File ID	uvapub:122791
Filename	Thesis
Version	unknown

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	PhD thesis
Title	Discovering value leaks and service imperfections in business processes
Author(s)	E.M. Peters
Faculty	FEB: Amsterdam Business School Research Institute (ABS-RI)
Year	2013

FULL BIBLIOGRAPHIC DETAILS:
http://hdl.handle.net/11245/1.394343

Edward Michael Peters

DISCOVERING VALUE LEAKS AND SERVICE IMPERFECTIONS IN BUSINESS PROCESSES

May 23, 2013

UNIVERSITY OF AMSTERDAM

Discovering Value Leaks and Service Imperfections in Business Processes

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Ámsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van de Boom
ten overstaan van een door het college
voor promoties ingestelde commissie.

Edward Michael Peters.

geboren te Bethlehem,
Verenigde Staten van Amerika

May 23, 2013

Discovering Value Leaks and Service Imperfections in Business Processes

Discovering Value Leaks and Service Imperfections in Business Processes

ACADEMISCH PROEFSCHRIFT

**ter verkrijging van de graad van doctor
aan de Universiteit van Ámsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van de Boom
ten overstaan van een door het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapal
op donderdag 23 may 2013, te 10:00 uur**

door

Edward Michael Peters.

geboren te Bethlehem, Verenigde Staten van Amerika

Promotiecommissie

Promotor:	prof. dr. Guido Dedene
Co-promotor:	prof. dr. ir. Rik Maes
Overige leden:	prof. dr. Aimé Heene
	prof. dr. Hans Jägers
	prof. dr. Stijn Viaene
	dr. Jonas Poelmans

Faculteit der Economie en Bedrijfskunde

ACKNOWLEDGEMENTS

The undertaking of a PhD thesis is journey through the wilderness that is guided by someone constantly encouraging you to stay on course and move toward the light. In my case I was lucky to have as my guide Professor Dr. Guido Dedene, to whom I want to express my deep and heartfelt gratitude. Our collaboration on the topics of technology and productivity began in November, 1987, during a conference in Edinburgh, Scotland. While our paths have, over the years, taken different directions, they have always re-connected at critical junctures. However no time was more critical than in 2008, during the beginning of my tenure at OpenConnect, when the desire to solve perplexing problems met the need for scientific research. It was Professor Dr. Dedene who convinced me that I could indeed solve some of those problems if I was willing to undertake the journey. I would also like to thank my co-Promoter, Professor Dr. Ir. Rik Maes, and Professor Dr. Stijn Viaene who patiently listened to my ideas and provided useful feedback on how to approach the topic. I thank sincerely Prof. Dr. Hans Jägers for stressing the importance of finding the appropriate Research Methodology. Prof. Dr. Aimé Heene was linking our work to Information and Communication Strategies and are the augmentation to ICT-based Services. I want to acknowledge in particular the University of Amsterdam, Faculty of Business and Economics, for allowing me to develop this work, which resulted in this thesis.

Additionally, I would like to thank two of my recent fellow travelers and, at times, collaborators on the journey, Dr. Jonas Poelmans and Dr. Paul Ezinga for sharing insights on areas of Data Discovery, C-K Theory, FCA and how they could be applied to Process Discovery. Our discussions were always interesting and led to a number of joint publications. The three PhD-thesis works, Jonas at KU Leuven, and Paul and myself at the University of Amsterdam, can be seen as a “tryptique” on Discovery, exploring the exciting frontiers of contemporary Business and Process Intelligence.

I would also like to thank all of my colleagues at OpenConnect and specifically, Stuart Burris, Dr. Chris Houck, Alex Moulas, Eric Armstrong, Kip Brewer, Kevin Culliton, David Parkes, Mike Darrish and Rene Schoenauer for their tireless efforts to ensure the success of the technology and the customer implementations described in the case studies. Additionally, their assistance and feedback on sections of this thesis were invaluable. Also, I would like to thank my assistant, Carla St. Leon, for scheduling my time, arranging travel, coordinating conference calls and assisting in the final production of this thesis.

Some of the case studies would have been impossible without the collaboration of the people that were driving these projects. In particular I would like to express sincere thanks to Gerda Verheyden and Herman Mussele at the GZA cancer clinic in Antwerp, Belgium for the last case study in chapter 5 (which was conducted in collaboration with Jonas Poelmans). Finally, I thank Rick Warnar from DENION, the Netherlands, for his kind cooperation in the second case study of chapter 5 (in particular for his drive for the “time and motion” measurements).

Additionally, I would like to express my deepest appreciation to Deborah Burney for the encouragement to remain on the path, especially during the darkest times; your love means more to me than I can ever hope to fully express.

Finally, I want to thank the KU Leuven, who enabled the OpenConnect Chair in Business Process Discovery and Analytics, under whose auspices, much of this work was accomplished.

Edward Michael Peters, Jr.
Dallas, Texas. USA
March, 2013

Table of Contents

Samenvatting XIV
Chapter 1 1
Fundamentals of ICT-Services and Processes 1
1.0 Introduction 1
1.1 Understanding ICT-Service Work 1
1.1.1 Knowledge-work versus data-work 1
1.1.2 Augmenting Human Intellect 3
1.1.2.1 Concept Structuring 3
1.1.2.2 Symbol Structuring 4
1.1.2.3 Process Structuring 4
1.1.3 Transactions versus Interactions 4
1.2 Approaches to Improving ICT-Based Services and Processes 6
1.2.1 Lean Management 7
1.2.2 Six Sigma 8
1.2.3 Lean Six Sigma 10
1.2.4 Automated Business Discovery 11
1.3 Emergence and Enterprise 2.0 13
1.3.1 Value Creation 14
1.3.2 Value Leaks 14
1.4 Services as Building Blocks of Business Processes 18
1.5 Value Creation and Value Leaks for ICT-based Services 21
1.6 Service Bundle Imperfections 24
1.6.1 Incompleteness in Bundles 25
1.6.2 Asymmetries in Bundles 27
1.6.3 Asymmetry of Meaning in Services 29
1.7 Problem Statement and Research Methodology 30
1.7.1 Research Objectives 30
1.7.2 Research Framework 31
1.7.3 Research Questions 32
1.7.4 Research Approach 32
1.7.5 Research Relevance 34
1.8 Overview of the rest of the thesis 38
Chapter 2 40
Discovery Techniques for Processes 40
2.0 Process Discovery Techniques for Value Leak Assessments 41
2.1 C-K Theory for Process Discovery 42
2.2 The start of Process Discovery 45
2.3 Input data for Process Discovery 47
2.4 Formal Concept view for Process Discovery 50
2.5 Probabilistic Finite State Machines 55
2.6 BPMN and other Petri Net Models 57
2.7 Hidden Markov Models 62
2.8 Normalized Process Representations 66
2.9 A hybrid approach for Process Discovery 78

2.10 Summary....84
Chapter 3....85
The Tooling and Process for Process Discovery....85
3.0 Introduction....85
3.1 The Origins of *Comprehend*....85
3.2 Architecture....88
3.2.1 Collectors....89
3.2.1.1 3270 Collector....89
3.2.1.2 HTTP Collector....89
3.2.2 *Comprehend* Analytics Components (AC)....90
3.2.3 *Comprehend* AC User Interface....91
3.2.4 *Comprehend* Process Intelligence Component (PIC)....91
3.2.5 Reporting....91
3.2.6 Security Provisions....91
3.2.6.1 Data Storage Protocol....92
3.2.6.2 *Comprehend* Analytics Components (AC)....92
3.2.6.3 Audit....92
3.2.6.4 Usage Log....92
3.2.6.5 System Change Logs....92
3.2.7 Standard Rules....92
3.2.7.1 System Administrator....92
3.2.7.2 Data Security Administrator....93
3.2.7.3 Security / User Administrator....93
3.2.7.4 AC-Developer....93
3.2.7.5 PIC - Analyst....93
3.2.8 *Comprehend* Artifact Capture....93
3.2.8.1 Session Capture....94
3.2.8.2 Mainframe Session Capture....94
3.2.8.3 Client-Server/Desktop Session Capture....95
3.2.8.4 Web Session Capture....96
3.2.9 Activity Definition....96
3.2.10 Event Generation....97
3.2.11 Process Intelligence Components....98
3.2.11.1 PIC Process Perspective....98
3.2.11.2 PIC Process Integration and Filtering....100
3.2.12 *Comprehend* Reporting....100
3.2.13 Summary....100
3.3 The Process of Using *Comprehend*....100
3.3.1 Roles....102
3.3.1.1 *Comprehend* Analyst....103
3.3.1.2 Solution Architect....103
3.3.1.3 Configure Developer....104
3.3.1.4 Business Subject Matter Experts....104
3.3.1.5 Business Approval Committee....104
3.3.2 Discover....104
3.3.2.1 *Comprehend* Analysis Report....115

3.3.2.2 Business Approval and Direction 116
3.3.3 Design 116
3.3.4 Build 133
3.3.4.1 Build Transaction(s) 133
3.3.4.2 Unit Test 134
3.3.4.3 Integration Test with Web Application 134
3.3.4.4 System Test 135
3.3.4.5 User Acceptance Testing 135
3.3.5 Execute 135
3.3.5.1 Deploy Configure Project to Production Environment 135
3.3.5.2 Execute New Edit Code 136
3.3.5.3 Run Reports 136
3.3.6 Monitor and Measure 136
3.3.6.1 Monitor 136
3.3.6.2 Create Attributes 136
3.3.6.3 Configure and Monitor 137
3.3.6.4 Measure 137
3.3.6.5 Create Attributes 137
3.3.6.6 Configure and Measure 138
3.3.7 Summary 138
Chapter 4 139
Case Studies I: Administrative Process 139
4.0 Introduction 139
4.1 Overview of Healthcare Claims Payment Process 139
4.1.1 Value Creation and Value Leaks in Health Insurance Claim Processing 141
4.1.2 Healthcare Claims Adjustments 143
4.2 Healthcare Administrative Processing Case One: Claims Adjustments 143
4.2.1 Context for this Process Discovery Case Study 143
4.2.2 Initial Process Analysis and Improvement 145
4.2.3 Additional Value Leak Identification and Reduction 148
4.2.4 Value Leak Recapture Summary 151
4.3 Healthcare Administrative Process Case Two: Determining the real costs of claims processing with a Business Process Outsourcer (BPO) 151
4.3.1 Context of this Case Study 152
4.3.2 Business Processing Outsourcing: Contract Renewal 152
4.3.3 Workforce Reporting 153
4.3.4 Examiner Productivity 154
4.3.5 Process Analysis ToS (Time of Service) 154
4.3.6 Process Analysis – Claims Impact to the Business 156
4.3.7 Process Reporting – Claims Impact to the Business 157
4.3.8 *Comprehend* Automation 158
4.3.9 Value Leak Re-Capture Summary 159
4.4 Summary 159
Chapter 5 163
Case Studies II: Care Processes 163

5.0 Introduction....163
5.1 Customer Contact Center Case One: Multi-National Bank 163
5.1.1 Goals 163
5.1.2 Speed....163
5.1.3 Quality....163
5.1.4 Revenue (Customer Churn) 163
5.1.5 Approach....163
5.1.6 Project Methodology and Timeframe 165
5.1.7 Analysis Approach....165
5.1.8 Technical Environment....166
5.1.9 Example Results....167
5.1.9.1 Speed 167
5.1.9.2 Account Balance Link 167
5.1.10 Benefit....167
5.1.10.1 Agent Improvement....168
5.1.10.2 Quality 169
5.1.10.3 Change of Address 169
5.1.10.4 Self-Service Site Improvement....170
5.1.11 Revenue....171
5.1.11.1 Leading Indicator of Churn 171
5.1.11.2 Cross-Sell 171
5.1.12 Justification and Rollout 172
5.1.13 Value Summary 172
5.2 Customer Contact Center Case Two: Municipality Contact Center 174
5.2.1 Customer Value in Contact Centers....174
5.2.2 Understanding the Current Process....176
5.2.3 Process Discovery of the Actual Contact Process 181
5.2.4 Value Leaks in the Discovered Contact Process....185
5.2.5 Estimated Value Contribution of Process Discovery in this Case Study....190
5.3 Clinical Pathways for Breast Cancer 192
5.3.1 Context for the Process Discovery Case Study....192
5.3.2 A First Discovery Analysis Examining Quality of Care 196
5.3.3 A Second Discovery Analysis in Search of Process Variations 198
5.3.4 Discovering the "Natural" Pathways 201
5.3.5 Discussion....205
5.4 Summary....206
Chapter 6....209
Conclusion 209
6.0 Thesis Conclusion....209
6.1 Research Approach Debriefing 214
6.2 Future Research....215
BIBLIOGRAPHY 217

Preface

Improving process performance has always been of utmost importance in my professional life. From my first position as a supervisor of back-office lottery ticket processing and accounting to my current day efforts at improving process efficiency for large organizations through packaged software products, I have been focused on reducing waste and improving efficiency. At the same time, I have always been amazed at the number of decisions that are made, risking untold millions of Euros, US dollars, and other currencies, based on best guesses, media-driven industry myths or beliefs that are held without any supporting evidence. One hundred and one years ago, in 1911, Fredrick Winslow Taylor published his groundbreaking monograph; *Principles of Scientific Management (Taylor 1911).* Taylor pioneered the application of the scientific method to the area of work (described as human labor or skill) to improve what he saw as inefficient methods of accomplishing a task. His approach, which is sometimes called task management, was one of observation, recording of data about how something was being done, analysis (finding a hypothesis about the "one best way" to perform a task, testing (measuring), and then implementing and re-measuring to ensure that the result was obtained. Taylor focused on eliminating *guesswork,* what he called *rule of thumb*, and replaced it with systematic analysis of work tasks that could be optimized, thereby improving productivity. His analysis attacked the very notion of what was valued, that of a supposed difference in skills possessed by various laborers. Through observation and analysis, Taylor refuted the idea of "skill" and showed that the efforts did not contain any specific knowledge but rather could be seen as a series of repetitive motions, that when analyzed, could yield a "one best way" to perform the task. That best way, when replicated by many workers not only produced increased output per worker, but also allowed for the task to be completed with fewer workers, thus reducing the overall labor components of the task.

His work was mostly performed for large manufacturing organizations (e.g., Bethlehem Steel) and often seen as an icon of the industrial era where the production for manufactured goods with little or no service component, was the primary output. Since then, the world has changed in ways Taylor could never have imagined; ICT-enabled processes have revolutionized both productivity and the role of labor in new and dramatic ways ranging from manufacturing to services. A central premise of this PhD Thesis is that just as in Fredrick Taylor's day, it is the systematic and scientific lack of knowledge about the current work process, what I refer to as *knowledge gaps,* and the intrinsic nature of service-based work itself, that must be re-examined in order to establish a new paradigm for productivity improvement.

While "observations" about the current environment remain central to the establishment of a baseline about the current work reality, making those observations in an ICT-enabled work process, where the actions cannot readily be seen, requires a new approach than those in common practice. Specifically, current improvement methods either use manual observations about digital environments or capture digital data to build a generalized model of the work environment. The first instance relies on techniques that place the veracity of the data in question while the latter factors out “noise” or exceptions in order to gain an overall understanding. It is precisely these two issues that need to be addressed in order to capture information in an ICT-enabled, digital method as well as to have the correct mathematical techniques to most effectively and unambiguously discover both the source of and reason for loss of process value.

With my work at OpenConnect, my initial hypothesis was that since the actual observations were indeed being recorded and captured by the technology itself, if it could be stored and re-created, it would be an actual or fact-based representation of the actual activity rather than what amounted to a best guess. Also, the recorded sequence of actions would represent an actual flow of the process, not an idealized flow or something based on inference, speculation or opinion. In essence, the event logs generated could be used as the basis for reconstructing actual processes around applications.

As this effort was to be based on captured data, it would be a “bottom-up” approach. This would be different from the normal “top-down” modeling effort where the taxonomy is given at the start the process. I was clear the effort needed to be based on an “emergent” paradigm since the data would be at the lowest (atomic) levels and the groupings would need to be found or “discovered” rather than imposed at the start. At this point the work of Cook and Wolf (Cook and Wolf, 1995) became important. They had coined the term “process discovery” while working to uncover processes in the software development effort. They suggested three techniques that could be used; Neural Networks, algorithm-based techniques and Markov Chains. The use of Markov Chains was the best fit since it showed the probabilistic nature of the work process as well as the issues inherent in the model. Specifically, the model contained a level of ambiguity in truly understanding the root cause of the any deviation from the main path, (e.g., process anomalies). This was a key moment since it was now clear that it was more than the main process that had to be “discovered” but also, it was the exceptions (or the noise) that were of key interest. Just as noise leaks information from a communication channel, process variations (or noise) are the source of leaks from an ICT-enabled process. Understanding from field experience that it was the process exceptions (noise) that were the root cause of process productivity issues led to the focus on finding which mathematical techniques were most suited to understand and represent value leaks in business processes. This understanding would form the basis in providing actionable information to management on how to deal with these value leaks and improve process performance.

The structure of this PhD Thesis can be understood as follows; chapter 1 provides an introduction to the subject of process improvement and describes issues with its implementation in ICT-enabled environments. In chapter 2, mathematical techniques are explored. Additionally, the methods for unambiguously discovering and representing process variations are described. Chapter 3 discusses the methodology and tooling that implements the concepts described in chapters 1 and chapter 2. It describes the platform that will be used to collect, analyze and represent process and their variations. The technology was developed by OpenConnect Systems, Dallas, Texas, USA, and is named *Comprehend*. With one exception, it is used as the platform for the case studies described in chapter 4 and chapter 5. Here it is interesting to note that two different environments for case study research were chosen. Specifically, chapter 4 addresses productivity issues in administrative environments where the activities are highly transactional in nature whereas chapter 5 addresses situations that are highly interactional as well as transactional. In all cases, discovering the true nature of value leaks led increased process perform7ance. In chapter 6 conclusions are discussed as well as areas for future research.

Overall, while I learned much through the researching and writing of this PhD Thesis, I am continually humbled when I think of how much there is yet to be done and, in that context, how little I actually know. It would be egotistical to think this is even a first step. At best, I hope it has cracked open a small window into an area where further research, and other researchers, can shed light into an area of darkness. I feel there is much to be done and sincerely hope others may find this work useful in their efforts to improve businesses processes as we move into the post-industrial, ICT-enabled service dominated economy.

Samenvatting

"Discovery" betekent letterlijk "ontdekken", waarbij geen a-priori structuur vooropgesteld wordt. Dit is een uitbreiding van de "mining" benadering waarin diverse statistische modellen worden gecombineerd om structuur te vinden in gegevens, die op zich ook geen a-priori structuur hebben. In deze thesis worden technieken samengebracht, vorm gegeven en geïllustreerd aan de hand van relevante gevallenstudies teneinde een beeld te krijgen van Discovery voor processen, en in het bijzonder voor processen die gebruik maken van ICT-gebaseerde diensten (zoals het geval is met meer dan 50% van de economische activiteit op dit ogenblik). Het uiteindelijk doel is een bijdrage te leveren tot het verbeteren van deze sort processen. Uiteraard bestaan industriële processen reeds meer dan een eeuw, en zijn talloze benaderingen ontwikkeld voor het verbeteren van deze – vaak materiaalgerelateerde – processen. De vraag rijst terecht hoeveel van deze technieken relevant blijven in de context van processen die gebruik maken van ICT-gebaseerde diensten als processtappen. Precies op dit vlak wil deze thesis een toegevoegde waarde leveren. Vertrekkend vanuit het strategische begrip van "waarde-lek" wordt aangetoond hoe welbepaalde procesvoorstellingen (op basis van Hidden Markov Modellen en Formele Concept Analyse) toelaten om potentiële waardelekken te ontdekken. De benadering maakt mee onderdeel uit van het vakgebied van de Process Intelligence. Aldus wordt het werk ook ingebed in de algemene benadering rond Intelligence die vooropgesteld in door Engelbart in de zestiger jaren, met gebruik van de C-K-methodologie als leidraad voor het innovatiegebeuren dat plaatsvindt bij de ontwikkeling van Discovery technieken. Uiteindelijk, en dat blijkt ook uit de gevallenstudies, weet men a priori niet wat men zal ontmoeten als mogelijke verbeteringselementen in processen (de fameuze managerial "unknown unknowns"). Discovery lijkt op dat vlak ook een bijzondere bijdrage te leveren.

De thesis vertrekt van de basisdefinities en begrippen die gekend zijn rond industriële processen en hun verbetering, waarbij het werk van Taylor het vertrekpunt was, en recentelijk Lean en Six Sigma een prominente plaats hebben verworven. In het eerste hoofdstuk wordt een nauwkeurige definnitie ontwikkeld voor de ICT-gebaseerde stappen in eigentijdse processen: deze stappen zijn betekenisvolle bundels van Informatie & Communicatie Technologie, Middelen en Materialen die samen een processtap onder bepaalde randvoorwaarden tot stand brengen. De stappen zijn de ICT-gebaseerde services, en de randvoorwaarden zijn de fameuze Service Level Agreements. Bij de processen wordt – in voorbereiding van de soorten gevallenstudies in deze thesis – een onderscheid gemaakt tussen transactionele (administratief verwerkende) processen, die vaak rechtlijnig en weinig interactief zijn, in tegenstelling met interactie (verzorgende) processen waarbij interactiviteit een grotere rol speelt. De laatste groep van processen wordt steeds belangrijker naarmate het impact van ICT 2.0 toeneemt. Vervolgens wordt het begrip "waardelek" ontwikkeld en gerelateerd aan imperfecties in ICT-gebaseerde diensten. In uitbreiding van de gekende imperfecties voor Informatie en Communicatie worden hier Onvolledigheid en Asymmetrie uitgewerkt als imperfecties voor ICT-gebaseerde diensten.

Daarmee wordt da nook de probleemstelling voor de thesis duidelijker: Process Discovery word teen zoektocht naar mogelijke waardelekken in ICT-gebaseerde processen, waarbij de waardelekken vaak overeenkomen met imperfecties in de ICT-gebaseerde processtappen. Het eerste hoofdstuk sluit af met de positionering van de onderzoeksbenadering die aangewend is, en grotendeels gefundeerd is op gevallenstudies. De benadering is holistisch, omdat niet alleen op één soort parameter (zoals kost, of tijd) de nadruk ligt. Ondermeer in de laatste gevallenstudie komen diverse patientgerelateerde waardeparameters aan bod. De relevantie van het onderzoek mag blijken uit het spectrum van publicaties en lezingen die reeds voorafgingen aan de publicatie van deze thesis.

De ontwikkeling van Intelligence toepassingen wordt in deze thesis gebaseerd op de C(oncept)-K(knowledge)-Theory die haar oorsprong vond in de Parijse Universiteiten in hun zoektocht naar methoden voor Innovatie in industriële processen. Vooreerst wordt in deze thesis deze theorie uitgebreid naar ICT-gebaseerde processen. Voor de disjunctiestap (de conceptualisering) speelt Process Discovery een centrale rol. In hoofdstuk 2 worden de voornaamste actuele procesvoorstellingen onderzocht op hun deugdzaamheid voor Process Discovery. Als vertrekpunt wordt een formaat vooropgesteld voor de invoergegevens dat comform is aan moderne standaards in de auditing: het Event – Actor – Object formaat. Vervolgens werd in de eerste plaats gekeken naar de Formele Concept Analyse FCA, die haar nut reeds uitgebreid had bewezen in voorafgaande thesiswerken rond Data Discovery (Poelmans, Elzinga). Er wordt aangetoond hoe FCA meer systematiek aanbreng in de ontdekking van hogere orde transities en actornetwerken rond processen. Het voornaamste onderzoekspunt van dit hoofdstuk is echter de bruikbaarheid van de traditionele – vaak Petri-Net gebaseerde – voorstellingen van processen, die actueel erg populair zijn. In dit hoofdstuk wordt uitgewerkt hoe deze voorstellingen – alhoewel ze vaak toereikend zijn voor een beleidsgerichte voorstelling van de processen – onvoldoende zijn om in het algemeen desefficienties en waardelekken in processen te ontdekken. In het bijzonder is het gebruik van parallelle operatoren (fork en join) een belangrijke oorzaak voor het onvermogen voor deze voorstellingen om diepgaande kwantitatieve analyse toe te laten. Alternatieven zijn Probabilitische Eindige Toestandsmachines, en Hidden Markov Modellen (HMM). Deze laatste blijkt in de praktijk veel meer mogelijkheden te bieden voor de ontdekking van potentiele waardelekken. Er is echter nog een bijkomende stap nodig: in analogie met data normalisatie, die de oplossing gaf voor anomalieën in de omgang met data, wordt in deze thesis de notie van genormaliseerde HMMs voorgesteld. Normalisatie kan gebeuren op twee technieken die aan elkaar gerelateerd zijn: enerzijds clustering, en anderzijds cloning van niet-genormaliseerde plaatsen. Vooral een hybride benadering waarbij beide gecombineerd worden leidt tot een adequate conceptualisatie, zoals de C-K-theorie die vereist. Tenslotte wordt in deze thesis duidelijk hoe eerste-orde technieken, die reeds aangewend worden in Process mining tools, kunnen leiden tot wiskundig incorrecte inschattingen van de probabiliteiten in de procesmodellen. De eerste-orde benadering is wel nuttig om de potentiele waardelekken te vinden onder de vorm van niet-genormaliseerde plaatsen in de Hidden Markov Modellen.

Daarmee is het kernpunt van de thesis, zoals het vervolgens in de gevallenstudies geillustreerd en verkend wordt, neergezet: niet-normalisatie is een belangrijke indicator voor potentiele waardelekken in ICT-gebaseerde processen, en geeft ook aanduidingen voor potentiele conjunctie van (mogelijks innovatieve, in de zin van C-K) verbeteringen van die processen.

Bij Process Discovery moeten grote gegevensverzamelingen verwerkt worden om te komen tot praktische toepassingen: het is een echte “big data” toepassing. De meeste open source tools zijn onvoldoende schaalbaar om de praktijk van tientallen miljoenen records aan invoergegevens toe te laten. In hoofdstuk 3 wordt de opbouw en architectuur besproken van een professionele instrumentatie, die over de jaren is gegroeid binnen de onderneming OpenConnect in Texas, U.S. Oorspronkelijk ontstaan als onderdeel van een toolset voor SOA (Service Oriented Architecture) voor – ondermeer – de analyse van services, bleek rond 2008 dat deze toolset, Comprehend genaamd, om te buigen was tot een werktuig voor Process Discovery, met een vrij verregaande ondersteuning van de C-K-benadering. Dit heeft geleid tot een grondige ommekeer binnen het bedrijf OpenConnect, dat nu meer en meer complexe projecten rond Process Discovery aanpakt in zijn portfolio. In het hoofdstuk wordt niet alleen de technische opbouw besproken, maar tevens het gebruik van Comprehend, en de verschillende analyse-rollen die daarbij spelen. Het geheel wordt – scherm per scherm –toegelicht met een beperkte gevallenstudie.

Vervolgens komt het tweede voornaamste deel van de thesis: de toetsing aan de hand van relevante en realistische gevallenstudies. In hoofdstuk 4 worden twee administratieve gevallen van transactionele ICT-gebaseerde processen behandeld. Niet toevallig betreft het gevallenstudies in de sfeer van de administratie rond gezondheidszorg, namelijk de administratieve verwerking van “medical claims”. In de U.S. is dit een omvangrijke private activiteit, die mede in de discussie rond “Obama-care” centraal staat in veel discussies. Beide gevallenstudies tonen verschillende soorten waardelekken, die op het einde telkens samengevat worden in overzichtelijke tabellen. Alhoewel geanonimiseerd, zijn de verbeteringsresultaten die beschreven worden reëel en verifiëerbaar. Tijdens de analyse komt de bevestiging van het feit dat de waardelekken mee aangeduid worden door het observeren van niet-genormaliseerde plaatsen in de HMM-voorstelling van de transactionele processen.

Interactieve processen zijn uiteraard complexer van aard, en worden geïllustreerd in hoofdstuk 5. Ook de soorten procesverbeteringen die nagestreefd worden zijn ingewikkelder: het gaat vaak om problemen die vroeger als “intangible” werden beschouwd, zoals “klantenverloop”… In het hoofdstuk van eerst twee gevallen van Contact Centers behandeld. Eerst voor een commerciëel bedrijf in de financiële sector, en vervolgens voor een non-profit overheidsorganisatie. Verschillende praktische aspecten komen naar voor in deze gevallenstudies, zoals ondermeer het opvangen van het ontbreken van systematische contactgegevens, en de manier waarop deze via “time & motion” experimenten kunnen achterhaald worden (zeer analoog aan de manier waarop Taylor zijn eerste experimenten rond industriële processen moest uitvoeren). De laatste gevallenstudie raakt evenzeer het echte leven en toont de analyse van zorgprocessen in een gespecialiseerde kankerkliniek. De nadruk ligt op de omkaderende processen (en niet de operatieve ingreep zelf). Hier wordt aangetoond hoe de eenzijdige financiële benadering van gezondheidszorg kan leiden tot suboptimale verblijfsduur in ziekenhuizen, met de bijhorende risico's, die zich uiteraard vertalen in waardelekken. Via de clustering, in samenhang met Data Discovery, wordt aangetoond hoe voor een bepaalde pathologie een "natuurlijke" verblijfsduur naar boven komt, en zich manifesteert in een procesverloop (klinisch pad) dat veel minder niet-genormaliseerde plaatsen bevat dan de processen voor een kortere of langere verblijfsduur. Het wordt duidelijk dat bepaalde vormen van zorg zich niet éénduidig laten dicteren door één parameter, zoals verblijfsduur. De gerealiseerde verbeteringen vertalen zich in een toenemende patiententevredenheid. Als rode draad doorheen de gevallenstudies vallen opnieuw de niet-genormaliseerde plaatsen op als indicatoren voor waardelekken.

Deze thesis is één ontdekking rond "Discovery". Het is zeker geen uitputtende verkenning, en er zijn nog talloze onderzoeksproblemen die overblijven. Sommige worden aangegeven in de conclusie. Vooral een verdere systematische verkenning van het rijk vermogen van de C-K-benadering kan nog vele bijkomende ontdekkingen opleveren…

CHAPTER 1

Fundamentals of ICT-based Services and Processes

1.0 Introduction

Information and services sit at the very core of economic activities in today's society. They permeate in the economy of every nation and every business in the marketplace. Presently, 84% of the GDP of the United States economy is driven by services. Services are driving now the majority of the economy of which information-based services makes up 53%. Services can either be classified as material-related services or information-based services. Material services are tangible services directed at people's tangible assets or possessions. These services can include, for example, tourism, retail, transportation and restaurants. Information-based services on the other hand, are intangible services that are directed at a person's mind and services that are directed principally to people's intangible assets. Typical examples of such services are education, banking, legal services, insurance, financial services, healthcare, telecommunications, accounting and consulting. Brian Arthur, a visiting researcher with the Intelligent Systems Lab at the Palo Alto Research Center (PARC) calls this the *second economy* (Arthur 2011). He describes the second economy as being digital and existing next to the "physical economy" as one where

… "processes in the physical economy are being entered into the digital economy, where they are "speaking to" other processes in the digital economy, in a constant conversation among multiple servers and multiple semi-intelligent nodes that are updating things, querying things, checking things off, readjusting things, and eventually connecting back with processes and humans in the physical economy".

Here is important to note that the activities and events that comprise the processes in the second economy are digital and unseen. This is in marked distinction to the physical sector where necessarily, process components are readily observable. Additionally, as material goods increasingly become commoditized, service is becoming the key differentiator, even in the manufacturing sector (Rust, 2006). With the exception of China, which may only be a temporarily situation (Karmakar, 2005), the predominance of service is similarly true for most of the other large economies worldwide. Commensurately, the knowledge of theory and best practices concerning the creation and delivery of services has not only intensified competition, but has simultaneously posed new challenges opportunities for service managers.

1.1 Understanding ICT-based Service Work

1.1.1 Knowledge-work versus data -work

Situations where ICT-enabled service-based work processes dominate the organizations operational activities are good examples of what are called knowledge-worker environments. Peter Drucker describes six key factors in knowledge-worker productivity (Drucker 1999):

1. Knowledge-worker productivity requires that we are able to describe a task as well as what it should be and what it is expected to contribute to the overall activity/process. The knowledge-worker is also expected to be task aware, understanding what can be

changed to improve the task rather than simply understanding how to do the task itself. This is a major differentiator between knowledge-work and what will later be described as data-work.
2. Knowledge-workers are autonomous and therefore both self-managing and responsible for their own productivity.
3. Knowledge-workers are expected to be continually innovative.
4. Knowledge-workers are expected to be continually learning as well sharing their knowledge with other knowledge-workers.
5. For a knowledge-worker, quality is as important of a measure as productivity (e.g. increasing output).
6. Knowledge-workers are seen as an asset rather than a cost to be eliminated.

This shows that knowledge in organizations exists through people but as a "process" rather than as an "artifact" (Coles et al., 2007) Grantham, et.al.) and suggests that there is a distinguishing set of characteristics that define knowledge-work from other types of ICT-based service work which we shall call *data-work*. In contradistinction to the above, this research proposes the following regarding data-work:

1. Data-workers execute a set of tasks that have been defined for them (e.g., editing a claim) rather than have the latitude to change the way a task is performed.
2. For data-workers, productivity standards are set by management hence they have little or no autonomy.
3. Innovation is not the responsibility of the data-worker; it is usually delegated to a staff group.
4. While a certain amount of continual learning is required, there is no expectation of information sharing with others.
5. Productivity of the data-worker is primarily a matter of quantity with quality being a secondary issue.
6. Data-workers are seen as a cost that is expected to be reduced if not totally eliminated.

The relationship between these two types of work and their distinguishing characteristics can are described in Table 1.1

Characteristic	Knowledge-work	Data-work
Task Improvement Aware	High	Low
Autonomy	High	Low
Innovation Ability	High	Low
Learning/Sharing	High	Low
Quality Focused	Quality	Productivity
Asset versus Cost	Asset	Cost

Work versus Characteristics Table 1.1

These differences show that there exists two different types of work and workers, both with differentiated tasks, knowledge and expectations. This is consistent with the definition of [Hopp et al.] who describes knowledge work as a subset of white collar work since knowledge intensive tasks are classified as white collar (Hopp et al., 2009). [Awad and Ghaziri] describe knowledge as that which is gained through experience or study that enables a person to perform a specific task (Awad and Ghaziri 2004). Clearly from the previous Table 1.1, if the workers have little autonomy, hence if they are seen as a cost with production standards being set for them with a focus more on productivity than quality, then they are data-workers. In an ICT-enabled service-based work system, data-workers are often assisting the ICT application (e.g., reviewing claims, payments or applications) in its data validation and editing tasks. Knowledge-workers on the other hand are using the data in some self-directed and possibly creative way to enhance the outcome of the ICT-based process in ways it cannot intrinsically perform.

This type of description is similar to the one used by Doug Engelbart (Engelbart, 1962). In this definition, automation is where one use of ICT is to eliminate the activities performed by humans, thus an economic substitution (technology for human labor) whereas, a second use, augmentation, focused on how ICT- enabled service-based technologies can extend the performance of the human in accomplishing the specific tasks. This lends itself to the observation that data-work is the candidate for automation and thereby substitution, along with the possible elimination of the data-workers who are seen as a cost, whereas ICT-enabled augmentation is applicable to the knowledge-worker (who can seen as an asset).

1.1.2 Augmenting Human Intellect

A question that arises is, “How can ICT-tools assist in this effort”? In his 1962 seminal paper “*Augmenting human intellect, a conceptual framework*”, Douglas Englebart described his research objective as understanding the limitations of man’s information handling capabilities and determining a framework to assess and develop new capabilities that would augment them so as to help improve society. Engelbart postulated that the human and his intellect could be augmented through a technology interface thus increasing his potential impact on the world. In terms of work, this type of augmentation by a machine is seen as an increase in productivity. It is through increases in productivity that societies create wealth and improve standards of living. In Englebart’s discussion of intellect augmentation, three issues are specifically relevant to this research:

1.1.2.1. Concept Structuring

A concept structure is *“something which we might try to develop on paper for ourselves or work with by conscious thought processes, or is something which we try to communicate to one another in serious discussion. We assume that, for a given unit of comprehension to be imparted, there is a concept structure (which can be consciously developed and displayed) that can be presented to an individual in such a way that it is mapped into a corresponding mental structure which provides the basis for that individual's "comprehending" behavior”* (Engelbart 1962). A concept structure is a representation of knowledge in a pre-defined format to enable a better understanding of its content. A Markov model is an example of a pre-defined format. In essence, a concept structure evolves within a culture (e.g., work) and can be purposely designed and modified. Indeed, one can observe how a formal concept analysis lattice representation contributed significantly to the understanding of ICT-based processes (Wille, 2005). As will be shown later in this thesis, a concept structure can be represented as a Markov model, formal concept analysis (FCA) lattice or other representations such as a Petri- nets.

1.1.2.2 Symbol Structuring

Englebart explained that a concept structure can be represented internally by instruments with symbols that are understood by those instruments. Symbols have inherent meaning. In a data related concept structure, symbols could be “key-words” which objects are mapped against. In a process related concept structure, symbols include events which may incorporate “key-words” as well as other data (e.g., time and resource consumption). Indeed, these symbols may be a better representation that what the person may have produced. Also, since a person may modify his “view”, instruments may be able to store those modifications without affecting the external and conceptual view held by the human. Hence capturing and storing those structures and concepts while allowing the user to maintain an independent view of the concept from which the modification can be judged is a worthwhile objective.

Additionally, Hubbard (Hubbard 2007) described the advantage of instruments over humans as providing for:

- Better Detection
- More Consistency
- Calibrated for Error
- Deliberate anonymity
- Complete recording
- Cost advantageous

This sets the stage for the development of an automated tool that can assists in this effort. In chapter 3 the description of an automated tool that meets this description will be offered.

1.1.2.3 Process Structuring

Concept and symbol structuring represent the language component of augmentation, whereas process structuring represents the methodology component. Methodology issues range from sequence of task execution to the organizational process for implementation and supervision. Engelbart states that many of them are applied to the formation and manipulation of symbol structures, the purpose of which will often be to support the conceptual labor involved in process structuring. Engelbart did not mention which method of process structure is most suitable and it is interesting to note than 50 years later, we are still having the discussion about the best representation for process structuring. Currently, Markov models, hidden Markov models, and Petri-nets are the most commonly used representations for process structures.

1.1.3 Transactions versus Interactions

A next characterization of work is the degree of transactional versus interactional characteristics that are involved. These notions can be seen in Figure 1.1 that places various types of processes in a matrix that positions transaction – interactions against ICT automation – augmentation objectives. As can be readily seen, processes that lend themselves to a high level of automation are transaction oriented whereas processes that demand a high level of interaction are appropriate for ICT-based augmentation, making the human more effective rather than a subject for replacement.

When analyzing the types of service work discussed above, it is important to determine which activities lend themselves to automation and which lend themselves to augmentation. In the case of many administration processes, such as claims processing, the tasks are mainly clerical (e.g., data validation) and can be completed through automated procedures. Here the desire is for the lowest cost possible that provides acceptable quality, in essence, it is a transaction activity. In the case of caring types of processes, (e.g., hospital post-surgical activities, customer contact centers) the issue of cost, while important, is off-set by issues of interaction quality. Specifically, the quality of the communication between the care-provider and the customer/patient, as well as its perception, is key issues as to the overall effectiveness of the process. As an example, imagine a situation where a customer contact center is managed with cost as the driving factor. In such a context, the agents would be measured by the amount of minutes spent on a call to achieve customer resolution with the lowest amount being the predefined desired goal. In such a context, agents will often work to achieve the solution quickly rather than completely, requiring the customer to call back an additional time. This can result in a significant perceived lack of quality causing customer irritation (Gronross, 2004).

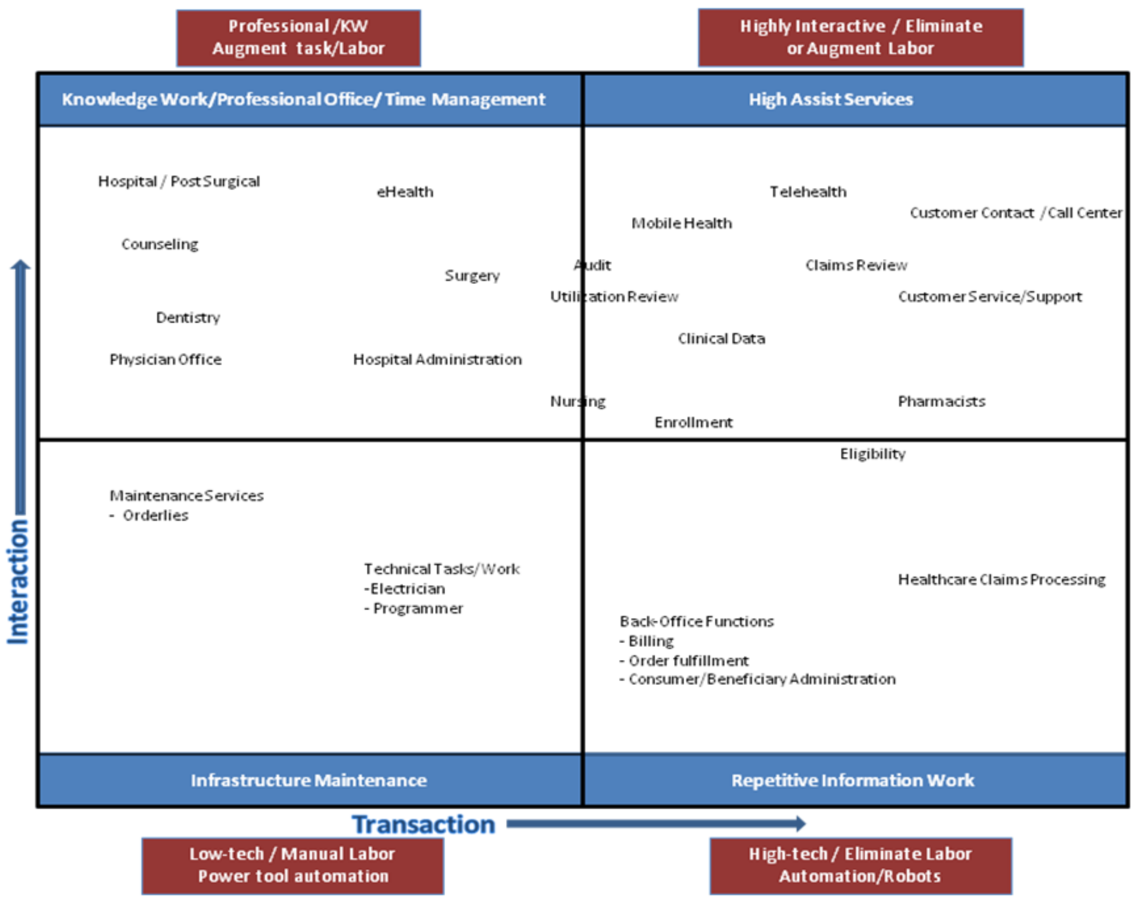


Figure 1.1 – Matrix

1.2 Approaches to Improving ICT-Based Services and Processes

In order to improve an ICT-enabled, service-based process, and gain a true understanding of which tasks lend themselves to automation or augmentation, an organization needs to capture data or facts about the true nature of the recurrent states of the process. Unfortunately, a majority of ICT managers state that improving business productivity is their highest priority while at the same time they report a lack of understanding of current processes either through documentation or a lack of subject matter expertise. This situation, where improving business process is the highest priority, is accompanied by a lack of understanding about the processes themselves, and demonstrates that a significant *knowledge gap* exists in today's organizations.

On February 12, 2002, former U.S. Secretary of Defense Donald Rumsfeld made the following pronouncement at a press briefing:

> *"As we know, there are known knowns. These are things we know we know. We also know there are known unknowns. That is to say, we know there are some things we do not know. But there are also unknown unknowns, the ones we don't know we don't know"* (Rumsfeld, 2002).

These statements are about knowledge and uncertainty. In the first case, there should be little uncertainty since we have knowledge of both the situation as well as what we know. In the last case, we are totally unaware hence we cannot be in a state of uncertainty, simply ignorance... But, in the second case, since we know that we do not know, we must do something to reduce the uncertainty before making a decision or we can choose to move ahead and let fate take its chances. Should I get more information or should I just take my chances? This is the central question faced by all decision makers whether it is in military, commercial or non-profit enterprises. While it would be great to say that in the majority of cases decision makers (from now on referred to as managers) always searched for more information to increase the likelihood of making a correct decision, the research shows this *not to be the case* (Pfeiffer & Sutton, 2006). Stanford University management professor Jeffery Pfeiffer states, "beliefs, half truths, opinions and ideology rather than hard facts" are often the main drivers in management decision-making. Drawing from research in the medical community, specifically evidence-based medicine; Pfeiffer uses the term Evidence-based Management. Evidence-Based Management or fact-base decision making, enhances the overall quality of organizational decisions and practices through reflective use of relevant and best available scientific data. This approach combines conscientious, judicious use of best evidence, or set of gathered facts, with individual expertise; ethics; valid, reliable facts; and consideration of impact on stakeholders. Much of Pfeiffer's discussion is about how management lacks the discipline to focus on this paradigm leading to less than optimal results.

In terms of business process improvement, it is the lack of hard facts, combined with the lack of a culture that supports gathering evidence and managing by evidence, which accounts for many of the unproductive issues faced by management today. Ironically, this is much like it was in the time of Frederick Taylor in that each person was seen as possessing a set of skills which made them productive and overall; one would use "rules of thumb" or best approximations, as a guide to managing output, or work process expectations. One of Taylor's main contributions was to debunk the idea of skill as a unique trait and instead view the work activity as a series of repetitive motions that could be recorded and analyzed. When the set of motions that yielded the best output was "discovered", the workforce could be trained in its application, thus increasing overall output via the "one best way". Similarly, in an ICT enabled service-based process, a productive worker is often viewed as skilled without any detailed analysis of the actual work performed. While this is often due to the nature of the digital work environment where much of the effort is being augmented by ICT, it will be shown that new techniques and technologies will once again show how it is often repetitive and augmented motion, rather than skill that is the defining factor. The following sections will discuss the major approaches used today understand service and processes.

1.2.1 Lean Management

The essence of Lean Thinking is identifying and eliminating waste in processes, whether manufacturing or services, speeding up process flow and reducing costs. Lean accomplishes business transformation through one to two week mini-projects, called Kaizen events. Starting with a value stream map to identify activities that are valuable to the customer or consumer and to document the overall process, Lean proceeds to identify and eliminate activities, which are not value added. Non-value added activities are called waste. The Japanese termed this as "*muda*" which means waste or any activity that does not add or create value (Womack & Jones, 1996). Eliminating *muda* speeds up the process and reduces its cost, making the producer more competitive. Lean Thinking defines value-added activity as possessing three specific attributes:

- The customer is willing to pay the producer to do it;
- It transforms the work in progress from its current state toward its eventual final state; and
- It is done correctly the first time.

