Pareto multi-objective optimiser, and should thus serve the purpose of a good baseline algorithm against which others may be compared. It is suggested that it may also serve well in real-world applications when local search seems superior to or competitive with population-based methods (Knowles and Corne, 1999). Recent results indicate that PAES is able to generate a diverse set of good solutions and it does so in significantly less time. These two advantages can prove crucial in the proposed optimisation framework for two reasons:

1. The search space is consisted of a number of islands (each corresponding to a business process design with different size) thus has multiple 'local' fronts (one in each island).
2. Producing a series of optimised business process designs in a timely fashion could be an additional strength of the proposed optimisation framework taking into account the complexity of the problem.

NSGA2 is known not to perform well in problems with multiple local fronts (Tiwari, 2001). The fitness assignment strategy of NSGA2 ceases to produce the driving force towards the global front once most of the solutions of the population share the shame non-domination level. This is further augmented due to the use of elitism and NSGA2 suffers from the tendency of getting trapped in local fronts (pre-mature convergence). However, NSGA2 is expected to provide diverse results to the business process optimisation problem. The diversity among non-dominated solutions is introduced in NSGA2 by using the crowded comparison operator that is used in the tournament selection and during the population reproduction phase. The crowded comparison operator states that between two solutions with different non-domination ranks, the point with the lower rank is preferred. Otherwise, if both the points belong to the same front, then the point that is located in a region with lesser number of points is preferred. In this way, the crowded comparison operator guides the selection process at various stages of the algorithm towards a uniformly spread-out Pareto front.

The region-based selection mechanism may be a key factor for PESA2 to outperform the other three EMOAs. Dividing the objective space in hyper-boxes (for multiple objectives) or squares (for two objectives) creates what is called the 'squeeze factor'. PESA2 uses this 'squeeze factor' both in selection and in archive update of solutions. If we assume that the algorithm will accurately create at least one hyper-box for each island (region) in the search space, then PESA2 will be capable of locating optimal solutions in most islands of the search space.

SPEA2 uses a novel selection strategy in which a 'strength' is associated with each member of the archive. The 'strength' of a solution is based on the number of solutions in the internal population which it dominates. This method relies on having population members around which are not in the current approximation of the Pareto front. Selection is biased towards minimising the strength of the solution thus preferring the exploration of less populated regions of the objective space. Taking into account that in the business process optimisation problem the search space is discrete and fragmented, it is expected that SPEA2 with its strength selection mechanism will demonstrate flexibility in converging to optimal solutions across the search space.

Finally, all the four evolutionary algorithms that are employed by the proposed optimisation framework are *elitist*. Elitism ensures that the search is driven towards the global Pareto front. The elitism approach of NSGA2 is through a selection operator that creates a mating pool by combining child and parent populations, and selecting the best (with respect to fitness and spread) N solutions. In SPEA2, PESA2 and PAES elitism is present through an archive of non-dominated solutions. This elitism ensures that the 'good' solutions of the population are not lost, thereby creating a selection pressure towards the global Pareto front.

D.7 Summary

This Appendix discussed the four EMOAs employed by the business process optimisation framework (bpo^F): NSGA2, SPEA2, PESA2 and PAES. For each algorithm and overview and its main steps are provided. The Appendix concluded with a brief discussion on the expectations of the performance of each algorithm on the context of business process optimisation.

APPENDIX E

Supplement on Real-life Scenarios

This appendix is a supplement to chapter 9 as it provides information about the real-life scenarios. It demonstrates the questionnaire that was used in the workshops (E.1). Also, it lists the on-line resources of web services (E.2) and the task libraries for scenarios A (E.3), B (E.4) and C (E.5).

E.1 Workshop questionnaire

PARTICIPANT INFORMATION:

Name: …………………………………………….

Email: ……………………………………………..