By way of example in a service, proofreading, to ensure correct data entry before pressing the enter key, would not constitute value added activity. It might be a necessary waste, but it is a waste nonetheless, by the above definition. The original wastes in manufacturing were identified as transportation, waiting, overproduction, defects, inventory, motion and extra processing. A recent addition is under-utilizing a worker, having a manager do menial tasks which an unskilled, lower paid worker could do, or not promoting a worker along his career path once he achieves the skills and experience needed to do the higher valued work. While Lean methods were initially developed for manufacturing activities (Liker, 2004), they have been re-interpreted for the elimination of wastes in services as follows (Bicheno and Holweg, 2009):

1. *Delay* on the part of customers waiting for service, delivery, in queues, for response, not arriving as promised. While the customer's time is seen as free to the supplier, it is a source of dissatisfaction and account turnover.
2. *Duplication*. Having to re-enter data, repeat data on forms, duplicate information, and answer questions from several sources within the same organization.
3. *Unnecessary Movement*. Queuing several times, lack of one-stop, poor ergonomics in the service encounter.
4. *Unclear communication* and the wastes of seeking clarification, confusion over product or service use.
5. *Incorrect inventory*. Out-of-stock situations, expectation mismatches, or service substitutions.
6. *Lost customer retention opportunities*, failure to build relationships, rudeness and impolite service attitudes.

7. *Service transaction errors,* product /service bundle defects, lost, misplaced or damaged goods.

While the benefits of using the lean approach have been well documented (Swank, 2003) its results are dependent upon analysis of data which in many cases does not exist in digital form. This issue had been described in an SAP white paper as follows (SAP, 2005):

"…teams end up spending a significant amount of time manually collecting data that is too latent and inaccurate. If any gathered data has bias at the source or is subject to interpretation, analyses based on this data may contain errors. Furthermore, manual data collection is prone to mistakes. Therefore, you need a systematic process of data collection that answers three basic questions:

- *How accurate are the data?*
- *How real-time and how relevant are the data?*
- *How scalable is the data collection process?*

…avoiding failures from inaccurate data requires stable and scalable information systems to collect data and a functional application system that allows you to analyze and distribute the information. Ideally, these systems should gather data automatically at the source of an activity. For example, they should gather data about yields, conversions, efficiencies, wait times, material availability, and production directly from the shop floor. Direct data gathering minimizes data inaccuracies and errors that result from manual interpretation. "

As mentioned above, many efforts rely on manual data collection, often, using worker self-reporting on paper forms. This type of approach to data collection is extremely error prone and unless subjected to rigorous calibration effort, may be useless in the production of valid analytics (Hubbard, 2007). If the veracity of the data is in question, then necessarily, so is the resulting analysis thereby diminishing the ability for successful change management and productivity improvement.

1.2.2 Six Sigma

Six Sigma is heavily based on statistical process control and uses has a large number of tools, each of which may be used in one or more of the phases of a process improvement project, though it is uncommon for all the be used in any one project. As mentioned earlier, Six Sigma improvement projects generally follow the DMAIC model (see Figure 1.3). It is very difficult to collect the required data for Six Sigma projects because it is difficult to accurately observe exactly how people spend their time unless production applications are instrumented. Six Sigma projects are oriented to eliminating defects. Statistically, a 'sigma' (s) refers to the **standard deviation** from the mean of a population. Standard deviation describes the probability of your next data point deviating from the mean of the data set. The sixth sigma refers to the likelihood that only 3.4 out of every 1 million data points will appear outside the sixth standard deviation. That translates into less than 4 errors per million transactions or 99.9997% of transactions are of good quality. Six Sigma (Pande, 2000), (Brue, 2002) has its roots in manufacturing quality management and is all about **variance reduction.**

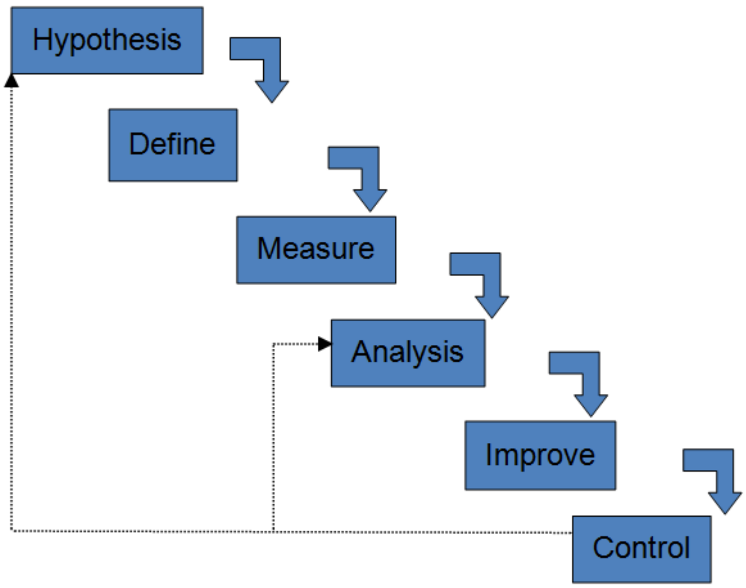


Figure 1.2 Six Sigma DMAIC Method

In the *Define* phase, a business case drives the creation of the following deliverables: a project charter stating the business problem, project goal, project team and high level schedule. Other deliverables include a SIPOC model, which is the highest level process diagram showing the Supplies, Input, Process, Output and Customers, and the Voice of the Customer. The voice of the customer lists the key issues, critical to quality metrics, and the major attributes of the product or services.

In the *Measure* phase, the existing process is described in more depth and precisely. Since Six Sigma aims to improve a process by reducing variation and defects by controlling the key performance indicators and centering them on their prescribed mean (numeric average). Six Sigma practitioners (Liker 2004) state that the result of a process comes from transforming inputs to outputs, or in the form

$$Y=f(X_1, X_2, \ldots, X_N),$$

where Y, is the main desired result, and X_1, X_2… is the set of inputs - those which are critical to quality (CTQ), and "f" is the function that expresses how the process transforms the inputs to create the outputs. Most often, a-priori assumed functions are used for “f”: time-series, linear regressions, cluster analysis, etc… Moreover there is no a-priori indication which function is most suited to quantify and interpret the measures.

Another deliverable in the Measure phase is the data collection plan, in which operational definitions are nailed down as well as what, how and when CTQ attributes will be measured. Since the whole idea is to eliminate defects, it is important to understand how the existing process behaves. There are various types of numbers that are used to describe the process performance, relative to how many defects it produces.

The *Analyze* phase is devoted to understanding which attributes, or main factors, have the most influence on how the process varies. In any complex process, there are many attributes involved.

One of the most important tasks of an improvement project is determining, which are the critical few and then ignoring the trivial many. One tool used to measure these phenomena is called the Pareto chart. Graphing the number of problems by cause will often illustrate which are the key sources of errors, and help the improvement team drill down to root causes. There are additional tools used to understand the main factors of process variation, including stratification, in which individuals are grouped by some similar characteristic, in order to find patterns of variation, and potential high impact improvements. Improvement efforts must focus on root causes, not just symptoms, so another deliverable in the analyze phase is the Ishikawa, fishbone or cause and effect diagram, in which the large categories of variation (the effects) form the skeleton of a fish-like diagram, and the individual instances of variation, and hence problems, form smaller and smaller bones, until the finest bones represent root causes

In the *Improve* phase, the project team creates solutions to the business problem. They typically continue to work with subject matter experts, front line associates, who are familiar with the process and often have useful ideas to improve it, as well as the project sponsor, who owns the process. In complex situations, teams often use brainstorming sessions to produce creative ideas. Additionally, various tools are used to ensure that the solutions are practical and sound economically. These tools include benchmarking, cost-benefit analyses, and similar filtering for the best approach. Since it is important to have real world experience, the team typically works with operations to pilot the proposed change

In the final phase, called the *Control* phase, the new production process is implemented, using the previous detailed planning. The control phase is also where we ensure that the new process stays in control and where we eliminate old behaviors. Since additional issues can surface when moving from pilot to full scale production, the improvement team and those who operate the process use the classical improvement approach, called the Plan, Do, Check, Act (PDCA) cycle, where they again plan for, implement, measure and bake the new methods into the full scale production process as well as actively checking for changes which affect quality. Once there is enough data, the process is re-evaluated for how often it produces desired results and its defects as well as current variations. In order to document the results of the project and to pass the new found knowledge on to others, the team publishes its key “learning’s” from the effort.

1.2.3 Lean Six Sigma

The characteristics of Six Sigma and Lean can be summarized as follows (George, 2002):

Six Sigma…

- emphasizes the need to recognize opportunities and eliminate defects as defined by customers
- recognizes that variation hinders our ability to reliably deliver high-quality services
- requires data-driven decisions and incorporates a comprehensive set of quality tools under a powerful framework for effective problem solving
- provides a highly prescriptive cultural infrastructure effective in obtaining sustainable results when implemented correctly, promises and delivers $500,000+ of improved operating profit per Black Belt per year (a hard dollar figure many companies consistently achieve)

Lean …

- focuses on maximizing process velocity
- provides tools for analyzing process flow and delay times at each activity in a process
- centers on the separation of "value-added" from "non-value added" work with tools to eliminate the root causes of non-value-add activities and their cost
- provides a means for quantifying and eliminating the cost of complexity

Lean Six Sigma (George, 2002), (Sproull, 2009) blends Lean, with its primary focus on *process speed through waste reduction* and Six Sigma, with its primary focus on *process quality through variation reduction*. Yet, as discussed in the section on lean management above, the same issue of the veracity of the data exists in Lean Six Sigma projects since often the data that will drive the design of experiments is non-existent and must be manually created thus introducing the above mentioned bias.

1.2.4 Automated Business Process Discovery

With the exception of process that are machine controlled, as in a manufacturing process, all information gathered in both six sigma and lean management projects are based on manual data collection through, or in combination with interviewing people. As mentioned above, this is both costly and error prone. An approach that overcomes these issues is called automated business process discovery (Cook & Wolf, 1995), (Verner, 2004). It is based on technology that monitors the interaction between the user and the information system during the actual execution of a business process. It can also be used to monitor and measure the activities and performance of various components of the ICT-enabled process itself, specifically, where it monitors the effectiveness of software robots. In essence, its aim is to derive a formal model from the event data produces during the execution of the process (Cook and Wolf, 1995). In this regard, the model must be both a representation of the process and able to be understood by an engineer or business analyst with domain expertise. It has the following characteristics:

- *Emergent Paradigm* - As opposed to the DMIAC methods used in Six sigma, automated discovery process relies on collecting data from the process over a given period of time. It is not based on a "design of experiments" approach that is meant to prove or disprove a given hypothesis. It is based on actual event data collect form the process itself in real time or through event logs. This data can then be analyzed to form a process model based on the observed reality. It is important to note that it is not a top-down requirements model or a specification of some desired future state; it is an actual instrumentation of state of the process when instrumented. Through automating the analysis of the data, the subjectivity of the design of experiments approach techniques can be removed. The process can have an ingrained methodology that has been shown, through repeated trials, to accurately discover processes and process variations without bias.

- *Grammar Inference* - As automated process discovery is an ICT-enabled function, it relies on the ability of tools; it relies on process behavior to be described as sentences of a language that contain a specific grammar. This is in essence a form of augmentation as described by Engelbart (Engelbart, 1962). Specifically, symbols and symbol structuring are the actual language of events. Process structuring then deals with how the events are involved both sequentially and temporally as described by the grammar. Actors in the process are concepts and, therefore subject to concept structuring.

- *Accurate Information* - Since the information is collected from the actual source is cannot be inaccurate through second party representation.

- *Complete Information* - An automated process captures all the information that is occurring within the ICT-enabled service-based process and represents them by time, date, user, etc…. Since the information is collected from the real-time interactions, it is not subject to lost or selective memory issues. This includes completeness on the information regarding exceptions in the processes. Often, in particular in statistics driven six sigma studies, exceptions are treated as statistical "noise". This may ignore important inefficiencies in business processes.

- *Standardized Process* - The automated collection of information will provide data on the process, which can be grouped quantified and classified. This provides the basis for the development and monitoring of both a current and new process to which benchmarks can be assigned. This provides the basis of both new process design and the determination of problem root cause. Additionally, it can set the stage for efforts at continuous process improvement.

As explained in the above section and completely in Chapter 3, the methodology for automated process discovery (Figure 1.3) differs significantly from the DMIAC methodology applied by Six Sigma and Lean efforts. It is based on the concept of emergence whereby the entire process can be viewed as a stream of complex events that are continually changing. Hence, it is in the idea of emergence, with its approach to learning and adaptation, which goes beyond the limitations of simple rules found in more rigid approaches (Berkowitz, 2003).

As opposed to the six sigma approach, it is important to note that the process does not begin with a hypothesis followed by a design of experiments (using a-priori assumed models, such as regressions, time series, correlations,…) as in the case of six sigma, rather it collects data from the process itself and then follows *emergent propositions* that are generated from the data. The discovered propositions then become the basis for analysis rather than any pre-determined hypothesis.

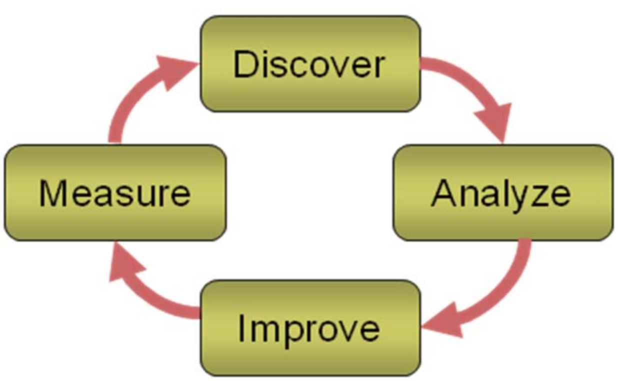


Figure 1.3 Automated Process Discovery Analysis Method

1.3 Emergence and Enterprise 2.0

The idea of emergence and the term Enterprise 2.0 were introduced by (McAfee, 2006, 2009) to describe social-network based modifications to company intranets and other classic software platforms used by large companies to organize their collaboration, communication and knowledge work. The term 2.0 seems at first sight just a software "version number", although something more fundamental happened: 2.0 is primarily emphasizing and pushing 2-way interaction. It took its form when the Web 2.0 emerged and started user-driven social media applications, such as blogs, wikis , RSS feeds and the like where users virtually participate in dialogues and contribute to user-generated content. In and around enterprises, Web 2.0 platforms have been professed to support a profound change in intra- and inter-enterprise communication patterns (DeHertogh et al, 2011).

In a two-way interacting network, also known in game theory as a collective decision game, the concept of Nash equilibrium prevails. Intuitively, a Nash equilibrium is a strategy profile (a collection of strategies, one for each player in the game) such that no player can do better by deviating (Haplern, 2008). The insight behind this concept is the representation of possible steady states in the play where each player maintains correct views about what the other player is doing and within such view, the player renders the best response. In the concept of a Nash equilibrium it is demonstrated how in such exchange both the client and provider are able to realize their local optimal value which may be different from a global optimum. An instance of such may happen when the client is able to personalize his experience to a level that allows him to perform his job efficiently while, at the same time, the provider is able to derive greater value from its service investment in the form of higher profit, customer lock-in or new forms of knowledge.

Given such two-way interactions, it is obvious how the local optimal interactions constitute communities that emerge around shared values. From a business perspective, two-way interactions urge business operations to consider seriously issues that deviate from "global optima". Typical examples of such issues are rich user experiences, long tail effects, agile design and development, self-service, trust-establishing initiatives, and so on. Moreover in most cases, information nowadays is freely available anytime, to anyone, from anywhere. As the degree of information increases more and more, and the service clients and providers are getting more informed, there is more potential for value co-creation.

1.3.1 Value Creation

The notion of "value" should be understood in a holistic way and as more than simply financial value (such as profit or revenues) (Bhardwaj, 2007). Value in a service interaction is anything that is considered/perceived as being meaningful. A concrete example is the new competition in Healthcare (Teisberg, 1994, Porter & Teisberg, 2006). Typically, governmental healthcare programs try to optimize, by minimizing, the costs of the healthcare services. Of course the obvious question is: why are patients choosing a particular hospital? Are they choosing it for the best medical care with the most competitive price? On the contrary, as they are mostly insulated from the price, how they perceive value in non-financial terms, patients specifically choose hospitals with manageable (minimal) waiting times, good pain comfort and good socio-emotional care. This corresponds with the definition espoused in the field of competence-based management where value is defined as the creation of benefits for the participants in the organized activities (Sanchez, Heene & Thomas, 1996).

The value co-creation is based on the joint perception of value, both from the point of view of the service providers (creating superior value propositions), and the service clients which are able to give a (potential non-financial) meaning to what happens in the services. In a traditional "market-driven" approach, suppliers are supposed to provide and sell the services and clients purchase such services. Today, also thanks to the 2.0 evolution, clients can engage in dialogues with providers during each stage of the service design and delivery. Both the client and provider have the opportunity to create unique values through customization and experiences (Payne & Frow, 2005). Hence the 2.0-based world of ICT-based services is not "flat", it is "bumpy" and the local optima are precisely the communities that gather around shared values. A careful inspection of one of the foundation texts on e-Business is a nice confrontation: reading the first lines in the Cluetrain Manifesto (Levine et al, 2001) and changing the original term "markets" into "communities" results in the following:

"Networked *communities* are beginning to self-organize faster than the organizations that have traditionally served them. Thanks to the web (*2*.0) *communities* are becoming better informed, smarter and more demanding of qualities missing from most organizations. … *Communities* are conversations. … *Communities* consist of human beings, not demographic sectors. … To speak with a human voice, companies must share the concerns of the *communities*."

Again, in all these sentences the original text contains the word "markets" instead of *communities.* This implies that some of the traditional competitive strategy analysis techniques are complemented by 2.0 effects. Complementary to Porter's "tying up the customer" comes "invest in understanding the customer *communities*". It should also be noted that 2.0 is not restricted to the Web: any 2-way interaction generates the potential for communities. Digital television and radio allow media companies to move beyond traditional broadcasting and to foster communities of emotion, experience, interest in their 2.0 business.

1.3.2 Value Leaks

Since Nash's theory is about optimizing local values, the optimum may be influenced by both the creation of value as well as the loss of value, meaning value leakage. In this thesis, the perspective of "opportunity value" will be examined in deep and proposed as a major driver for strategic decisions on ICT-based services. Opportunity value means the perception of "lost value". As previously mentioned, ICT-based services are the building blocks of business processes. Therefore, it is not only interesting to see how value is created in the services, but also where value gets lost at some points, in other words, where the value is "leaking" from the processes. There is of course a link to quality management issues: consider the example of a typical oil refinery set-up. A traditional quality management program will try to minimize the time to detect and repair leaking pipes in the refinery. However, the more fundamental question is; "Why are some pipes leaking?" in the first place. From this mindset, the alternative perspective to value creation is to prevent value loss, to avoid value leaks. Peters has defined an example of how the healthcare system in the United States "leaks" as much as 50% of its defined value or $1.2 Trillion per year (Peters, 2012). Preserving value in a process is perhaps as important as the process itself (Gronross, 1990).

Using the same example of healthcare, potential patients often choose a hospital that offers the highest perceived value, meaning in reality that they choose the hospital with the lowest perceived value leaks: hospitals with the shortest waiting times, fewer number of pain incidents and a better socio-emotional caring support program. It is already a fact that patients have self-organized into blog-communities, sharing good and bad experiences on waiting times, pain incidents, socio-emotional comfort, and other issues, in such a way that these blogs have become highly influential in the patient choice for or against certain hospitals.

While ICT-enabled, service-based processes are creators of value, service-based processes are inherently leaky. Specifically, as the environment changes, the ability of the service-based process to adapt comes into question. The mismatch between environment and the ICT-enabled, service-based process to adapt leads to instances of value leakage. Value leaks, by definition, can also emerge where the value perception is attacked or even denied. Value leaks occur when the technologies, resources and assets create insufficient value relative to either absolute or potential expectations. Typical examples are unnecessary processing steps, long waiting time, lack of communication, wrong interpretation, simplistic rule implementation, long search times, unnecessary production or use of resources, avoidable transport of people or materials without purpose, and so on. It is obvious that this relates to the concept of *muda* described earlier in this chapter. Whereas this concept has been applied successfully as described above in Lean production systems (for example, at Toyota), the first application of value leaks in ICT-based services was mentioned in [Peters & Dedene, 2011].

Value leaks represent opportunity values: if they are not contained, value continues to drain out of the business processes. Stopping value leaks in a service may involve new techniques, capabilities and instruments to extend beyond its current service configuration. The opportunities are really opportunities for service improvement and innovation: implementing these innovative improvements allows service customers as well as providers to detect and close value leaks. The main challenge can be seen as one of *discovering muda.* One obvious source of *muda* is all forms of impreciseness in the service contract or imprecise interpretations of a service contract.

Value leaks may also appear in an emergent way in an ICT-enabled, service-processes. In Enterprise 2.0, McAfee points to the "dawn of emergent collaboration". He describes how traditional top-down planning and analysis techniques are insufficient and must be complemented by bottom-up techniques, such as the pure "discovery" of *actual workflows as they happen on the digital work environments*. Automated process discovery is an innovative, analytics-intensive way of observing processes, based on sophisticated technologies and resources to collect and interpret the observations (Peters & Dedene, 2011). It is important to distinguish discovery from traditional lean and six-sigma thinking: as mentioned previously, typically in lean manufacturing, process improvement starts with the posit of a pre-assumed production process model against which experiments and measurements are performed to reduce *muda* and improve the quality results. In *discovery* there is no pre-assumed model hence the analysis includes the detection of the apparent model, without any assumptions. In fact, case studies have shown how pre-assumed models (such as a process manual describing the "standard" process) may be very misleading and actually strongly deviate from the real practices on the work-room floor.

Consider the process case example shown in Figure 1.4 (Peters and Dedene, 2011), illustrating a highly simplified process to handle Customer calls in a Customer Contact Center (Call center). Incoming calls are registered, next the call agent tries to handle the call immediately, in some cases there may be a need to search for additional information and then finally the call is closed (which typically involves storing of all relevant monitoring information regarding the call).

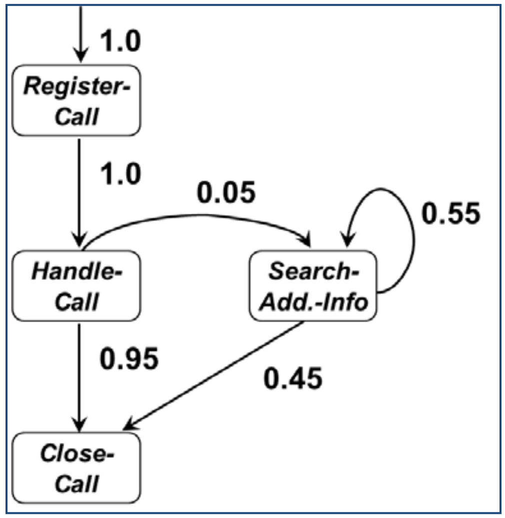


Figure 1.4
Discovery of a Value Leak in Customer Call Handling Process

In this case, the Call handling Business Process consists of 4 ICT-based services:

- A registration service for the incoming calls
- A handling service (typically based on a FAQ-type of documentation service for the call agents)
- A search service to look for additional information, when required
- A closing service which stores all relevant call data in the Customer contact database

Assume that the above processes, and the probabilities attached to the “next step” arrows in the process model have been obtained from analyzing the Customer Contact Logging Database. At first sight, this process seems fine. Indeed, a traditional Key Performance Indicator (KPI) for a Customer Contact Center is the First-Time-Resolution-Rate for the Customer Calls: the fraction of the calls that can be handled immediately without searching for additional information, or without repeating some steps. In this case, the rate is 95%, which is very high. However, from the probabilities it becomes clear where there is a value leak: once additional information is required, only in 45% of the cases the required information is found immediately. So this creates a long-lasting “loop” in the search service step of this process. This loop translates obviously in annoying waiting times for the customer, who might not know what is going on, and who feels neglected on the long run. In this example approaches to stop the value leak might be the following:

- Improving the efficiency and quality of the search engine that is underlying the search service (for example, provide more refined Boolean search facilities)

- Better understanding of the type of calls that (exceptionally) lead to searches for additional information
- Further enhancements to the FAQ document system in the handling service (for example, by providing a richer glossary of keywords)

This example shows how an apparent value leak is the waiting time during the search for additional information. Indeed, another typical KPI that is used for Customer Contact Centers is the call handling time, where the goal is to minimize the call handling time. In a way, this might be irrelevant, as clients don't mind waiting if at the end their call is handled successfully. What matters is avoiding "unexpected time" that is perceived as waste time. Another side effect of reducing waste time is of course the reduction of the occupancy of the call agents, which should typically be below 85% in general: if not, call agents become stressed, which translates in unfriendly, and possibly less efficient behavior, which becomes another value leak.

Other sources of value leaks are in the availability and reliability of services.

Availability - Anticipating and avoiding value leaks is also the responsibility of the service provider, who is supposed to bundle technologies, resources and assets to meet the needs of clients. If the bundle has insufficient capacities (e.g. during the peak times), certainly additional value leaks may occur. Sometimes, service providers may protect themselves by including such conditions in the invariants articulating that they cannot meet the post-conditions during peak periods. This is however not really a very strong provider position as it easily opens opportunities for competition.

Localizing the presented process in a customer contact center, the call handling time should have been better managed and thus less waiting time for client and service time for provider if the comprehensive presentation of information and requirements have been done, if time is not lost due to inefficient searching. In the example of the Customer contact center, it is the responsibility of the service provider to allocate sufficient (for example variable) additional capacity (for example, call agents, human resources) to handle peak periods. Service providers are supposed to become experts in optimizing their bundles, to provide as accurately as possible the required capacity and performance in the services.

Reliability - Services bundle together technologies (often ICT-intensive), resources and assets which must be reliable as well as available together in the collaboration in the bundling. Unreliable information, as well as unreliable human resources may be examples of significant value leaks. Discovering unreliable behavior (such as fraud) in business processes is highly non-trivial (Viaene et al., 2005) and requires appropriate analytics. Again this illustrates how value is co-created: uniform correct fraud detection in insurance claims would typically allow to lower insurance policy rates by 15 to 20%! This is a joint opportunity value for both the insurance providers as well as their customers.

1.4 Services as Building Blocks of Business Processes

For a better understanding of the matching of services with the business on the demand side, an explicit link with business processes can be considered. Using the concepts of Service-oriented Architecture (SOA) such as expressed in Archimate (Lankhorst, et al., 2005) as a reference, in the SOA-approach to business processes, the granular (indecomposable) processing steps in business processes are considered as being business services. Some of the services may be “external”, open to customers of the business, whereas others are “internal” within the business. In the processing of an insurance claim, for example, the steps of registering the claim, confirming the claims validity and the actual payment of the claim may be examples of external processing steps, whereas the actual adjudication of the claim be an internal processing step, hence an internal business service. Apart from the processing steps, business processes may also include material flows, material handling steps (such as material combinations or inspections), document flows and document handling (such as signatures). Of course these materials and documents may become required to perform a next processing step (hence acting as preconditions for the next processing step). Archimate as well as UML/BPMN provide graphical representations of business processes, including their external as well as internal processing steps.

The issue of granularity is relevant here. Granular services are services that can no longer be decomposed into subservices. The granularity requirement does not put a limitation to the fine-grained or coarse-grained functionality encapsulated in the service (Beisiegel, 2007). Services such as a “banking service” and a “customer service” are too broadly defined to be researchable in practice, although they may be large ensembles of (external) business processing steps. Fine-grained services can include technical services, (e.g. check logs), infrastructure services (e.g., check the status of the network) or atomic business logic tasks (e.g. do check sum validation on an account number). Fine-grained services are very often ICT–centric in nature since they consist of computational steps followed by a decision. Nevertheless they may involve multiple business concepts or business parties (actors in a particular role) to perform the service. They encapsulate a meaningful business functionality that hides complexities of the underlying implementation, which is precisely the challenge for the service systems that bundle all what is required to render the service. In UML, (universal modeling language) an ICT-based processing step in a process is also called a “Use Case” and granular indecomposable use cases are also known as “User Stories” in Agile approaches to Systems Development. In Figure 1.5, a business process is shown, together with its elementary processing steps as well as the business parties that are involved.

In this case, the business process is a combination of 4 granular services, all of which are ICT-based services:

- A registration service to create and annotate the order
- A logistics service to deliver the order
- An accounting service to invoice the order
- A financial service to pay the order.

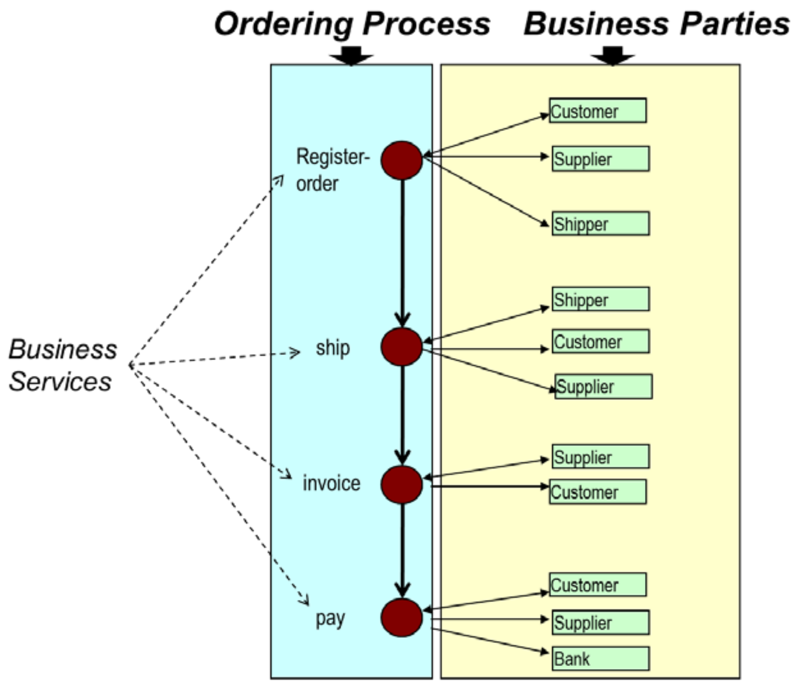


Figure 1.5 - Process vs. Services

It is obvious how the different granular (but coarse-grained) services involve different business parties.

In the past, often times information systems aimed at automating entire processes, which were often standardized throughout the organization. In the SOA approach to Business – ICT alignment, it is increasingly required that the ICT department is the provider (or at least the facilitator) of the information-based services in business processes. The problem with standardized processes (often in the way they are implemented in ERP systems) is the rigidity in the standardization, without sufficient consideration of exceptions and process variations.

Frequently, ICT-departments defend themselves by stating that they cannot handle all variations and exceptions, and try to bring "simplicity and order" in the business by means of standardized processes. Unfortunately, cybernetics argues how this approach may not result in sustainable systems. Ashby's Law of Requisite variety states that for a system is to be stable, the number of states of its control mechanism must be greater or equal to the number of states in the system being controlled, also known as the environment or the context of the system (Ashby, 2011). Ashby states that "only variety can destroy variety". In other words, a system can only survive if it has more variability than its context. Systems, and in particular, information systems that have less variability than their environments are doomed to be destroyed in the long run by their demand for variations and dealing with exceptions. And, it is the exceptions that cannot be ignored in today's economic environment: eliminating the "exceptional" orders in Amazon would destroy more than 30% of their turnover (Anderson, 2006). In other words: information systems cannot make the world simpler: in the best case, our expectation should be that they are able to deal with the complexity of the world "as it is".

Ashby's law of requisite variety is a very fundamental law for organizations. An implication of this law for business organizations is that they must not only develop sufficient information management, but also variability and decision-making capabilities in their systems to cope with the complexity in the environment in which they operate (Lewis, 2003). An examination of the services as building blocks in the ordering process from Figure 1.6 shows four sequential steps which, in this case, are the information-based services. Clearly, if this would be the only way to order products in an organization, this process would severely limit the flexibility of the customer to enable variations in the business process. Such instance may prohibit a customer from paying upfront and receiving an invoice before a shipment is made as in the book ordering protocol of Amazon.com. If, instead of automating the entire standardized process, the ICT department could merely be the provider of the four building blocks (the information-based services), the business would be enabled to develop alternative workflows (e.g., different ways to interface with the customer) combining the four services. The interfaces of services are reflected in the service contract that states under which preconditions service can be invoked. Of course, if the preconditions of services are broader, there can be more flexibility to configure and reconfigure them in business process variations. Atomary services, which cannot be further decomposed into "sub-services" are typically "independent" services that are not hard-coded to interface to other services. If services would be too tightly coupled together, they are indeed considered as one single service.

Variations could occur in different usage situations and conditions. Since the specific usage of a service and the usage conditions may change, the subsequent variations in service performance can influence the client preference for service as well as influence their satisfaction with the service. Sometimes variations in the process are favored by the client, while also too much variability might confuse the customer and create a negative impact.

In designing service configurations, the various sources and types of variability should be understood (Kannan and Proenca, 2010). For example, we see how clients are looking now for solutions in "the cloud" because they want to take complete (internal) abstractions of what they are doing. Some prefer the service bundles as completely virtualized systems. In such cases, the interest on the client side is severely focused on the service contract, which should include sufficient measures for data security, integrity and privacy considerations in the service cloud systems.

In conclusion, the design of information-based services as building blocks of business process requires them to be robust enough to take into account variations not only in usage, but also in variations in the competences of the client and provider side, perception of service complexity, client preference and satisfaction. Thus, exceptions and variations are a reality. Emerging technologies (such as Web 2.0) allow for additional flexibility in services, taking advantage of the two-way interaction that is enabled in those technologies (Anupam, 2010). In the next section alternative aspects of the meaningfulness of information-based services will be discussed in detail.

1.5 Value Creation and Value Leaks for ICT-based Services

ICT-based Services are the ICT-enabled building blocks for processes. These services should contribute some value for the client of the service, but also for the provider of the service. Apart from the client-provider interaction in a service, Rust (2006) elaborates the service definition in the context of the value that is created and captured. Value creation means to not only generate new value in elements where none existed before, but also to improve/augment existing values. Also, it can enable the parties involved in the services to detect/discover incremental values that may already exist but not perceived as such (e.g. trust). This implies that value creation on its own is insufficient; true value emerges only if it is also "perceived" as such. This illustrates the subjective nature of value, as well as the fact that "value" is not simply some "good" that is transferred from one party to another one.

There are various techniques to determine the potential for value creation in the service systems where suppliers and providers interact. A first view is to see a service from a supply side as a configuration, a system that is bundling technologies, resources and assets to meet the value proposition that is formulated in the service contract (Spohrer, 2007.) In the competence-based view, a system is a combination of not only the resources and assets, but it also includes the actions and decisions enacting on the strategic logic (Sanchez, Heene & Thomas, 1996). Clearly value is created in optimizing the configurations, providing not just a bundle, but indeed optimized systems addressing the specific needs of the demand side with a clearly motivated view on issues such as excess capacity (Ng et al., 1999), duplication and variations. An optimal configuration may try to optimize the resources (e.g. human resources that are involved) but might also attempt to minimize the stress levels of the human resources in the service system (e.g. the utilization rate of customer agents in a call center service). A good service configuration is intended to make use of its own choice and mix of technologies, resources and assets to improve its circumstances for both the providers and the users of the services (Cai, 2007; Vargo et al ,2008).

Vargo (2008) posits that service configurations are established through the proposition, acceptance, and evaluation of (perceived) value. Hotel booking services, for example, may realize value based on their competences, capabilities and reputation. Potential clients in need of such service make a decision to either accept or reject the proposition of value in exchange of money. This is the key of value-based pricing. Value can be made very explicit when the post-conditions in the service contract are very precise (as strong as possible). Value-based pricing will typically benchmark the potential financial appreciation of the post-conditions, in order to determine the acceptable service rates. Even in the case where the post-conditions are not met, there are potential opportunity values for the customer. Clients in this case can demand a refund or claim another type of compensation. In a hotel reservation service for instance, the hotel may opt to book you into a higher class category than when you made the initial reservation after determining that upon your arrival, all the rooms in the category you have earlier booked are. The higher class category room compensates for a potential loss of reputation. This shows that in value creation, the "opportunity values" as introduced earlier should not be ignored.
In the acceptance or rejection of an ICT-based service (where the information/communication constitutes the basis for the appreciation of the service) the resources and competences of the customer must be understood and optimally integrated to realize value. Meaningful post-conditions for one type of community may be meaningless for other communities. Service value theories typically make the distinction between "value-in-use" and "value-in-context" (Vargo, 2008). The "value-in-use" is what is perceived or experienced by the client, and in many cases, this usually happens after the purchase of the service. The "value-in-context" is made explicit in the formulation of invariants. In the same example of hotel reservation, the invariant can be "under the condition that no weather- related or natural catastrophes exist".

Co-creation of value between the suppliers and the consumers is not just a pure "exchange" such as in the economies of goods. A service configuration without suitable customers is meaningless, as well as a customer community that doesn't find the right service solution. Hence there are also issues of access, adaptation and integration versus (security) protection that contribute to the welfare created in a service configuration. These access issues not only play a role on the customer side. As a service supplier is building bundles of technologies, resources and assets, it is the challenge for the supplier to find the right technologies, resources and assets, in a network of co-suppliers, co-producers and pure suppliers of parts of service configurations. So a service provider is operating in a competitive network, both on the side of the suppliers for technologies, resources and assets as well as on the side of the consumers who may encounter alternative service providers with (perceived) similar bundles. This view brings together the two dimensions of the competence-based management view on competition: on the one hand there is the competition for the technologies, resources and assets, and on the other hand the market competition for services (Sanchez, Heene & Thomas, 1996).

In Figure 1.6 this value network is represented. In the competitive network the right technologies, resources and assets are acquired, rented or leased to bundle them in meaningful configurations for communities of customers.

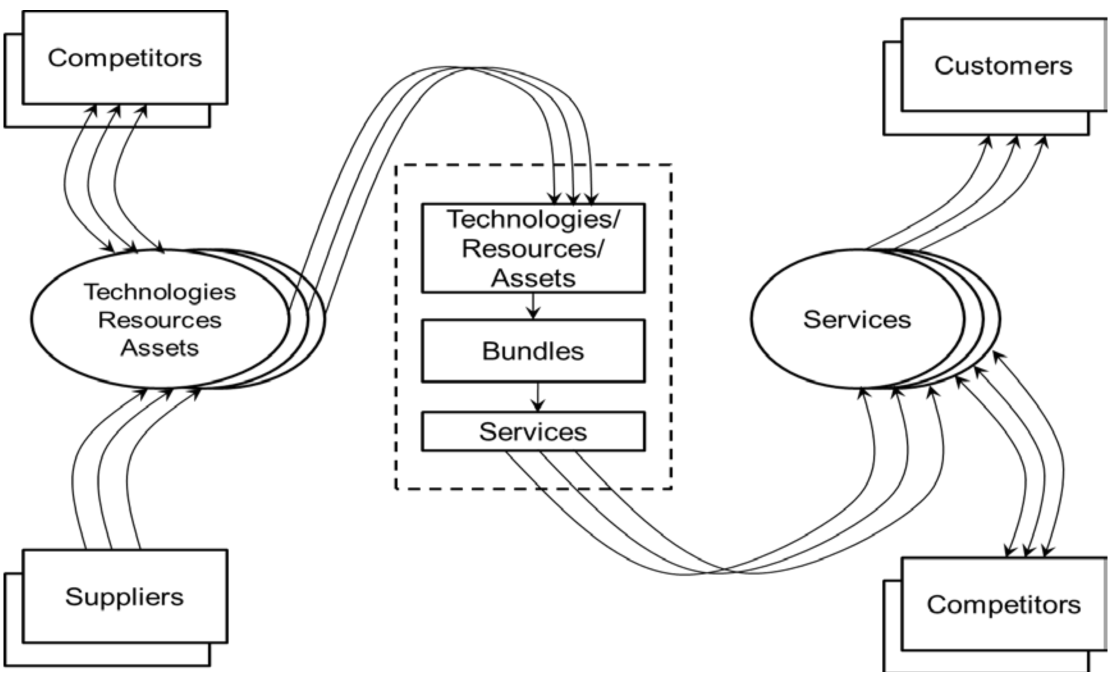


Figure 1.6 - Co-creation of value of interrelated service systems in a value network.

Value-in-exchange is another technique to create value in service interactions. In the case when there is an effective exchange of technologies, resources and assets between the service provider and the service customer, value-in-exchange is the conferred measurement of that degree of exchange. Customers may help to beta-test a new software service, for example, whereby they receive the software at better conditions (or even for free) and in return devote their time and expertise to the developer who provides the software service. Another example is the use of loyalty cards, where the customers exchanges private information on her/his behavior in exchange for loyalty benefits. Again, the degree of value-in-exchange that can take place is context dependent: one-time customers may not be interested at all in loyalty cards. Anyhow, in the value network, service providers try to "possess" their technologies, resources and assets, and moreover try to generate sustained added-value on top of the ownership of the elements in the bundles. When value creation sustainable, it is called rent (Mosakowksi & Kelvey, 1997).

Rent, means additional returns in excess of a resource owner's opportunity costs, which is defined as the revenue that a resource (bundle) can generate when put to an alternative use in the firm or the price which it can be sold or rented for (Grant, 1991). When firms consistently earn higher net profits in comparison to the industry average this indicates that the firm is indeed earning rents (Truijens, 2004). Sustained added value co-created in services is "*service rent*". It will become clear later in this thesis that service rent is highly influenced by the way in which service providers deal with the input and output elements in their value networks. Service rent can occur in the co-creation of value by means of a deep interaction with the clients. It can also occur in the co-creation of value by means of a deep integration with co-providers (co-suppliers).

There are various types of rent that can be explored and exploited by firms in order to earn rents. These can include but are not limited to monopoly rents, Ricardian rent, Penrosian rent Schumpeterian rent, and quasi-rent.

Monopoly rent results when the provider can dominate potential customers, by the exclusive ownership of some technologies, resources or assets. Firms can realize monopoly rents by exercising market power and raising barriers to competition (Truijens, 2004). Monopolies can emerge accidently, or by purpose. Monopoly rents stem from the firm's ability to somehow impede the competitive forces that tend to drive economic returns to zero by either the firm defending itself against competitive forces ('defensive' effects, e.g. as shown by Microsoft in workstation services), or to influence them in its favor ('offensive' effects, e.g. as shown by IBM in mainframe services) (Spanos *et al.*, 2001).

Ricardian rents occur when providers exploit temporal shortages. Such type of rent maybe achieved by owning a valuable resource that is (potentially temporarily) scarce (Mahoney et al, 1995). Building up Ricardian rents may for example be realized by creating "buffers" in the value networks to handle temporal shortages. An example of this is the current ICT off-shoring industry which substitutes lower labor costs of technical resources and infrastructure in Asia for costs in Western Europe and the United States.

When the firm employs better quality management, has distinctive competences, or just makes better of their technologies, resources and assets, then the firm has achieved Penrosian rents (Penrose, 1959). A firm may convey certain customer communities their trustworthiness in order to create Penrosian rent. Again, it must be noted that this may just only be a "perception" of trustworthiness, which nevertheless, allows the firm to charge higher rates.

When the firm has better risk management, entrepreneurial insight or stricter corporate governance, then it has achieved Schumpeterian rent (also known as "entrepreneurial rent") (Schumpeter, 1934).

Lastly, the quasi-rents are created when there is a high cost involved in making a particular bundle available to the others (Klein, 1978). A typical example might be a complicated service that requires a steep learning curve to be able to use it. In that case the service provider realizes additional value in providing the required education, adding that to the bundle. However, this also exposes a weakness, in that the technology requires an additional cost, and therefore exposes an opening for competition to offer better price and value to the customer through innovation or other means.

A firm can engage in various rent types simultaneously, if it is capable to understand the basic drivers for creating sustainable rent. In the next section, imperfections in service bundles will be identified as a major driver for service rent. Of course, service rent should not only be studied in a "positive" way: imperfections in service bundles may give rise to perceived value erosion, hence value leaks in services. Service rent may also consist of the "right" measures to stop value from leaking in services. Service improvements and innovations are additional ways to create service rent.

1.6 Service Bundle Imperfections

ICT-based services are defined as meaningful bundles of technologies, resources and assets where the bundles are able meet the perception of value by the client and can generate positive rents for the provider. Since a lot of research has been devoted to value creation in services, the focus here is on discovering, understanding and handling value leaks. The maxim that is put forward is that value leaks are primarily created by imperfections in the services, meaning imperfections in bundles of technologies, resources and assets in view of the service contracts.

Since ICT-based services are the economic majority amongst the services, it makes sense to follow the work of Stiglitz (1985) theory of information imperfections and the extensions that have been added afterwards by Truijens (2004). Hence three (3) categories of service imperfections will be considered:

1) Incompleteness of bundles of technologies, resources and assets
2) Asymmetry in bundles of technologies, resources and assets
3) Asymmetry of meaning for bundles of technologies, resources and assets

As a reminder, the concepts of incompleteness and asymmetry are fairly different. Incompleteness, in this context, means that some technologies, resources and assets are missing and the service is used to deal with this incompleteness, for example by detection of the missing elements, or even a potential realization of the missing elements in the service. These choice options will be discussed in the next section of this. Asymmetry on the other hand means that the technologies, resources and assets are effectively available but they are not symmetrically or equally distributed among the different parties involved in the services.

1.6.1 Incompleteness in Bundles

The increasing complexity of modern markets, alongside the broadening of the expectations of the customer‘s role in the co-production of services implies that typically information problems are more likely to occur. While some questions about services have concrete and available answers, many don't have the answers, leading to client confusion in areas where the right choice is imperative (Lindley, 2007). The most fundamental reason that markets with incomplete information differ from those in which information is complete is that, with incomplete information, market actions or choices convey information. All parties involved in the market interactions implicitly know this and behave accordingly (Stiglitz, 2000).

Incompleteness as earlier defined means that some technologies, resources and assets are non-existent, creating the opportunity for services to deal with these missing elements. Given the definition of services that is used in this thesis, there may be incompleteness for each of the elements in bundles that constitute services: there might be an incompleteness of technologies (in particular information and communication, or even knowledge related), incompleteness of appropriate resources as well as assets.

A first instance of incompleteness may occur when client requests may require services that are heavily dependent on a particular technology or the client being in the middle of development or deployment of a service may think of possible ways to enhance the service which might not have been considered before. As an example: A client desires a voice enabled service that would translate voice data into animated signs for the deaf, however the provider, having no such technology available, is confronted with the option to engage in such a service opportunity or not.

As an extension of incomplete technologies, incompleteness of information and communication is a very important category of incompleteness in service bundles. Even with the right resources and assets available, services may completely go wrong because of lacking/missing information. *Intelligence* is – by definition – an approach to find the knowledge gaps in an organization as well as possible ways to handle these knowledge gaps. Business Intelligence tries to identify missing business information, very often managerial information. The same definition applies for other intelligence services: workflow intelligence, process intelligence, police intelligence, and so on. One might argue that perhaps in some cases the data are already there, although this is misleading: the data might be there, but only the data that point the organization to the missing information. In most intelligence services, one of the first steps is indeed the identification of the missing information elements that must be collected, sometimes even created (Viaene & Van den Bunder, 2011). A typical example is police work, whereby often abundant evidence material is present for some types of crimes, whereas the precise definition (and interpretation) of the crime really creates new definitions (see Elzinga, 2011, for an instructive case on searching the right definition for domestic violence).

In the 2.0 society, the majority of information that must be handled is unstructured in nature: blogs, emails, SMS, tweets, video fragments, pictures, and so on. There is no apparent a-priori structure in these information elements, so again innovative techniques might be required to deal with information incompleteness situations (such as data discovery (Poelmans et al., 2011). Missing information elements may also include meta-information, such as patterns of (buying) behavior, preferences, workflow habits… Again, in police work, a lot of unstructured data is available, but it is the core of the police services to discover in the data patterns of communication. In research work on understanding Human Trafficking, it turned out that conversation patterns on “coffee” were actually the key to discover the relevant information to stop the trafficking gangs (Poelmans et al., 2010).

The incomplete information may include monitoring information, which is typically handled in monitoring services. In general, such types of services usually prevent downtimes or system failures and allow quick responses to problems as they arise. Example of monitoring (system) security monitoring, inventory monitoring, track and trace etc. Monitoring services can also be used for doing performance analysis, reporting on which quality of service the system has given the users and identifying those components that generate the most errors etc. An example would be bringing together customer experiences in hotels and restaurants through blogs/reviews to get a better perspective of the quality of the hotels/restaurants, in other words to know the value leaks in the hotels/restaurants. Just as in the hospital example discussed before, incompleteness of information prevails without appropriate feedback from the clients, which participate (as well as contribute) to the hotel/restaurant evaluation service. Very often, pricing and tariff structure deliberately deal with incompleteness of information. Clients of information services compare the opportunity costs of the lacking information with the value-based pricing rates they are potentially willing to pay for the service.

Again, incompleteness should not be confused with asymmetry: warning services for police enforcement are not creating new information that was missing before: the police know very well when and where they do the controls: so the information is available, but obviously not equally distributed amongst all involved parties. On the other hand, real time traffic information is clearly a service related to incompleteness of information.

Incompleteness can apply to all constituents of a service bundle, in particular to assets and resources. Typical examples for incompleteness in assets are borne from assets that are initially non-existent but are created through a request of such service. An example would be a 3D printing service (Materialize, 2011). Essentially such material is initially not yet present, however on a client request, a 3D design can be realized in a material that the client might prefer, thus creating the desired asset. From a client perspective, all kinds of services that try to do innovation-work or designs of technologies or materials constantly try to identify potentially missing bundles that may attract specific client demands. An example in the banking sector: banks are finding creative ways to meet the specific fund management demands of clients in order to allow them to invest their savings under the best circumstances. To do so, they sometime create customized versions of standardized banking and insurance services tailored to meet the specific needs and preferences of the client. Increasingly, pharmaceuticals see their business being transformed into the design of non-existing medicines, thereby inventing even new molecular structures. These are services where the providers basically create something that was not there before.

Incompleteness arises also in all situations where the individual is not capable to bundle the technologies, resources and assets in an optimal way (from various value perspectives). This is the main challenge for governmental services, which are supposed to bundle technologies, resources and assets for communities (such as regions, nations, municipalities, and so on) in a way that is beneficial to all parties involved. Again there is obvious co-creation of value: a good government service creates not only value for the citizens and business organizations under the government, but the service also creates jobs, better resource and asset allocation, and so on.

In general, any intermediate “channel” organization that is capable of gathering and bundling technologies, resources and assets in a unique, unprecedented way, exploits incompleteness of bundles from the client perspective. This applies to organizations such as travel agencies, party organizers, attendants for museums, and even in the ways that they can create unique "experiences", exclusive hotels and restaurants in reality fulfill incomplete bundles.

An essential element to support "experiences" is to build explicitly memory in the service bundles: in an experience driven restaurant, the "chef" remembers the customer, who is far from anonymous. The maître knows what the customer likes and dislikes, which – as explained before – anticipates potential value leaks: if you know the customer is a vegetarian, you will not recommend goose liver as the daily vegan specialty.

1.6.2 Asymmetries in Bundles

Asymmetries basically represent in-homogeneity of the distribution of (existing) technologies, resources and assets. In its simplest occurrence, which is asymmetry of information, it can simply be described as given the fact that different people know different things: the workers know more about their skills than the firm, the person who subscribes a health insurance policy knows more about his health than the insurance agent, the owner of the car knows more about the car than the potential buyer, the owner of the firm knows more about it than the potential investor, the borrower knows more about the risk and risk taking than the lender, and so on (Stiglitz , 2003).

According to Lindley (2007), information asymmetry exists when consumers have uncertain or (at least in their perception incomplete) complex information by means of which they make a decision. While there is a single way in which information can be perfect, there are many ways in which information can be imperfect (Stiglitz, 2002). Obviously, there are scenarios where one party in a transaction has more or more advanced information compared to the other. This might apply to both clients as well as suppliers in service interactions. Imperfections can occur on both ends. Obviously this may lead to imbalances in negotiation or transactions. It may even impact the notion of service contracts. Sometimes the client may propose a service contract, knowing better than the potential service supplier the probability to meet the post-conditions, or in how many cases the invariants effectively apply.

Again value co-creation happens. Clients as well as suppliers may try to benefit from developing opportunistic behavior (such as fraud), or may benefit from anticipating and preventing such behavior. The presence of asymmetry of information gives rise to value leaks: in an un-trusted service interaction, additional steps are needed for the validation of the client, the potential fraud scanning, and so on. In a business process they may represent additional, sometimes superfluous steps. Consider again the example of a Customer Contact Center. A lot of contact centers apply the policy to validate the client identity upfront early in the contact handling process. It might be understandable to have such a process step in the case the client is asking personal, private information. But for general questions (such as inquiring the opening hours of some office) client validation is *muda*: it is an unnecessary step in that particular variant of the call handling process. It creates unnecessary long handling times, increasing the call agent occupation, annoying the client who doesn't understand why the validation is needed in the case of such simple questions…

An important aspect of asymmetries is the fact that the technologies, resources and assets already exist, but are not homogenously distributed amongst all parties in the transaction. This goes beyond asymmetry of information: any situation where a provider has unique resources or assets that are not available to the customer creates the potential for valuable service interactions. Whenever the service supplier has unique skills, tools, competences available that are not readily accessible by customers, an asymmetry with a potential service interaction is created:

- Repairmen have the unique tools and experience that is not available in general to customers.
- Healthcare services develop clinical pathways to deals with pathologies in ways that are in general unknown to patients.
- Professional services (such as auditors, specialized consulting) provide to the customers missing skills and techniques.
- In sales services often sales persons don't disclose all information about the products they are selling (for example, there may be an excess stock or a newer version coming soon). This is not an incompleteness of information for the customer, since the information is effectively there. And the sales person doesn't lie: they just create an asymmetry.
- Some specialized food delivery services may serve the best and most fresh products in first instance to restaurants instead of individual customers, realizing in that case effectively asymmetries of assets in food related services. Of course this can only persist if the food ingredients that are provided are effectively "first-hand"…
- Travel agents may have access to unique, specialized reservation systems that are not available to the general public. They may prevent the client from losing significant time in searching the appropriate combinations of hotel, excursions, transportation and so on for particular traveling destinations on particular dates. It is again an example of co-creation of value.

It is obvious that there is a clear link with some forms of "service rent" that were discussed before: asymmetry exists whenever there are situations of monopolies on certain technologies, resources and assets, temporal shortages or differences in qualities on the provider side. (Mishra, 1998) elaborated the inability to assess quality represented in exposure to two different forms of supplier opportunism which are (1) the misrepresentation of an outlet's true characteristics and (2) actual quality cheating. These scenarios describe adverse selection and moral hazard problems, respectively.

From the provider's perspective, asymmetry means that the provider has no "access" to the technologies, resources and assets from the client side. This happens, for instance, when the client tries to hide or to show opportunistic behavior towards the service provider. It is the challenge for the provider to develop the right measures (for example in the service contract, or by including sophisticated technologies in the service bundle) to deal with this potential opportunistic behavior. The provider may force the client, for example, to use a particular technology to make use of the service, whereby this technology is only offered on a rental or license basis to the client. Any situation in which clients are forced to use particular technologies, resources or assets that are not exchanged by the provider to the client, but only rented, are examples of asymmetries in service bundles. The link with the notion of service rent is again obvious.

Asymmetries may even occur at community level. There are numerous examples of 2.0 communities that try to create deliberately asymmetries of information by not disclosing certain information, tools or assets unless one is a "validated" member of the community. Obvious examples are criminal networks, but also some specialized research communities.

1.6.3 Asymmetry of Meaning in Services

In this thesis a service was defined as a *meaningful* bundle of technologies, resources and assets. Although a service contract tries to make the "meaning" as explicit as possible, meaning remains something subjective in nature. Asymmetry of meaning occurs when different parties are giving a different (potential new) meaning/interpretation to the same bundle of technologies, resources and assets. Asymmetry of meaning is essentially an asymmetry in the semantics of the bundle that constitutes a service.

A first category of asymmetry of meaning happens when there is a lot of subjective interpretation possible in the bundle, making the bundles meaningful for one client, while giving another (or even no) meaning to other clients. Typical examples include leisure and entertainment services. The co-creation of value typically consists of changing the meaning of a bundle from being "less interesting" to become sufficiently interesting such that the customer effectively surrenders and engages in the service bundle. A successful leisure service must indeed seduce potential customers. Service providers will again try to create unique experiences, targeted at particular customers or communities, giving them (at least) the feeling of an individualized bundle. Education and training services provide another example. Apart from the asymmetry of resources and assets that is often present in training services, part of the service is indeed to change the meaning of class/training materials for the (potential) students.

A second category of asymmetry of meaning occurs in all situations where ambiguity is apparent in the presentation of technologies, resources and assets. Specifically, when information is not properly positioned in a particular context, there is a high possibility that it might be interpreted differently. Interpretations are by nature subjective and can be anchored on personal experience, personal settings, pre-conceived ideas based on what was previously heard or read and other contexts which may not necessarily be the same context that the other party has in mind.

This second category of asymmetry of meaning becomes increasingly relevant in ICT-based services. In the recent experiences with large scale usage of information, it became clear how "master data management" and "data governance" are insufficient to use the information and communication effectively throughout the firm. Information governance requires not only a uniform administration of the data, but also the full semantics, including the context and the correct interpretation of the data (Kooper et al, 2011). These are some typical examples:

- In a case on local government administration of agricultural administration, it was found that the term "cattle" had several incompatible interpretations in the legislation. From a data perspective, it could each time just be an INTEGER, but depending on the context the number could also be interpreted as square meters required to foster the cattle, liters of milk produced, and so on. It is obvious that an appropriate reporting on cattle becomes impossible without knowing the context.

- Actually, governmental legislation seems to contain in general a lot of asymmetry of meaning, which is of course largely in contrast with the intention of a legislation in general, namely to provide a general accepted textual representation of rules and procedures applicable to some community…

- Large scale organizations (production companies as well as financial institutes) are seeking increasingly to uniform their terminology used throughout the firm. Financial investment products may become so complex that there is an increasing risk for asymmetry of meaning. Proliferation of terminology in large- scale production companies may induce risks in the assembly and stock management services.

Again, in a perfect service economy, with perfect service contracts and bundles, no asymmetry of meaning is present. In reality however, an increasing emergence of asymmetry of meaning, stressing the need for data governance and information governance in general, is more than prevalent.

1.7 Problem Statement and Research Methodology

This section will outline the objective of this research effort, the framework that guided the effort and the research questions that are investigated. Lastly, relevance of this research will be addressed.

1.7.1 Research Objective

The research objective is the goal that is to be achieved through the process and completion of the research effort. As discussed in previous sections, traditional data mining and process improvement techniques such as Lean Management and Six Sigma often require augmentation (e.g., manual removal of anomalies from process models) in order to be fully capable of detecting value leaks and service imperfections when applied to ICT-enabled services. As such, these tools and methods have shown themselves to be insufficient to capture the required information to enable this level of analysis.

Research Objective: The objective of this research is to develop an appropriate set of methods, tools and techniques that can be used to automatically discover the root cause of value leaks and service imperfections in ICT-enabled business processes. A second objective is for these methods, tools and techniques to provide accurate and reliable guidance information for ICT-enabled business process improvement as defined as increases in productivity and/or quality.

1.7.2 Research Framework

The research framework followed in this Thesis is shown in Figure 1.7 It is a schematic diagram of the steps taken to achieve the research objective (Verschuren and Doorewaard, 2000).

The research model can be explained as follows:

The initial level of understanding started with a combination of methods and technologies (D1) developed at OpenConnect, along with field experiences gained during customer engagements (D2), and a review of literature on process mining (D3), automated process discovery and business process representation (D4).

Careful examination of these materials revealed both the shortcomings in current methods and tools for service process improvement as well as significant issues that needed to be addressed in order to achieve a reliable understanding of the actual root causes of service process inefficiencies and imperfections through automated means. Specifically, this required research on the applicability of a service-oriented architecture approach to business processes (C1) as well as a new approach to process discovery that included aspects of knowledge discovery (C-K Theory) (C2). Since this was also to be an improvement in the application of automated methods, research was required for data representation methods (such as formal concept analysis) and their applicability for the appropriate representation and understanding of processes (C3) as well as the relevance of current Process representation techniques (C4).

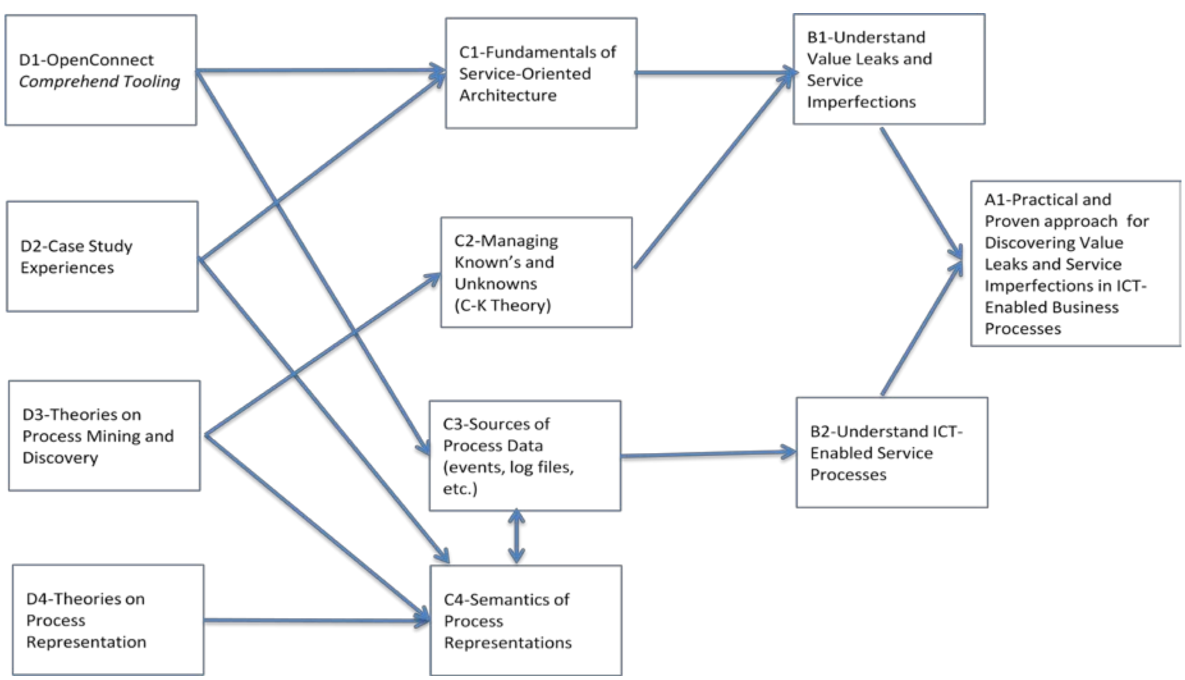


Figure 1.7 – Research Model

This research is embedded in the specific context of ICT-based Services as they are used in Business Processes. This required a fundamental examination of the notion of ICT-based

Services (B2) as well as a exploration of the applicability of state of the art techniques for dealing with unknown unknowns (B3).

The final research results provide management with a practical, accurate and reliable set of methods, tools and techniques that can be used to target value leaks and service imperfections in ICT-enabled business processes when attempting to improve service productivity and effectiveness (A1).

1.7.3 Research Questions

The research questions describe a set of issues that the research will answers in the course of the effort. The main question that will be addressed is:

"How, through the appropriate framing of methods, techniques and tools, can the root cause of value leaks and service imperfections in ICT-enabled service processes can be discovered so as to give guidance to management on increasing service productivity"?

As part of the process, we will explore the following related research questions:

"How can process models be improved so as to more accurately reflect the underlying reality of the events being represented"?

Mainstream process modeling techniques often result in representations that are not actionable either for analysis or improvement efforts. Specifically, by their intrinsic nature, process/event log data may result in models with anomalies regarding the unclear nature of relationship occurrences as well as the actual behavior of the data itself. This suggests that research in mathematical methods of process normalization is required to ensure accurate event representation at the atomic level.

"What type of trade-offs between volume of data and mathematical accuracy are required are required when working with real-life volumes of process log data?"

High volumes of process log data require commensurate scalability of discovery techniques and tools. This requires finding the correct compromise between absolute mathematical accuracy and sufficient applicability of the proposed techniques and technologies.

1.7.4 Research Approach

This research is based upon gathering data based upon automated observations of a real-time work environment. Based on this, the case study approach has been chosen as the most appropriate method to guide the investigation. Robert Yin describes three types of case studies; the descriptive case study, the explanatory case study and the exploratory case study (Table 1.2) Yin (2003). Yin explains that a case study approach is best used to answer "how" and "why" questions.

METHOD	(1) Form of Research Question	(2) Requires Control of Behavioral Events?	(3) Focuses on Contemporary Events?
Experiment	how, why?	Yes	Yes
Survey	who, what, where, how many, how much?	No	Yes
Archival Analysis	who, what where, how many, how much?	No	Yes/no
History	how, why?	No	No
Case Study	how, why?	No	yes

Table 1.2

According to Yin (2003), this PhD research can be qualified a holistic multiple case design approach. The unit of analysis is the search for value-leaks in ICT-based Services and Processes, for which multiple cases studies have been developed. The approach is holistic as there is no restriction of the research to a particular type of measurement or value leak.

To ensure sufficient variety for the cases, and to be compliant with the understanding of knowledge work as developed in this first chapter, the business processes studied in the cases are of two different types; administrative processes and caring processes. Administrative processes are focused on increasing the efficiency of both the resources involved in the processes and the overall processes itself. The two specific cases to be reviewed involve healthcare claims adjustments and healthcare claims processing costs. Caring processes are those that involve a high level of customer interaction. The three cases studies include two customer contact center situations; the first involving a multi-national bank and the second, a governmental agency. The third caring case involves the caring process that supports a hospital post-surgical breast cancer unit (Figure 1.8).

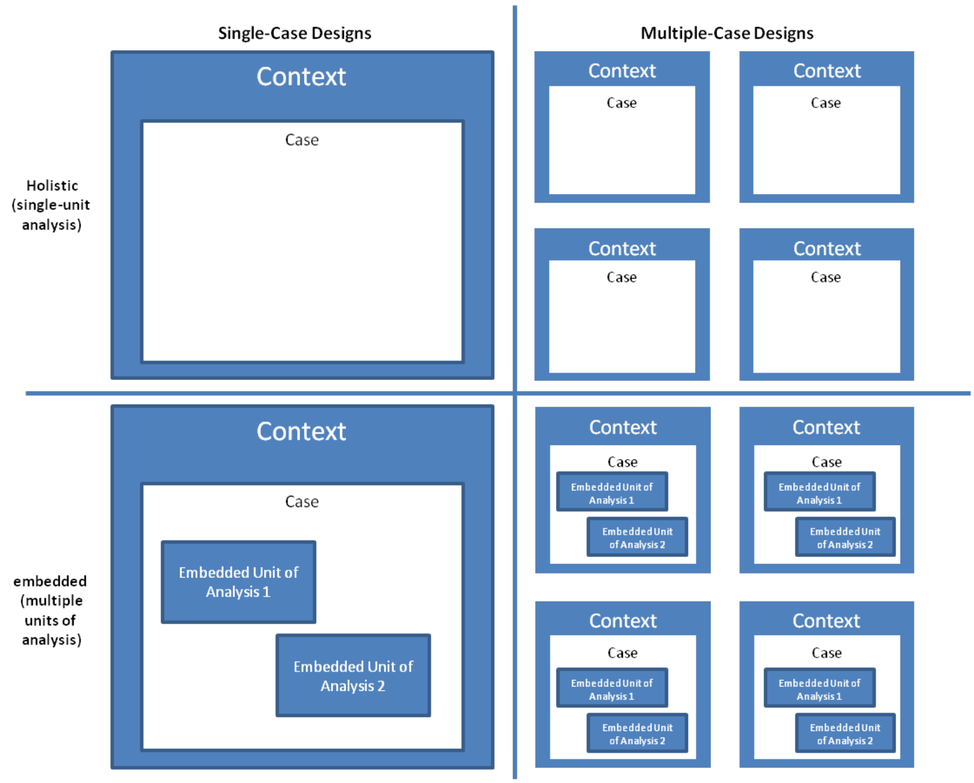


Figure 1.8 – Case Study Results

1.7.5 Research Relevance

A central premise of this research is that just as in Fredrick Taylor's day, it is the systematic and scientific lack of knowledge about the current work process, what refer to as *knowledge gaps*, and the specific nature of service-based work itself (service imperfections) that are the keys to solving the productivity issues faced by management. Additionally, an increasing amount of the actual work in services is ICT-enabled (e.g., digital), and has become largely invisible (George, 2000). This diminishes the capability of popular methods and techniques (e.g., six sigma, lean management) to provide actionable information about the underlying service process and the root cause of its productivity issues. This research aims to provide an appropriate set of methods, tools and techniques that can be used to reliably provide information on the performance of an ICT-enabled service/business process that can be used to guide management in its improvement.

The weaknesses of current improvement methods lie in four specific areas:

- ***Data collection based on manual methods***- automated process discovery, when modified and applied to ICT-enabled business processes, delivers accurate observations on ICT-service work as it actually occurs rather than on derivative methods such as manual activity logs which are based on what one records or remembers. This eliminates human error as well as inconsistencies caused by non-calibration among recordings/recorders.

- ***Analysis techniques based on a-priori biased models***- including hierarchical representations (e.g. knowledge trees), time series and regression models, which allow majorly the detection of first order linear relationships. In this thesis models that allow a maximum of semantic expression in conjunction with the required mathematical representation are explored, resulting in sound results based on lattice structures for concept structuring, hidden Markov models for process structuring and using the services as the primary symbols to analyze.

- ***Event representations that do not allow for detection and removal of anomalies-*** many of the representation techniques used today (e.g., Petri-nets) accommodate anomalies in the actual representation of the process. This issue results in a lack of actionable information since the precise cause root cause of any issue cannot be precisely be determined. As a result, productivity improvement efforts are often hampered. Improving this requires enhancing to the underlying mathematical techniques used for process representation.

- ***Service imperfection understanding that does not lead to actionable improvement recommendations-*** a method of categorizing discovered imperfections in ICT-based services that can be used as a guide to both understanding the underlying issues and possible improvement approaches from an organization and management perspective does not exist. The relevance here is from both an operational improvement perspective as well as an increase in additional management competencies required in an ever increasing ICT-enabled sector of the global service economy.

As previously mentioned, we will use the case study method to test the proposed methodologies, tools and techniques that address the identified issues.

Further relevance of the research presented in this thesis may be reflected in the fact that major parts of this thesis work have been published in Academic Journals and Conferences, as well as some White papers and Conference Presentations. These are the relevant publications for the chapters of this thesis:

Chapter 1:

Mercado, C., Dedene, G., Peters, E., Maes, R. (2012), Towards a Dynamic, Systemic and Holistic Theory for Strategic Value Creation in ICT-based Services, *in " A Focused Issue on Competence Perspectives on New Industry Dynamics", ed. R. Sanchez & A. Heene, Research in Competence-Based Management, Vol. 6, pp. 153 – 207, Emerald Group Publishing Limited, 2012.*

Peters, E.M.L., (2011), Workforce Productivity in Today's Service-Based Economy, Innovation Nation at Risk, Tech America Foundation, September, 2011, (http://www.techamericafoundation.org/content/wp-content/uploads/2011/09/OpenConnect090111.pdf

Chapter 2:

Peters, E.M.L., G. Dedene, C. Houck (2007). Hybrid Techniques for Business Process Discovery, *Proceedings of the INFORMS Annual Conference 2007.*

Peters, E.M.L., G. Dedene, C. Houck (2009). Business Process Discovery and Workforce Intelligence Techniques with Healthcare Applications, Proceedings of the INFORMS ORAHS 2009 Conference, Leuven Belgium 2009, http://www.econ.kuleuven.be/eng/tew/academic/prodbel/ORAHS2009/page5.htm

Peters, E. M., & Dedene, G. (2011). Business Process Discovery & Workforce Intelligence Techniques in Healthcare, *International Journal of Health Management and Information*, vol. 2, no. 1, pp. 25 - 39.

Peters, E. (2012). Processes are Concepts, aren't they ? *Invited Keynote for the 10th International Conference on Formal Concept Analysis ICFCA 2012, Leuven.*

Chapter 3:

Fleming, M., J. Silverstein, B. McDonough (2009). *OpenConnect Comprehend Work Event Monitoring and Analytics Provide End-To-End Process Transparency, IDC* Whitepaper

Chapter 4:

Fleming, M., J. Silverstein (2009). *Healthcare Insurer Uses Comprehend to Improve Claims Processing Productivity,* IDC Whitepaper

Peters, E.M.L., Cicchino, John (2011), Driving Administrative Cost Reductions by Improving Service Processes and Workforce Productivity, AHIP Fall Forum, November 15, 2011 Chicago, IL.

Ghanayem, D, Peters, E. (2012), Effective Strategies for Streamlining Administrative Processes to Meet MLR Requirements, AHIP Institute Conference, Salt Lake City, Utah, June, 2012.

Sinur J., (2013), *Using Process Mining for Increased Profitability on Insurance Claims: Success Snippet*, Gartner Research Blog Network, March 2013

http://blogs.gartner.com/jim_sinur/2013/03/11/using-process-mining-for-increased-profitability-on-insurance-claims-success-snippet/

Chapter 5:

Houck, C. (2007), Breaking through with Business Process Discovery, Business Intelligence Journal, Vol 12 (4), 9 pp. 2007

Poelmans J, Dedene G, Verheyden G, Van der Mussele H, Viaene S, Peters E, 2010, Combining business process and data discovery techniques for analyzing and improving integrated care pathways, in *Lecture Notes in Computer Science, Advances in Data Mining. Applications and Theoretical Aspects*, vol. 6171, Industrial Conference (ICDM) (Leipzig (Germany)), pp. 505 - 517,

Poelmans, J., Dedene, G., Verheyden, G., Peters, E., Van Der Mussele, H., Viaene, S. (2010). Concept mining and discovery in clinical pathways. INFORMS Data mining and health informatics (DM-HI) Workshop. Austin, Texas (US), 7-10 November 2010.

Peters, E.M.L., Payack, P.J.J. (2010), *The Paid for Option*, *Bending the Healthcare Cost Curve*, Tower Oaks Press, 2010, Dallas, TX.

Peters, E.M.L.,(2012), How to Slow the Healthcare Value Leak, Forbes, May 7, 2012, (http://www.forbes.com/sites/realspin/2012/05/07/how-to-slow-the-health-care-value-leak).

Peters, E.M.L., (2012), Greater Healthcare Payer Efficiency Through Process Intelligence and Workforce Analytics, Insurance and Technology, January 2012,

(http://www.insurancetech.com/business-intelligence/greater-healthcare-payer-efficiency-thro/232301494

1.8 Overview of the rest of this thesis

This chapter has addressed ICT-enabled services and discussed how they are an integral component of the knowledge and data-work environments. It has also surveyed the major approachs used to analyze and improve service-based work and included a new entry, automated business process discovery. ICT-enabled service work was then discussed within the framework of emergence and Enterprise 2.0. Additionally, this section discussed the major topics of value creation and value leaks. Value leaks and their discovery is a major work of this PhD thesis. Services were then discussed as the building blocks of business processes with a comparison to the principles of system oriented architecture (SOA). This included the concepts of granularity as well as a discussion of pre and post condition servicing.

This Chapter concluded with a discussion of service value as created through strategic value configuration. This included the competence-based view in which a system is a combination of not only the resources and assets but also includes the actions and decisions enacting on the strategic logic. Additionally, the concept was established that service configurations are established through the proposition, acceptance, and evaluation of perceived value. This requires a not only a consumer and provider but also a network or ecosystem involved in the value creation process.

Service bundles were then examined as to their possible imperfections. Three major types of imperfections were discussed. Specifically, they are:

- Incompleteness of bundles of technologies, resources and assets
- Asymmetry in bundles of technologies, resources and assets
- Asymmetry of meaning for bundles of technologies, resources and assets

These categories will form the basis of the type's service imperfections that will be automatically discovered and quantified as to their associated value leakage. This will provide actionable information to guide management in this service- business improvement effort.

This chapter also discussed the problem statement guiding this research as to the research questions, research model and the approach to the entire process. Finally, the relevance of the research was described as to its overall contribution.

Chapter 2 will discuss the mathematical methods for ICT-enabled service- process and value leak discovery. It is explained why the currently used Petri-Net Process representations are insufficient for the research goals, and a deeper understanding with Hidden Markov Models is required to find value leaks in Processes and Services (seen as Process steps). Apart from small enhancements to the current uses of Markov Models, the notion of Normalized Process Models is introduced. In the same manner as data normalization emerges from concept analysis and deals with data anomalies, process normalization aims to prevent ambiguities in the representation and understanding of processes.

Chapter 3 discusses the tooling, process and architectural components that are required to implement an automated business process discovery centric approach in an ICT-enabled service work environment. Here the technologies and the requirements for scalability based on actual data volumes will be described. The issues and requirements for the non-invasiveness of the collection methods, so as not to degrade the performance of an actual production environment, will also be discussed. It is important to note that this chapter also includes an embedded case study to show an example of the outputs and functionality of the tooling.

In Chapter 4, the first set of case studies, administrative processes, is discussed. Administrative processes are focused on increasing the efficiency of both the resources involved in the processes and the overall processes itself. The two specific cases to be reviewed involve healthcare claims adjustments and healthcare claims processing costs.

In Chapter 5 discusses caring processes. Caring processes are those that involve a high level of customer interaction. The three cases studies include two customer contact center situations; the first involving a multi-national bank and the second, a governmental agency. The third caring case involves the caring process that supports a hospital post-surgical breast cancer unit.

Conclusions of the research are discussed in Chapter 6 as well as recommendations for future research.

.

CHAPTER 2

Discovery Techniques for Processes

2.0 Process Discovery Techniques for Value Leak Assessment

In the first chapter it was argued how “intelligence” techniques augmenting the human understanding of Processes are needed to assess Value leaks in ICT-based processes. The guidelines of Engelbart will be followed to find out the appropriate techniques that enable process understanding. According to Engelbart, the underlying principles for techniques that augment Human understanding are based on:

- Understanding the right “symbols”
- Deriving concepts and concept structures from these symbols
- Put all of this in the right process structure.

In this chapter, the (meta-) process structure that drives Process Discovery is C-K theory [Hatchuel et al., 2006]. As will be explained more in detail later in this chapter, this theory emerged as one of the “rare” process models for guiding the innovation process. Indeed, many publications address quantitative aspects of Innovation, but little is known on “how to arrive at an innovate idea”. C-K-theory was – so far – successfully applied to industrial processes, similar to the Taylor approach.

In this thesis (in conjunction with the work of Poelmans and Elzinga) C-K-theory is specialized and applied to ICT-based processes. It is shown how the original C-K theory, as it was developed in Paris, can be extended with Formal Concept Analysis for the conceptualization step in the C-K theory. The original theory uses knowledge trees, which turn out to be too limited for a semantically rich conceptualization. So, in the first part of this chapter, the use of Formal Concept Analysis to conceptualize ICT-based processes is explored.

In the next part, complementary process representations are discussed. A vast majority of the current literature is concentrating on variations on Petri-nets (UML Activity Diagrams, the Business Process Modeling Notation BPML, various extensions to Petri-Nets). In this chapter it will become clear that Hidden Markov Models (HMM) are more suitable for Business process discovery purposes. A number of cases showing the advantages of HMMs over Petri-nets, at least for discovering potential value leaks in ICT-based processes. Observe that it is not the intention to create a negative impression on Petri-nets: it is a most valuable instrument for presenting and discussing Business Processes with Business Professionals (and management in particular). Hidden Markov Models may scare some people because of its mathematical nature.

Next, the notion of Process normalization is defined and discussed. Non-normalized process models may contain ambiguities and can mislead process mining algorithms in such a way that erroneous results are obtained. There are two ways to arrive at normalized models:

- Sequence clustering, which may become tedious for (very) large data sets
- Transformations to higher order transitions, a fairly simple algebraic technique, which may provide the same results in a simpler way.
-

At this point the compromise between absolute accuracy and the potential to process big data comes in place. When processing tens of millions of event data, first order transition based models, even being not normalized, may already be sufficient to find the major evident value leaks. Next normalization can be used to discover further details.

By the end of this chapter it becomes clear that there is no unique golden bullet technique for the complex issue of Process Discovery. A hybrid approach, combining the forces of alternate techniques may lead to the best results. In this way, this chapter is a kind of cook book, providing Process scientists with an overview of the relevant discovery techniques, illustrated with examples.

2.1 C-K Theory for Process Discovery

Process Discovery is Knowledge intensive work that requires augmentation to the human intellect. Innovative approaches are needed to develop these augmentations. As explained in chapter 1, the pioneering work of Engelbart provided the necessary ingredients for augmentation. In this chapter the ingredients for Process Discovery augmentations will be developed. Following Engelbart, the following is required:

- *Understanding the "right" symbols.* In the case of Processes, the primary symbols are event logs. When these logs are in an appropriate auditing format (which will be called *Actor-Event-Object* format later in this chapter) they become suitable for Process Discovery. Of course additional input can come from (informal) process descriptions, training materials, documentation sets, … For that type of input, Data Discovery techniques [Poelmans, 2010] can be applied in addition to Process Discovery.

- *Concept structure.* The notion of a "concept" is precisely defined in Formal Concept Analysis FCA. Hence it is straightforward to try to apply FCA as a primary conceptualization technique. Although this gives already interesting results, the FCA concept lattices must be complimented with other Process representation techniques. In this chapter Hidden Markov Models are used for this purpose.

- *Process structure.* Engelbart refers not only to the appropriate representation for process structure in general, but also specifically to the *innovation process itself.* It is not straightforward to find publications on "how to achieve innovations" in a methodological way (in contrast to the many software development methodologies, for example). Our attention was drawn to a series of publications from ParisTech, proposing a guideline for innovation based Knowledge Spaces versus Concept Spaces, referred to as the **C-K-**

Theory [Hatchuel et al., 2006]. Inspired by the successful application of this theory for Data Discovery purposes, with successful case studies in Police Intelligence Services [Poelmans et al, 2010], this theory will be used as the process structure for Process Discovery.

The original C-K-theory is using Knowledge spaces, where the current and future knowledge is located, and transformations form Knowledge spaces into Concept spaces ("Disjunction") and from Concept spaces into Knowledge spaces ("Conjunction"). The underlying idea is that the disjunction should transform the current symbols into concept spaces, which should be structured in such a way that they allow to detect potential knowledge gaps, and ways how to deal with these knowledge gaps. New the newly discovered, augmented, adapted or (re-)designed concepts result by conjunction in new knowledge. Of course this new knowledge only becomes effective if it is recognized by the relevant domain experts as true innovative knowledge (in contrast to "trivialities"). This whole process structure can be summarized as follows:

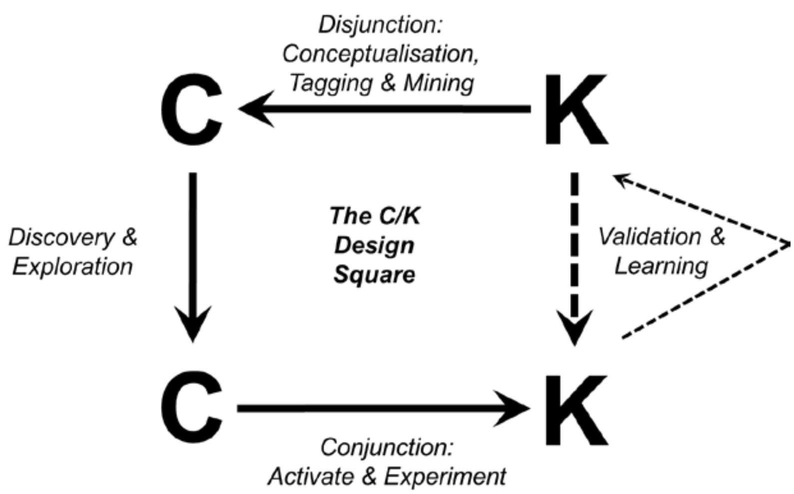


Figure 2.1 - Overview of C-K-Theory

A non-trivial element in applying C-K-theory is the choice of the "right" representation in the concept space. If the purpose is to discover knowledge gaps in the current situation, two elements are crucial for appropriate concept spaces:

- A concept space must have an ***open, semantically rich structure***. In the original C-K-publications most often hierarchical concept trees (also known as knowledge trees) have been used. In practice, a *hierarchy* is a fairly poor, restricted semantic ordering of concepts. The history of databases and data structures sufficiently illustrated the limitations of hierarchical representations. The limitations of hierarchies were already stressed by architects such as Christopher Alexander [Alexander, 1966]. Formal Concept Analysis is using partial ordering relations (lattices) which allow for a much richer semantics than hierarchies. This freedom is also required to deal with "unknown unknowns", whose structure is a priori unknown.

- Next, a "good" concept space should be ***free of potential ambiguities***. Examples of such potential ambiguities are many-to-many relations, optional-to-optional relations, many-way-relations… In the work of Poelmans it became clear that this issue is intimately related to the notion of ***normalization*** [Poelmans, 2012]. He demonstrated how FCA automatically leads to normalized data representations, extending the notion of database normal forms. In this chapter, the notion of normalization for process representations will be investigated. When first order representations of processes contain portions that require further normalization, this often reveals value leaks in processes! Furthermore, the usage of parallel constructions (fork and join) in Petri-Nets creates unnecessary ambiguities in these process representations. Again, let there be no misunderstanding: these are not ambiguities in the visualization of processes for managerial presentation purposes, but potential ambiguities for the quantitative analysis of the processes in searching for value leaks.

This shows how the original C-K-theory is specialized for Process Discovery purposes (Figure 2.2):

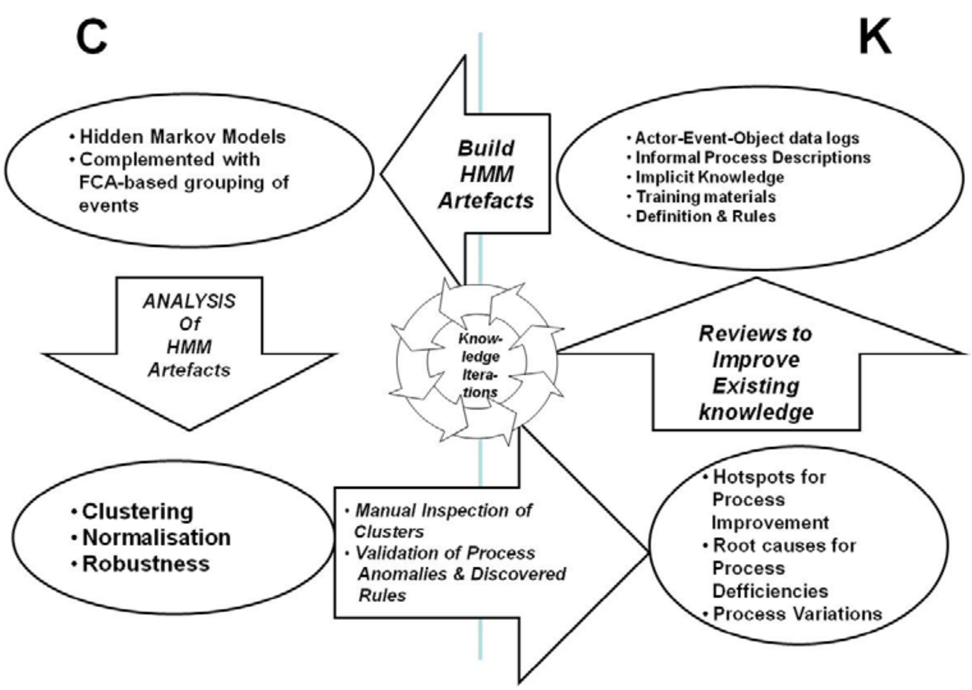


Figure 2.2 - Extending C-K-Theory for Process Discovery

This is the process structure that is followed not only in theory, but also in practice for Process Discovery. In the case studies various alternative examples of value leaks will be discovered in two major ICT-based process types: transaction (administrative) and interactive (caring) processes.

2.2 The start of Process Discovery

Process Mining and Process Discovery [Cook and Wolf, 1995] are both techniques that try to gain an *understanding* of the as-is processes, and in particular an understanding of the factual interactions between a workforce and information systems services which are the building blocks for the processes on the work floor.

Process Mining is a statistical approach to understanding the "general" behavior in Processes, in the way they are observed on the work floor. The statistical techniques in Process Mining often concentrate on finding the “main” or “average” process. This was illustrated in the main case study in [Maruster et al. 2006], where it was revealed that the statistical technique *“is able rather to capture the general process model rather than the process model containing exceptional paths”*.

Process Discovery tries to go further than Process Mining, by explicitly addressing process variations and process exceptions. In Process Mining, exceptions are often considered as nuisance information in the statistical analysis, and eliminated as “noise”. Process Discovery is incorporating the exceptions and variations, if needed (such as, for example, the distinction between experienced senior versus junior contact agents and their related process variations in the second contact center case study in chapter 5).

The term “Process Discovery” was first used at the end of the nineties by Jonathan E. Cook (New Mexico State University) and Alexander L. Wolf (University of Colorado) in their famous paper on “*Discovering Models of Software Processes from Event-based Data”* [Cook & Wolf, 1996]. Although their main interest was in SOFTWARE processes, it became rapidly clear how their work was equally applicable to workflows and modern Business Processes (which are of course increasingly supported by software services). They were the first to give a systematic overview of techniques that make it able to derive (in an automatic fashion) a formal model of a process from basic event data (collected on the actual process executions).

Their work laid the foundations that were significantly deepened in subsequent research. This is an overview of what they did put forward:

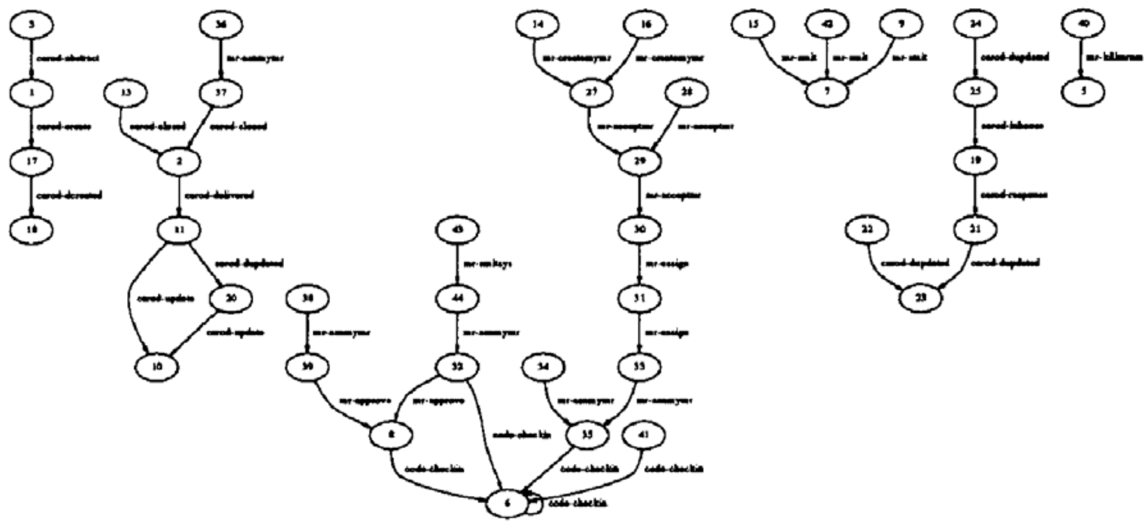
- Techniques for discovery may vary from purely algorithmic (and qualitative) to purely statistical (and quantitative). The families of techniques discussed by Cook and Wolf are:

    - Backward learning, whereby trying to characterize a current event in the process by a (mainly statistical) analysis of previous events. Neural Networks are highly suited for this (and have been deployed later in more sophisticated forms as techniques for emergent Data Discovery, such as Emergent Self Organizing Maps [Poelmans et al., 2010]. Drawbacks in these methods are sometimes the lack of hints for the optimal values of parameters, such as optimal number of layers in the neural networks.

    - Forward learning, whereby trying to characterize a current event in the process by (mainly algorithmically) analyze the future events. This involves the transformation of the events and transitions into first-order and higher-order transitions. An important question here is also: what is the “optimal” higher order length that is required.

- Markovian approaches, which combine in a hybrid way algorithmical as well as statistical techniques, with well-known machine learning techniques such as the Expectation Maximization Algorithm. Variations of this algorithm combine higher order transitions with their Markovian probabilities. Major challenge is to find on the one hand the optimal higher order, and on the other hand the maximal threshold value for the probabilities (meaning, up to which value the probabilities can be ignored).

- All techniques involve in some ways “looking back” or “looking forward” up to some higher order “length” and try to cluster in an appropriate way. It is not clear a priori how many clusters will be needed. Cook and Wolf (1999) present in a subsequent publication [Cook & Wolf, 1998] a model in which only one-to-many or many-to-one transitions are present in the clusters.