Organisation: …………………………………

The information provided will only be used for academic and research purposes.
If you agree please tick the box: □

AIM of the WORKSHOP:

To demonstrate the working and the benefits of a proposed Business Process Optimisation framework using a real-life business process scenario.

INSTRUCTIONS:

1. Please complete the following exercises and questions using the information related to the *Sales forecasting* business process
2. For the **process composition** exercises and questions, you need to create a business process design with the *requested number* of participating tasks ensuring that:
    a. All *process inputs* are utilised in the beginning of the design ,
    b. All the *process outputs* are produced in the end,
    c. Each task is connected with another based at least on a common resource or it is connected directly with the process input or output resources,
    d. Include process patterns (such as AND, OR, etc.) when a task receives inputs from more than one tasks.
3. For the **process optimisation** exercises and questions, you need to calculate the optimisation objectives *Service Delivery Price* (SDP) and *Service Fulfilment Target* (SFT).
    a. For a process design, the objectives are calculated by summing the SDP and SFT values of all the participating web services,
    b. The proposed framework attempts to minimise SDP while maximising SFT.
4. After the exercises, please answer all the questions provideE.

PROCESS COMPOSITION - EXERCISES & QUESTIONS

- **Exercise 1:** Based on the given example process (provided in a separate sheet), compose a business process design with 5 participating tasks from the library.

– Please mark the duration of the exercise in minutes:
– Rate the exercise on a scale 1-5 based on how challenging it is to come up with a process design (1 – easy, 5 – difficult):

- **Exercise 2:** Based on the example process (provided in a separate sheet), compose a business process design with 4 participating tasks from the library.

– Please mark the duration of the exercise in minutes:
– Rate the exercise on a scale 1-5 based on how challenging it is to come up with a process design (1 – easy, 5 – difficult):

---

1. *How efficient do you consider the proposed algorithmic composition approach based on the time it takes to create business process designs?*

○ Very efficient / fast

○ Satisfactory / Average

○ Inefficient / slow

○ I don't know / I am not sure

---

2. *Based on the steps of the algorithm and the demonstrated results do you consider the algorithm beneficial in terms of variety of results?*

○ Yes, the algorithm produces results that can be overlooked by a human designer

○ Satisfactory / Average variety of results

○ No, the number of alternatives and size of the process can be managed by a human designer

○ I don't know / I am not sure

---

3. *Do you consider the proposed algorithmic composition approach a significant shift from the current practice in web services composition?*

○ Yes

○ No, why? (*optional*) ……………………………………………...

---

4. *Any further remarks / observations on the proposed algorithmic composition of business process designs?*

…………………………………………………………………………………………………

PROCESS OPTIMISATION - EXERCISES & QUESTIONS

- **Exercise 3:** For the designs that you have created, calculate the SDP and SFT (optimisation objectives)

- **Exercise 4:** Create a business process design with 5 participating tasks and minimum SDP value.
  - Please mark the duration of the exercise in minutes:
  - Rate the exercise on a scale 1-5 based on how challenging it is to come up with a process design that has min-SDP (1 – easy, 5 – difficult):
  - If the exercise involved the task to also maximise the SFT objective how challenging would it be in your opinion (1 – easy, 5 – difficult)?

---

5. *In exercise 4, if there was no fixed number of tasks in the process design, how it would affect your preference regarding the number of participating web services?*
   …………………………………………………………………………………………
   …………………………………………………………………………………………

---

6. *Based on the results demonstration, how do you assess the capability of producing from 50 up to 500 optimised alternative designs?*
   - ○ Plenty / Abundant number of alternative designs
   - ○ Satisfactory / Average number of alternative designs
   - ○ Poor / Low number of alternative designs
   - ○ I don't know / I am not sure

---

7. *Based on the research results, the optimisation algorithms are capable of locating designs with optimal objective values (SDP, SFT). How do you compare this capability with the manual approach of exercise 4?*
   - ○ Only optimisation algorithms can ensure optimal results in a timely fashion
   - ○ Human designers can create optimal designs / little contribution of the algorithms
   - ○ I don't know / I am not sure

---

8. *The proposed multi-objective optimisation approach lies at the heart of this research. Based on your experience, is it a capability that can significantly benefit business processes?*
   - ○ Yes, a formal business process improvement approach is much sought after
   - ○ Not really.
   - ○ I don't know / I am not sure

---

9. *Any further remarks / observations on the proposed optimisation capability of business process designs?*…………………………………………………………………………

GENERAL QUESTIONS

10. *Are you familiar with the concept of web services?*

- ○ Yes, web services are current practice within my organisation
- ○ Yes, I am familiar with the concept
- ○ No, I have not come across it before

11. *Assuming that you are familiar with the concept, do you consider the proposed approach (business process composition and optimisation using web services) as reasonable / feasible?*
    - ○ Yes, web services are the future for business process composition over a network
    - ○ No, business processes should be traditionally implemented within an organisation
    - ○ I don't know / I am not sure

FINAL REMARKS / COMMENTS

…………………………………………………………………………………………………

E.2 On-line resources of web services

Table E.1 provides the list with the on-line libraries of web services that were used to compile the task libraries for each of the three real-life business process scenarios.

1. www.wsfinder.com This website is a community effort to create a list of all APIs and web services that are publicly available online. The web services are organised in categories such as advertising, finance, product search, web search, etc. In total there are 47 different categories, accommodating 233 web services (as of March 2008).
2. www.xmethods.net Provides a list of submitted web services. There are not organised but listed with the most recently submitted appearing at the top. The site has around 200 web services (March 2008).
3. www.wsindex.org/Web_Services/index.html This website provides information and links to web services. It has a search facility and the web services are organised into distinctive categories. It accommodates information for 75 different web services.
4. www.webservicelist.com This directory provides web services that can be search either by category, alphabetically or with keywords. It list most of the publicly available web services.
5. splice.xignite.com Xignite specialises in financial web services. Their prototype software tool, Splice, accommodates a library of 644 web services either implemented by Xignite or by a third party.

Table E.1. On-line resources of web services

E.3 Task library for Scenario A

Task 0
Name: Achworks Soap (T$$ - Rico Pamplona)
Input(s): Customer account details (1) and Order details (2)
Output(s): Payment details (3)
Provider (source): rpamplona (www.achworks.com)
Description: Web Services for ACH Processing and Payments
Task 1
Name: BAX Global Tracking Service
Input(s): Order details (2) and Payment details (3)
Output(s): Order tracking status (5)
Provider (source): BAX_Global (www.baxglobal.com)
Description: Retrieve shipment tracking information
Task 2
Name: CDYNE Death Index
Input(s): Customer account details (1) and Payment details (3)
Output(s): Payment confirmation (4)
Provider (source): CDYNE
Description: The CDYNE Death Index (CDI) Web service is used by leading government, financial, investigative, credit reporting organizations, medical research and other industries to verify identity as well as to prevent fraud and comply with the USA Patriot Act. The CDI is an effective weapon against financial fraud and other forms of terrorism, completely on the Internet and in real-time.
Task 3
Name: Credit Card Processor
Input(s): Customer account details (1) and Order details (2)
Output(s): Payment details (3) and Payment confirmation (4)
Provider (source): Payment Resources International
Description: Authorise, credit and void credit card transactions.
Task 4
Name: D&B Business Credit Quick Check
Input(s): Customer account details (1) and Payment details (3)
Output(s): Payment confirmation (4)
Provider (source): D&D (www.strikeiron.com)
Description: Perform low risk credit assessments and pre-screen prospects with D&B's core credit evaluation data. Information includes identification, payment activity summary, public filings indicators, and the D&B® Rating.
Task 5
Name: Drupal authentication
Input(s): Customer account credentials (0)
Output(s): Customer account details (1)
Provider (source): Drupal
Description: Distributed authentication in every site
Task 6
Name: ecommStats Web Analytics
Input(s): Website tracking request (6)
Output(s): Website statistics (7)