The one exception that seems to occur is state 6, in which there is a potential many-to-many transition, which can (according to their discussion) be ignored basically because it is an end-state in the model, and contributes nothing fundamentally new to the process model. It is this example that has inspired the definition of normalized process models in this thesis.

The decision rule that seems to emerge from their work is that the analysis can start from low thresholds and short lengths, and gradually drive up the thresholds and lengths until the model becomes normalized.

- Finally, Cook and Wolf were the first to reveal the deep connection between behavioral aspects of processes and the structure of ***languages*** “spoken” in terms of the events by concepts (objects) involved in (and participating in the processes).

Languages can be expressed as algebraic regular expressions, finding their structure in idempotent semi-rings. These viewpoint was partially developed parallel to their work in [Snoeck & Dedene, 1998] and later on elaborated in [Poelmans 2012].

2.3 Input data for Process Discovery

A major discussion element in the work of Cook and Wolf is their input data. They start from “anonymous” sequences on events, just mentioning the events irrespective of the “actors” that are participating, or the “objects” that are involved in the processes. A typical such sequence might be as shown in Figure 2.3:

```
RCERCECRECRERCECRECRERCECRERCERCECRECR
```

Figure 2.3 - A typical anonymous event sequence

The anonymity of the sequence makes it more difficult to find the most realistic starting and ending point for a single relevant sequence that represents a process. Incorporating Actors and Objects can make this more straightforward, by assuming, for example, that all events around one particular “object” constitute one instance of a process. Of course, if objects not the “key id” for individual process instances, the problem still remains. Only very recently, one approach for dealing with anonymous event sequences with actor and object ids has been proposed [Sundaravaradan et al., 2012]. Although their interesting results based on a clustering approach of subsequences, there is the non-trivial problem to verify that the discovered subsequences are actually valid – really occurring – processes or process fragments.

Auditing requirements however very often require that individual log records in database operations should have an ***individual*** responsible, a responsible Actor. That is why in practice the following generic input model has been used throughout this work to feed the Process Discovery mechanisms (Figure 2.4):

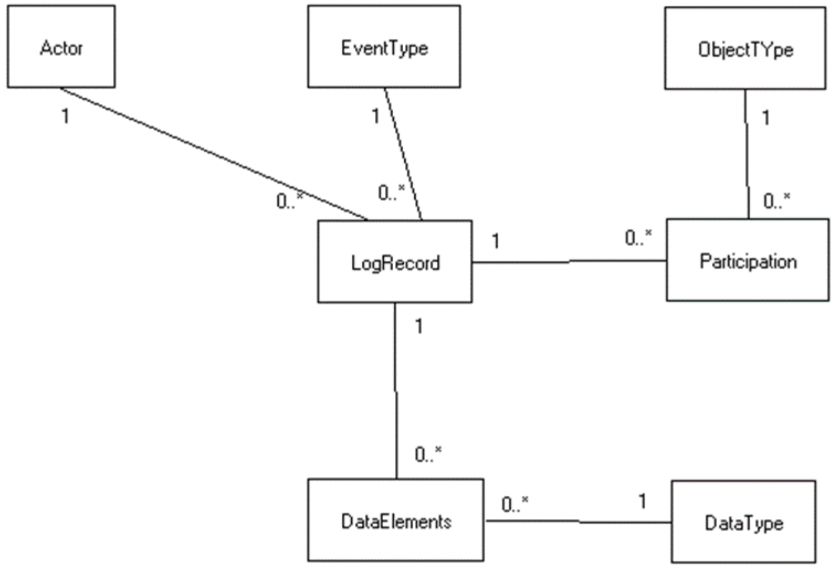


Figure 2.4 - A UML Class Diagram for the Event – Actor – Object Data Format

This is a UML-based class diagram description of the input datasets. Intuitively the input structural relationships are the following.

Each log record represents an instance of an Event Type, executed by a (responsible) Actor, and involving (potentially multiple) object type instances. Each log record may refer that various data elements of interest (however most often related to the value proposals that are challenged in the discovery studies), which could be:

- Time / time stamps
- Costs
- Quality indicators
- Satisfaction feedback
- …

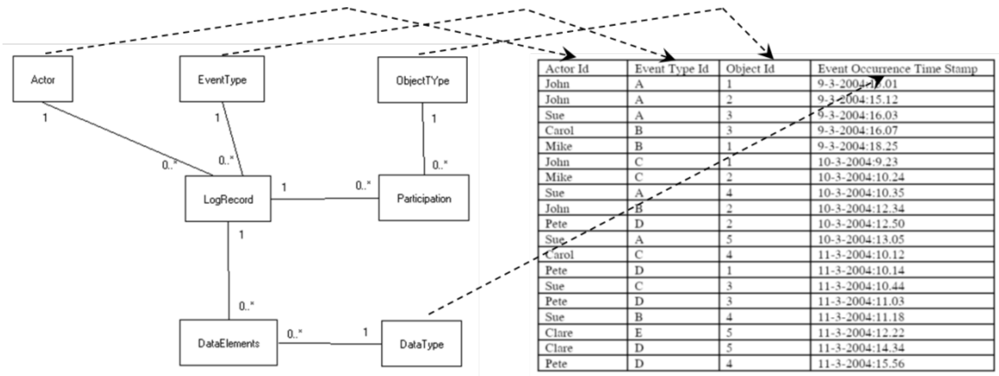
These data elements may turn out to be essential to check – for example – the pre- and post-conditions for services, to verify whether the service contracts are realized. In one of the case studies, an initial lack of information (information incompleteness) which was solved by Process discovery enabled to detect deeper (and actually fraud related) Value leaks.

An instance of such input log could be the following small but non-trivial set, which is also recurrently used for illustrative purposes in most of the publications of [Van der Aalst 2006], related to the Petri Net representations. (Table 1.1)

Actor Id	Event Type Id	Object Id	Event Occurrence Time Stamp
John	A	1	9-3-2004:15.01
John	A	2	9-3-2004:15.12
Sue	A	3	9-3-2004:16.03
Carol	B	3	9-3-2004:16.07
Mike	B	1	9-3-2004:18.25
John	C	1	10-3-2004:9.23
Mike	C	2	10-3-2004:10.24
Sue	A	4	10-3-2004:10.35
John	B	2	10-3-2004:12.34
Pete	D	2	10-3-2004:12.50
Sue	A	5	10-3-2004:13.05
Carol	C	4	11-3-2004:10.12
Pete	D	1	11-3-2004:10.14
Sue	C	3	11-3-2004:10.44
Pete	D	3	11-3-2004:11.03
Sue	B	4	11-3-2004:11.18
Clare	E	5	11-3-2004:12.22
Clare	D	5	11-3-2004:14.34
Pete	D	4	11-3-2004:15.56

Table 2.1 - An example Event Log

The relation with the input class model is obvious (Figure 2.5).



Actor Id	Event Type Id	Object Id	Event Occurrence Time Stamp
John	A	1	9-3-2004:15.01
John	A	2	9-3-2004:15.12
Sue	A	3	9-3-2004:16.03
Carol	B	3	9-3-2004:16.07
Mike	B	1	9-3-2004:18.25
John	C	1	10-3-2004:9.23
Mike	C	2	10-3-2004:10.24
Sue	A	4	10-3-2004:10.35
John	B	2	10-3-2004:12.34
Pete	D	2	10-3-2004:12.50
Sue	A	5	10-3-2004:13.05
Carol	C	4	11-3-2004:10.12
Pete	D	1	11-3-2004:10.14
Sue	C	3	11-3-2004:10.44
Pete	D	3	11-3-2004:11.03
Sue	B	4	11-3-2004:11.18
Clare	E	5	11-3-2004:12.22
Clare	D	5	11-3-2004:14.34
Pete	D	4	11-3-2004:15.56

Figure 2.5 - Relating the log to the Event – Actor – Object model

Observe how the above UML model has an XML-equivalent representation, which can be used to feed data manually to Process Discovery tools. For one of the cases, where a fully automatic logging of the events was impossible, this was the input format of the data as resulting from time and motion measurements and next forwarded into the Process Discovery techniques.

2.4 Formal Concept views for Process Discovery

Formal Concept Analysis (FCA) is a qualitative discovery technique that emerged to become very successful in the recent years [see all International Conferences on Formal Concept Analysis, and in particular Peters, 2012]. In its simplest format, FCA transforms association matrices into concepts, which are formally subsets of rows sharing the same subsets of columns associations. Next the “discovered concepts” are partially ordered in a concept lattice, which can be very helpful for the interpretation of the associations, and potential “rules” that are following from the association matrix. Since the input data for Process Discovery contain a number of iterations, it is interesting to explore the results of some straightforward association matrices, illustrated by means of the above input data.

The simplest association follows from assuming that each “object” corresponds to a process instance. The following shows the interaction matrix of the process instances and the events involved in them for the example from section 2.3:

	A	B	C	D	E
Obj 1	X	X	X	X	
Obj 2	X	X	X	X	
Obj 3	X	X	X	X	
Obj 4	X	X	X	X	
Obj 5	X			X	X

Table 2.2 - Association matrix for the previous Event Log

The following is the concept lattice that follows from this matrix. It tells us that all 5 processes have A and D as processing steps. Four (4) of them have B and C as processing steps, and one of them has E (instead) as processing step. E could be an example of an "exceptional" process step, replacing the combination of B and C in some cases…

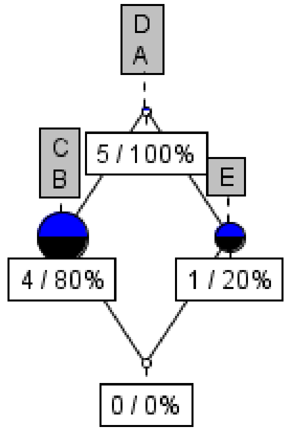


Figure 2.6 - The corresponding FCA concept lattice

The next logical step is to map the first order transitions (showing each process step with its immediate follow-on step). From the input data, the following is the association matrix

	AB	AC	AE	BC	BD	CB	CD	ED
Obj 1	X			X			X	
Obj 2		X			X	X		
Obj 3	X			X			X	
Obj 4		X			X	X		
Obj 5			X					X

Table 2.3 - The first-order transition association matrix

The resulting concept lattice reveals already the essentials of the actual process variations:

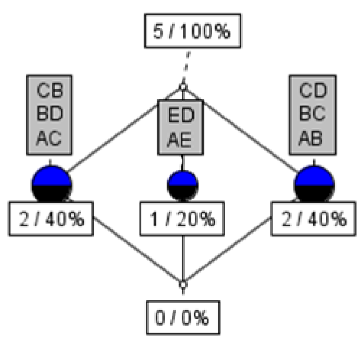


Figure 2.7 - The corresponding concept lattice for the first-order transitions

Reading the lattice bottom-up shows that in 2 of the 5 process instances A is followed by B, next C and D. In 2 other instances A is first followed by C and only next comes C and D. Finally one of the processes has A followed by E, and it is clear that all processes end by D.

A next interesting association relates Actors to the processing steps, which is for this example the following matrix:

	A	B	C	D	E
John	X	X	X		
Sue	X	X	X		
Carol		X	X		
Mike		X	X		
Pete				X	
Clare				X	X

Table 2.4 - The Actor – Event Association Matrix

The corresponding concept lattice outlines the responsibilities of the Actors for the different processing steps:

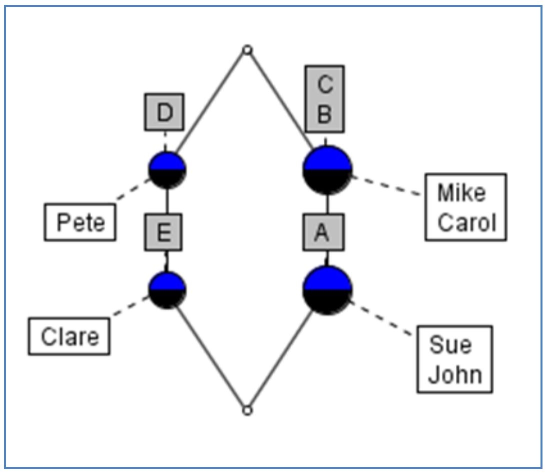


Figure 2.8 - The Concept Lattice for the Actors

Clare can execute D and E, while Pete can only execute D. Sue and John can execute A, B and C, while Make and Carol can only execute B and C. *Observe that this lattice corrects an imprecise representation* in [Dustdar et al, 2004]. The lattice shows the clear separation between the two teams in their responsibilities.

A last association matrix of interest is the relation between the actors and the processes. The starting matrix is the following (Table 2.5):

	1	2	3	4	5
John	X	X			
Sue			X	X	X
Carol			X	X	
Mike	X	X			
Pete	X	X	X		
Clare					X

Table 2.5 - The Actor – Process Association Matrix

The concept lattice shows the involvement of Actors in the various processes:

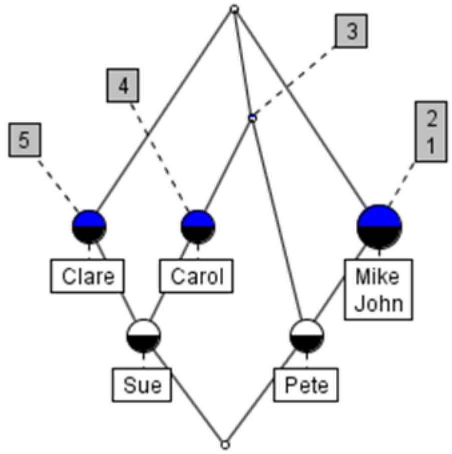


Figure 2.9 - The corresponding Actor – Process concept lattice

It shows how Sue is involved in processes 3,4 and 5, actually assisted by Clare for process 5 and Carol for process 4. Pete works together with Sue and Carol for process 3, while processes 1 & 2 are handled by Mike and John.

The following association matrix shows how one Actor is related to another one by means of "transfer of work", which means the second Actor is taking the next processing step.

	John	Sue	Carol	Mike	Pete	Clare
John				X	X	
Sue			X		X	X
Carol		X				
Mike	X					
Pete						
Clare						X

Table 2.6 - The Actor to Actor association matrix

The corresponding lattice is the following:

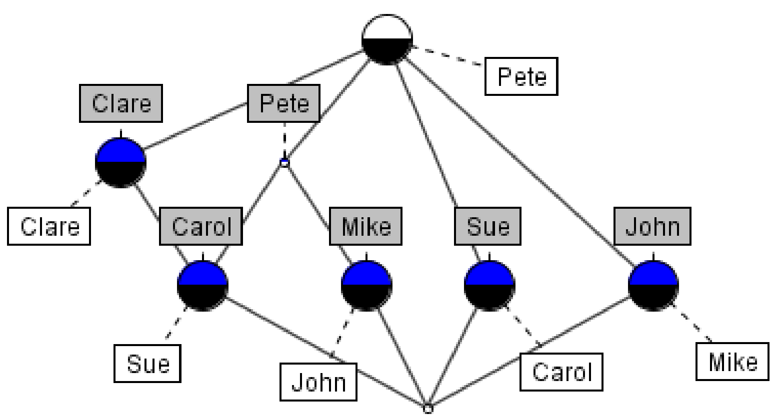


Figure 2.10 - The corresponding Actor network as a concept lattice

The white actors are the ones transferring work in the workflows, whereas the gray actors are the ones receiving work from someone else. In this lattice, it is clear that:

- Pete doesn't send work
- Sue and Carol both send work to each other and receive work from each other
- The same holds for Mike and John
- Sue and John are also sending work to Pete
- Sue is sending work to Clare
- Clare sends work to herself

This lattice formalizes the so-called "sociogram" based on transfer of work, as described in [Dustdar et al, 2004]:

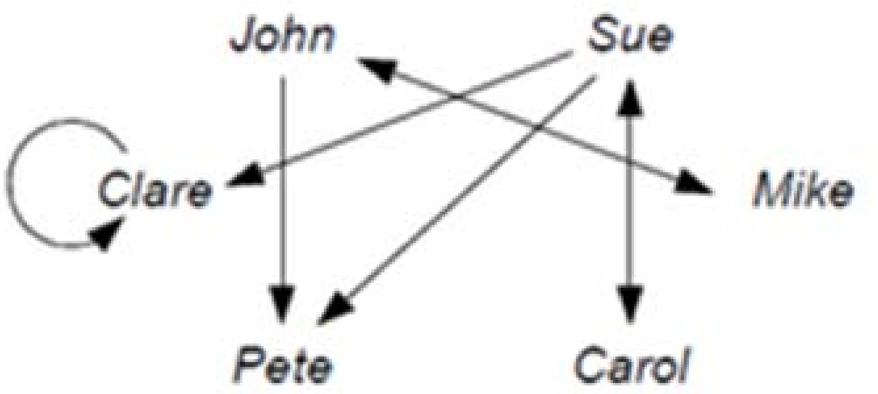


Figure 2.11 - An Actor sociogram example

Finally, traditional FCA turns out to be very useful to intercept visually outliers or exceptions in processes. Suppose that the "regular" standardized process is the one obtained so far in this example, and suppose that a new log of 7 processes with their processing steps is obtained:

	A	B	C	D	E
Obj 1	X	X	X	X	
Obj 2	X	X	X	X	
Obj 3	X	X	X	X	
Obj 4	X	X	X	X	
Obj 5	X			X	X
Obj 6	X			X	X
Obj 7	X	X		X	

Table 2.7 - An association matrix with an exception

The corresponding lattice shows immediately that there is one exception, in which the combination of B & C together is missing:

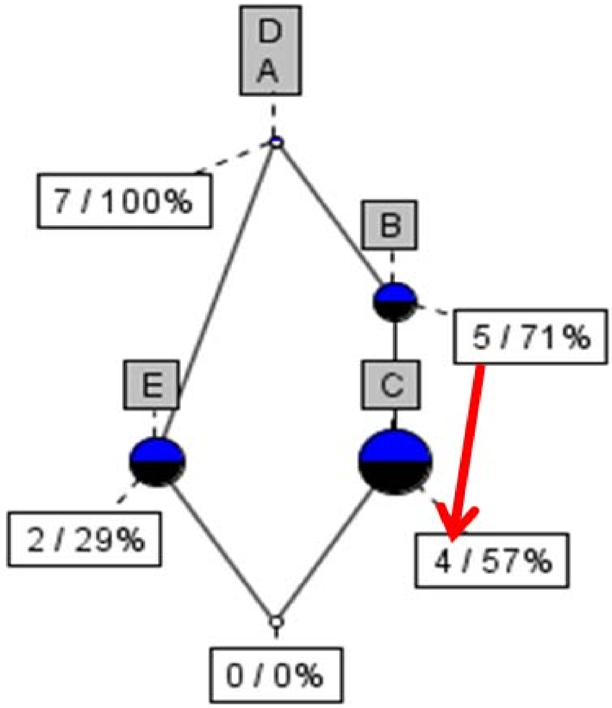


Figure 2.12 - Discovery of the exception in the concept lattice

This technique will effectively be used in the case study on clinical pathways in chapter 5 (caring processes), to detect some instances of value leaks in processes.

2.5 Probabilistic Finite State Machines

Basically there are two families of techniques for representing processes:

- Show the processing steps on the arrows: this is the main technique for probabilistic finite state machines (which were actually the first technique that were used in Process Discovery by Cook and Wolf).
- Show the processing steps in the nodes (places): the main two approaches using this technique are Petri Nets and Hidden Markov Models. Both techniques have advantages and disadvantages which are discussed in detail in this chapter.

A probabilistic Finite State Machine (PFSM) is the next level of representation of Process Discovery results [Vidal et al., 2005]. It is essentially an “activity on the arrow” technique, whereby each arrow also has the corresponding probability of the arrow happening. The arrows represent the process steps. The concept of a PFSM extends the classical FSM-concept as used in consistent UML modeling [Snoeck & Dedene, 1997].

More formally:

A Probabilistic Finite State Machine (PFSM) D is a tuple

$D = (E_D, Z_D, T_D, Prob_D, I_D, F_D)$, where

E_D is the Event Alphabet, Z_D denotes the set of states, I_D is the initial state and F_D is the set of final states ($F_D \subset Z_D$).

$$T_D \subseteq Z_D \times E_D \times Z_D$$

and

$$\mathrm{Pr}\,ob_D : T_D \rightarrow [0,1]$$

are the transition set and the transition probability function, respectively, where for all $z \in Z_D$:

$$1 = \sum_{x \in E_D, z' \in Z_D} \mathrm{Pr}\,ob_D(z, x, z')$$

This means that Prob_D induces a probability distribution on the set of transitions that are leaving for each state.

The PFSM for the example introduced in the beginning of this chapter is the following:

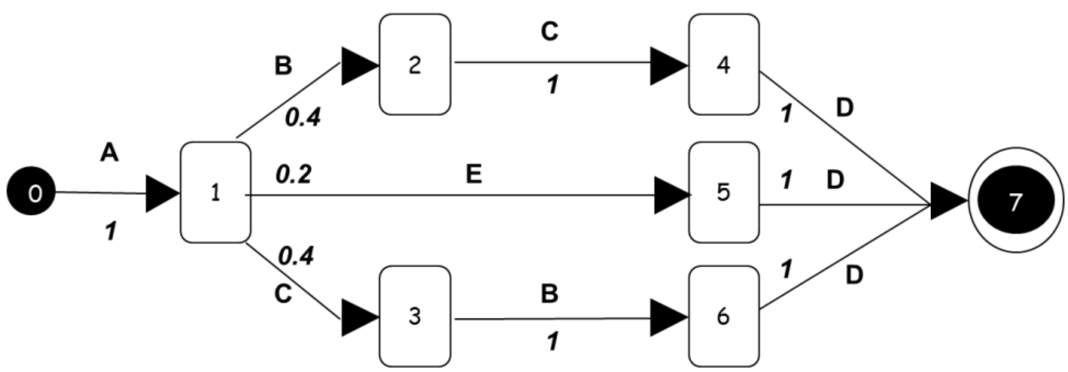


Figure 2.13 - A probabilistic FSM for the event log example

On the right we see again the equivalent, simple FCA representation.

To analyze the properties of the PFSM, it must be transformed into a Markov Model, enumerating all the states, and forming the Markovian transition matrix for the Markov chain in the PFSM.

$$a_{ij} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.4 & 0.4 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

If p_i denotes the steady state probability of being in a state i, and a_{ij} denotes the probability associated with the state transition from state i to state j in the FSM, the steady state probabilities are the solution of the following equations:

$$p_i = \sum_j p_j \times a_{ji} \qquad \text{(state balance equations)}$$

$$\sum_i p_i = 1 \qquad \text{(normalization condition)}$$

The main fundamental problem with PFSM as a Process Discovery representation is the aggregation of the state transitions to the same pairs of states, which often prohibits the actual discovery of the subtle details of value leaks.

The following is a concrete example of such a situation:

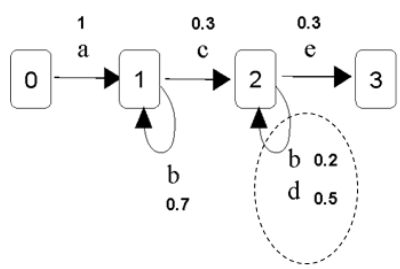


Figure 2.14 - A potential value leak in a Probabilistic FSM

In the transition matrix, the transitions from state 2 into itself are aggregated into ONE state transition with a probability 0.7. This problem can only be solved by switching to the Hidden Markov Model representation, which is discussed later in this chapter.

2.6 BPMN and other Petri Net Models

With the emergence of the Unified Modeling Language (UML) as a de facto standard for software design, the Petri Net representation became increasingly popular. In its simplest form, Activity diagrams are used. Activity diagrams are an “activity on the node” representation of Process Discovery results. The drawing technique seems similar to Finite State Machines, but a major difference is the presence of so-called *FORK (aka SPLIT)* and *JOIN* transitions, which may act as AND-gates or as OR-gates. Hence, in the case of an *AND*-FORK all steps after the FORK must be executed. In the case of an *AND*-JOIN, all steps preceding the join must have been executed before the JOIN-transition can occur. Typically, as a best practice, every FORK in process ends up either in different end points, or a JOIN.

This is the activity diagram representation of the above example in this chapter:

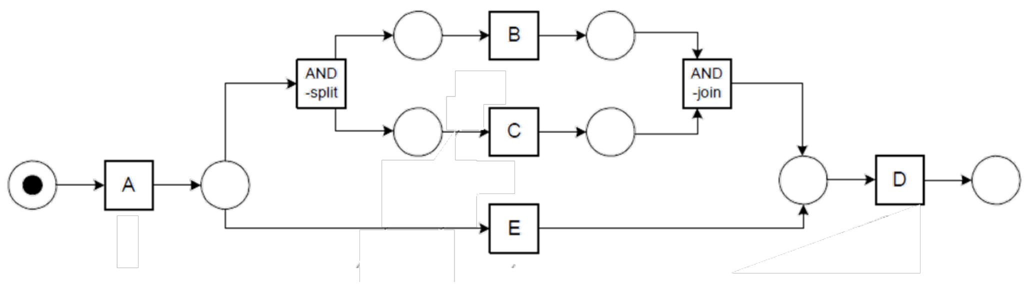


Figure 2.15 - An Activity Diagram for the event log example

There is one AND-split, which in this case forces that BOTH B and C are executed after A in the case when E is not chosen. Observe how this does not impose any sequence on the steps B and C, and basically expresses that these steps can be executed in parallel. The AND-JOIN also implies that BOTH B and C must have been executed (and finished) before the next step (in this case D) can be executed.

As shown in Figure 2.15, it suffices to track a “token” to see the possible progress paths of the process.

BPMN is mainly an extension of Activity diagrams, which a much richer (graphical) semantics in the steps and transitions, including the possibilities for including default steps, timers, exceptions, etc… But, formally, BPMN diagrams are Petri Nets, just as Activity Diagrams.

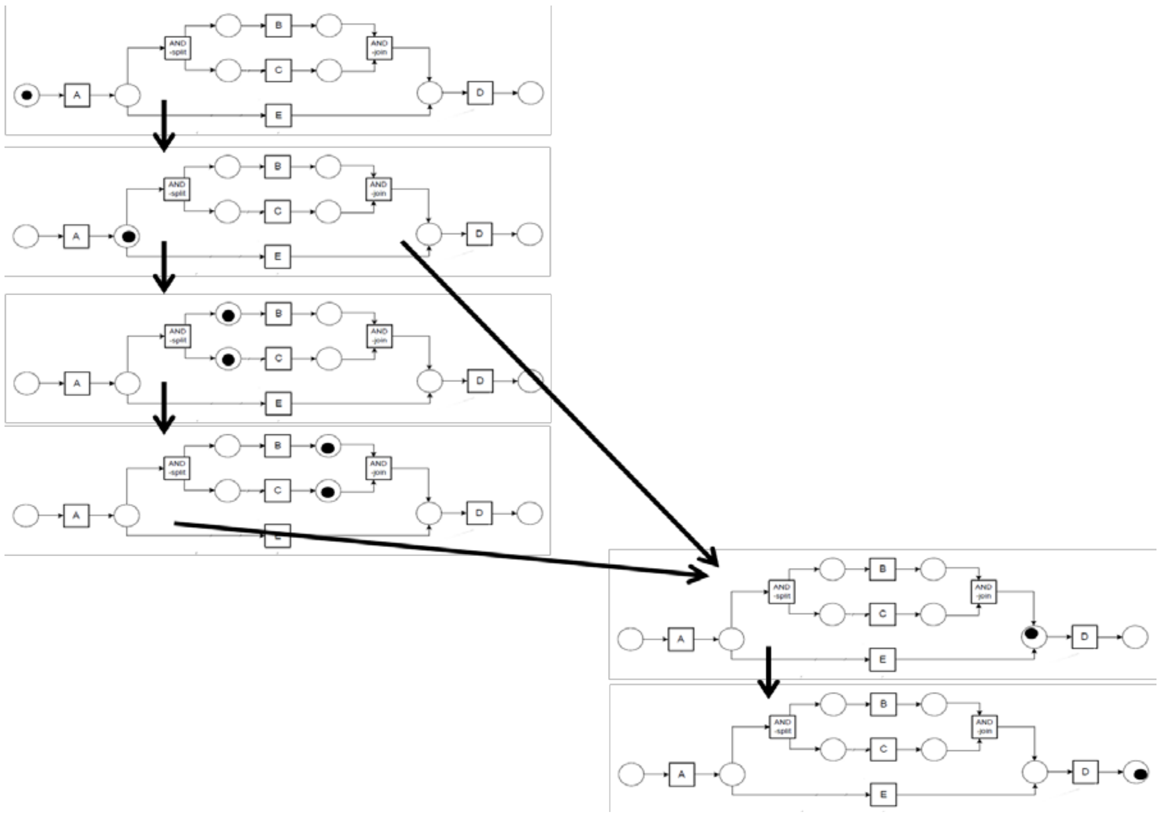


Figure 2.16 - The execution of the Activity Diagram

The Petri Net notation in its original pure mathematical format regained substantial interest because of the work of the huge research group, supervised by Prof. Dr. Will Van der Aalst at Eindhoven University [Van der Aalst et al., 2011].

In the "pure" Petri-net notation, the notion of transitions, places, flows and tokens is more precisely defined, as follows.

A Petri net is a triple (P, T, F):

- P is a finite set of places,

- T is a finite set of transitions $(P \cap T = \phi)$*,*

- $F \subseteq (P \times T) \cup (T \times P)$ *is a set of arcs (flow relation)*

A place p is called an input place of a transition t if and only if there exists a directed arc from p to t. Place p is called an output place of transition t if and only if there exists a directed arc from t to p.

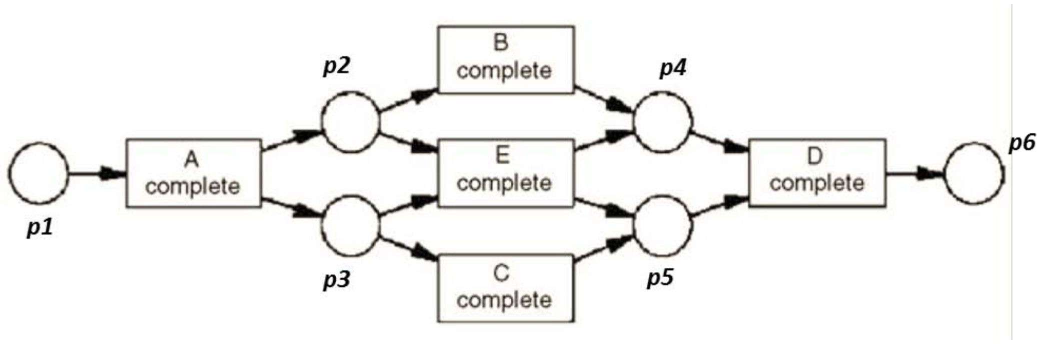


Figure 2.17 - The Petri-net representation of the event log example

This Petri Net representation of the ongoing example in this chapter has 5 processing steps as transitions (A, B, C, D, E) and 6 places ($p_1,\ldots,p_6$). At any time a place contains zero or more tokens, drawn as black dots. The number of tokens may change during the execution of the net. Transitions change the number of tokens in their input and output places according to the following “firing” rule:

(1) A transition t is said to be enabled if and only if each input place p of t contains at least one token.

(2) An enabled transition may fire. If transition t fires, then t consume one token from each input place p of t and produces one token for each output place p of t.

This is the execution of the above Petri Net:

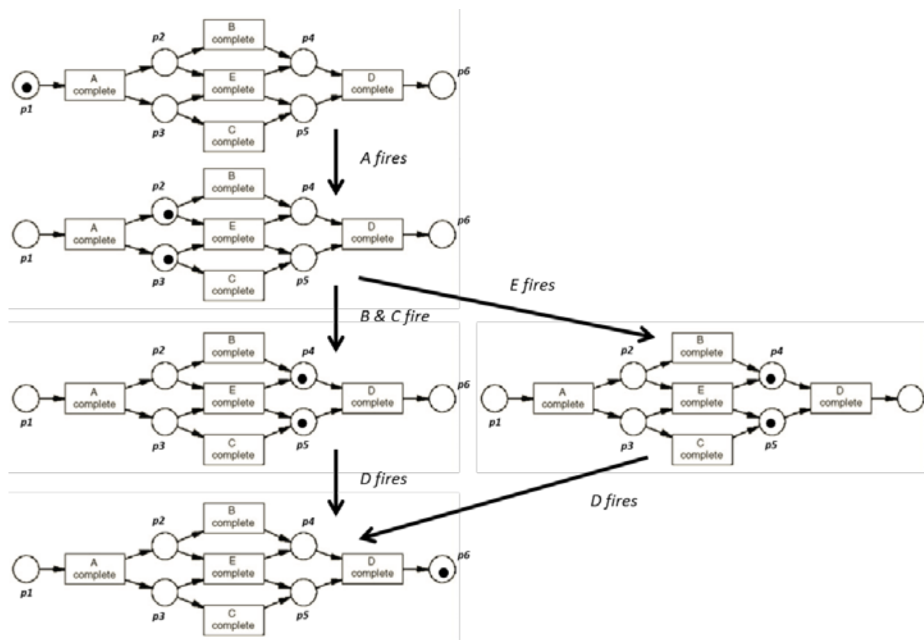


Figure 2.18 - The execution of the Petri-net

One clear disadvantage of Petri Nets is the lack of probability indications. In the above diagrams it is not clear in how many of the process instances A is followed by B first or A is followed by

C first. This makes a quantitative analysis in search of Value Leaks in Services and Processes difficult. One example is the complex integer programming that is needed [Van der Aalst, 2006] to match modeled and actually observed process samples. One variant of Petri Nets allows using transitions with multiple "weights", requiring – for example – that 3 tokens are present before some transition can fire. However, we have not seen a lot of usage of it in practice so far.

Petri Net techniques (in all their variations) are also often preferred because of their ability to deal with concurrent activities. For sure concurrent (interleaving) parallel activities may occur a lot in caring processes. However, if the main target is the discovery of Value Leaks and Imperfections in Processes, parallelism is mostly not helping a lot.

Consider the following example, in which B, C and D are parallel activities. However, the sequence of B and C is fixed: C comes after B. This could be an Activity Diagram for this process:

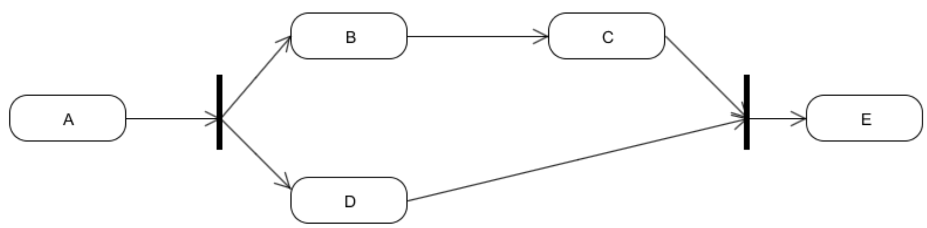


Figure 2.19 - Parallelism in Activity Diagrams/Petri-nets

But to find out that maybe there is a Process deficiency in all process variations where B precedes D requires detailed probabilities on the various operational execution paths of the activity sequences BCD, BCD, CBD, and so on. The representation that is really needed is the following:

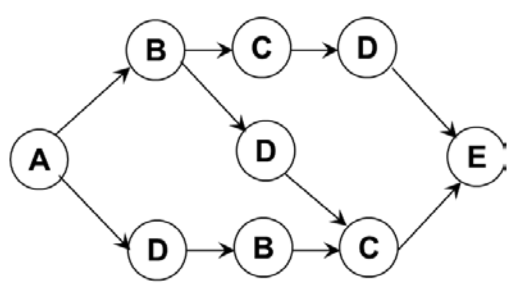


Figure 2.20 - The corresponding Hidden Markov Model representation

which is nothing but the Hidden Markov Model (HMM) representation for the same Process Discovery.

Concurrent activities are nice semantically to aid in simplifying a Process model to present it in an understandable fashion to management personnel who have the neither the need nor desire to understand the mathematical details of the other representations. However, what is pointed out here is the fact that a nice management representation is good for that particular purpose, whereas other representations are needed to detect deficiencies and Value Leaks in Processes. In saying this, please understand the respect and high admiration held for the very important work on Process Mining (later extended into Discovery [Van der Aalst et al., 2011] produced at the University of Eindhoven. Their main contribution is the reverse engineering of event logs into Petri Net representations of Business Processes that are especially useful in the context of managerial discussions. Additionally, BPMN provides Process representations that are equally easy to understand and useful by managers. Additionally, BPMN provides Process representations that are equally easy to understand and useful by managers.

2.7 Hidden Markov Models

A Hidden Markov Model (HMM) provides an "activity on the node" representation of Processes. It extends the classical Markov Model, whereby each node has a ***unique*** activity that is the id of that node, implying that the behavior in the model is uniquely defined by the places: knowing the name of the current activity allows an unambiguous positioning in the Process model. In a HMM the same activity may appear in several nodes in the model. In fact there may even be multiple "symbols" in each node. As a consequence, knowing the current activity is insufficient for a precise positioning in the model, and additional information, such as the higher order transitions, may be needed for that purpose.

There is an overwhelming amount of literature on Hidden Markov Models. In particular, HMMs have been used extensively in pattern recognition in strings of symbols, with obvious applications in speech recognition, web usage mining, customer behavior analysis and biological sequence modeling. [Rabiner, 1989]

A (Discrete) Hidden Markov Model with state emission (HMM) D is a tuple

$D = (E_D, Z_D, A_D, e_D, i_D)$, where

E_D is the Event Alphabet, Z_D denotes the set of states,

$A_D : Z_D \times Z_D \to [0,1]$ *is a mapping defining the probability of each transition,*

$e_D : Z_D \times E_D \to [0,1]$ *is a mapping defining the emission probability of each element of the alphabet on each state,*

$i_D : Z_D \to [0,1]$ *is a mapping defining the initial probability for each state.*

The following is a simple example borrowed from [Dupont et al., 2005]:

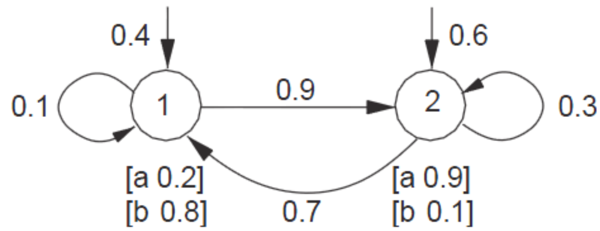


Figure 2.21 - A Hidden Markov Model example

$E_D = \{a,b\}$,

$Z_D = \{1,2\}$,

$A_D(1,1) = 0.1$; $A_D(1,2) = 0.9$; $A_D(2,1) = 0.7$; $A_D(2,2) = 0.3$;

$e_D(1,a) = 0.2$; $e_D(1,b) = 0.8$; $e_D(2,a) = 0.9$; $e_D(2,b) = 0.1$;

$i_D(1) = 0.4$; $i_D(2) = 0.6$

There are also variants of HMMs that allow emissions on the transitions.

It is clear that the general definition of an HMM is very generic, and allows many alternative combinations of alphabets and state names. ***For Process Discovery purposes only a more concrete (but still very powerful) type of HMM is required, namely HMM's with***

a) The Alphabet consisting of the Process Steps
b) Every place/node having only ONE element of the Alphabet that it can emit *(which makes the emission probability distribution e_D trivial)*
c) One initial step *(which makes the initial distribution i_D trivial)*

This special type of HMM is sometimes indicated as a POMM (Partially Observable Markov Model [Callut & Dupont, 2005]).

The following is the HMM representation of the recurring example in this chapter:

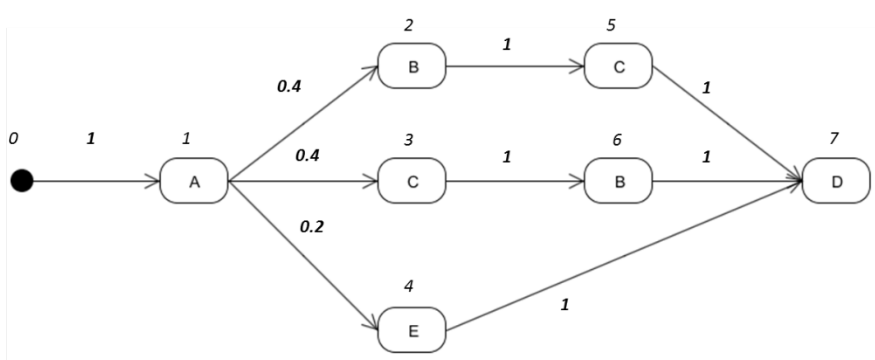


Figure 2.22 - The HMM for the event log example

The easiest straightforward analysis of an HMM is the calculation of the visit ratio's, which expresses for every process instance, the average number of visits to each of the transitions.

$$v_0 = 1$$

$$v_i = \sum_{j=1}^{n-1} v_j \times A_{Dji} \qquad \text{i = 1,…,n-1}$$

In case of multiple occurrences of the same step at different transitions, the maximal visit ratio of all transitions where the step occurs gives the visit ratio of that Process step. The solution in this example is

$V_0 = 1.0$

$V_1 = 1.0$

$V_2 = 0.8$

$V_3 = 0.4$

$V_4 = 0.2$

$V_5 = 0.4$

$V_6 = 0.8$

$V_7 = 1.0$

and

$V_A = 1.0$

$V_B = 0.8$

$V_C = 0.8$

$V_E = 0.2$

$V_D = 1.0$

It implies that in this Process representation the average Process length is 3.6 steps.

Since this Process representation is the equivalent of the Petri Net representation, it is now easy to verify the matching of the Process model with actual token configurations in the Petri Net. Following the example in [Van der Aalst, 2006], where he asks the question whether the following token frequency table is feasible:

Process Step	Frequency
A	3
B	2
C	2
D	3
E	2 or ?

Table 2.8 - An example of a token configuration of actual processes and their steps

can easily be solved by using the visit ratios, and rounding the results to integers (because tokens are not in parts):

$V_A = 1.0 \times 3 = 3$

$V_B = 0.8 \times 3 = 2.4 \sim 3$

$V_C = 0.8 \times 3 = 2.4 \sim 3$

$V_E = 0.2 \times 3 = 0.6 \sim 1$ (which solves the ?)

$V_D = 1.0 \times 3 = 3$

which exactly matches the results of [Van der Aalst, 2006].

How about the equivalence with the PFSM representation for discovered Processes? Basically this is the classical problem of translating an “activity on the node” network into an “activity on the arrow” network schema. The oldest technique to do so is to build a higher-order representation of the same network. Higher-order transitions are discussed in the section of Hybrid techniques at the end of this chapter. This technique was nicely further investigated and presented in [Dupont et al., 2005]. Their main research objective was to be able to apply some machine learning techniques from PFSMs to HMMs, for a broad range of applications. They demonstrated the following results on equivalences between PFSM and HMM (in particular POMMs for the sake of our research).

a) An HMM can be transformed into a PFSM that has no final state (and hence no probabilities for reaching the final state). This type of PFSM, which can be non-deterministic and give rise to infinite sequences which is not practical for our research.

b) A PFSM can be transformed into an HMM, but not necessarily with the same number of states, whereby some (crucial) information may be lost. In particular in nodes with multiple transitions to the same node, the probabilities are aggregated into one probability number, which disables the possibility to trace back to individual transitions. An example was already given in the section on PFSMs.

This discussion makes it clear how HMM are indeed for our discovery and comprehension of Value and Value leaks in Services and Processes a preferred representation technique.

Other arguments for preferring HMMs as a representation technique for Process Discover were explained in a brilliant way in the very inspiring works of [Blum et al., 2008] on the discovery of workflow models for surgery. We quote:

"Most workflow mining methods use Petri-Net like models, as these share many similarities to workflow models that are established in business sciences. In this work, a statistical approach, using Hidden Markov Models (HMMs) is taken to model the workflow inside the Operation Room. Unlike Petri-Nets, HMMs do not allow for concurrent actions, but can be used to model properties like the average duration of actions, the variance of their duration or transitions probabilities between single actions. They are also very flexible and can be used with continuous data, like positions of the surgeon or tools, obtained by a tracking system. Another reason for choosing HMMs is that usually the process logs contain only a relative low number of different high-level events. Each of these events can be represented by one Petri-Net state. In our case, the surgical workflow is modeled at a finer level of detail and many different combinations of tools might be used. Using Petri-Nets, we would have to represent each of these combinations with one state, resulting in a huge model. Using HMMs we can represent different combinations of tools using only one state."

Observe how the Value Leaks discussed in the paper of Blum are again obviously non-financial in nature: they consider

- Length of the operation time
- Risks of repeating some steps
- Blood loss

2.8 Normalized Process Representations

The topic of normalization is well known in database design. Database normalization is the reorganization of database tables and columns in a relational database to minimize redundancy and dependency in the database. One of the side-purposes is to cure anomalies that might occur in updating, and in particular in deleting rows from tables. The most important four normal forms for databases were developed by Codd (Codd, E.F.,1970), at the end of the sixties and the beginning of the seventies.

Recently, it was discovered how normalization is by no means something particular for SQL-based databases, and how actually normalization should already occur in the higher order conceptual representation of data (e.g. Entity-Relationship-models, or Class Diagrams). In

[Poelmans et al. 2012] the notion of normalized UML class diagrams was introduced. In short, normalized associations are characterized by means of the following criteria:

- No many-to-many relationships
- No optional-to-optional relationships
- No reflexive relationships
- Only binary relationships

In [Put 1989] it was already demonstrated that this type of normalized association diagram automatically translates into a fourth normal form relational database design. In [Poelmans et al. 2012] it was demonstrated that normalized association diagrams are actually concept lattices (related to the links between use cases and classes, which was put forward in [Snoeck and Dedene, 1997]). During this research the real meaning of normalization of data became apparent: ***normalization is in the first place a means to reduce the degree of potential ambiguity in a data model.*** All the elements that were mentioned in the above criteria (many-to-many, optional-to-optional, reflexivity and multiparty relationships) are potential sources for misinterpretations, ambiguities, and the related Value Leaks (often Asymmetries of Meaning). The potential ambiguities in the conceptual data model are also the root causes for potential anomalies in update/delete operations in databases (often also leading to incompleteness of information).

Following the same reasoning for Process Discovery representations, HMMs may also contain potential ambiguities under the form of many-to-many transitions, such as the following situation in Process step D.