Provider (source): ecommStats (www.ecommstats.com) **Description:**
Task 7
Name: Entrust login **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** Entrust (www.entrust.com) **Description:** Entrust develops Internet security services that provide identification, entitlement, verification, privacy, and security management capabilities.
Task 8
Name: Fed Ex Tracker **Input(s):** Order details (2) and Payment details (3) **Output(s):** Order tracking status (5) **Provider (source):** FedEx **Description:** Ship, manage and track your FedEx packages
Task 9
Name: FedEx / UPS Package Tracking **Input(s):** Order details (2) and Payment details (3) **Output(s):** Order tracking status (5) **Provider (source):** FedEx **Description:** Ship, manage and track your FedEx and/or UPS packages
Task 10
Name: FraudLabs Credit Card Fraud Detection **Input(s):** Customer account details (1) and Payment details (3) **Output(s):** Payment confirmation (4) **Provider (source):** FraudLabs **Description:** The FraudLabs Credit Card Fraud Detection Web Service is a hosted, programmable XML Web Service that allows instant detection of fraudulent online credit card order transactions. The FraudLabs Credit Card Fraud Detection Web Service helps the Internet merchant avoid lost revenue, wasted productivity, and increased operation costs in chargeback and higher reserved funds as a result of online frauds.
Task 11
Name: Google Analytics **Input(s):** Website tracking request (6) **Output(s):** Website statistics (7) **Provider (source):** Google (www.google.com/analytics/) **Description:**
Task 12
Name: Google Checkout **Input(s):** Customer account details (1) and Order details (2) **Output(s):** Payment details (3) **Provider (source):** Google (checkout.google.com) **Description:**
Task 13
Name: GUID Generator **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1)

Provider (source): GUID **Description:** Generator for unique user identification
Task 14
Name: Internet Payment Systems **Input(s):** Customer account details (1) and Order details (2) **Output(s):** Payment details (3) **Provider (source):** Internet Payment Systems **Description:** Online internet payment gateways, payment online, credit card Internet shopping.
Task 15
Name: LID login **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** NetMesh **Description:** Multiple implementations and hosted service available
Task 16
Name: OpenID login **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** OpenID **Description:**
Task 17
Name: Paypal online payment **Input(s):** Customer account details (1) and Order details (2) **Output(s):** Payment details (3) **Provider (source):** Paypal.com **Description:**
Task 18
Name: Real Time Check Verification (T$$ - Rico Pamplona) **Input(s):** Payment details (3) **Output(s):** Payment confirmation (4) **Provider (source):** rpamplona (www.achworks.com) **Description:**
Task 19
Name: Rich Payments NET **Input(s):** Order details (2) **Output(s):** Payment details (3) and Payment confirmation (4) **Provider (source):** RichSolutions (richsolutions.com) **Description:** e-Payment Web Service that supports credit cards, debit cards and check services.
Task 20
Name: SAINTlogin users validation **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** SAINTlogin (www.saintlogin.com) **Description:**
Task 21
Name: Servicetrack

Input(s): Website tracking request (6) **Output(s):** Website statistics (7) **Provider (source):** www.bindingpoint.com/servicetrack/ **Description:** Powerful software solution which adds essential logging, analysis, monitoring, reporting, and firewall abilities to operational web services.
Task 22
Name: Smartpayments Payment **Input(s):** Order details (2) **Output(s):** Payment details (3) **Provider (source):** richsolutions.com **Description:** Payment Web Service that supports credit cards, debit cards and check services.
Task 23
Name: Smartpayments CardValidator **Input(s):** Payment details (3) **Output(s):** Payment confirmation (4) **Provider (source):** richsolutions.com **Description:** Credit card validation and card type Web Service
Task 24
Name: StrikeIron Global Address Verification **Input(s):** Customer account details (1) and Payment details (3) **Output(s):** Payment confirmation (4) **Provider (source):** StrikeIron **Description:** The StrikeIron Global Address Verification Web Service instantly validates and enhances addresses from over 240 countries.
Task 25
Name: SXIP login **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** SXIP **Description:** Commercial Identity Provider
Task 26
Name: Typekey authentication service **Input(s):** Customer account credentials (0) **Output(s):** Customer account details (1) **Provider (source):** six apart (www.sixapart.com) **Description:** TypeKey is an authentication service that allows distributed applications to handle log-ins in a simple and secure way, so that users only need one login across many sites.
Task 27
Name: UPS Tracking **Input(s):** Order details (2) and Payment details (3) **Output(s):** Order tracking status (5) **Provider (source):** UPS **Description:** Ship, manage and track your UPS packages
Task 28
Name: VeriSign Payment **Input(s):** Customer account details (1) and Order details (2) **Output(s):** Payment details (3)

Provider (source): VeriSign Inc. (www.verisign.com)
Description: To help Internet merchants process a broad range of Web-based payment types (including credit and debit cards) for B2B and B2C e-commerce, VeriSign has created this service for sending payment requests and responses through financial networks.

E.4 Task library for Scenario B

Task 0
Name: D&B Business Verification **Input(s):** Company name (3) and Business query (1) **Output(s):** Business details (0) and Financial data (5) **Provider (source):** D&B (www.strikeiron.com) **Description:** Provides programmatic access to D&B's many business reports, including rich business and credit information reports that can greatly enrich any business intelligence initiative.
Task 1
Name: Fax.com **Input(s):** Time-series forecast (8) and Chart / graph (2) **Output(s):** Fax (4) **Provider (source):** www.fax.com **Description:** On-line faxing service
Task 2
Name: Gale Group Business Information **Input(s):** Company name (3) and Business details (0) **Output(s):** Financial data (5) **Provider (source):** Gale Group **Description:** Based upon a company name and address, this service will provide standard financial and corporate information for approximately 440,000 international private and public companies.
Task 3
Name: Gale Group Business Intelligence **Input(s):** Company name (3) and Business query (1) **Output(s):** Business details (0) and Financial data (5) **Provider (source):** Gale Group **Description:** This service will provide in-depth financial and corporate information such as revenue history, key executive contact information, product information, and a broad range of additional data for approximately 440,000 international private and public companies when queried by company name.
Task 4
Name: GraphMagic's Graph & Chart Web Service API **Input(s):** Financial data (5) and Time-series forecast (8) **Output(s):** Chart / graph (2) **Provider (source):** GraphMagic **Description:** This web service generates charts with the choice of language and platform. No need to worry about server load because images are generated on our server, and you get seamless free upgrades.
Task 5
Name: interfax.net **Input(s):** Company name (3), Business details (0), Financial data (5) Time-series forecast (8) and Chart /