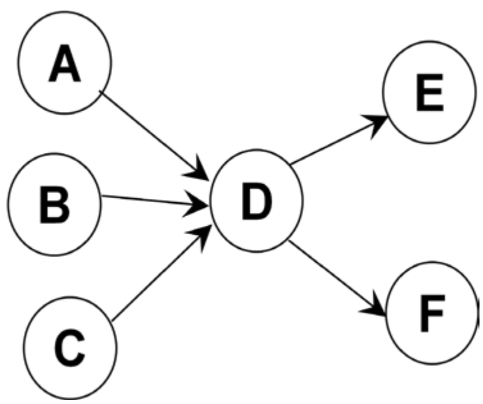


Figure 2.23 - An example of a not-normalized place in a HMM

Indeed, it might be very difficult to relate a Value Leak in E, for example, to a potential root cause in A, B or C, because of the ambiguity in the many-to-many transition in D. In this case – when needed – subsets of Process instances for which the discovered Process representation is normalized must be constructed. It could well be that the above many-to-many transition falls apart into 3 Process variations, as follows:

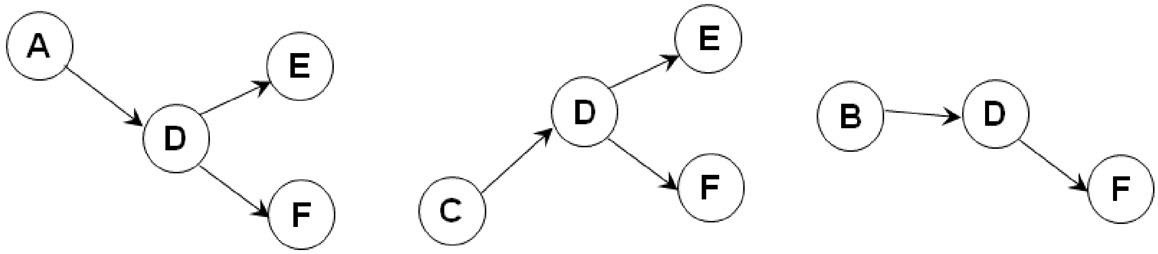


Figure 2.24 - Examples of normalized transitions

In the above representations there are no more many-to-many transitions.

Many-to-many transitions are indeed also confusing for the machine learning techniques for Process Discovery by means of HMMs, because they act as “joint” (bifurcation) steps in Process Variations. The simplest approach to achieve normalized Process representations is the use of ***cloning (splitting)*** of the many-to-many transition steps. Splitting was also considered to solve some representation problems when the discovered Processes turned out to be too general ([Callut & Dupont, 2005b],[Herbst & Karagiannis, 2000]).

A first example is well-known in the literature as the driving license problem. Consider the following list of Process steps:

4 Apply for driving license
5 Attend classes Car driving
6 Attend classes Motor driving
7 Do Theoretical Exam
8 Do Practical Exam Car driving
9 Do Practical Exam Motor driving
10 Get Result

and consider the following event logs:

4 start, a, b, d, e, g, end
5 start, a, c, d, f, g, end
6 start, a, b, d, e, g, end
7 start, a, c, d, f, g, end

In many tools, the discovered Process representation will be the following un-normalized model.

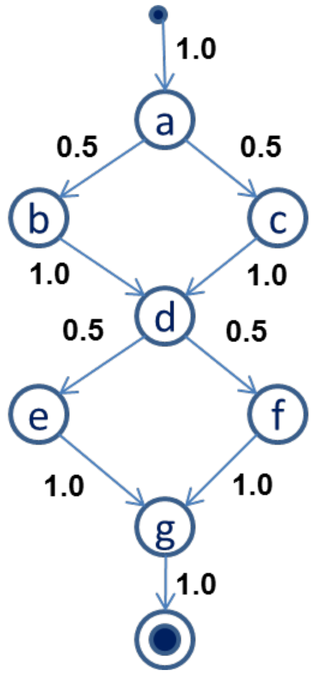


Figure 2.25 - The HMM for the driving license example in not-normalized form

The problem is Process Step **d**, which is a many-to-many transition. It is obvious that the discovered Process representation is too general: one could do the Practical Exam for Car driving after following the classes for Motor driving…

At the end of this chapter the maximum number of incoming versus outgoing transitions will be discussed. In this case there are 2 incoming, and 2 outgoing transitions for d, which suggest that splitting d into two clones d1 and d2 (which are just synonyms with the same meaning) should solve the problem. Indeed, transforming the logs into

O1: start, a, b, d1, e, g, end
O2: start, a, c, d2, f, g, end
O3: start, a, b, d1, e, g, end
O4: start, a, c, d2, f, g, end

results in the following correct discovered Process:

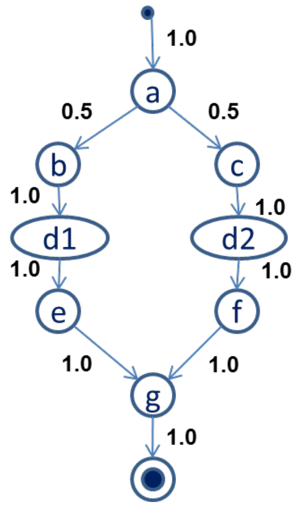


Figure 2.26 - A normalized version of the driving license example

In the section on Hybrid techniques it will also become clear how higher order transformations also help in allocating the correct cloning.

The cloning of Process steps may even be needed within the same sequence of steps, as shown in the following slightly more complicated example that is adapted from [Herbst & Karagiannis, 2000]. Consider the following event logs for Process Steps a,b,c :

- a,b,c,a,c
- a,b,c,c
- a,b,c,a,b,c,c
- a,b,c,a,b,c,a,c

The discovered Process model that typically emerges from these traces is the following:

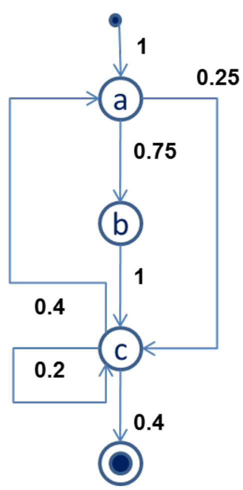


Figure 2.27 - A HMM with un-normalized places [Herbst 2000]

Again this Process representation is too general: it allows several event sequences that are not provided in the input data. The root cause for the problem is again the fact that this representation is not normalized: a and c are many-to-many transitions. Splitting both of them in 2 clones solves the problem. Transforming the event logs into

- a1,b,c1,a2,c2
- a1,b,c1,c2
- a1,b,c1,a2,b,c1,c2
- a1,b,c1,a2,b,c1,a2,c2

results in the following correct discovered Process representation (which is in normal form):

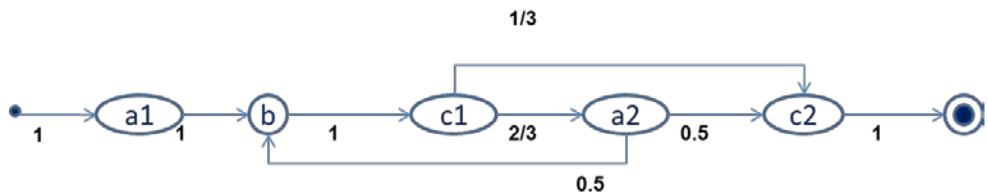


Figure 2.28 - The normalized HMM for the above example

The impressive, but difficult work of [Greco et al., 2006] contains the notion of "discriminant features" for Event logs. Careful reading of their paper and in particular the examples reveals that these "discriminate features" actually describe potential anomalies in the Process event logs (such as the fact that in the driving license example we have on one hand the sequences abd and acd on the one hand, and the sequences deg and dfg on the other hand, while there is no sequence abdfg and no sequence acdeg, which could in theory be composed from the previous sequences). Hence, implicitly they deal with the same phenomenon of un-normalized Process representations. Their first and simple example starts from the following event logs:

O1: a,b,c,d,e

O2: a,d,b,c,e

O3: a,e

The typical discovered Process representation would be the following:

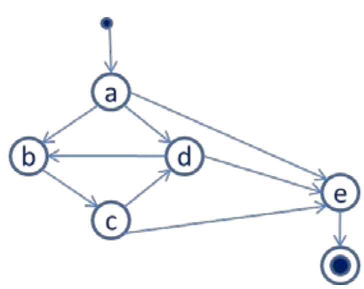


Figure 2.29 - The simple but un-normalized example of [Greco 2006]

where the many-to-many transition d makes the model un-normalized. Splitting d into 2 clones and transforming the event logs into the following:

- a,b,c,d1,e
- a,d2,b,c,e
- a,e

results in the following correct and normalized discovered Process representation:

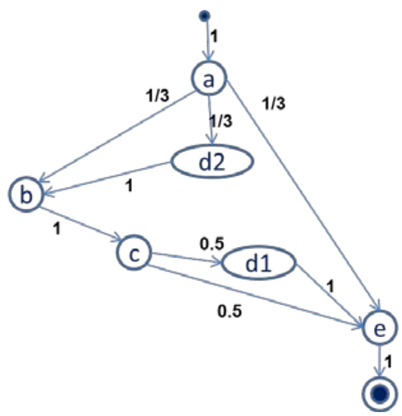


Figure 2.30 - The normalized version of the above process

The main example in their paper is the discovery of an Order Management Process, which considers the following Processing steps:

3.2 receiving the order
3.3 authenticating the client
3.4 checking the availability of the required product in the stock
3.5 verifying the availability of external supplies
3.6 registering the (new) client
3.7 evaluating the trustworthiness of the client
3.8 evaluating the production plan
3.9 rejecting the order
3.10 accepting the order
3.11 contacting the mail department to speed up shipment
3.12 apply some discount
3.13 prepare the bill

The following event logs are the input for the Process Discovery:

O1: ACBGFH
O2: ABFCGH
O3: ACGBFH
O4: ABCGFIL

O5: ABFCGIKL
O6: ACBFGIJL
O7: ACBGFIJKL
O8: ABCEGFIJL
O9: ABEFCGIL
O10: ACGBEFIJL
O11: ABCEDFGIL
O12: ACDBEFGIL
O13: ABCFDGIKL
O14: ACDBFGIKL
O15: ABCDGFIKL
O16: ACBFDGIL

From these event logs, the following simple facts can already be discovered:

- A is the starting state,
- L and H are ending states,
- if D occurs, J cannot occur

  which actually means:

  "Faster shipping cannot be asked for whenever external suppliers are asked"

- if E occurs, K cannot occur

  which actually means:

  "Discounts can never be applied to new customers"

The paper of [Greco et al., 2006] uses a BPMN-like notation, denoting explicitly the OR, XOR and AND logic in the transitions. For the above event log, they obtained the following discovered Process representation:

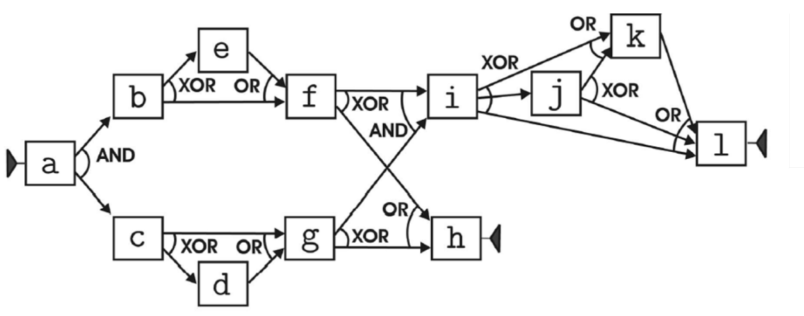


Figure 2.31 - The complex Petri-net example from [Greco 2006]

It is not trivial to discover the place where the representation is un-normalized. It is actually the Process step **I** (accepting the order). It is actually (from an HMM perspective) a four-to-four transition, which can be summarized as follows:

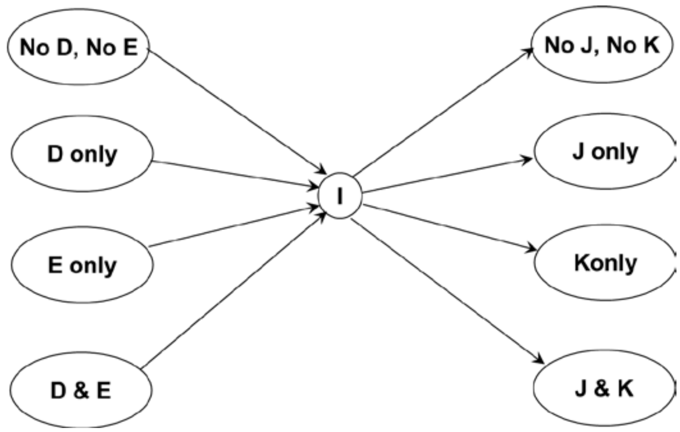


Figure 2.32 - Finding the not-normalized place in the model

This time, the Process step I is split into 4 clones, and the event logs are transformed into the following:

O1:	ACBGFH
O2:	ABFCGH
O3:	ACGBFH
O4:	ABCGFI4L
O5:	ABFCGI4KL
O6:	ACBFGI4JL
O7:	ACBGFI4JKL
O8:	ABCEGFI2JL
O9:	ABEFCGI2L
O10:	ACGBEFI2JL
O11:	ABCEDFGI1L
O12:	ACDBEFGI1L
O13:	ABCFDGI3KL
O14:	ACDBFGI3KL
O15:	ABCDGFI3KL
O16:	ACBFDGI3L

which results in the following 4 Process variations that are discovered.

Variant 1: both D and E occur, so NO J and NO K

O11: ABCEDFGI1L
O12: ACDBEFGI1L

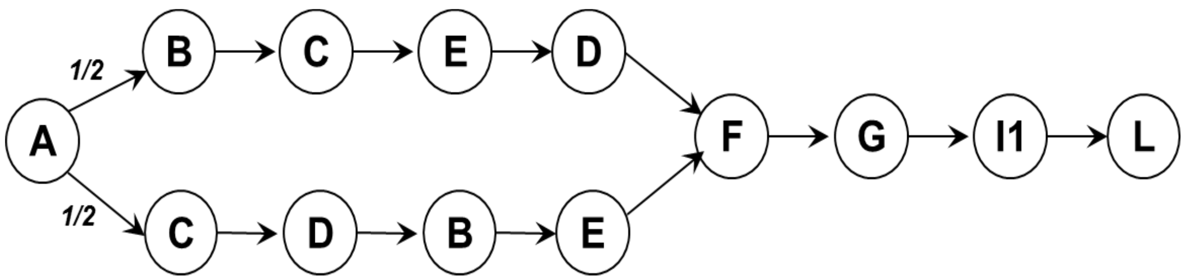


Figure 2.33 - The normalized HMM for the first process cluster

This is the process variant for ***new clients*** *whereby also* ***additional supplies*** *are needed: after accepting the order, it is just billed.*

This must be compared to the following cluster found in [Greco et al., 2006]:

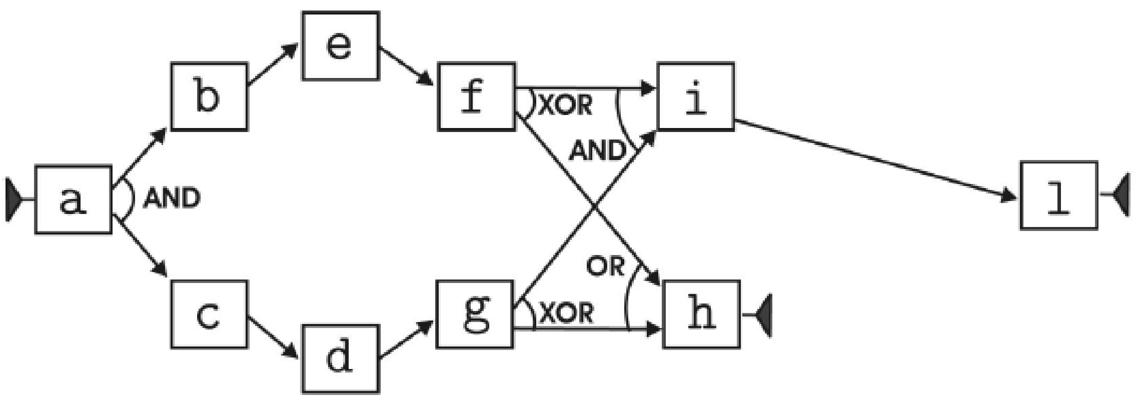


Figure 2.34 - The corresponding Petri-Net in [Greco 2006]

Variant 2: E occurs but NO D, so NO K and (possibly) J

O8: ABCEGFI2JL
O9: ABEFCGI2L
O10: ACGBEFI2JL

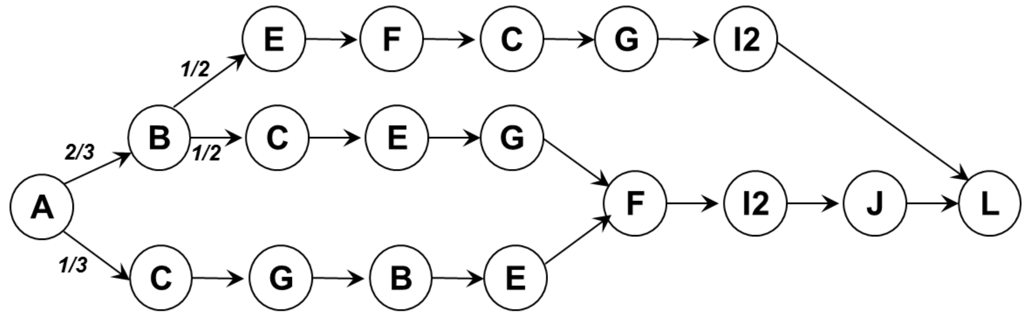


Figure 2.35 - The second cluster HMM in normalized form

This is the process variant for ***new clients*** *whereby* ***no additional supplies*** *are needed: after accepting the order, it is billed, and a speed-up in the logistics is also possible.*

This must be compared to the following cluster found in [Greco et al., 2006]:

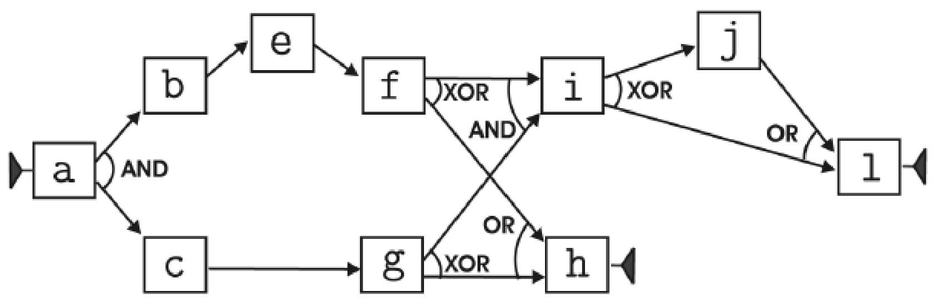


Figure 2.36 - The corresponding cluster in [Greco 2006]

Variant 3: D occurs but NO E, so NO J and (possibly) K

O13: ABCFDGI3KL
O14: ACDBFGI3KL
O15: ABCDGFI3KL
O16: ACBFDGI3L

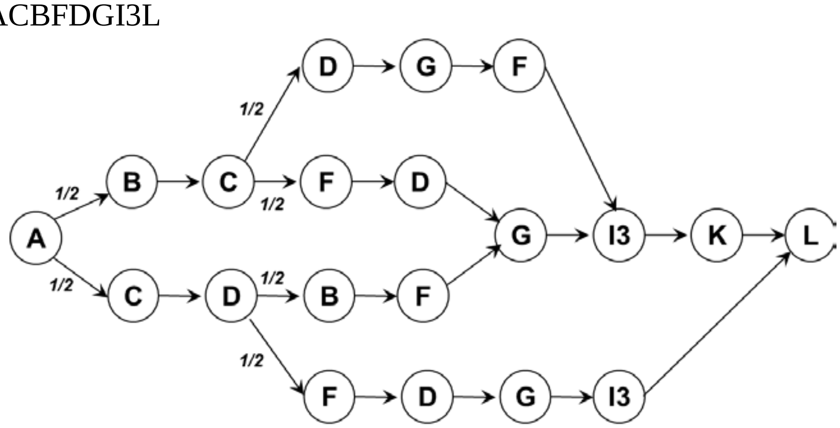


Figure 2.37 - The third cluster in normalized form

*This is the process variant for **regular clients** whereby also **additional supplies** are needed: after accepting the order, it is billed, and possibly a discount is granted.*

Actually this process variant shows a nice detail: in case multiple additional supplies are needed, NO discount is granted…at least in the event log…

This must be compared to the following cluster found in [Greco et al., 2006]:

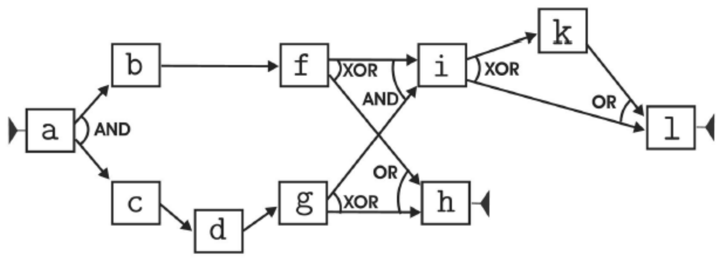


Figure 2.38 - The corresponding Petri-net in [Greco 2006]

Variant 4: NO D and NO E occur, but (possibly) J and (possibly) K

O1: ACBGFH
O2: ABFCGH
O3: ACGBFH
O4: ABCGFI4L
O5: ABFCGI4KL
O6: ACBFGI4JL
O7: ACBGFI4JKL

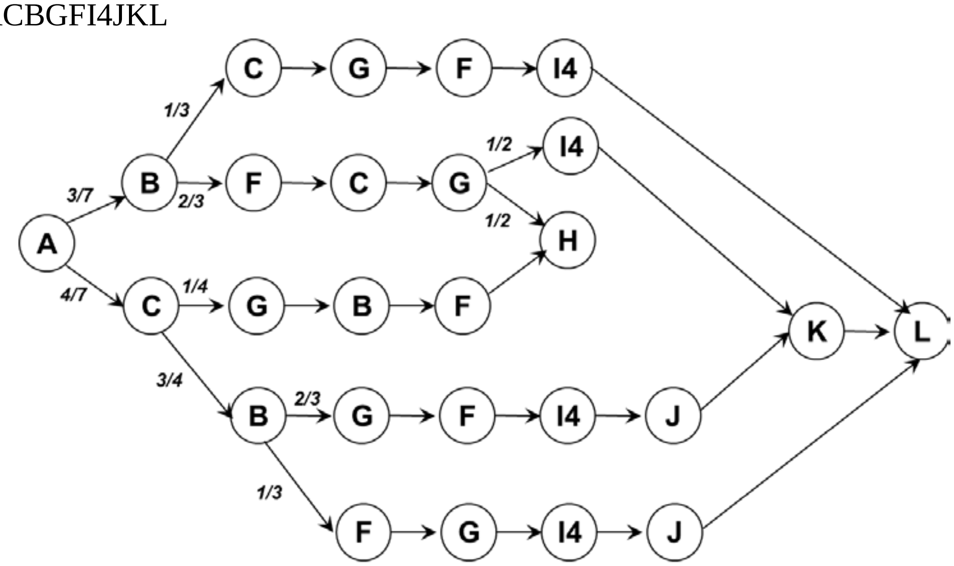


Figure 2.39 - The last cluster in normalized form

*This is the process variant for **regular clients** whereby also **no additional supplies** are needed. Orders may be rejected, and after accepting the order, it is billed, and possibly a discount and/or a delivery speedup may be granted.*

Actually this process variant shows another nice detail: in this case, the "occurrence" of C and G immediately after A is an "early warning" indicator that the order will be rejected.

This must be compared to the following cluster found in [Greco et al., 2006]:

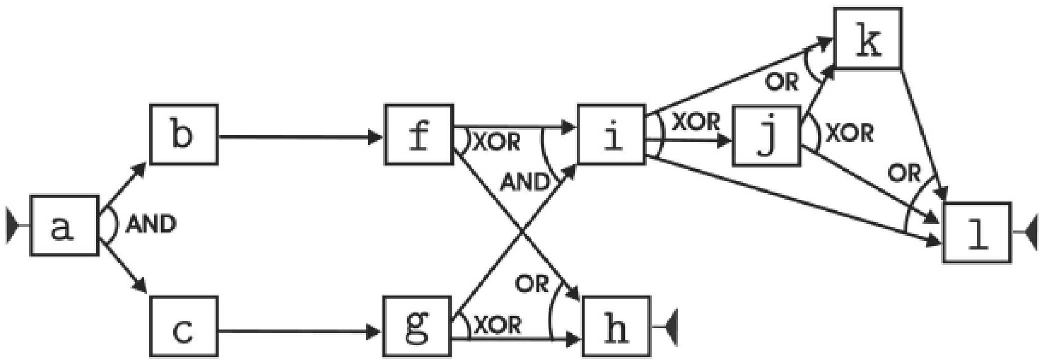


Figure 2.40 - The corresponding Petri-Net in [Greco 2006]

This example illustrates how the normalization of discovered Process representations is very instrumental in discovering process variations. In the next section on Hybrid techniques, further approaches (apart from the splitting into clones that was discussed in this section) to obtain normalized representations will be presented.

2.9 A hybrid approach for Process Discovery

During this research it became rapidly evident that one single heuristic is in general insufficient for detecting deficiencies and Value Leaks in Services and Processes. This implies that in practice combinations of heuristics and techniques are required to achieve results. There is no "general recipe" because the generic problem of the "unknown unknowns" that are typically encountered in this type of projects. This will also become clear in the next chapters, exploring the supporting technology *Comprehend* and the case studies. In this section we bring together our experiences for the following techniques/heuristics and combine it at the end in a hybrid algorithm that should help Process Discovering projects to achieve optimal results:

- clustering
- higher-order transitions
- reverse engineering from the object-event-actor model

of course have Process Normalization as a central element for achieving reliable results.

Many publications employ clustering of the sequences that make up the input for Process Discovery. The main intension of the clustering is to detect Process variations directly. The parameter that is most often used for guiding the clustering is the ***maximum likelihood*** which estimates how accurately an HMM predicts a next state from a previous state. The stronger the likelihood, the more the model fits the reality of the input sequences (and only those sequences, meaning, not allowing other sequences). Saying this, it is obvious that the likelihood is also a supervising parameter for the machine learning techniques that can be used to reconstruct the HMMs from the input sequences of events. The simplest algorithm is just looking at the first order feed-forward probabilities, which can be achieved easily by sorting all couples off events in the input in an appropriate way. The most complex algorithm is combining feed forward and backward learning, and is well-known in the literature as the Baum-Welch algorithm (being a special case of the expectation maximization algorithm). Remark that this algorithm may be very complicated for HMMs in general, but is reduced significantly in complexity if restricted to the particular type of HMMs that is used in this research (the POMMs [Dupont et al., 2005]).

A nice overview of applying maximum likelihood for clustering was provided in the work of Liu ([Liu 2004]. At the same time her thesis shows some of the weaknesses of just using clustering. First of all, clustering should not be used to "rediscover" how sequences of events go together with some actors and objects. In the basic example in [Liu, 2004] the following event logs are used to feed the clustering:

Actor1: a, b, a

Actor1: a, a, b, b, c, b, b, a, a, b, b, a, a

Actor2: d, e, e, e, f, d, e, e, f, d, d

Actor3: a, a, a, b

Actor3: a, a, a, a, b, a, b, b, b, a, a

Actor4: d, d, e, d, e, e

It is already clear from the input data that Actor1 and Actor 3 form one cluster, and Actor2 and Actor4 are the second cluster. Combining however all elements in one discovered model results in HMMs that are not really reliable for discovery purposes.

The second fundamental difficulty with clustering is the estimation of the minimal number of clusters that are needed to really capture the Process variations. Our suggestion is of course to follow the ***Normalized Process Representation*** as a guiding principle for the clustering algorithm. Therefore, before going into clustering, anyhow a general HMM must be reconstructed. Most often it will not be in normal form, meaning that there are k transitions $\{p_1,p_2,\dots p_k\}$ in the HMM with more than one input combined with more than one output. The next step is to try K-means clustering based on maximum likelihood with at least K clusters, whereby

$$K = \sum_{j=1}^{k} Max(input, output) TransitionsForp_j$$

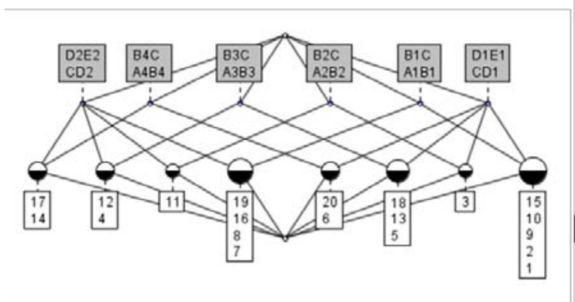
So, in the following example, consider the following event log for 20 process instances:

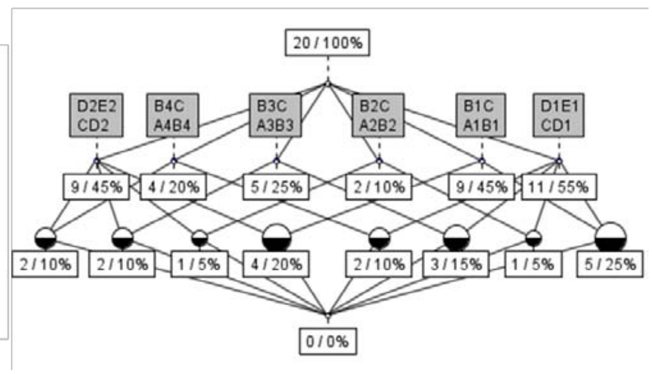
Timestamp	Obj	LogEvent
1	1	A1
2	2	A1
3	1	B1
4	3	A2
5	3	B2
6	2	B1
7	3	C
8	4	A3
9	5	A3
10	4	B3
11	4	C
12	6	A4
13	2	C
14	5	B3
15	1	C
16	1	D1
17	1	E1
18	6	B4
19	6	C
20	6	D1
21	6	E1
22	2	D1
23	7	A1
24	7	B1
25	5	C
26	7	C
27	2	E1
28	4	D2
29	4	E2
30	8	A1
31	9	A1
32	8	B1
33	8	C
34	8	D2
35	8	E2
36	9	B1
37	10	A1
38	9	C
39	11	A2
40	11	B2
41	10	B1
42	11	C
43	12	A3
44	13	A3
45	12	B3
46	12	C
47	14	A4
48	10	C
49	13	B3
50	9	D1
51	9	E1
52	20	A4
53	14	B4
54	14	C
55	14	D2
56	14	E2
57	10	D1
58	15	A1
59	15	B1
60	13	C
61	15	C
62	10	E1
63	12	D2
64	12	E2
65	16	A1
66	20	B4
67	16	B1
68	16	C
69	16	D2
70	16	E2
71	15	D1
72	15	E1
73	7	D2
74	5	D1
75	5	E1
76	7	E2
77	13	D1
78	13	E1
79	11	D2
80	11	E2
81	3	D1
82	17	A4
83	17	B4
84	3	E1
85	18	A3
86	17	C
87	18	B3
88	17	D2
89	18	C
90	19	A1
91	20	C
92	20	D1
93	20	E1
94	18	D1
95	18	E1
96	17	E2
97	19	B1
98	19	C
99	19	D2
100	19	E2

Figure 2.41 - A complex event log example

The analysis of the first order transitions by means of formal concept analysis shows the following results:

	A1B1	B1C	A2B2	B2C	A3B3	B3C	A4B4	B4C	CD1	D1E1	CD2	D2E2
1	X	X							X	X		
2	X	X							X	X		
3			X	X					X	X		
4					X	X					X	X
5					X	X			X	X		
6							X	X	X	X		
7	X	X									X	X
8	X	X									X	X
9	X	X							X	X		
10	X	X							X	X		
11			X	X							X	X
12					X	X					X	X
13					X	X			X	X		
14							X	X			X	X
15	X	X							X	X		
16	X	X									X	X
17							X	X			X	X
18					X	X			X	X		
19	X	X									X	X
20							X	X	X	X		





2

Figure 2.42 - The FCA analysis of the transitions

Substituting “A followed by B” for a symbol “U” AND “C followed by D” for a symbol V, the following Markov Model is discovered:

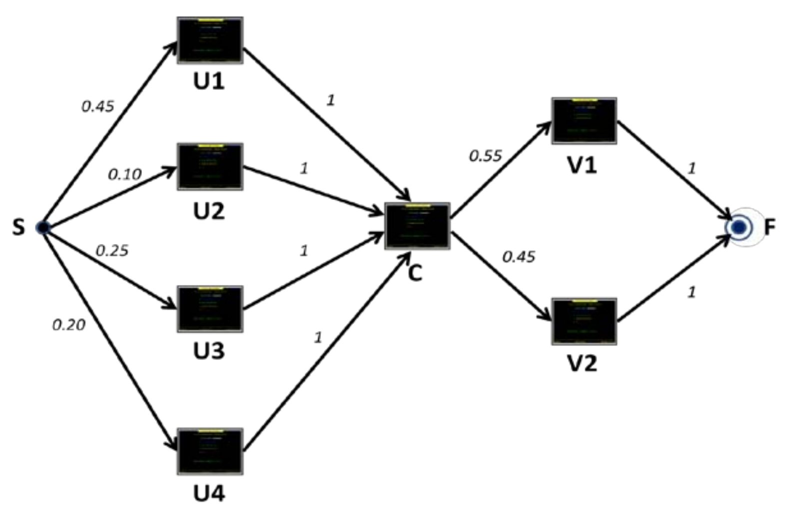


Figure 2.43 - The first order transition model with the not-normalized place C

This process representation, based on first order transitions, is not normalized in the service C, which has a many to many transition. There are 4 inputs and 2 outputs, so 4 clusters (or 4 splits of the service C) should be enough to arrive at a normalized process representation, as follows:

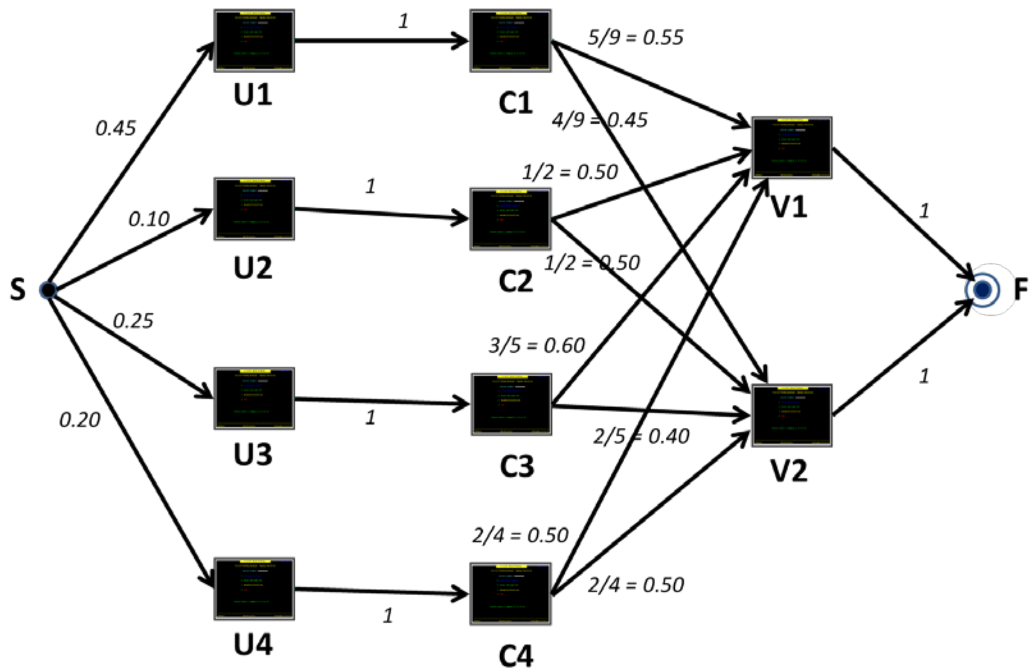


Figure 2.44 - The corresponding normalized HMM

Not only this is the correct representation of the process, the first order transitions induce ambiguity about the real process path probabilities. Indeed, they result in – for example – the following path probabilities:

P(S,U3,C,V2,F) = (5/20)x(5/5) x (4/9) = 20/180 = 11,11%

P(S,U3,C,V1,F) = (5/20)x(5/5) x (5/9) = 25/180 = 13,88%

whereas the real path probabilities are the following:

P(S,U3,C,V2,F) = P(S,U3,C3,V2,F) = 10%

P(S,U3,C,V1,F) = P(S,U3,C3,V1,F) = 15%

which are exactly the path probabilities that are obtained from the normalized process representation.

The idea to represent the clusters by means of “clones” of particular places in the HMMs is inspired by the PhD-work of Levene on Web Analytics [Borges & Levene, 2005]). They have, however, no systematic discussion of normalized models.

What has been shown in this example extends into all process representations that are not normalized: ***irrespective the algorithm that is used to obtain the transition probabilities, the first-order transition probabilities are incorrect.*** This is not a consequence of the algorithm, but just the result of have a not normalized representation. It is obvious that this has far reaching consequences for all papers that don't use normalized process representations. A systematic overview of them is out of the scope for this PhD-work, since the main objective is not the ultimate mathematical representation of the process, but practical ways in getting insights in potential ambiguities and value leaks in the processes, which will be illustrated in the chapters 4 and 5.

As a last remark, note that the further clustering some events that always belong together (such as was done in the above example with A and B, and C and D, respectively, can also be useful to reduce the complexity of large scale models. This point will be touched in the next chapter again, when dealing with automated Process discovery based on intercepting all possible screen and keyboard interactions in an end-to-end processing, typically around a server environment. The event log that follows from scanning all events in the network typically contains a relative high number of entries that have represent in reality the ***sequence interaction*** (also known as the user interaction sequence of steps) that is needed to actually perform a particular Process step.

The following schema gives an overview of one hybrid approach of techniques that may be combined together for Process Discovery [Peters & Dedene 2011]:

a) Assign a unique id to each sequence of services/activities, based on the actor and object(s) that are involved in the Process.

b) Construct an initial Hidden Markov Model HMM1 by means of the (standard) Expectation Maximization algorithm applied to the above sequences. Observe that this gives more precise results than other clustering methods (e.g. sequence clustering), because the sequences are not anonymous.

   Observe in HMM1 the number of many-to-many transitions (=M) and the maxima of their input-output transitions (K_M).

c) Perform a first-order transformation in which all transitions in the original HMM1 become places in a transformed HMM2.

   Transform all sequences according to the transitions in HMM2.

d) Perform a clustering analysis on HMM2, searching for MxK_M independent clusters, represented by a HMM.

   Retrofit the clustered sequences to the original HMM1.

e) Repeat c) and d) until perfect normalized HMMs are discovered in the clusters.

f) Perform a final Expectation Maximization analysis to reconstruct the detailed HMM for each of the clusters, also to discover probability variations within the original clusters resulting from c), d) and e)

Various combinations are possible, depending on the volume of the event logs and the complexity of the process that must be discovered.

2.10 Summary

The main original contributions to Process Discovery research in this chapter are the following:

- the specialization of the C-K-theory to the problem of Process Discovery. This thesis work is the first that relates C-K-theory in detail to Process Discovery, after the successful applications of C-K-theory to Data Discovery.
- although there have been indications in other thesis works that the Petri-net representation is less suited for Process Discovery and Process Analytics in general, this work illustrated how several problems (such as matching an actual event sequence with a process representation) is easier to deal with when using the Hidden Markov Model representation.
- the notion of Process Model Normalization, stressing the potential ambiguities induced by many-to-many transitions in Process representations. It is demonstrated how several complicated extensions to the Petri-net representation (such as in [Greco et al., 2006]) are merely needed because of models that are not normalized. Moreover, it is clearly shown how models that are not normalized can lead to erroneous analytical results.

In the next chapter, a complete technology toolset for practical Process discovery is presented and discussed. Although that toolset is already capable of dealing with very large volumes of event data, compliant to the Actor-Event-Object pattern; the current chapter sets the agenda for the further evolution of this toolset.

Chapter 3

The Tooling and Process for Process Discovery

3.0 Introduction

This section provides an overview of the *Comprehend* product suite developed by OpenConnect Systems (OC), Dallas, Texas, USA, in conjunction with research performed at KU Leuven. OC designed *Comprehend* to perform automated business process discovery primarily for service sector operational processes where the service is performed on a large scale and where data-workers operate business-critical systems. Some ideal operational environments are manual claims processing, contact center workflow analysis and productivity improvement, telecom order entry and tracking, and so on healthcare insurers in the United States use the product extensively to reduce administrative costs associated with processing claims and adjustments, a particularly expensive form of rework. *Comprehend* users routinely improve operational first pass rates (a measure of process productivity) used by healthcare claims rates by several percentage points. These business users have spent years improving their processes and have reached diminishing returns for their efforts because they did not have the ability to measure worker activity in detail and/or process variations. Therefore, they could not calculate the true costs of their workflows. This section describes the technology and methodology at a high level. The technology description represents the architecture of *Comprehend*. The methodology section discusses each major step in the method including roles and process steps; it also includes actual system screenshots and describes inputs, transformations and outputs.

3.1 The Origins of *Comprehend*

Comprehend is comprised of several software components working together to provide data collection, business process discovery, process mapping, analysis and reporting of operational process information.

The journey, whose outcomes to date are presented in this Thesis, began in 2005. I had recently joined at OpenConnect Systems, Dallas, TX, which at that time was the producer of a software tool named WebConnect, a web to host (IBM 3270 Series) terminal emulator. The product allowed a person to use an internet web browser to connect to a "green screen" mainframe system thus allowing the legacy technology to utilize a modern web interface without the operator having to do anything more than what they did on the old system. By 2005, this market (known as terminal emulation) was no longer growing and had fully commoditized.

The WebConnect product consists of a number of components, the main grouping being a "listener" or data collector which can intercept the signals (packets) as they are transferred over the mainframe 3270 communications protocol. It moves across the wire from the mainframe to the desktop machine and finally a screen painter. The data in these packets shows the order in which the user accessed the application at the screen level starting with the log-on procedure. It then shows how they moved from data field to data filed, as well as from screen to screen, as they performed various tasks. Since the technology was based on the overall 3270 standard for terminal emulation, it captures the data regarding any 3270 (mainframe) emulator session, not just those of OpenConnect. This led me to believe that if this data could be captured and stored, it would represent a true sourced data about "work" as it "*actually occurred*" since it provided an activity log of what the person "*actually did*" as they worked with an application to perform some defined task. In essence, the data were captured from the work performed by the employee at the user interface itself and, therefore provide source data about the actual instance of work. My thought was that if one had this type of information, it could provide valuable insights into how time and resources were being spent to accomplish some defined task as well as show whether they were being used effectively. As a young industrial engineer, I had spent time performing task analysis and productivity studies for a healthcare insurance company. I used work activity analysis and other direct observation and self-reporting data collection methods to track the productivity of medical claims examiners using ICT-based systems. In this approach, one of the major concerns is the integrity of the data and its collection, ranging from the role of the observer to the veracity of the data when self -reporting methods are utilized. This was a shortcoming in the process since the actual work was in essence manually "unobservable" and the recording methods were based on manual observation.

With my work at OpenConnect, my initial hypothesis was that since the actual observations were indeed being recorded and captured by the technology itself, if it could be stored and re-created, it would be an actual or fact-based representation of the actual activity rather than what amounted to a best guess. Also, the recorder sequence of actions would represent an actual flow of the process, not an idealized flow or something based on inference, speculation or opinion. In essence, the event logs generated could be used as the basis for reconstructing actual processes around applications.

As this effort was to be based on captured data, it would be a "bottom-up" approach. This would be different from the normal "top-down" modeling effort. Specifically most commonly used to top-down modeling efforts provide a taxonomy at the beginning of the process. I was clear the effort needed to be based on an "*emergent paradigm*" since the data would need to be at the lowest (atomic) levels and the groupings would need to be found or "discovered". During this time I became familiar with the work of Cook and Wolf (Cook & Wolf, 1995/1996). They had coined the term "process discovery" while working to automatically uncover processes in the software development effort. They suggested three techniques that could be used; Neural Networks, algorithm-based techniques and Markov Chains. After much discussion, it was decided to use Markov models. This technique discovered clusters and provided a representation that can be best described as a probabilistic finite state machine. This was a revelation in two ways; first, it showed the probabilistic nature of the work process and second, it showed that there were many issues in the model itself. Specifically, while we could initially understand the model, it contained a level of ambiguity in truly understanding the root cause of any deviation from the main path of the model (e.g., process anomalies). This was a key moment since we now understood that it was not the main process that had to be "discovered" but rather, it was the exceptions (or the noise) that were of key interest rather than a generalized process. What we understood from our field experience was that it was the process exceptions that were root cause of problem that drained productivity from the environment hence "leaked value" from the process. We focused on finding which techniques were most suited to understand value leaks in business processes and assist in understanding how to deal with these value leaks (e.g., close the *knowledge-gap* on the true cause of value leaks and service imperfections thereby providing guidance to management on ways to improve process performance)?

It was at this point that I contacted Professor Dedene at KU-Leuven. We discussed his research in probabilistic finite state machines and its applicability to automated process discovery. This resulted in an OpenConnect sponsored Research Chair at KU Leuven to focus on the issues of automated process discovery and analytics. As an initial step, we sponsored an open, roundtable discussion that included various academic and industry representatives with products and interest in the subject. The session included presentations from the Gartner Group, BMC, Software AG, OpenConnect and KU Leuven. To the best of my knowledge, it remains the only multi-vendor, industry analyst and academic session devoted to this topic as of this day.

After the aforementioned industry and academic session, we worked with Professor Dedene and his team to explore different approaches to the problem of anomaly detection and elimination in Process Discovery. I remember clearly a dinner with Professor Dedene, where I mentioned that the problem of process anomalies reminded me of the data normalization issue where ambiguity was removed through the process of normalization. This triggered other thoughts leading to the development of the hybrid model described in Chapter 2 of this Thesis. This was clearly the major scientific breakthrough of the effort.

As stated earlier, the primary goal of automated business process discovery (process discovery) is to "understand" processes rather than to "control" processes. It is the increased understanding that leads to the closing, or at least narrowing, of the management knowledge-gap regarding the root cause of process inefficiencies.

Remember from chapter 2 how Process Discovery tries to go further than Process Mining, by explicitly addressing process variations and process exceptions. Process Discovery incorporates the exceptions and variations and investigates them as the root cause for value leaks in the process.

3.2 Architecture

Comprehend's architecture is depicted below (Figure 3.1). *Comprehend* adds the critical time dimension to the usual volumes that previous analysis products use.