graph (2) **Output(s):** Fax (4) **Provider (source):** www.interfax.net **Description:** On-line faxing service
Task 6
Name: Lokad Business time-series forecasting and analysis **Input(s):** Financial data (5) and Recent market trends (7) **Output(s):** Time-series forecast (8) **Provider (source):** Lokad **Description:** Business time-series forecasting and analysis. Time series forecasting is the activity of applying statistical models to financial data such as sales or demand to generate a forecast for the future.
Task 7
Name: Midnight Trader Financial News **Input(s):** Market update request (6) and Company name (3) **Output(s):** Recent market trends (7) **Provider (source):** Midnight Trader **Description:** Determine how a stock is likely to react to published news events (i.e., earnings, analyst upgrades and downgrades, etc.) in the future based on how it has reacted to similar events in the past.
Task 8
Name: StrikeIron Company Search **Input(s):** Company name (3) **Output(s):** Company name (3) and Business details (0) **Provider (source):** StrikeIron **Description:** Perform Company Search for a particular Company Name
Task 9
Name: StrikeIron Get Business Prospect **Input(s):** Company name (3) and Business query (1) **Output(s):** Business details (0) and Financial data (5) **Provider (source):** StrikeIron **Description:** Get basic business prospect information, including DUNS Number, Address, Telephone Number, CEO Name, Line Of Business, SIC Code, Year Started, Annual Sales, Number of Employees, Business Structure, and Corporate Family Relationships for the business you are interested in.
Task 10
Name: StrikeIron Lookup Business **Input(s):** Company name (3) **Output(s):** Company name (3) and Business details (0) **Provider (source):** StrikeIron **Description:** Perform Company Search for a particular Company Name
Task 11
Name: Wall Street Horizon Real-Time Company Earnings **Input(s):** Company name (3) and Business query (1) **Output(s):** Financial data (5) **Provider (source):** Wall Street Horizon **Description:** Access comprehensive earning calendars providing earnings announcements, investor conference call dates and times, dividend announcements, and split announcements to track and evaluate over 5,000 U.S. companies.

Task 12
Name: Xignite Get Balance Sheet **Input(s):** Company name (3) and Business query (1) **Output(s):** Financial data (5) **Provider (source):** Xignite **Description:** Returns a company's balance sheet.
Task 13
Name: Xignite Get Chart **Input(s):** Time-series forecast (8) **Output(s):** Chart / graph (2) **Provider (source):** Xignite **Description:** Returns a chart for a time-series data. The chart is returned as a url.
Task 14
Name: Xignite Get Chart Preset **Input(s):** Time-series forecast (8) **Output(s):** Chart / graph (2) **Provider (source):** Xignite **Description:** Returns a preset chart for a time-series data. The chart is returned as a url.
Task 15
Name: Xignite Get Growth Probability **Input(s):** Financial data (5) and Recent market trends (7) **Output(s):** Time-series forecast (8) **Provider (source):** Xignite **Description:** Returns the probability for a stock growth, price, or market value to reach a certain value based using a Monte Carlo simulation.
Task 16
Name: Xignite Get Market News Headlines **Input(s):** Market update request (6) **Output(s):** Recent market trends (7) **Provider (source):** Xignite **Description:** Get most recent market news headlines from Reuters.
Task 17
Name: Xignite Get Market Summary **Input(s):** Market update request (6) **Output(s):** Recent market trends (7) **Provider (source):** Xignite **Description:** Returns the current market level for the Dow, Nasdaq and S&P indices as well as the NYSE and NASDAQ volumes and the 10 Year Bond index.
Task 18
Name: Xignite Get Topic Data **Input(s):** Company name (3) Financial data (5) and Recent market trends (7) **Output(s):** Time-series forecast (8) **Provider (source):** Xignite **Description:** Returns time-series data for a topic.
Task 19
Name: Xignite **Get Topic Chart** **Input(s):** Company name (3) Financial data (5) and Recent market trends (7)

Output(s): Time-series forecast (8) and Chart / graph (2)
Provider (source): Xignite
Description: Returns detailed data and a chart for a time-series data.

E.5 Task library for Scenario C

Task 0

Name: Address Doctor Global Address Verification
Input(s): Customer ID details (2)
Output(s): ID verification outcome (3)
Provider (source): StrikeIron
Description: The Address Doctor Global Address Verification Web Service instantly validates and enhances addresses from over 240 countries. Simply provide basic (even incorrect) address information and it will be validated with accurate data. The advanced features provide additional address formatting options like specifying country of origin, preferred language, capitalization and much more.

Task 1

Name: cbarron bankValidate
Input(s): Customer credit details (1)
Output(s): Credit assessment (0)
Provider (source): cbarron (www.unifiedsoftware.co.uk)
Description: The bankValidate service checks bank sort codes against the latest BACS Industry Sort Code database. It then validates sort code / account number combinations using each banks own modulus checking rules.

Task 2

Name: CDYNE Death Index
Input(s): Customer ID details (2) and Customer credit details (1)
Output(s): ID verification outcome (3) and Credit assessment (0)
Provider (source): CDYNE
Description: The CDYNE Death Index (CDI) Web service is used by leading government, financial, investigative, credit reporting organizations, medical research and other industries to verify identity as well as to prevent fraud and comply with the USA Patriot Act. The CDI is an effective weapon against financial fraud and other forms of terrorism, completely on the Internet and in real-time.

Task 3

Name: CDYNE Email Verifier
Input(s): Customer ID details (2)
Output(s): ID verification outcome (3)
Provider (source): CDYNE
Description: Email address hygiene plays a role in effective and efficient email delivery. The CDYNE Email Verifier (CEV) will check the validity of email addresses from a mailing list or in real-time as a Web service. CEV will verify 80-90% of invalid mail addresses.

Task 4

Name: CDYNE Phone Verifier
Input(s): Customer ID details (2)
Output(s): ID verification outcome (3)
Provider (source): CDYNE
Description: The Phone Verifier identifies the phone numbers in your list that have new area codes

following a split and replaces incorrect area codes. The Phone Verifier will reduce data entry errors in Batch or Real-time mode. The Web service does not check the last 4 digits of the phone number.
Task 5
Name: D&B Business Credit Quick Check **Input(s):** Customer credit details (1) **Output(s):** Credit assessment (0) **Provider (source):** D&D (www.strikeiron.com) **Description:** Perform low risk credit assessments and pre-screen prospects with D&B's core credit evaluation data. Information includes identification, payment activity summary, public filings indicators, and the D&B® Rating.
Task 6
Name: D&B Business Verification **Input(s):** Customer ID details (2) and Customer credit details (1) **Output(s):** ID verification outcome (3) and Credit assessment (0) **Provider (source):** D&B (www.strikeiron.com) **Description:** Provides programmatic access to D&B's many business reports, including rich business and credit information reports that can greatly enrich any business intelligence initiative.
Task 7
Name: Dimple Email Address Validator **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron (Dimple Software) **Description:** Dimple Software's Email Address Validator Web Service is one of the most powerful, robust Email Address and MX Validator Web Service available on the market today for .NET.
Task 8
Name: Drupal authentication **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** Drupal **Description:** Distributed authentication in every site
Task 9
Name: Dun & Bradstreet Business Credit Quick Check **Input(s):** Customer credit details (1) **Output(s):** Credit assessment (0) **Provider (source): StrikeIron** **Description:** Perform low risk credit assessments on-demand with D&B core credit evaluation data. Information includes company identification, payment activity summary, public filings indicators and the D&B Rating.
Task 10
Name: Dun & Bradstreet Business Verification **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** Verify a business identity and its location. Validate a company and its location with background information such as primary name, address, phone number, SIC code, branch indicator and D&B D-U-N-S Number.
Task 11

Name: Entrust login **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** Entrust (www.entrust.com) **Description:** Entrust develops Internet security services that provide identification, entitlement, verification, privacy, and security management capabilities.
Task 12
Name: FraudLabs Credit Card Fraud Detection **Input(s):** Customer ID details (2) and Customer credit details (1) **Output(s):** Credit assessment (0) **Provider (source):** FraudLabs **Description:** The FraudLabs Credit Card Fraud Detection Web Service is a hosted, programmable XML Web Service that allows instant detection of fraudulent online credit card order transactions. The FraudLabs Credit Card Fraud Detection Web Service helps the Internet merchant avoid lost revenue, wasted productivity, and increased operation costs in chargeback and higher reserved funds as a result of online frauds.
Task 13
Name: Google Docs **Input(s):** ID verification outcome (3) and Credit assessment (0) **Output(s):** Risk assessment report (4) **Provider (source):** Google (docs.google.com) **Description:** Free web-based word processor and spreadsheet, which allow you share and collaborate online.
Task 14
Name: GUID Generator **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** GUID **Description:** Generator for unique user identification
Task 15
Name: LID login **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** NetMesh **Description:** Multiple implementations and hosted service available
Task 16
Name: OpenID login **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** OpenID **Description:**
Task 17
Name: Real Time Check Verification (T$$ - Rico Pamplona) **Input(s):** Customer credit details (1) **Output(s):** Credit assessment (0) **Provider (source):** rpamplona (www.achworks.com) **Description:** Web Services for Online Verification of Bank Accounts (ACH) .T$$ Check Verification is a