Major *Comprehend* Components:

- **Collectors**
  - Passive 'Sniffers' for mainframe and web traffic
  - Agent for desktop client / server applications
  - Custom data import and normalization
- **Analytics Component (AC)**
  - Logic for discovering, defining, delimiting and labeling activities, aka logical units of work
  - Builds events from click stream data,
  - Augments input data with process metadata, and provides input to automated process mapping
- **Process Intelligence Component (PIC)**
  - Process mapping, rather than modeling, based on empirical data
  - Process analytics based on event data
  - Interactive Analytical workflows
- **Reporting**
  - PIC analysis is exposed to standard reporting engines
  - Production style reports, aimed at operational leads, managers.
  - Input in standard formats for ETL to existing data marts or data warehouses
  - Input to operations dashboards
  -

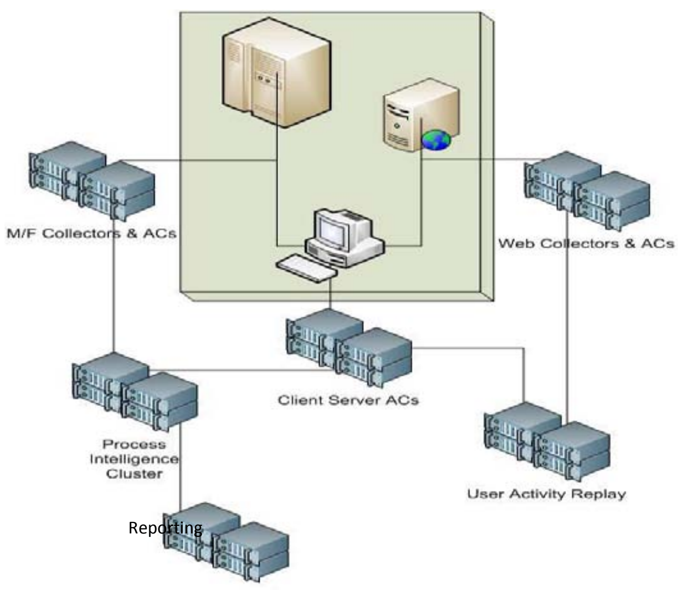


Figure 3.1 – *Comprehend* Architecture

3.2.1 Collectors

3.2.1.1 3270 Collector

The 3270 Collector is a network "sniffer" (i.e. makes copies of packets as they traverse the network) that observes TCP/IP network traffic and records only TN3270 sessions (all other traffic is ignored). These session files are stored as collector specific files locally on the collector hardware. Because of the low-level access to the Ethernet device, the collector component stores and encrypts data on the fly. Additionally, since the collectors passively observe and record traffic as it flows across the communication channel, it makes no changes to the monitored servers or software, and does not cause any performance degradation. Additionally, password fields are redacted in the recorded screen shots and not stored in Comprehend.

On installation, the collector is provisioned through a local HTML wizard that requires the installer to provide the collector with an encryption key phrase that is shared between the Analytics Cluster and the Collector, and used to encrypt/decrypt the collector specific files. This key phrase is only used to generate a shared key for Collector-AC Server communications and is never required or used again. The installation also requires the administration password for the Analytics Cluster to register the collector with the AC Server. Any change to the credentials requires a re-installation of the Collector.

All further configuration of the collector is done through the AC server. The collector can be configured to monitor specific internet protocol (IP) & port ranges on the client side or the server side.

The 3270 collector captures all TN3270 traffic, and provides a temporary encrypted collector specific store of these sessions. However, all non-display modifiable fields (typically password fields) are never stored on the Analytics Cluster; they are discarded during initial analysis.

3.2.1.2 HTTP Collector

The HTTP collector is a network sniffer that observers and records web-specific TCP/IP traffic and stores steams as user sessions. These sessions are encrypted and stored locally as collector files. Because of the low-level access to the network device, the collector runs as root. On installation, the collector is provisioned through a local HTML wizard that requires the installer to provide the collector with an encryption key phrase that is shared between the Analytics Cluster and the Collector, and used to encrypt/decrypt the collector specific files. This key phrase is only used to generate a shared key for Collector-AC Server communications and is never required or used again. The installation also requires the admin password for the Analytics Cluster to register the collector with the AC Server. Any change to the credentials requires a re-installation of the Collector. The configuration of the collector is done through the local interface; the local interface provides three logical functions traffic filtering, data exclusion and HTML stream configuration. Traffic filtering allows the collector to monitor only specific IP addresses & port ranges on the client side or the server side. Data exclusion rules strip data out of the captured HTML stream ensuring that unwanted data is never stored locally, or transmitted to the AC. Traffic filtering is often used to strip sensitive information, (e.g. passwords, etc.,) out of the data stream. Finally HTML stream configuration is a set of rules that encompass session detection, page naming and content extraction.

3.2.2 *Comprehend* Analytics Component (AC)

The Analytics Component (AC) analyzes collected data, and stores it in a relational database. The analyst interacts with AC to determine when events start and stop, what data will be kept and discarded, groups collected data into activities which are logical units of work, and timestamps each event. The collector specific files are pulled from the collectors via a HTTPS connection, and are analyzed and then stored into the relational data store. The user interaction with the AC is via an Eclipse workflow connecting to the server process via HTTPS connection. On installation, three sets of credentials are required, database administrator, *Comprehend* database user and *Comprehend* AC administrator. The relational schema is created in the database using administrator credentials to the database. The *Comprehend* Server uses the database user credentials to access the database; this can be a normal user role without any data definition language specific (DDL) privileges. The *Comprehend* AC administrator is the system administrator for the AC and is responsible for configuration and installation. All of these credentials are available for modification through the administrative interface of the *Comprehend* AC. All data collected is stored in the relational data store. The “raw” data collected is stored encrypted as binary large objects (BLOBs) within the database. Users with access to the encryption key phrase can view this “raw” data.

This data includes:

Session level replay – the system will be able to decrypt the collector specific file and replay entire sessions.

Field level data – For any Web Page or TN3270 panel, the system can show all values observed for any particular field.

The following data is available to any user of the *Comprehend* AC:

- Statistical data relating to usage paths through the system
- Statistical data relating to page / panel use
- Statistical data relating to think time / system time for individual transactions

3.2.3 *Comprehend* AC User Interface

The *Comprehend* AC User Interface (UI) is a set of Eclipse plug-ins. These provide the interaction between Eclipse and the AC server. All communication is done via HTTPS and users are authenticated locally via challenge by Eclipse. A secondary challenge prompts for the data encryption key phrase. On 15 minutes of inactivity, the challenge expires and upon the next invocation of data secured functionality, the user is re-challenged.

User credentials are managed through the administrative user interface of the *Comprehend* AC. Credentials can be expired, deactivated, created or changed via this interface.

3.2.4 *Comprehend* Process Intelligence Component (PIC)

The Process Intelligence Component provides fast *ad hoc* analytics and reporting on captured process events. All user interaction with the PIC is via a web browser, with a variety of user specified roles and credentials. The Process Intelligence Cluster is a set of java processes and an XML database that provides the analytics and reporting on the process events. All user interaction with the PIC is via an applet or HTML browser interface over a HTTPS connection.

On installation, an internal shared key is generated used that is used to encrypt/decrypt any secured data. The shared key is not used externally to the application. All user administration and provisioning is done via the web server, and authenticated users (with an appropriate role definition) are passed to the PIC UI. The PIC data security role controls what data elements within PIC are encrypted vs. unencrypted. This is done at an attribute level. User roles within PIC can be created such that these roles have one of three levels of access to the PIC data, decrypt, view, none. Decrypt access provides the ability to see the attribute values unencrypted. The server uses the shared key to decrypt the data and pass it over the HTTPS to the application. View access allows users to view the data, and if the attribute is encrypted, the user is able to see encrypted values of the attribute values, but not the values themselves. No access provides no visibility to the user that the attribute even exists within the PIC.

3.2.5 Reporting

Reports can be extracted from *Comprehend* through any standards-based (e.g., ODBC, JDBC, XML) report writing software. While LogiXML is the embedded reporting engine, *Comprehend* data can also be exported to Microsoft Excel and other analytical tools (e.g., Tableau. LogiXML) where data further examined or combined with external data for other types of analysis. Project teams routinely roll up data into business objects; say lists of all orders touched in a day, or all users who touch each order. This information can then be sent to existing reporting engines, and analysts can use extract transform and load (ETL) tools to integrate it into their environment, though it is also available on a standalone basis and integration is not required.

3.2.6 Security Provisions

The following sections summarize the security provisions and default security roles built into *Comprehend:*

3.2.6.1 Data storage protocols

All data that is stored is encrypted using 256-AES encryption, where legal. No captured data is ever stored or transmitted in the clear. At the AC/Collector level, there is the ability to configure *Comprehend* to strip data out of the event stream ensuring sensitive data is never stored.

3.2.6.2 Data transmission protocols

All network communications between major components of the *Comprehend* solution is performed over SSL. Where possible, the data stream within the SSL wrapper is also encrypted. Additionally, *Comprehend* is generally deployed on the customer campus so that sensitive data is less subject to interception.

3.2.6.3 Audit

Comprehend provides an audit log of system usage and configuration changes.

3.2.6.4 Usage Log

System usage can be logged and time-stamped, providing an audit log of tasks executed in the system.

3.2.6.5 System Change Logs

Whenever there is a change to the configuration of *Comprehend*, a system change record is generated. Specifically, changes that relate to:

4 Changing a role definition (add/delete/modify)
5 Changing the configuration of a collector (source monitored system)
6 Changing a data exclusion rule (omit data from a collector data stream)
7 Changing the encryption passphrase
8 Changing the level of security on an attribute in PIC
9 Adding new attributes to PIC

3.2.7 Standard Roles

While in any implementation there is configuration of roles. *Comprehend* includes the following default roles:

3.2.7.1 System Administrator

The system administrator has only rights to install and do base level configuration of *Comprehend*. This would include data retention policies, server start/stop/status, system upgrades, etc. The System Administrator does not have any rights to view any of the data collected by *Comprehend*.

3.2.7.2 Data Security Administrator

This role controls access & visibility rights to the data collected by *Comprehend*. This includes the ability to configure the systems the collector is monitoring, any data exclusion rules, as well as the control of what attributes are encrypted/viewable by role.

3.2.7.3 Security / User Administrator

The Security or User administrator provides user level administration to *Comprehend*. This role is responsible for managing the user to role mapping (done in the existing user management solution integrated to *Comprehend*). As well as would modify roles based on new use requirements by the user community.

3.2.7.4 AC – Developer

The Analytics Cluster developer is responsible for analyzing the system usage data, and creating processes. This role typically has access to view the unencrypted session level data. However, this is not required.

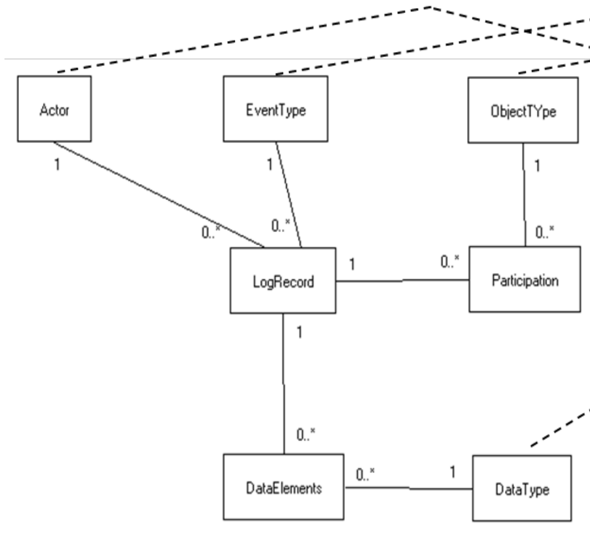
3.2.7.5 PIC – Analyst

The PIC analyst role(s) are the power users of the process analytics in PIC. These analysts investigate the process variations; examine data attribute values that are driving the variations.

3.2.8 *Comprehend* Artifact Capture

The key benefit of automated business process discovery is, as the name suggests, the ability to automatically capture the elements of work in a digital environment from the work process itself and re-create the activities and events as they actually occurred during the execution of the process. Artifacts (e.g., data. time, user id, etc...) are captured automatically by the previously described collectors from ICT-based work sessions on multiple platforms (TN3270, client-server/desktop, web) or from various log or audit files. Additionally, *Comprehend* has facilities to augment collected data with relevant metadata, to query databases in order to collect process changes prior to deployment, and in cases where manual capture is the only available method, an xml data input mechanism is also available.

As shown previously in Chapter 2, the basis for the input data for analyzing the artifacts is the Actor – Event – Object model, which is open-ended to any type of process characteristics (such as time, risk probabilities etc) that desire to be analyzed. The UML representation for this model is shown in Figure 3.2. This representation easily accommodates the transformation of the data from logging components into process representations that can be reviewed for accuracy by process owners or domain experts.



Actor Id	Event Type Id	Object Id	Event Occurrence Time Stamp
John	A	1	9-3-2004:1[illegible].01
John	A	2	9-3-2004:15.12
Sue	A	3	9-3-2004:16.03
Carol	B	3	9-3-2004:16.07
Mike	B	1	9-3-2004:18.25
John	C	1	10-3-2004:9.23
Mike	C	2	10-3-2004:10.24
Sue	A	4	10-3-2004:10.35
John	B	2	10-3-2004:12.34
Pete	D	2	10-3-2004:12.50
Sue	A	5	10-3-2004:13.05
Carol	C	4	11-3-2004:10.12
Pete	D	1	11-3-2004:10.14
Sue	C	3	11-3-2004:10.44
Pete	D	3	11-3-2004:11.03
Sue	B	4	11-3-2004:11.18
Clare	E	5	11-3-2004:12.22
Clare	D	5	11-3-2004:14.34
Pete	D	4	11-3-2004:15.56

Figure 3.2 – UML Representation Model

Each set of data and the process and methods used for session capture, activity definition and event formation are described in the following sections.

3.2.8.1 Session Capture

All information about work is discovered through capturing the data in real time from the process itself. Specifically, data is captured from the actual work session that is being performed by a resource (person) as they interact with the application through the user interface (screen). In this manner, every interaction (keystroke, data change, etc..) the resource makes with the screen during the session is extracted and captured along with a time-stamp of when the page was viewed, the data change made as well as the IP address of the client. In essence, the session is described by who the resource was, the time the session took place, a view of what was viewed as well as what was changed. Information can be captured from mainframe/3270 platforms, client -server /desktops or web-based sessions.

With session capture every page or screen image is captured and stored. Specifically, data is extracted from each data stream, time stamped as to when the page was viewed and includes the Internet Protocol (IP) address of the client workstation on which it appeared. For IBM mainframe-based systems, the full TN3270 user payload is extracted for both pages (if working under a terminal emulator) and a screen image is recreated. In a desktop based client-server/environment, any desired data can be captured and stored and the data is extracted from the screen and stored as data in the database. A similar process is used for web-based systems. Each type of session is described in more detail below.

3.2.8.2 Mainframe Session Capture

Mainframe session capture is based on intercepting and copying the signals sent between the mainframe itself and the client workstation (Figure 3.3). In this manner, all TN3270 newtork traffic is captured. User workscreen information is captured based on analyst-specified information transmitted between the user and the application. After initial capture, the appropriate panel rules are determined based on the observed consistent placement of text on the panels of interest. Next, data/field tags are created to extract information from the TN3270 stream based on field tag name, panel position regular expressions, or other criteria. These felds can then be manipulated through JavaScript or regex (Linux/UNIX regular expression parsing).

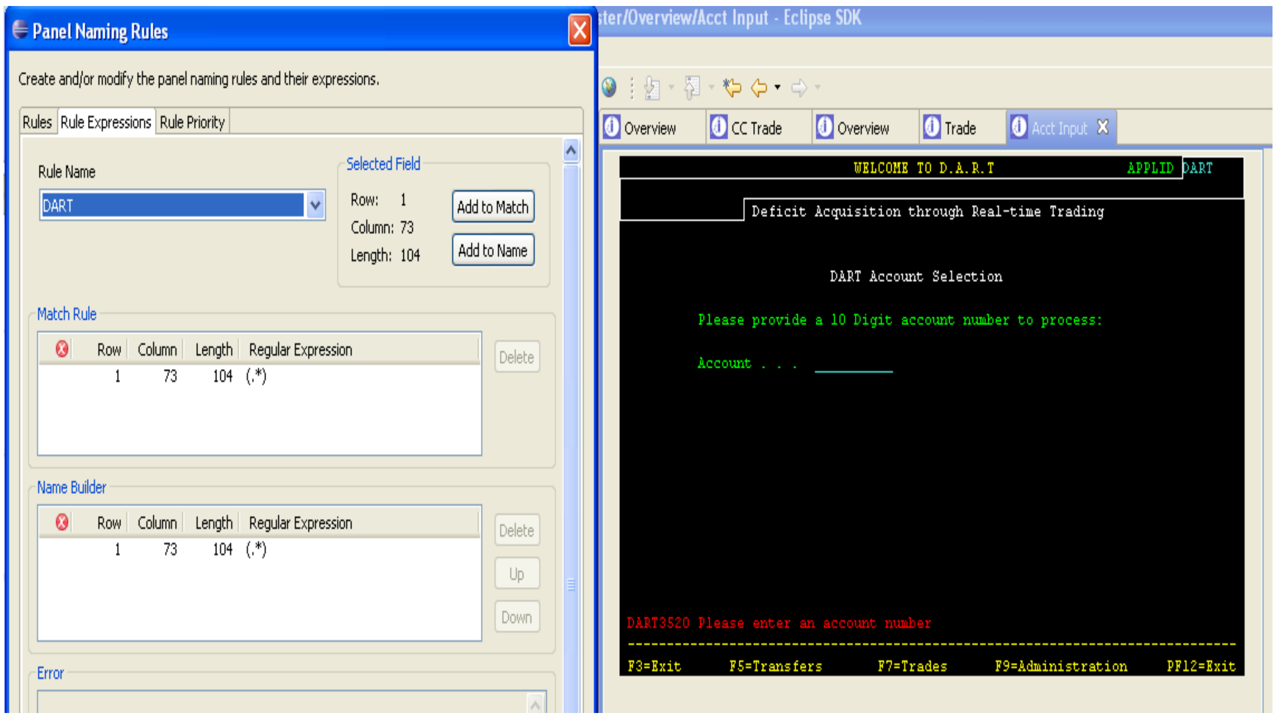


Figure 3.3 – Mainframe Session Capture

3.2.8.3 Client- Server/Desktop Session Capture

Client-server/desktop session capture differs from mainframe session capture in that the information is not intercepted from the network communication transmission but rather from the actual changes on the desktop device itself (Figure 3.4), by a resident agent with a small memory footprint. This is necessary as the important data may never traverse the network. In this manner, user interactions, process name, window titles, before and after data, focus change; Windows handles (for differentiating identical windows) are all available for capture and analysis. In this environment, field level data values are extracted from Windows controls through the use of code embedded rules or “triggers”. Triggers are implemented on a central server environment and “pushed” to the desktop by standard enterprise methods. All data is logged locally on the client/desktop and sent or “pushed” to the desktop collector server in near real time, in order to reduce the memory footprint that would occur by storing data locally.

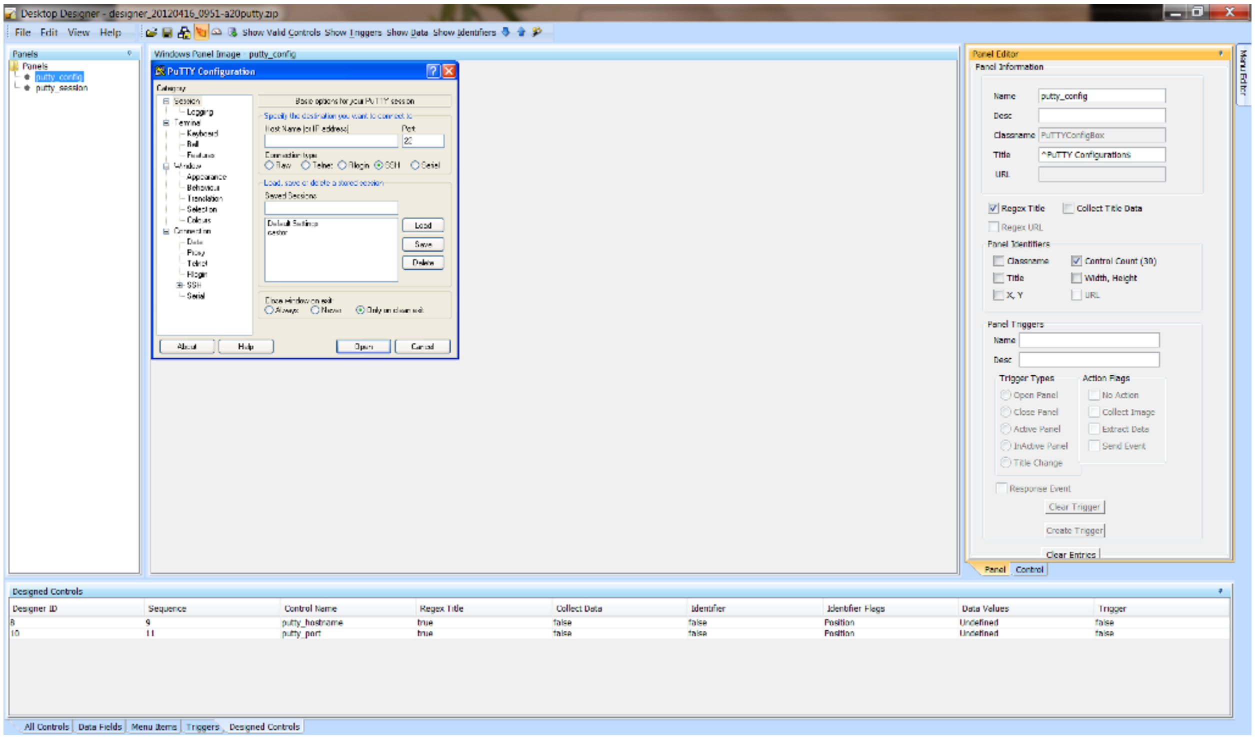


Figure 3.4 – Client-Server/Desktop Session Capture

3.2.8.4 Web Session Capture

Session capture in terms of web/internet data is similar to the mainframe oriented methods described earlier, or can be captured by the desktop collector agent should the project require more detailed analysis. In one former case, it is pure, passive "click-stream" information that is of high interest. In web data capture, page naming rules are defined based on matching URL stems, page content and page tags (Figure 3.5). In the latter case, *Comprehend* can collect and report on detailed workflow analysis among web and desktop applications

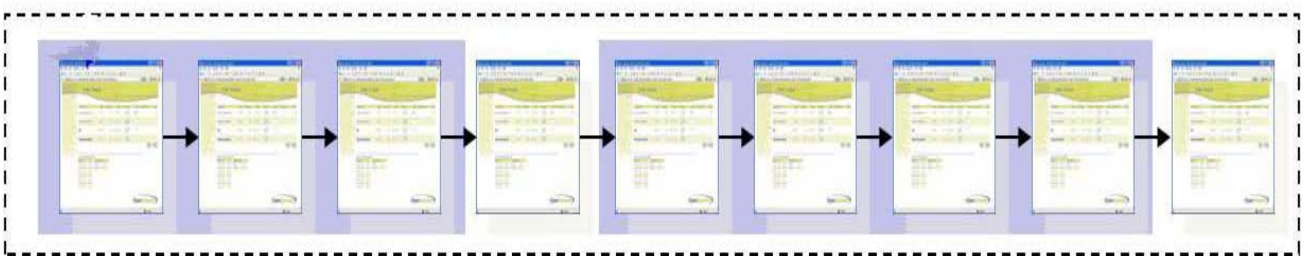


Figure 3.5 – Web Data Capture

The rules for content extraction are described in XPath. XPath is a programming language that XML uses path expressions to select nodes or node-sets in an XML document, hence the name.

3.2.9 Activity Definition

The *Comprehend* Analytics Component is where activity definition takes place. Activities consist of a set of low-level actions comprising a logical unit of work that is being observed.

Activities generally consist of several user interactions and may navigate several screens, panels or desktops. In the definition of activities, screen-to- screen transitions are combined based on observed actual user usage patterns. Groups of activities form the basis of "events" (e.g., input claim) that are generated by *Comprehend* as a component of the discovered business process. All information is based on data that are extracted from activities that are defined by member pages of the activity itself (Figure 3.6).

3.2.10 Event Generation

Analytics Component – Event Generation

As described in Figure 3.6, activities define the core elements associated with business events. Once defined, data associated with all pages, screens, clicks and keystrokes are collapsed into an event. Each event is described by an event descriptor that includes both attributes in XML documents (e.g., screen name, user id, and IP address) and actions as attributes of the process. This also includes the source/destination of the information as well as a timestamp, and an activity name associated with the event.

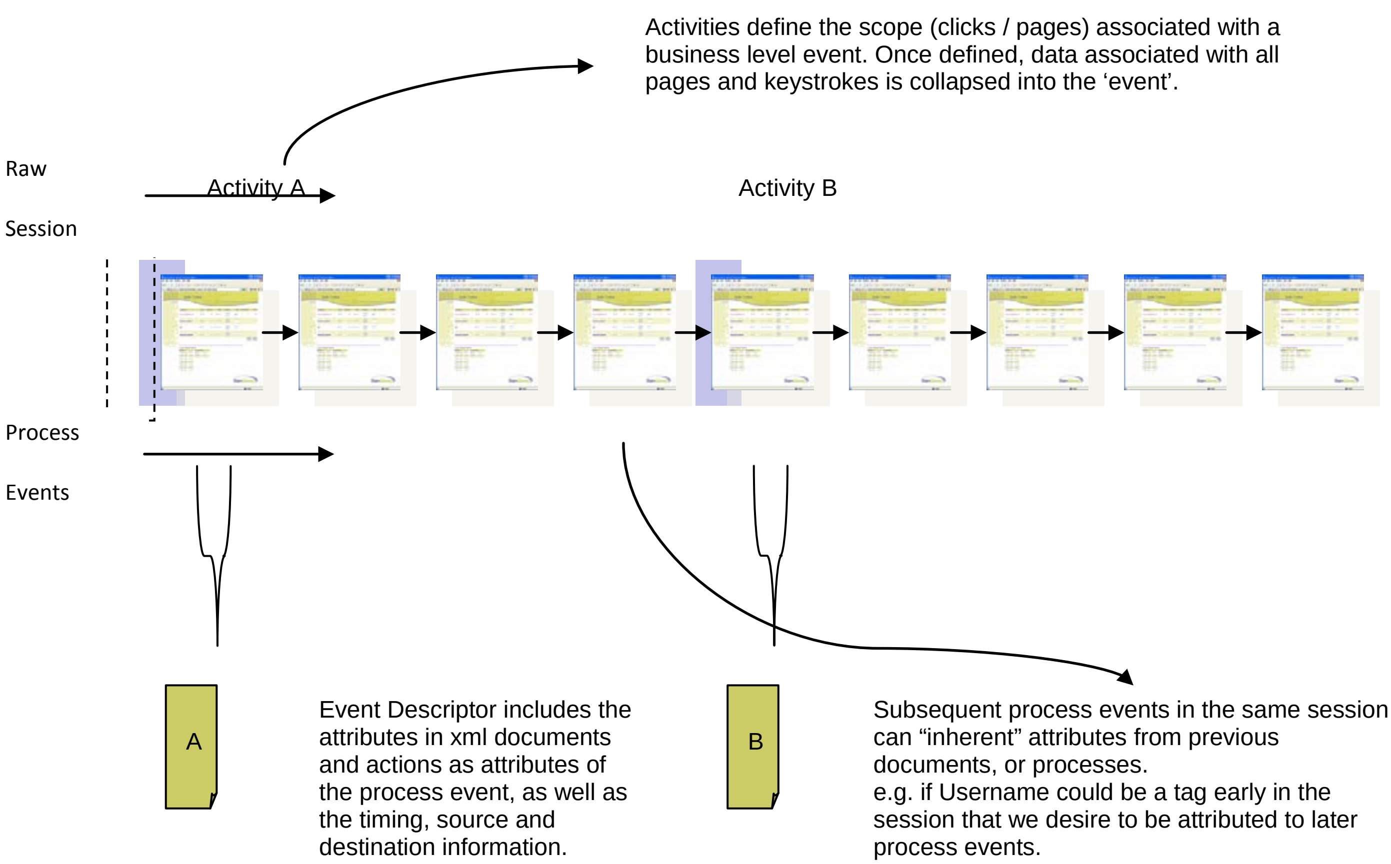


Figure 3.6 – Definition of Core Elements

3.2.11 Process Intelligence Component

The *Comprehend* Process Intelligence Component (PIC) is a process intelligence engine enabling more abstract business processes analysis of activities defined by the *Comprehend* Analytics Component (AC). Where AC analyzes logical units of work (activities), PIC provides the ability to perform analysis on an entire process (a collection of activities) based on any attribute or combination of attributes contained in the processes. Additionally, PIC allows for multi-system analysis by combining real-user interaction on disparate systems and augmenting this information with data from existing system interfaces, such as log files, databases, or essentially any user defined data feed, for a full end-to-end business process analysis.

3.2.11.1 PIC Process Perspective

The PIC can perform multiple analyses on the same set of data. Each analysis is called a perspective. Each perspective is based upon the presence of a single designated attribute that is selected by the user. Every event in the system that contains the designated attribute is grouped by unique attribute value, ordered by time, and passed to the analysis script for processing. Of course, additional attributes are also typically present, such as critical data fields to be analyzed, userid, timestamp, and activity name.

The grouping and ordering of events creates a series of traces of which each trace is associated with a specific value of the specified attribute in the perspective. Perspective analysis generates the nodes and transitions on the activity map, trace attributes, and rollup events. Trace attributes are data elements annotated onto each event that can be filtered and charted just like a normal event attribute. Rollup events are synthetic events created by the script, contain only trace attributes, and are used to summarize output from the analysis of a trace.

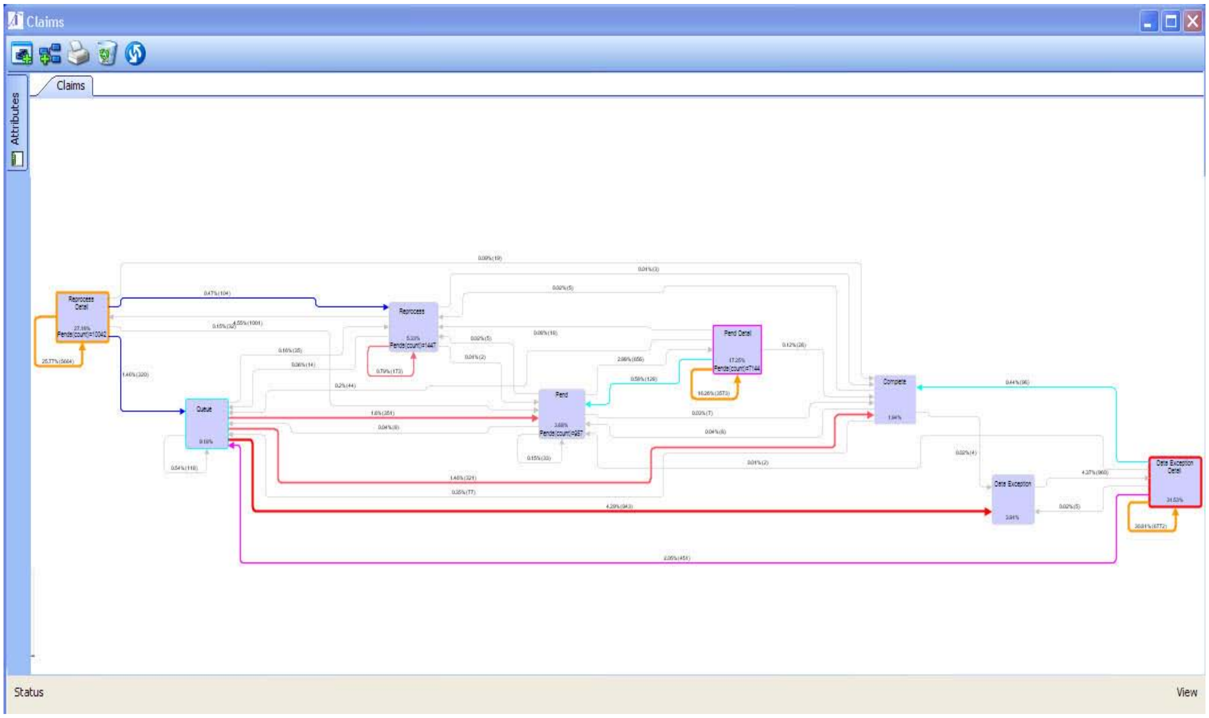


Figure 3.7 – Example Values of User IDs used with Defined Processes

PIC Process Perspectives are data views based on a specific attribute. They are created through an analysis of the event data where the events are used to construct processes consisting of activities. Activities are a particular type of event, (i.e., login).

An example of an attribute is a *Comprehend* field tag such as UserID. This attribute (Figure 3.7) represents the values of the User IDs used with the defined processes. In this way, the perspective provides the analysis of how users move from activity to activity.

As PIC acquires new events from AC, the perspective is updated. The Perspectives panel allows the analyst to create Perspectives based on attributes to create a specific view of the data. They can be edited to change properties associated with it. The use of open ended user defined scripts allows new trace attributes (often called business objects), the combination of existing attributes into trace attributes, or the suppression of events relating to existing attributes.

PIC Analysis Scripting allows the user to customize the data created by the analysis process. The default analysis performed by the PIC server generates the transition graph from activity to activity, and calculates two default trace attributes. The default attributes are **deltaT** and **deltaP**. These attributes contain the time in milliseconds between the current event and the next or previous event respectively (Figure 3.8).

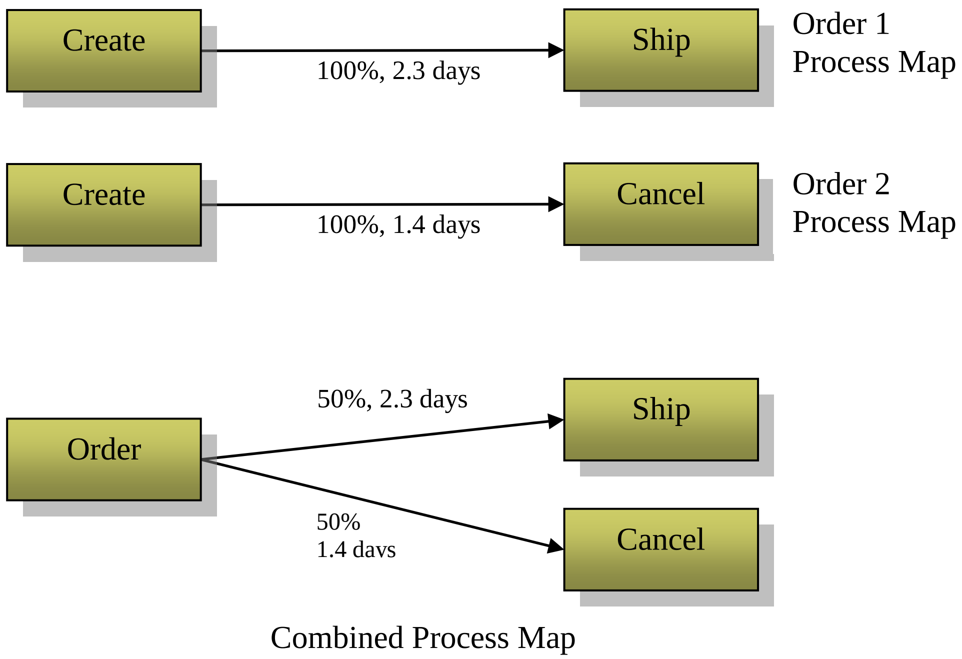


Figure 3.8 – Combined Process Map

The following picture shows the example from Chapter 2.8 in the *Comprehend* PIC. In this example, the input data was provided through the Event-Actor-Object data format. It shows how PIC adds more information to the models presented in chapter 2.

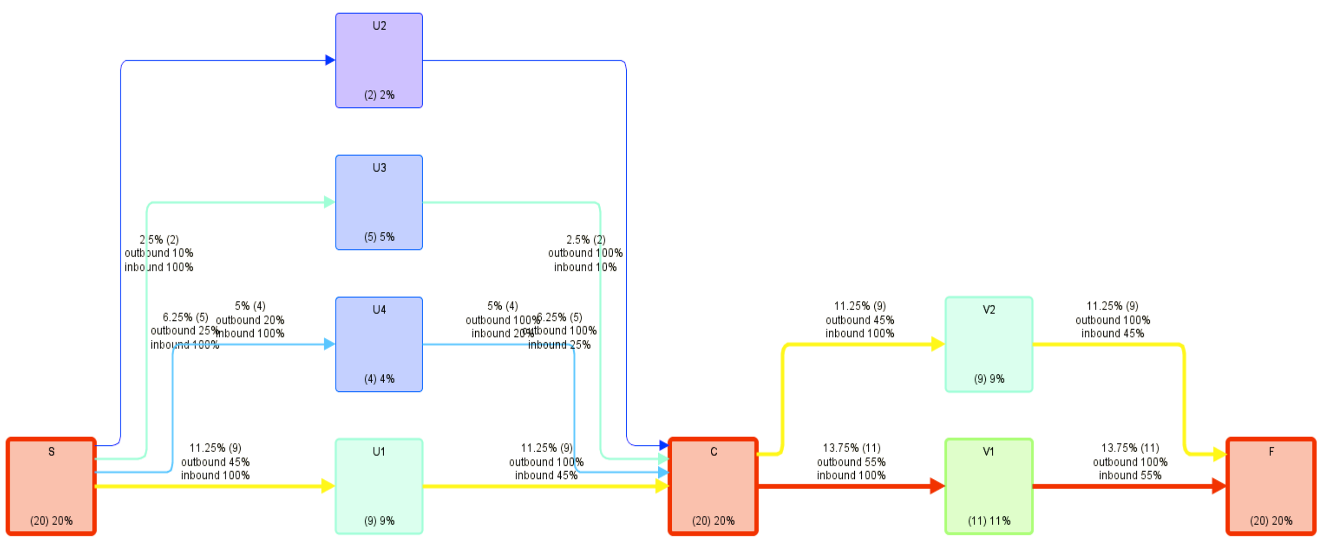


Figure 3.9 – Example *Comprehend* PIC

3.2.11.2 PIC Process Integration and Filtering

Filters are another dimension (beyond X and Y axes) that allows an analyst to test a hypothesis, based on the empirical data in the system can also be set to refine a process map based on any other vales in the tag set (e.g., user-id, claim number, etc…). Thus filters provide additional views to be generated that can examine the timing of events, sources, or other items of interest. This function allows for the interrogation of the data in order to examine and understand root cause process variations from any perspective on the process**.**

3.2.12 *Comprehend* Reporting

In addition to drawing process maps, *Comprehend* includes a number of graphing and charting capabilities (Figure 3.10). These charts show descriptive statistics on any attribute of interest (e.g., claim edit codes by examiner by date, etc...). A standards-based interface also exists for exporting data to popular reporting and analysis tools such as Microsoft Excel, export to an included relational database, synthesis of events, or any number of user defined extensions to *Comprehend's* out of the box capabilities.

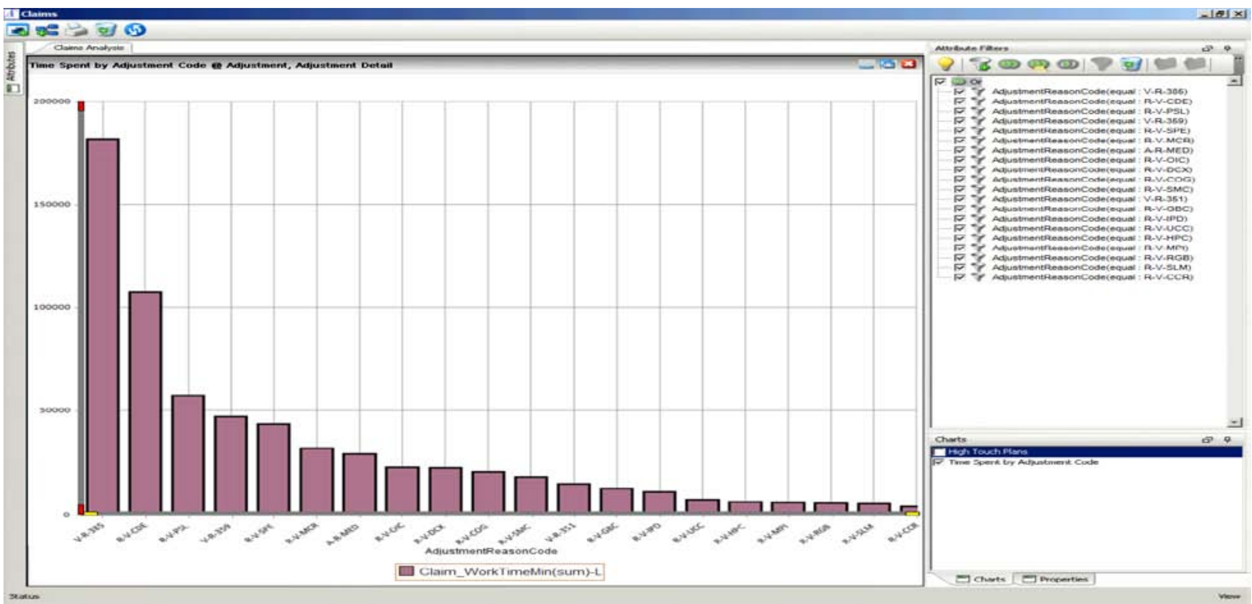


Figure 3.10 – Example Graphing and Charting Capabilities

3.2.13 Summary

Comprehend was built to scale to accommodate data capture and analysis in the largest enterprises in the world. It is common to store information on thousands of users with just a handful of industry standard servers. Additionally, *Comprehend* is flexible. Should the requirements change, it is simple to change the data collected, analyze the process in a different way and determine whether the hypotheses necessary to improve a process can be accepted or rejected, based on empirical data, not assumptions and guesswork.

3.3 The Process of Using Comprehend

The concept of automated business process discovery is based the analysis and instrumentation of data that is automatically captured from the work activity itself rather than by human observation of the process. Automated Business Process Discovery tools capture the required data and transform it into a dataset that can be used for analysis. This differs from other process analysis methods where the work activity is observed by a person, often with the assistance of tools (stop-watch, video camera, etc…) and then recorded for analysis. As discussed in Chapter 1, the standard method is often error- prone, costly and leads to sub-optimal conclusions. In contradistinction to the afore-mentioned approach, the methodology used with *Comprehend* centers on the automatic capture of data from the process itself as the starting point for the analysis of the current environment. Its use supports a discover, design, build/implement approach which is based on an emergent, bottom-up approach to understanding what is known as the "as-is" environment. Whereas most top-down modeling methods move quickly past this step to get to future solution design, the use of *Comprehend* facilitates a deep understanding of the "as-is" state and looks to uncover the root cause of the process variations. Each variation is seen as a potential *process value leak* that, when corrected, has the potential of recovering significant economic value for the organization.

Comprehend's approach, while software-centered, is collaborative. The *Comprehend* analyst works closely with both a solution architect and business subject matter expert to validate the discovered process and its variations as well as propose solutions. Also, there are no parallel activities in this methodology where one group focuses on a specific set of tasks while another works on different set. This lack of parallelism maintains unity of focus and common shared understanding. It also prevents errors due to misunderstanding of the same facts.

The following sections describe both the roles required for a *Comprehend*-based automated business process discovery project as well as the actual process steps to complete an effort. An example is presented that shows the use of *Comprehend* including process steps, input/output of each step and tool screen-shots for illustrative purposes. The case of an adjustment made to a healthcare claim is the subject of analysis in the example. The full detail of the healthcare claims processing environment can be found in Chapter 5, Case Study l.

3.3.1 Roles

The following roles are defined for this process:

a. *Comprehend* Analyst
b. Solution Architect
c. Configure Developer
d. Business Subject Matter Expert (SME)
e. Business Approval Committee

Each role is defined in more detail below in Figure 3.11.

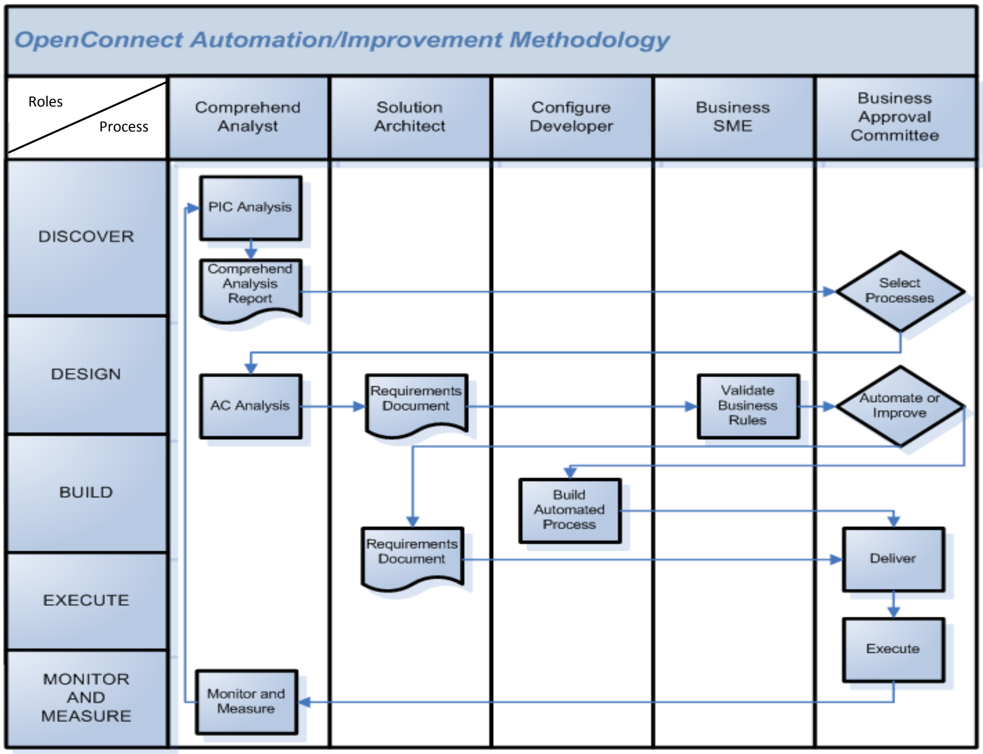


Figure 3.11

3.3.1.1 *Comprehend* Analyst

The *Comprehend* Analyst has two primary roles: to analyze and report on the operation of the claims processing function using OpenConnect's *Comprehend* Analytics Cluster (AC) and Process Intelligence Cluster (PIC) analyses; and to perform a detailed analysis of existing, manual workflows used in creating and supporting the automation process.

The *Comprehend* Analyst creates a *Comprehend* Analysis Report (CAR) summarizing PIC's analysis. The Business Approval Committee uses the CAR to select edit codes to be automated or improved.

The *Comprehend* Analyst also creates Initial Requirements Documents (IRD). Each IRD contains edit code-specific workflows based on AC analysis. These workflows provide edit code business process details used to discover potential improvement areas or automation opportunities.

3.3.1.2 Solution Architect

The Solution Architect works with the Business Subject Matter Experts (SME) to convert the IRD to a full Business Requirements Document (BRD) by adding existing business logic obtained in SME interviews. The Solution Architect passes the BRD to the Business Approval

Committee for final approval. The Business Approval Committee passes the final approval version of the BRD to the Configure Developer to code the transaction.

3.3.1.3 Configure Developer

The Configure Developer uses a BRD to implement claims automation transactions using the Configure Development Environment. The Configure Developer also performs unit testing before deploying a transaction to the runtime environment for user acceptance testing.

3.3.1.4 Business Subject Matter Experts

Business SMEs are critical to the improvement process and are involved at multiple stages. They consult with the *Comprehend* Analyst and Solution Architect to insure correct design of the claims business workflows are properly set up specific to their line of business, if needed. After the BRD has been created, the Business SMEs validate the business rules. The validated BRD goes back to the Business Approval Committee for acceptance.

3.3.1.5 Business Approval Committee

The Business Approval Committee is responsible for selecting the edit codes to automate or improve based on the CAR and their extensive business knowledge. Once the BRD is created and the Business SMEs have validated the business rules, the BRD goes for final approval with the Business Approval Committee.

3.3.2 Discover

The Discover stage is executed by the *Comprehend* Analyst to discover potential process automation or improvement opportunities. The *Comprehend* Analyst uses PIC to analyze events generated by AC. The analysis is performed using the following criteria:

a) Impact to the business by Edit Code
b) Distribution by Line of Business
c) User efficiency by Edit Code

The *Comprehend* Analyst generates the CAR which includes the top ten edit codes, the top three edit codes that follow each top ten edit codes, and the top three edit codes that follow those. This CAR also includes a graph for each top ten edit code showing the distribution by line of business. The process to create the *Comprehend* Analysis Report is provided in more detail below.

- Edit Code Analysis

This section describes the Edit Code analysis steps using PIC (see Chapter 5 for a complete description of the healthcare claims process, edit code analysis activity).

- Impact to the Business by Edit Code

Business Impact analysis can help determine which edit codes to automate or improve based on the actual cost to the business.

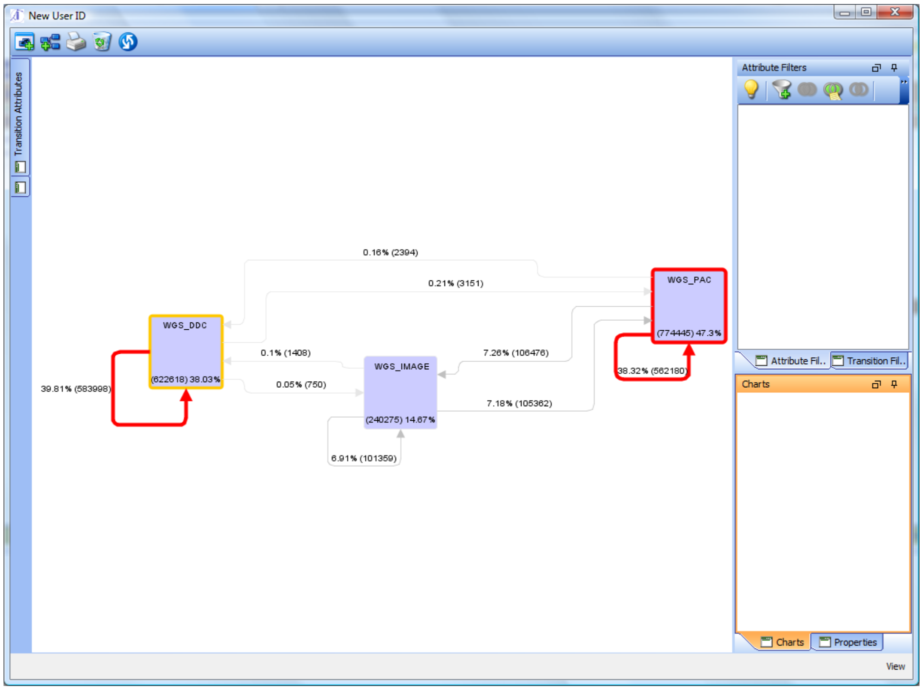


Figure 3.12 – Chart Group

- Process Intelligence Cluster Process Map

The business impact is determined by the total time spent by users to work specific edit codes.

The actual impact to the business is calculated using the formula:

(total time in hours / hours of work each day) * hourly rate

The following steps create the top edit code by user time chart:

c) Place your mouse over one of the activities
d) Right click and select 'New Chart'
e) If this is the first chart, you will be prompted for a chart group (Figure 3.13)
    a. Enter 'Edit Codes' and click 'OK'
f) You will now be prompted for a chart name and properties (Figure 3.14)
    a. Chart Name: Edit One by User Time
    b. Chart Type: Line/Bar Chart
    c. Chart Group: Edit Codes
    d. Check the 'Floating' check box
    e. Click on 'Next'

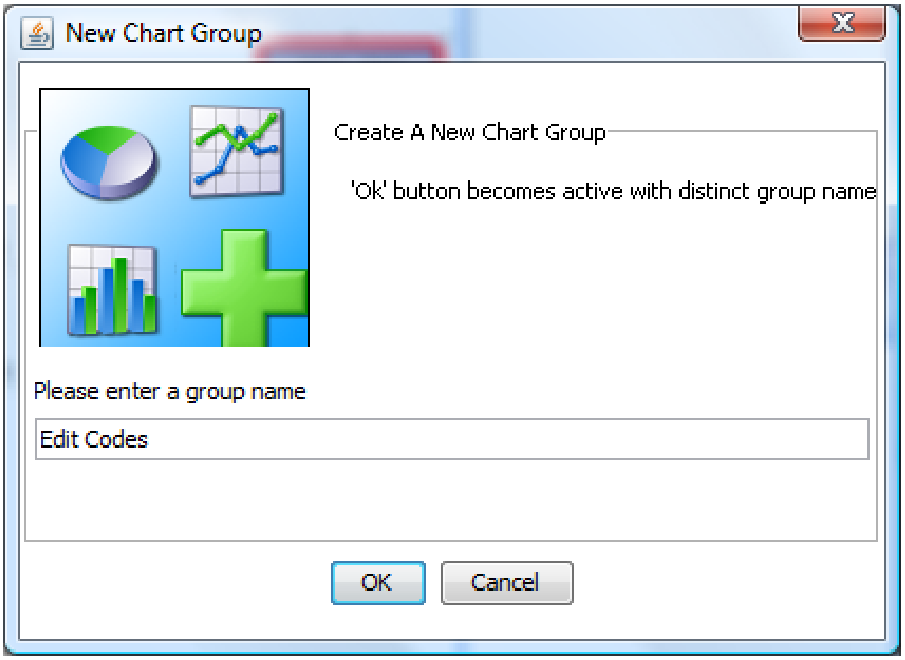


Figure 3.13: New Chart Group

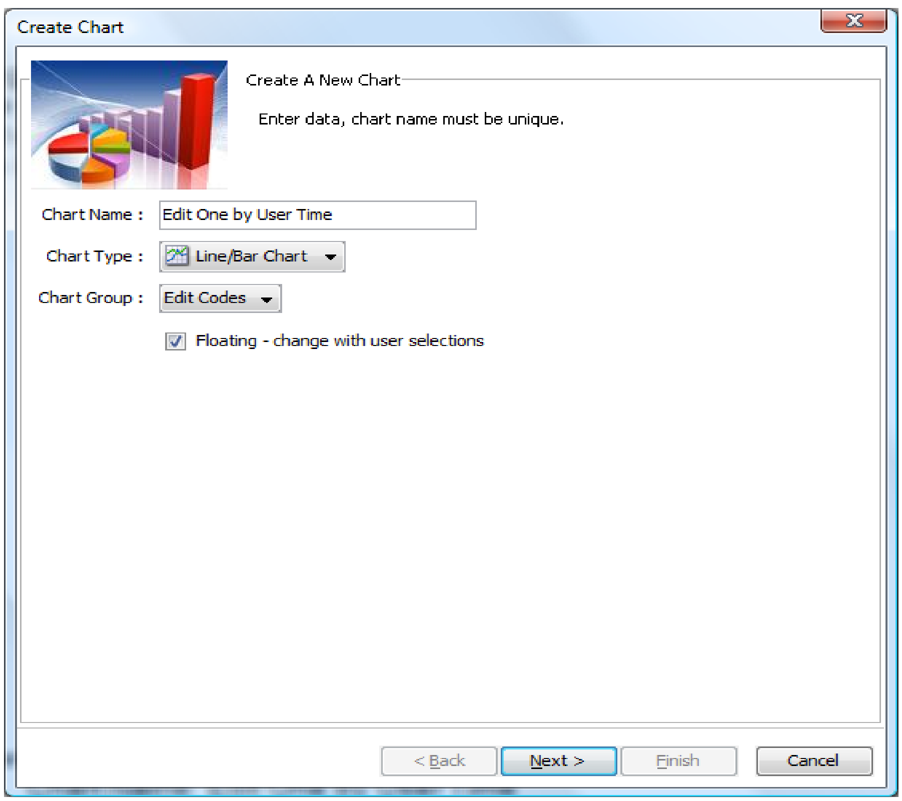


Figure 3.14: New Chart Dialog

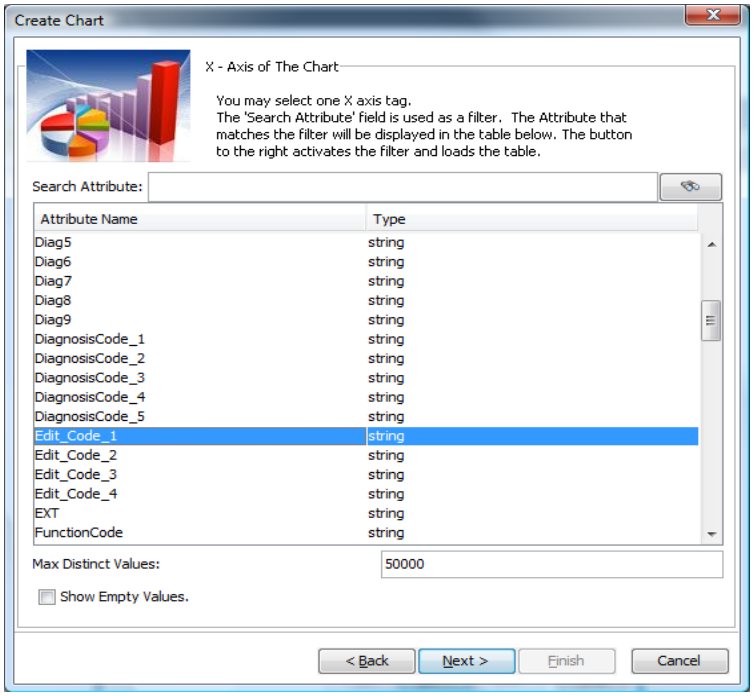


Figure 3.15: Add X-Axis Attribute

g) Configure the X-Axis attribute as shown in Figure 15
   a. Attribute: Edit_Code_1
   b. Max Distinct Values: 50000
   c. Click 'Next'

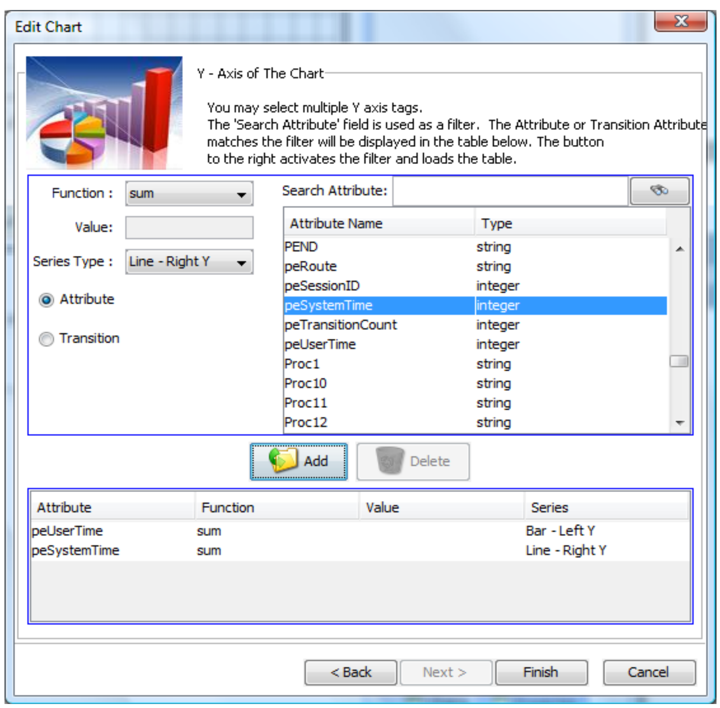


Figure 3.16: Add Y-Axis Attributes

h) Configure the Y-Axis values as shown in

i) Figure 3.16
   a. peUserTime - Sum - Bar - Left
   b. peSystemTime - Sum - Line -Right
j) Click on 'Finish'
k) You now have a chart showing User Time and System Time by Edit Code One
l) Sort the chart and resize the view to see the data better
m) You may notice the top value, or one of the top values, is 'No data'
n) To remove this value from the results we need to add a filter as follows in Figure 3.17
   a. In the Attribute Filters section click on the 'Add an Attribute' button
   b. Select Edit Code 1
   c. Click on 'Next'
   d. Select 'Greater than' (Figure 3.18)
   e. Enter a '!' exclamation point in the value input box
   f. Click on 'Finish'

Select the filter you just added.

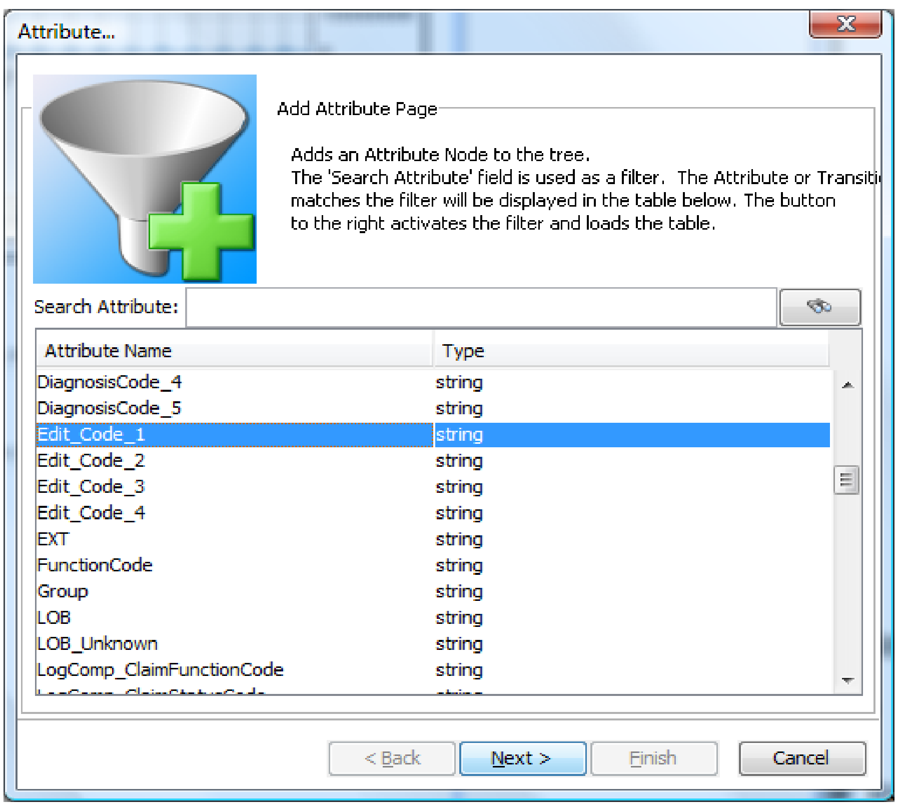


Figure 3.17: Add Attribute Filter

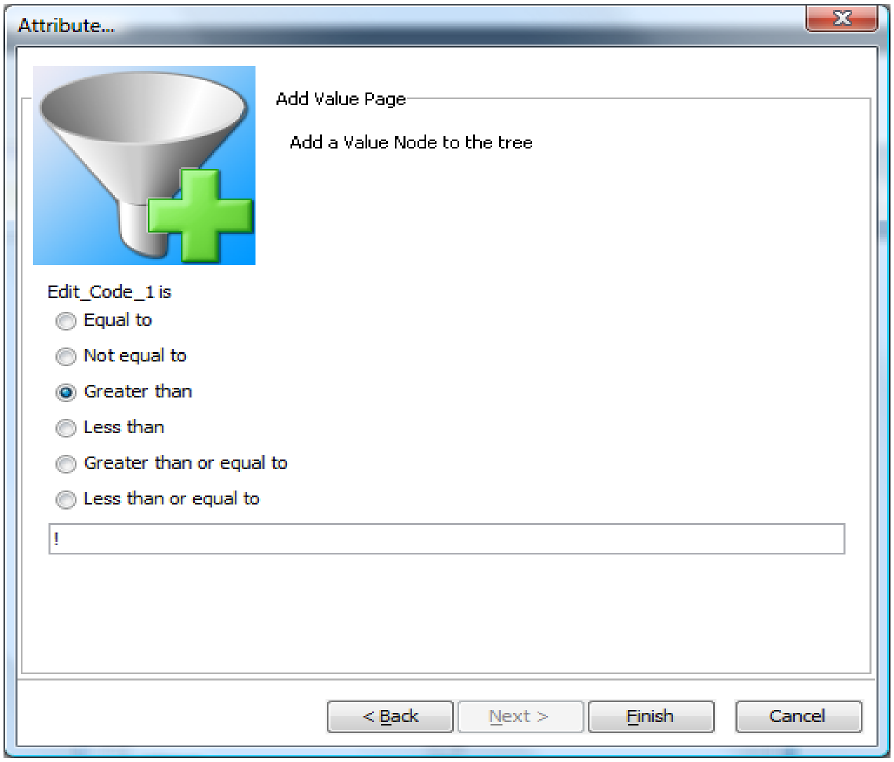


Figure 3.18: Add Filter Value

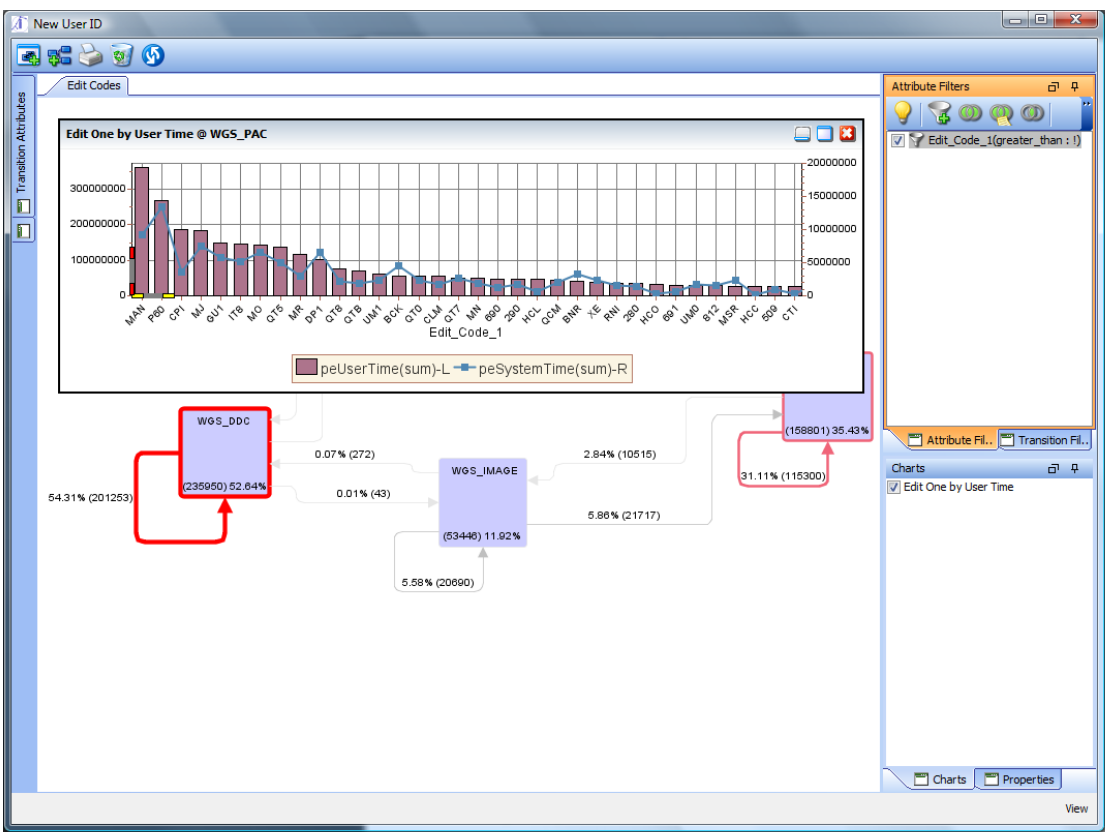


Figure 3.19: Edit Code One by User Time

By clicking on each of the activities you can see which edit codes have the highest user time and system time associated with them (Figure 3.19). However, this is only part of the analysis. It is also necessary to create a graph showing the total user time spent outside of the activities while working claims. We do this by creating a chart showing the deltaT time for the transitions (see Chapter 4, Section 4.1.17.1 for a description of creating PIC process perspectives).

The following steps create the deltaT analysis chart:

1. Right click on a transition and select 'New Chart'
2. Provide the following values:
    a. Chart Name: Edit One Delta Time
    b. Chart Type: Line/Bar Chart
    c. Chart Group: Edit Codes
    d. Floating: Checked

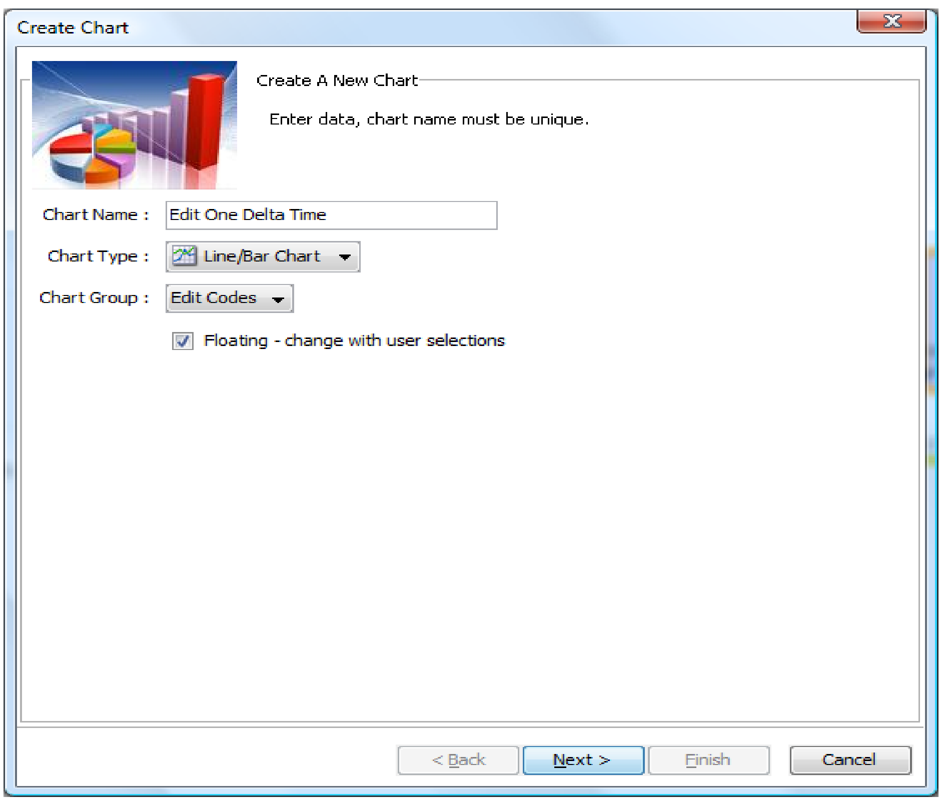


Figure 3.20: Create Edit One Delta Chart

3. Click on 'Next'

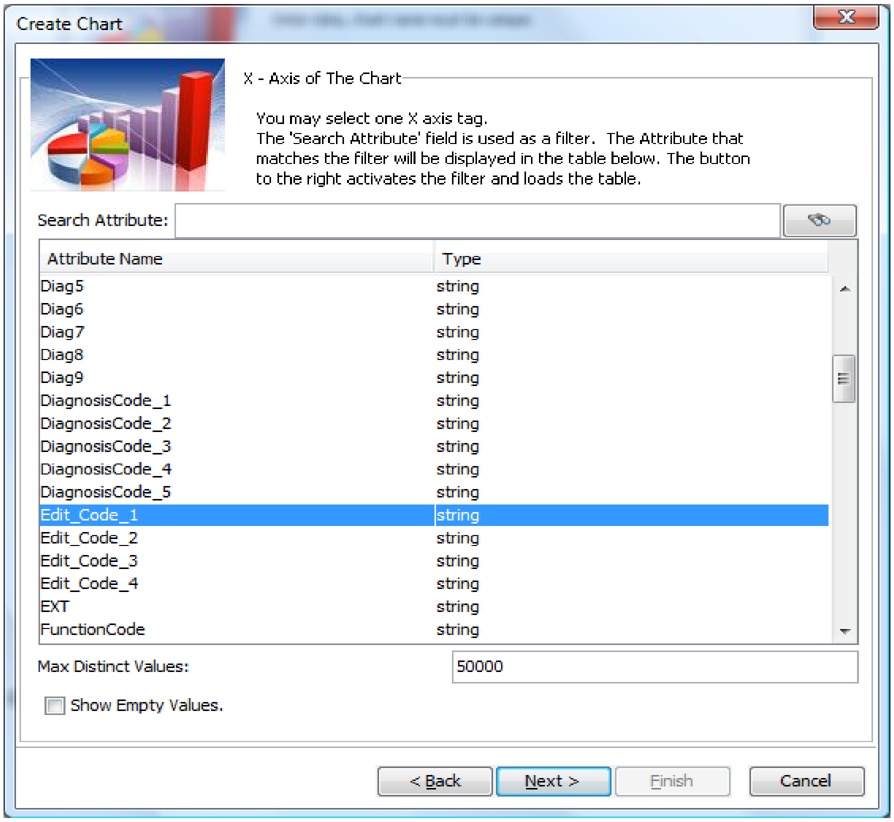


Figure 3.21: X-Axis Attribute for Delta Chart

4. Select Edit_Code_1 for the X-Axis attribute (Figure 3.21)
5. Enter 50000 for the Max Distinct Values
6. Click 'Next'
7. Select the following (Figure 3.22):
    a. Transition: checked
    b. Y-Axis attribute: deltaT
    c. Function: sum
8. Click on the 'Add' button
9. Click on 'Finish' (Figure 3.23)

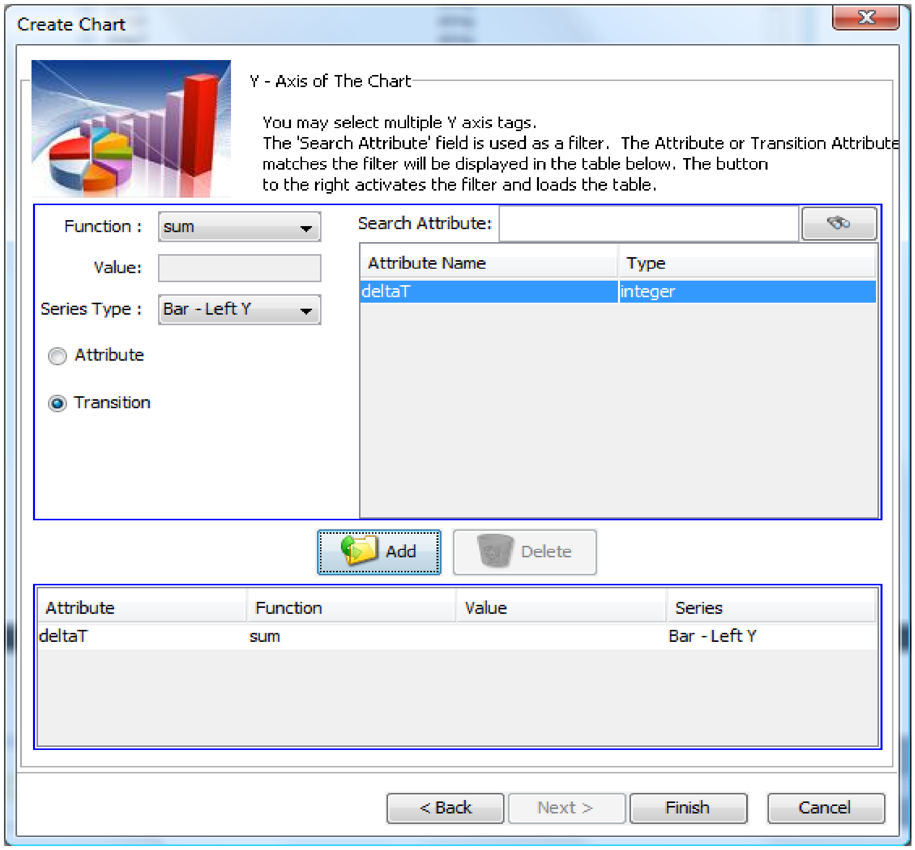


Figure 3.22: Y-Axis Attribute for Delta Chart

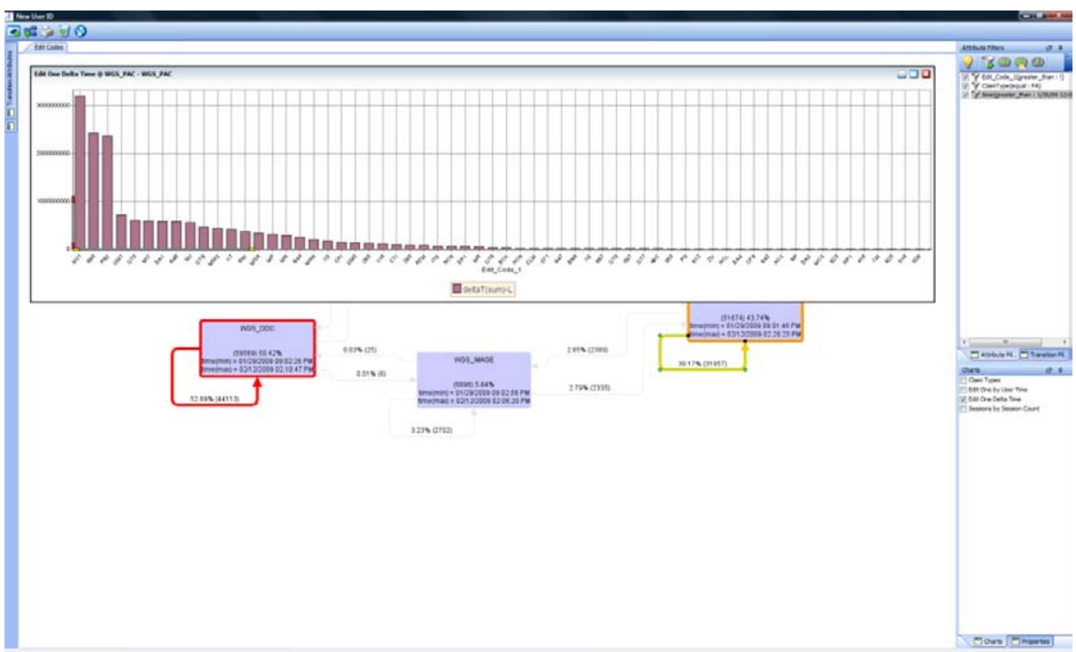


Figure 3.23: Edit Code One by DeltaT

Clicking on the various transitions provides the total time users spent working on claims outside of the standard claims activities. For example, we do not have a membership activity defined so if users navigate to any of the membership screens while working on an edit, the time spent in membership will be included in the deltaT for the loopback transition.

Distribution by Line of Business

Next, we want to discover the line of business distribution for each edit code. That requires another chart. The following steps create the line of business distribution chart:

1. Right click on an activity and select 'New Chart' (Figure 3.24)
2. Provide the following values:
    a. Chart Name: Line of Business
    b. Chart Type: Pie Chart
    c. Chart Group: Edit Codes
    d. Floating: Checked
3. Click on 'Next'

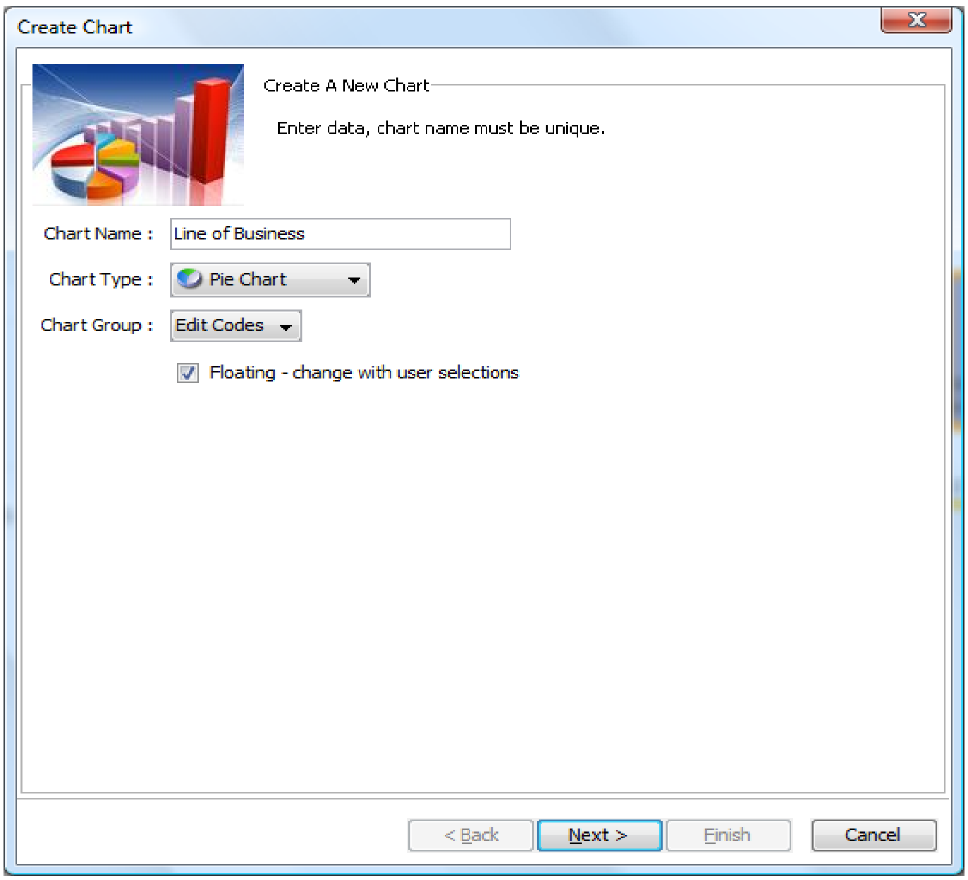


Figure 3.24: Line of Business Chart Creation

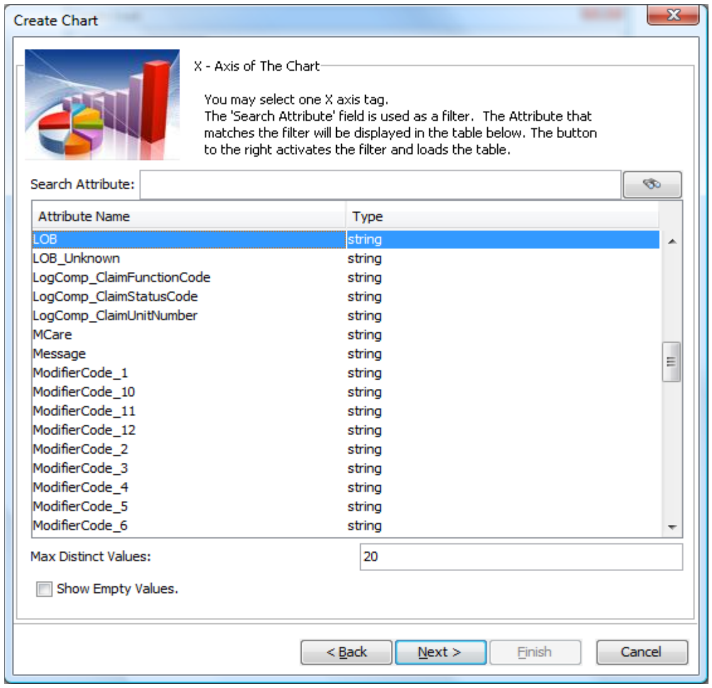


Figure 3.25: Line of Business X-Axis Attribute

4. Select the Line of Business attribute for the X-Axis (Figure 3.25)
5. Click on 'Next'

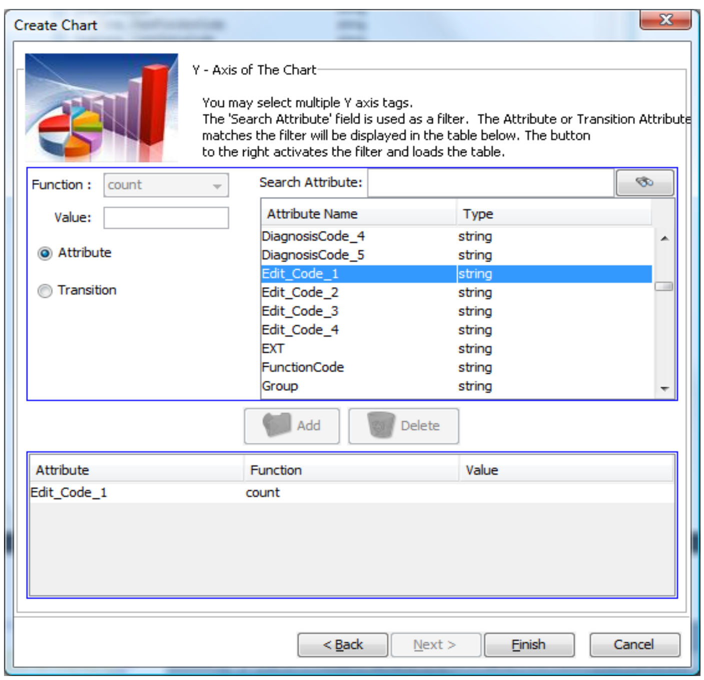


Figure 3.26: Line of Business Y-Axis Attribute

6. Select Edit_Code_1 for the Y-Axis Attribute (Figure 3.26)
7. Click on the 'Add' button
8. Click on 'Finish' (Figures 3.27)

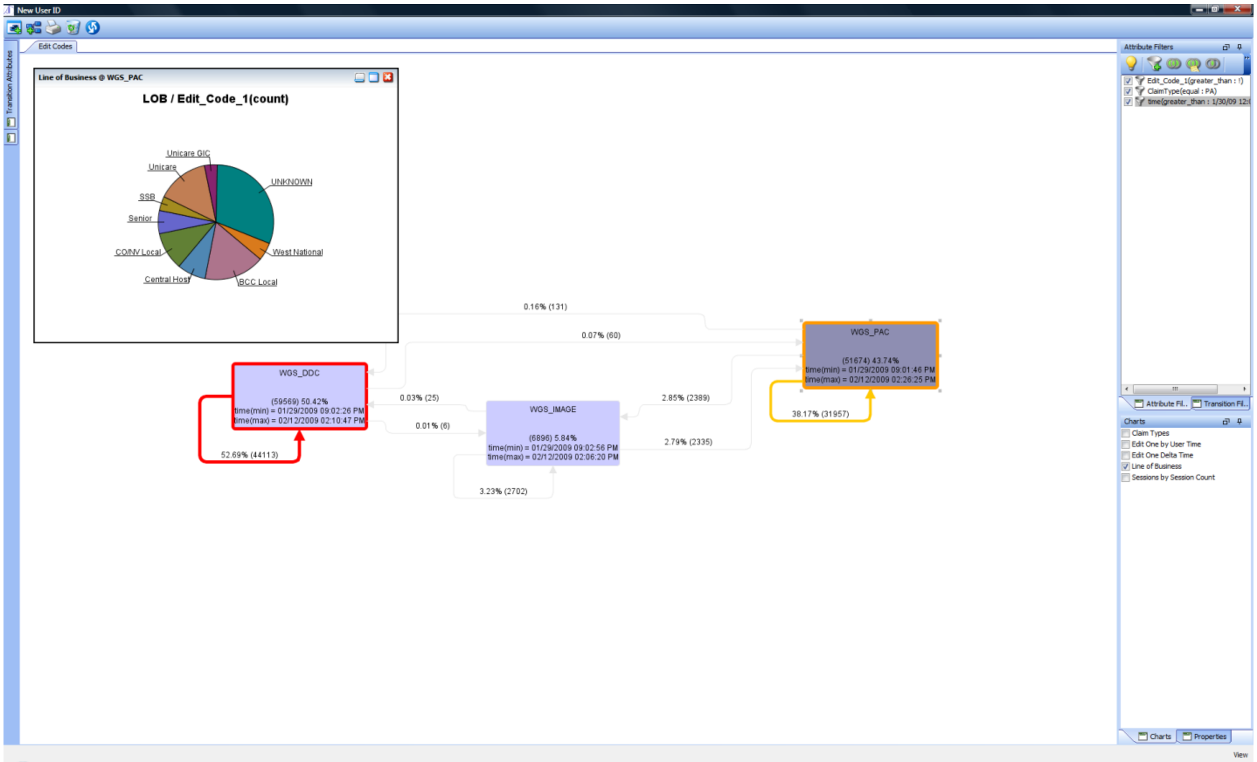


Figure 3.27: Line of Business Distribution by Edit Code

The next step in the process is to create the same charts for the other edit codes (Edit Code 2-3). These charts are used to create the CAR described next.

3.3.2.1 *Comprehend* Analysis Report (CAR)

The CAR displays the top ten edit codes in position one based on the impact to the business. It also display the top three edit codes in position two for each edit code in position one, and the top three edit codes in position three for each code in position two.

The purpose of the CAR is to show the top edit codes for potential automation and to highlight edit code patterns (Figure 3.28). Many edit codes fall into patterns, so that working one may fix others. The Analysis Report allows the business to see the relationship between edit codes and perhaps to see certain edit codes appear in the list more often than others.

	Edit One	Edit Two	Edit Three	Hours	Percentage of Edit	FTEs	Yearly Cost	Estimated Automation Percentage	Estimated Yearly Savings
3	Edit Group 1								
4	GU1			3159.11	100.00%	46.46	$1,545,362.28	20%	$309,072.46
5	GU1	STL		1445.37	45.75%	21.26	$707,041.00	20%	$141,408.20
6	GU1	STL	HP1	356.75	24.68%	5.25	$174,513.71	20%	$34,902.74
7	GU1	STL	CLM	277.32	19.19%	4.08	$135,658.42	20%	$27,131.68
8	GU1	STL	MC3	208.93	14.46%	3.07	$102,203.64	20%	$20,440.73
9	GU1	HP1		574.73	18.19%	8.45	$281,144.39	20%	$56,228.88
10	GU1	HP1	261	187.96	32.70%	2.76	$91,945.61	20%	$18,389.12
11	GU1	HP1	619	167.9	29.21%	2.47	$82,132.73	20%	$16,426.55
12	GU1	HP1	686	113.58	19.76%	1.67	$55,560.66	20%	$11,112.13
13	GU1	ITS		418.5	13.25%	6.15	$204,720.35	20%	$40,944.07
14	GU1	ITS	HP1	235.2	56.20%	3.46	$115,054.31	20%	$23,010.86
15	GU1	ITS	STL	162.01	38.71%	2.38	$79,251.48	20%	$15,850.30
16	GU1	ITS	VS	6.81	1.63%	0.10	$3,331.29	20%	$666.26

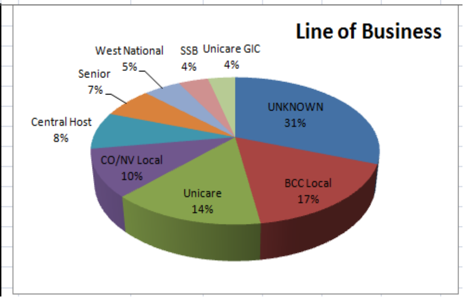


Figure 3.28: Sample *Comprehend* Analysis Report

3.3.2.2 Business Approval and Direction

The CAR then goes to the Business Approval Committee. The committee will analyze the report and choose a group of edit codes to automate or improve. The list of edits will be provided to the *Comprehend* Analyst for further analysis. The chosen edit codes will be used in the Design stage of the process.

3.3.3 Design

In the Design stage, *Comprehend* discovers details of edit code-specific or edit code group-specific claims processes, then feeds them into the automation or improvement plan for the top edit codes. This stage uses PIC to determine which session IDs to import into AC for session level analysis.

Detailed Design Stage Description

The improvement process Design stage consists of the following steps:

- Identify Attributes for Analysis
- Create Filters
- Add Edit Code Filter
- Add Data Filter
- Identify Power Users
- Create Session ID Charts
- Export Data from PIC
- Import Data into AC
- Analyze Sessions
- Build IRD
- SME Sessions
- Build Business Requirements Document
- Finalize BRD
- Business Approval

Identify Attributes for Analysis

It is necessary to determine which few attributes are critical to claims process performance before doing deeper analysis. The most common attribute is the edit code or list of edit codes found in the Discover stage. Additional attributes should be selected based on edit code or lines of business specifics. For example, if the edit code is a membership-related edit, then it may be desirable to look at some of the membership attributes to help determine the most common reasons for the pended claims.

Additional critical claims processing attributes commonly include:

- Edit Code 1 - 4
- Line of Business
- Claim Type
- Subscriber ID
- Provider ID
- Group Number
- COB

Create Filters in PIC

After determining which attributes to use it is necessary to add PIC filters for them. For example, it may be important to know which edit code has the biggest impact to the business by claim type. The chart created in section 0 can show the impact to the business by edit code. However, before any analysis by edit code can be done, it is necessary to create a filter for the specific edit code to be analyzed.