Web Service for verifying an account for use in electronic fund transfer (EFT) transactions. The service allows you to submit a single set of data (amount, account type, account number and ABA routing number) and respond to tell you if the bank account whether it is existing, open/valid account, closed, in a bad list, fraudulent, a member of ACH participating banks, with available balance, etc.
Task 18
Name: SAINTlogin users validation **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** SAINTlogin (www.saintlogin.com) **Description:**
Task 19
Name: Smartpayments CardValidator **Input(s):** Customer credit details (1) **Output(s):** Credit assessment (0) **Provider (source):** richsolutions.com **Description:** Credit card validation and card type Web Service
Task 20
Name: StrikeIron 24-hour Accurate Residential Lookup **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** This Web service looks up and validates information on any residential phone number or address with 24-hour accuracy.
Task 21
Name: StrikeIron 24-hour Accurate Reverse Phone Lookup **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** The StrikeIron 24-hour Accurate Reverse Phone Lookup Web Service provides a programmatic interface to name and address data associated to any telephone number, including residential, business,. This data is updated nightly making them the most accurate and up to date resource of their kinE.
Task 22
Name: StrikeIron Email Verification **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** Indicates whether or not an email address actually exists or not, without actually ever sending an email message. For any email address, the Web service will simply return a true or false as to whether or not the email address is valid or not. This goes far beyond the structure of the email address, actually identifying bad email addresses and non-existent domains.
Task 23
Name: StrikeIron Gender Determination **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** The StrikeIron Gender Determination Web Service takes any contact record and, using a

name analysis database, provides the gender of almost any full name.
Task 24
Name: StrikeIron Global Address Verification **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** The StrikeIron Global Address Verification Web Service instantly validates and enhances addresses from over 240 countries.
Task 25
Name: StrikeIron Reverse Phone Residential Intel **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** Based on a residential phone number, this Web service appends personal information such as name, address, head of household, dwelling type, length of residency, homeowner probability, and more. Updated files are processed within a 24–hour period or less to ensure the quality of the appended phone number.
Task 26
Name: StrikeIron Reverse Residential Lookup **Input(s):** Customer ID details (2) **Output(s):** ID verification outcome (3) **Provider (source):** StrikeIron **Description:** Based on a residential name, address and phone information, this Web service verifies and appends personal information such as name, address, head of household, dwelling type, length of residency, homeowner probability, new connection information and more. Updated files are processed within a 24–hour period or less to ensure the quality of the appended phone number.
Task 27
Name: SXIP login **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** SXIP **Description:** Commercial Identity Provider
Task 28
Name: Typekey authentication service **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** six apart (www.sixapart.com) **Description:** TypeKey is an authentication service that allows distributed applications to handle log-ins in a simple and secure way, so that users only need one login across many sites.
Task 29
Name: Web Services Security Monitor **Input(s):** Security login credentials (5) **Output(s):** Customer ID details (2) and Customer credit details (1) **Provider (source):** esynaps (www.esynaps.com) **Description:** Authentication and logging service for web services
Task 30
Name: webba E-Mail validator

Input(s): Customer ID details (2)	
Output(s): ID verification outcome (3)	
Provider (source): webba (www.wsdirect.net)	
Description: This web-service checks up both SMTP server and user existence.	