Add Edit Code Filter

Follow the steps below to create the Edit Code filter:

1. Click on the "Add an Attribute" button shown below

2. Select Edit_Code_1 for the Search Attribute as shown below (Figure 3.29)

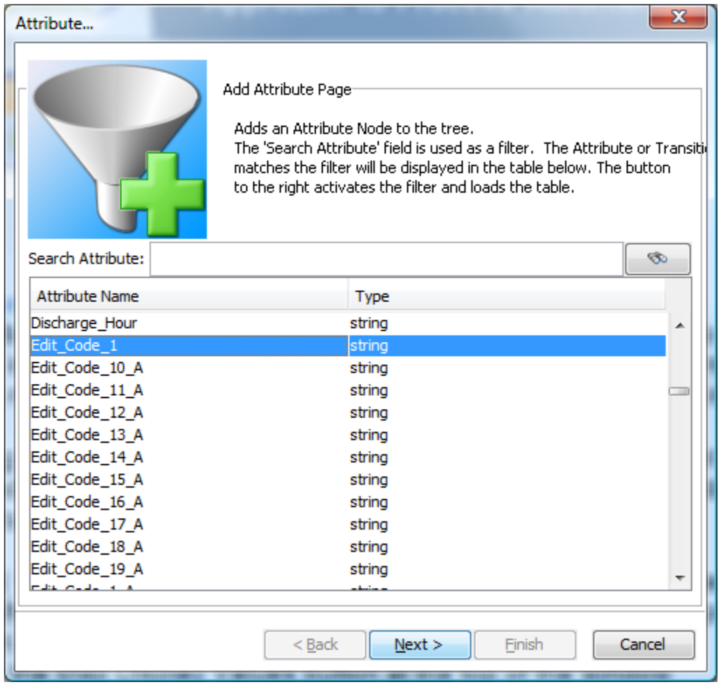


Figure 3.29 - Sample *Comprehend* Analysis Report

3. Click 'Next' (Figures 3.30)

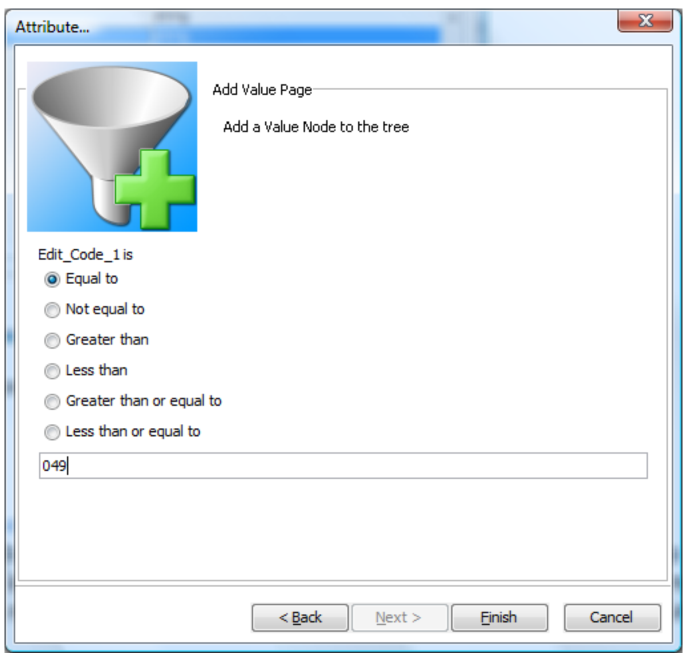
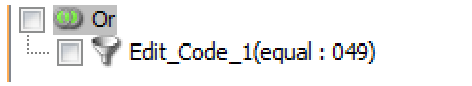


Figure 3.30 - Sample *Comprehend* Analysis Report

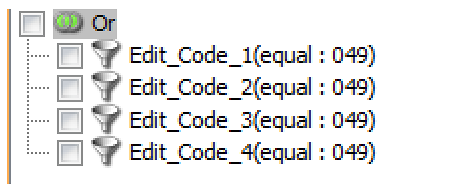
4. Click on the "Equal to" radio button
5. Enter the edit code to filter on in the text box
6. Click Finish
7. The newly created filter now appears in the Attribute Filter list
8. Single click on that filter to select it

9. Click on the "Add an Or node" button

10. The newly created filter is shown inside an OR statement



11. Click on that Or statement, then click on the Add an Attribute button
12. Add filters for Edit_code_2 through Edit_Code_4
13. The filter should now look like the one shown below



14. Check the Or check box to select the filter

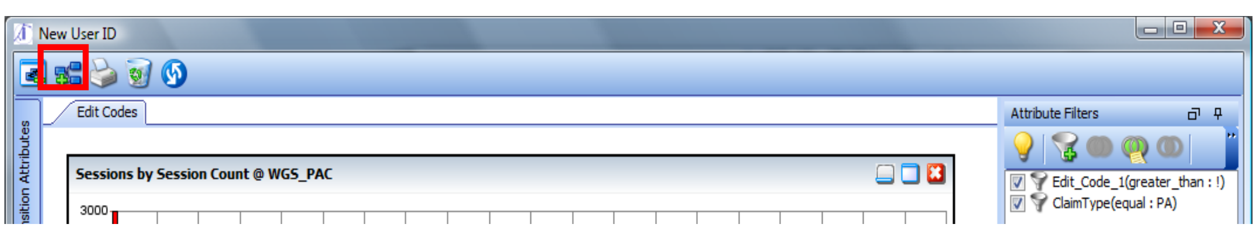
Add Date Filter

Due to size constraints, it is possible, and, in fact, most likely, that the raw data in AC will be limited to one or two weeks while the data in PIC could span four weeks, six weeks or more. Due to that situation, it is necessary to add a filter by time to ensure the session IDs in the chart represent sessions still retaining in AC.

Before creating this filter it is necessary to know the data retention policy set in AC. If you do not have administrator rights for AC it will be necessary to ask the AC administrator for that information.

Next, it is necessary to determine how much data PIC retains. It may require adding activity properties to show the maximum and minimum times using the following steps:

- Click on the Map Display Values button at the top of the window



- Add the time attributes as shown in Figure 3.31
    - time - min
    - time - max
    - Click on the 'OK' button

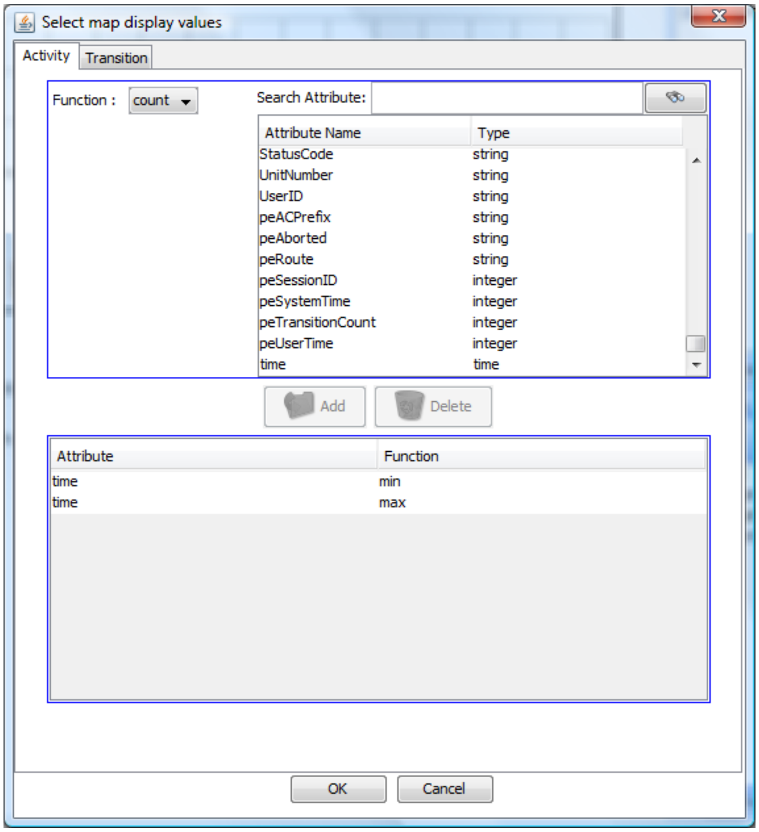


Figure 3.31 - Add Time Properties

The process map will update and display the date and time range of the data in PIC. Then, compare the range with AC's data retention policy. In this example, the data retention policy is 14 days. There is also a minimum date/time in PIC of 1/23/09 4:04 PM, and a maximum date/time of 2/12/09 2:26 PM. This tells us that the latest date in PIC is 2/12 and the data retention policy in AC is 14 days, showing that the filter should be set for 14 days back from 2/12. So, the filter should be Greater than 1/30/09 12:01:00 AM as shown in Figure 3.32.

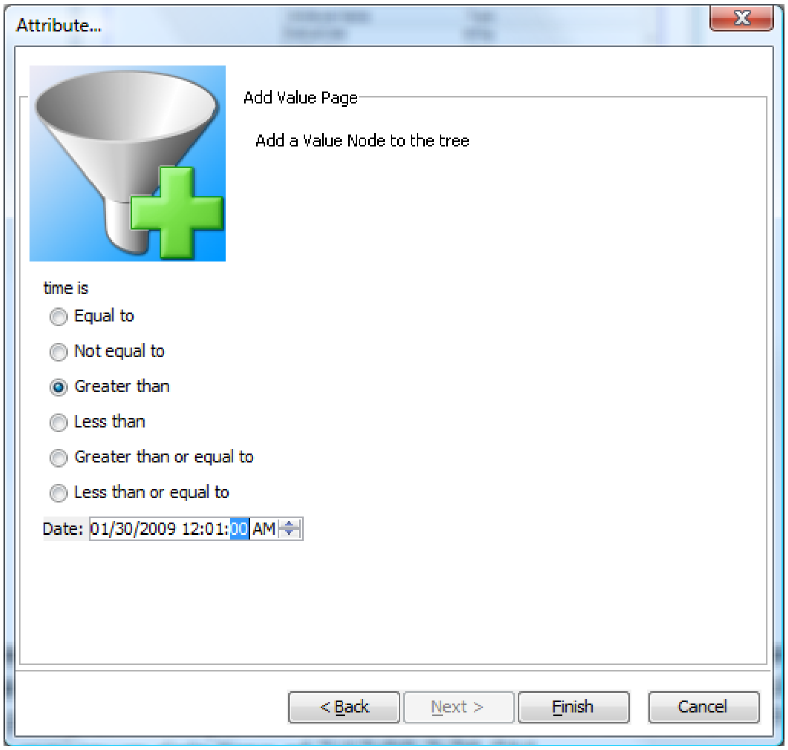


Figure 3.32 - Time Filter

All data displayed in the charts will now be updated to display only the previous 14 days which matches the amount of raw data in AC. We only need the time filter when we are retrieving session IDs for AC session analysis. For all other types of analysis make sure you remove this filter.

Identify Power Users

Power claims users can often be identified by the number of claims processed in the shortest amount of time. While that is not always true, it is a good starting point. To find the power users, create a chart showing user ids by count and user time in the following manner:

1. Right click on the main claims processing activity and select 'New Chart'
2. Fill in the information as follows
    a. Chart Name: Power Users
    b. Chart Type: Line/Bar
    c. Floating: Checked
    d. Click on 'Next'
3. Select UserID for the X-Axis attribute
4. Set Max Distinct Values to 50000
5. Click 'Next' (Figure 3.33)

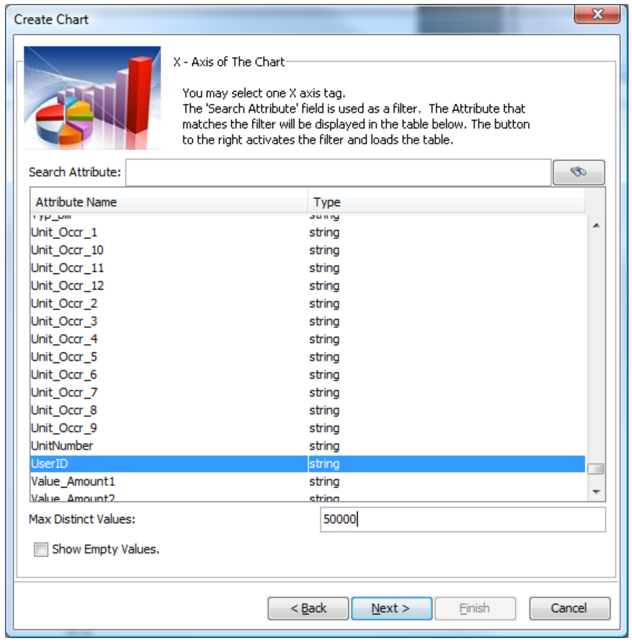


Figure 33 - User ID X-Axis Definition

6. Select DCN for the first Y-Axis attribute
7. Select count for the function
8. Click the 'ADD' button
9. Select peUserTime for the second Y-Axis attribute
10. Select sum for the function
11. Select Line - Right Y for the Series Type
12. Click on the 'ADD' button
13. Click on 'Finish' (Figure 3.34)

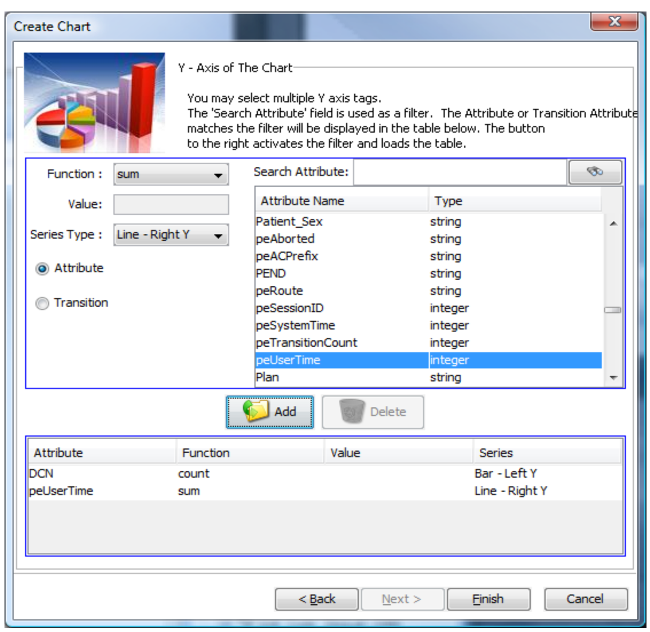


Figure 3.34 - User ID Y-Axis Definition

The result is a chart showing users by the count of claims worked and the amount of time used to work those claims (Figure 3.35).

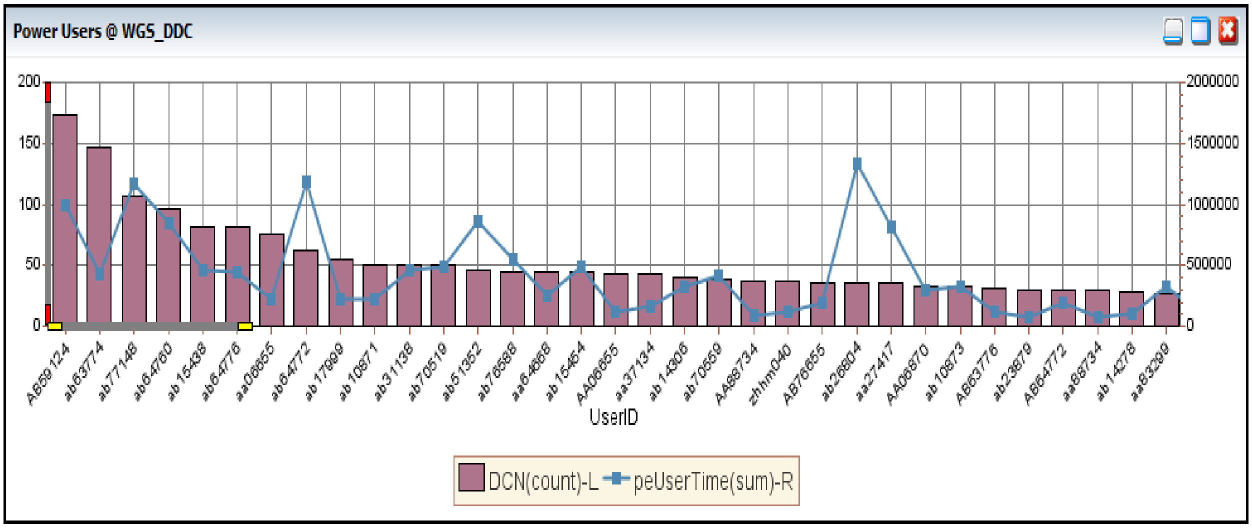


Figure 3.35 - Power User Chart

While the chart is not very useful to look at, it contains the necessary information. The next steps are to export the data, filter user ids by the power user index, then import the filtered user ids to use as a filter in PIC. The steps below illustrate the method:

1. Right click on the chart and select Export -> Comma Separated Values…
2. Select a location to store the file
3. Open the file in Excel
4. You should see a file similar to the one below in Figure 3.36

	A	B	C
1	UserID	DCN(count)	peUserTime(sum)
2	2744mu	0	0
3	3919mu	5	22988
4	9294mu	0	0
5	9294MU	0	0
6	aa05678	0	0
7	aa06655	75	228157
8	AA06655	43	114371
9	AA06870	33	293010
10	aa08647	0	0
11	aa08652	0	0
12	aa27155	0	0
13	AA27155	0	0
14	aa27417	35	817831
15	AA27417	22	139004
16	aa27454	9	220322
17	aa33199	0	0
18	aa36987	0	0
19	aa37134	43	159518
20	aa38054	0	0
21	aa38569	0	0
22	aa39765	0	0
23	aa40809	0	0
24	AA40809	0	0
25	aa44819	0	0
26	AA48627	0	0
27	aa49055	0	0
28	aa50474	0	0

Figure 3.36 - Sample *Comprehend* Analysis Report

5. Next, sort the data by DCN
6. Delete all rows where DCN is 0
7. Find a DCN cutoff number **-** the number of DCNs that really matter
    a. The number that matters will be dependent on the data
    b. This example uses 10
8. Delete all rows where DCN is lower than the cutoff number
9. Next, create the power ranking
    a. In the cell to the right of the user time create the formula
    b. (B2/C2)*1000
10. Sort the data by the power index column
11. The result should be similar to the following output in Figure 3.37

UserID	DCN(count)	peUserTime(sum)	Power Rank
AA88730	11	15344	0.72
ab23895	13	30552	0.43
aa57765	19	48648	0.39
AA88734	37	94738	0.39
aa88734	29	74808	0.39
AA06655	43	114371	0.38
ab23879	30	80691	0.37
ab63774	147	433800	0.34
ab18015	10	29518	0.34
aa06655	75	228157	0.33
AB51345	10	31481	0.32
aa66956	21	66613	0.32
AB26804	17	54265	0.31
zhhm040	37	124965	0.30
aa98164	17	58195	0.29
AB76590	25	86505	0.29
ab15456	12	43233	0.28
ab14278	28	102229	0.27
AB63776	31	114302	0.27
aa37134	43	159518	0.27
ab31132	20	76915	0.26

Figure 3.37 - Sample *Comprehend* Analysis Report

12. Save the file with a different name for future use
13. Next, select the user ids from that file to use as a filter
    a. Select the user ids
    b. Select copy
    c. The user ids are now ready to be imported into PIC
14. The next step is to import the user ids into PIC so we can use them as a filter
    a. Select the 'Add an Or node and Attributes from a list' button

    b. The following dialog appears:
    c. Select User ID for the Search Attribute Figure 3.38

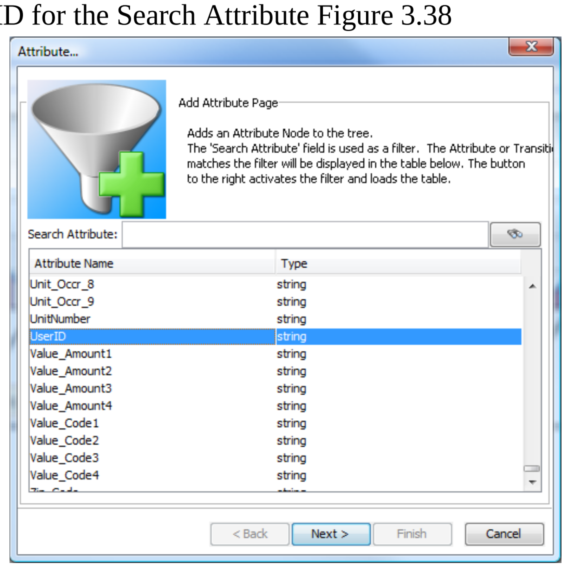


Figure 3.38 - Power User Filter X-Axis

d. Click 'Next'
e. Paste the user ids into the import box (Figure 3.39)

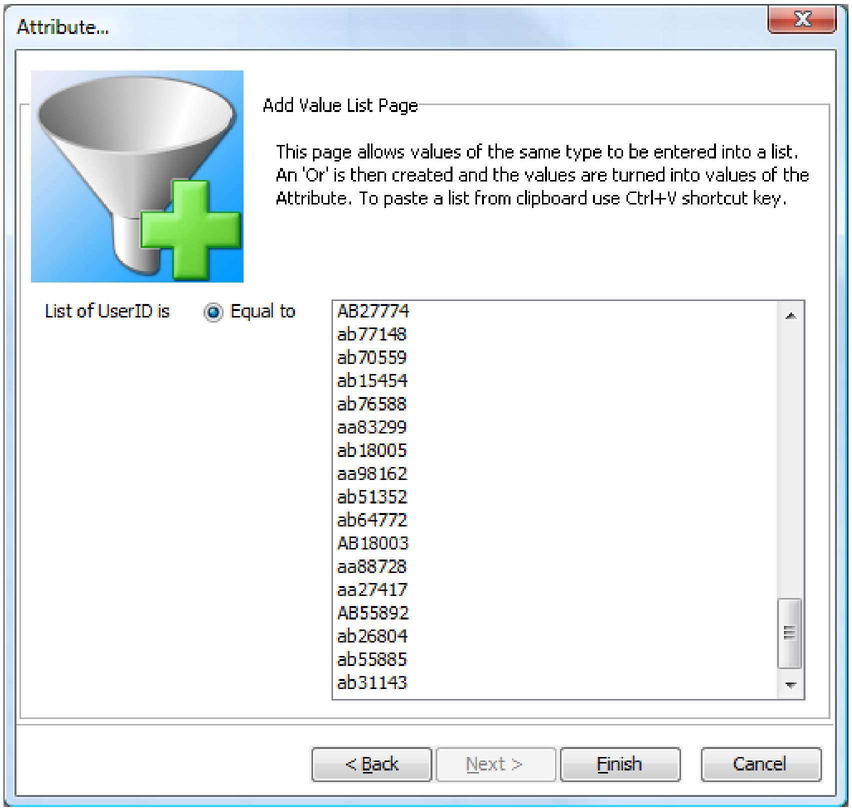
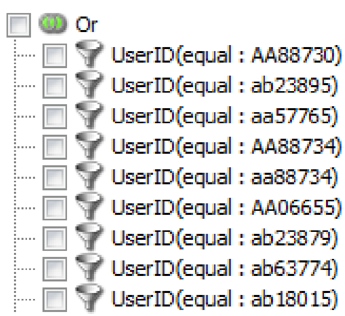


Figure 3.39 - Power User Filter Import

f. Click 'Finish'

15. At this point, an OR filter with all of the user ids should exist



16. Click on the Or check box to select the filter
17. That completes building all filters needed to create the session id chart, described next

Create Session ID Chart

It is then necessary to create a chart to get a list of session IDs based on the filters set in previous steps. The following steps describe how to create the required chart:

- Right click on the main claims processing activity and select 'New Chart'
- Fill in the information as follows
    - Chart Name: Sessions by Session Count
    - Chart Type: Line/Bar
    - Chart Group: Edit Codes
    - Floating: Checked
    - Click on 'Next'
- Select peSessionID for the X-Axis attribute (Figures 3.39)
- Set Max Distinct Values to 50000
- Click 'Next'

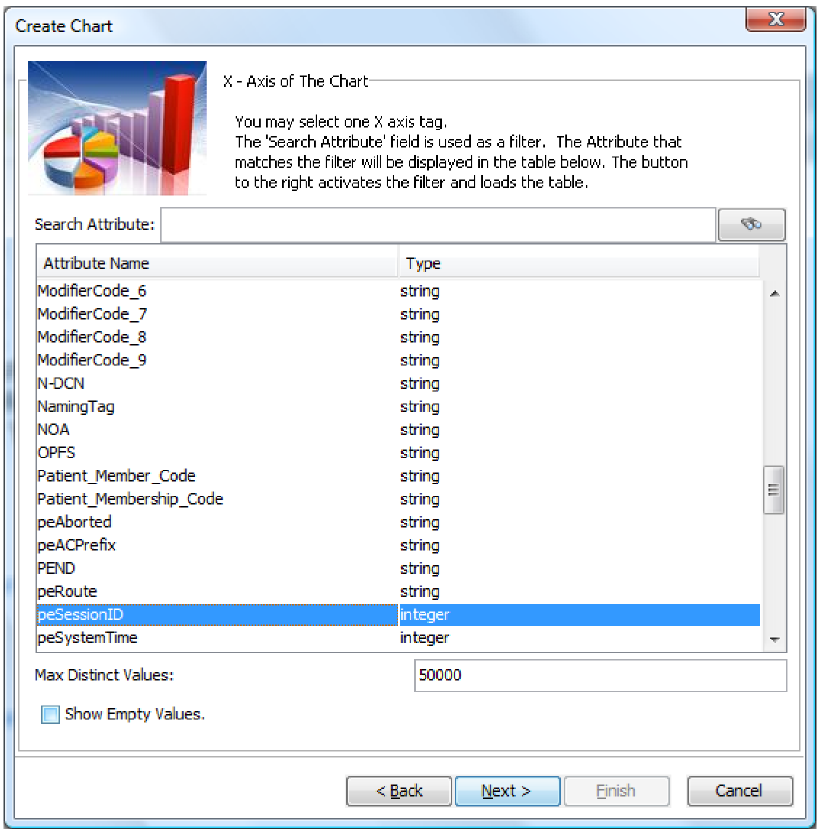


Figure 3.40 - X-Axis Attribute Definition for Session IDs

- Select peSessionId for the Y-Axis Attribute (Figure 3.41)
- Click on 'Add'

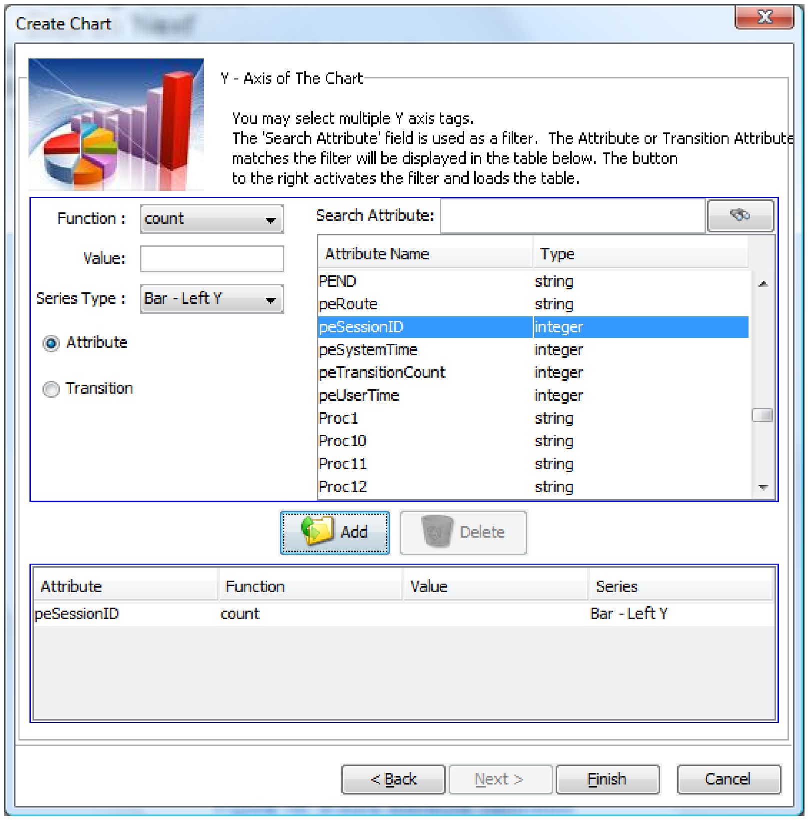


Figure 3.41 - Y-Axis Attribute Definition for Session IDs

- Click 'Finish'

That produces a chart of session ids based on the chosen filters.

Export Data from PIC

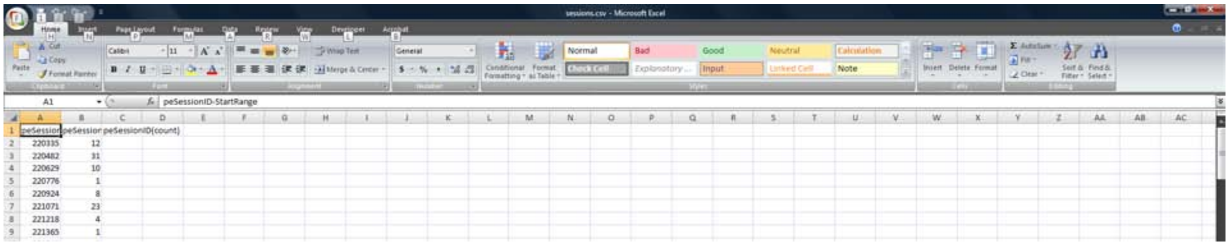
The chart of session ids are now exported for use in AC.

- Right click on the chart and select Export -> Comma Separated Values
- See a file dialog box appear
- Navigate to the desired folder, provide a file name and click on 'Save'

There is now a file with ids for sessions matching the criteria set in PIC. Next, import those session ids into AC.

Import Data into AC

Open the csv file created in 0. See a file similar to that listed below:



The file should have a column with session ids and a column with counts. Sort the data by count in descending order so that the session ids with the highest count are displayed at the top. That sets the most relevant sessions at the top of the list.

Next, copy the top session ids. A good rule of thumb for choosing the actual number of sessions to walk through in AC is to take the top 25 percent or the top 300 sessions, whichever is higher.

Open the Eclipse interface to AC and open the base report project.

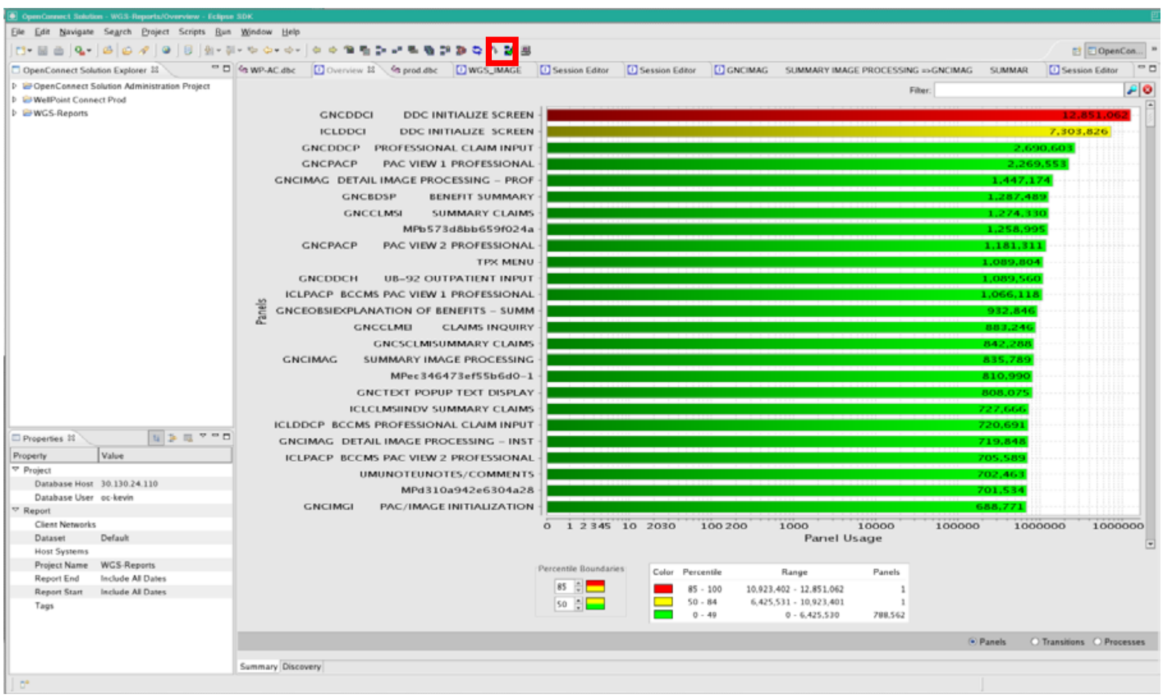


Figure 3.42 – View Sessions

Click on the View Sessions (Figure 3.43) button on the tool bar. Place the cursor in the text box and paste the session ids into the text box.

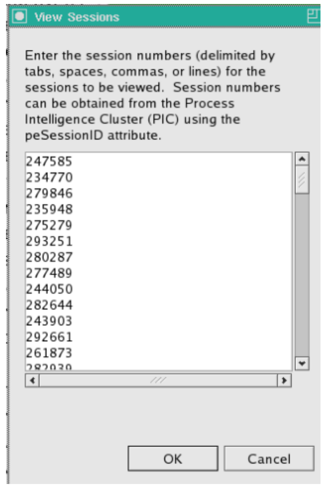


Figure 3.43 – View Session

Click on the 'OK' button to get a prompt for the Security Key. Enter the key and click on the 'OK' button to view the sessions. If you do not have the security key, please contact the AC administrator.

Analyze Sessions

Analyzing sessions is an art and takes practice. The more one knows about the claims processing system, sub systems and the specific edit to be analyzed, the easier it is to analyze the sessions.

The first step is to select the first panel in the currently displayed session. Next, select the search button on the toolbar. Enter the search criteria, such as the edit code to be searched, and select String search. If that search does not return relevant information, the Regex search may be useful.

Walk through the user's session from this point on by clicking on the next panel name in the list on the right (Figure 3.44).

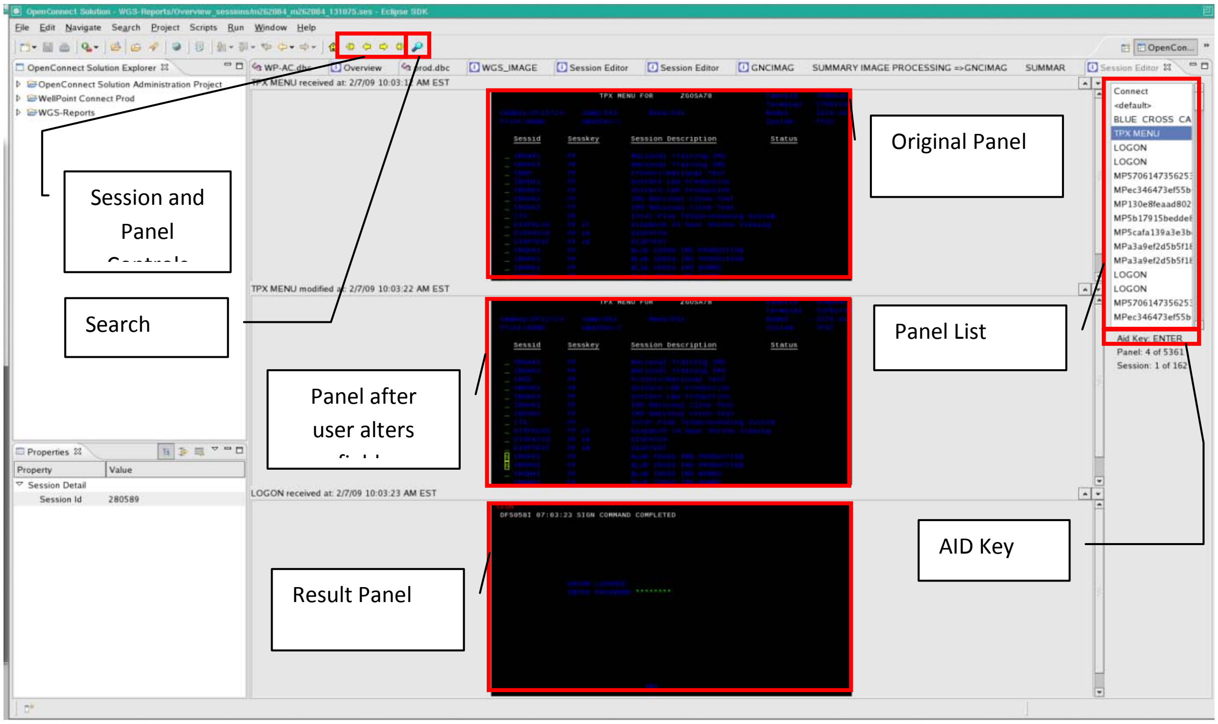


Figure 3.44: Session View in AC

Keep track of the claim number to continue looking only at screens related to the same claim. Pay close attention to the AID key used to move to the next panel and the fields modified by the user. The modified fields are highlighted with a yellow box and appear in the second panel.

Walk through the session to find the end of the claim. Note the final outcome to working the claim (finalize, save, etc.). After determining that the claim was actually worked for the desired edit code, go back and take screen shots of the panels used in the workflow. Walk through as many sessions as possible and take note of the various paths the users took to work the claims, the lines of business, and the results of working the claims.

Build Initial Requirements Document

The Initial Requirements Document (IRD) has two purposes. First, it documents the various workflows used to work a specific edit code. Second, it determines the most accurate and efficient workflow to use as a standard for process automation or improvement.

The various workflows should be documented in an IRD. If any of the workflows are specific to a line of business that should be noted. Also take note of the results of each workflow and the data elements involved.

Walking user sessions helps to understand what the users do to work claims but does not provide the reasons that the user made certain decisions. For example, some workflows may show users entering an 'S' in a particular field and other workflows show the user entering an 'N'. It is easy to document those occurrences but still not know why the user chose an 'S' or an 'N'. It is a best practice to make a note in the document about that situation. The SMEs can provide the business rules behind the users' choices later in the improvement process..

It is probable that there are not enough sessions nor time to see all of the potential values in a field. Take note of the fields that have multiple values so that additional analysis can be done in PIC to determine all of the values used. This analysis should also be provided in the IRD.

Subject Matter Expert Sessions

The next step in the process is to meet with SMEs for the specific edit code, or group of edit codes, and walk through the IRD. If the IRD provides different paths based on line of business or claim type, then the SMEs should be from the specified line of business or be familiar with the claim type.

Walk through the IRD and ask for specifics not fully explained by walking the sessions. Document the answers given by the SMEs and complete the requirements document.

Build Business Requirements Document

After working with the SME to understand the business rules it's time to create the BRD. The BRD should include the screen shots from the IRD and the business logic from the SME. The format of the BRD can be different for each customer but should contain an outline of the workflows. For example:

1. If line of business = Senior
    a. Enter an 'S' in the LOB field
2. Else
    a. Enter an 'E' in the LOB field
3. Press [Enter]

All 'IF' statements must be closed with an 'ELSE' statement. The idea is to make sure all logic paths are accounted for so the Configure developer can create "bullet-proof" transactions.

Finalize Requirements Document

Additional analysis might be required to help fill in the gaps. After all requirements are documented and it is believed to be complete, schedule a meeting with all participants to do a final walkthrough of the BRD. All lines of business associated with the specific edit code should be represented in this meeting.

Take notes and update any information appropriate based on team consensus. After all changes are made and everyone agrees on the contents, that constitutes a final version. Then, deliver the final BRD to the Business Approval Committee.

The BRD should provide a recommendation to either automate the process or improve the process. The recommendation should be based on the percentage of claims that can be worked by a robot. That percentage can be determined by analyzing the various paths and determining if all data is available to the robot through the standard methods (3270 screen, database access, web service, properties file, etc).

Business Approval

The Business Approval Committee reviews the BRD to verify that it meets the needs of the business. The Committee determines whether the process should be automated or improved and documents its decision by signing the acceptance page.

If automation is chosen, the document is delivered to a process automation team. If the process is to be improved, then the document is delivered to a process improvement team.

3.3.4 Build

In the Build stage, the process automation team creates Configure transactions based on the BRD created in the Design stage of the improvement process.

Design Transaction(s)

The new transaction(s) should conform to the Transaction Framework developed specifically for the customer by OpenConnect. Before the automation team can build any transaction, it is necessary to design it. Follow the step below to design the transaction:

1. Review the BRD
    a. If there are any questions, send them to the owner of the BRD
    b. If there are still questions, ask for sample claims that meet the requirements in order to personally walk through the claim to verify the business rules.
2. Determine how the new transaction will interact with other transactions already available in the Transaction Framework
3. Determine how to create the Panel Handler
    a. Panel Handler behavior will depend on the panels with which the new transaction will interact
    b. It is also necessary to plan how to handle unknown panels
4. Determine whether Perform or Set-state will be used to call other areas of the new transaction

3.3.4.1 Build Transaction(s)

The new transaction(s) will be created using the OpenConnect Configure Development Environment. Open Configure and create your new transaction. Develop the transaction based on the BRD and the considerations in the Design Transaction phase.

3.3.4.2 Unit Test

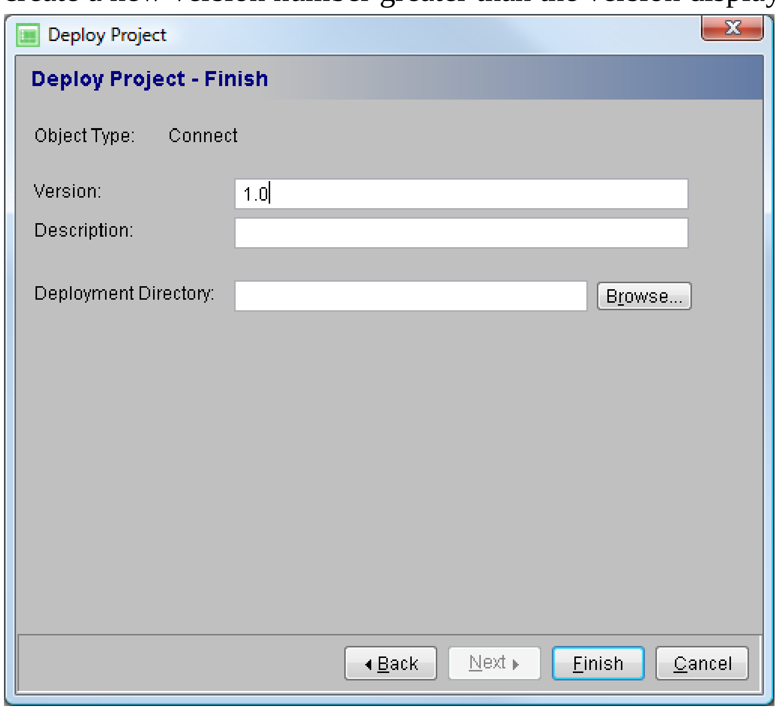
All unit testing is performed within the Configure development environment. Use the personally created test cases or those created by others. The test cases should match the various scenarios outlined in the BRD.

3.3.4.3 Integration Test with Web Application

After all test cases have been successfully executed and validated, deploy to the development environment to test with the web application. The steps for deploying are shown below:

1. Click on the Deployment Wizard button on the toolbar
2. Click 'Next'

3. Accept the default package name or create a new one
4. Click 'Next'
    a. This will generate the source files for the project
5. Click 'Next'
    a. This will compile the source files
6. Accept the version number displayed, which should be one greater than the current version, or create a new version number greater than the version displayed



7. Enter a meaningful description of this version
8. Click 'Finish'

3.3.4.4 System Test

Before system tests can be executed test claims have to be available. Make sure there are enough test claims to test all of the various transaction paths and scenarios. Configure the web application to execute the test claims. This could mean pointing it to the correct mainframe environment or it could be entering a search string.

Once the claims are available, execute the transactions through the web application. That execution depends on the environment, customer and web application. After execution, run the web application report to validate the results.

If the environment supports the setVisible feature of Connect, make sure to have setVisible set to 'true'. This will allow observation of the emulated 3270 screen as claims are executed.

3.3.4.5 User Acceptance Testing

After all tests are successfully executed it is time to send the project for user acceptance testing. That may require deploying to a different environment (database) or marking the new version available to the UAT environment depending on the configuration.

The most important considerations for UAT are the test claims and the expected results. That should come from the UAT team. Once the test claims become available it is possible to execute the web application pointing to the UAT environment.

As before, create a report from the web application and provide this report to the UAT team for validation. If there are claims that did not get the expected result a ticket will be opened and it will be necessary to work with the UAT team to determine the root cause.

3.3.5 Execute

The Execute stage is used to run the web application in production after User Acceptance Testing is complete and all tests have successfully executed.

3.3.5.1 Deploy Configure Project to Production Environment

The version of the Configure Project that was tested in UAT should now be deployed to the production environment. This could mean to physically deploy to another server or it could mean to just mark the version as current for the production environment depending on the configuration.

Also, depending on the environment, it may be desirable to disable the web application before making this change. In some environments that may require synchronizing the deployment with the mainframe parameters used to provide claims to the web application.

3.3.5.2 Execute New Edit Code

Add the new edit code to the web application so that claims with this edit code can be worked. The required parameters will depend on the environment and the web application.

If the web application is disabled, make sure all synchronization issues are resolved before enabling it.

3.3.5.3 Run Reports

Depending on the environment one can run reports either at the end of a run, or at any time before or during a run. OpenConnect recommends creating a schedule, having the web application automatically create the report and email it to the defined recipient(s).

3.3.6 Monitor and Measure

The Monitor and Measure stage is used for two purposes. The first is to monitor the execution of the web application and Connect transaction to make sure they are working properly. The second is to measure the solution's success rate.

3.3.6.1 Monitor

Monitor refers to observing the automated processes to determine if they are executing properly. This is an important step in the improvement process because the underlying business process changes over time. This is called process drift and is caused when business rules change, new functions are required or new data is entered, to name a few examples. If the automated process is not continuously monitored then errors can occur which could result in a large negative financial impact to the organization.

3.3.6.2 Create Attributes

Before it is possible to monitor the automated processes it is necessary to identify which attributes to use. To monitor the web application and Connect transactions it is necessary to ensure that the following attributes are defined:

1. User ID
2. Edit Code(s)
3. Success (this can be a finalized flag or a transition to a finalized activity)
4. Adjustments (this is usually defined as a separate activity)

Sometimes the proper execution of a transaction requires access to specific functions. Those functions can usually be provided either by an attribute on a claim screen or by another activity. If the information can be provided through an attribute then there is no need to create the activity. However, it may be required to create the activity to determine if the transaction followed the correct procedure.

3.3.6.3 Configure and Monitor

Create a perspective in PIC by Claim Number (DCN). This will expose the paths taken for each claim as it is executed. Create a View from this perspective and give the view name of 'Monitor'.

Add the filters for the appropriate attributes. For example, add the filter for the user id(s) used by the Configure transactions. Also add a filter for the specific edit code to be monitored. The method is to create filters for each of the edit codes then only apply one filter at a time.

After all of the filters are applied, PIC shows the paths taken by the automated process for the filtered edit code. The paths will help determine if the transaction is following the correct procedures. In addition to the paths it is desirable to create charts to help understand how the transactions are running. These charts are edit code specific but at a minimum create the following charts:

1. DCN by User Time
    a. This chart helps determine how long a process executes for a specific edit code
    b. Create the baseline right after deploying the edit code to production
    c. Monitor the execution time and keep track to determine trends
    d. If the execution time increases or decreases significantly over a short period of time there may be a problem with the transaction
2. Edit Code by Count
    a. Create this chart and watch the various paths in the process map
    b. Keep track of the edit code distribution for each path
    c. If there is a significant increase or decrease of a specific edit code for specific paths over a short period of time, there may be an issue with the transactions for the specific edit codes

3.3.6.4 Measure

Measure refers to the process to measure the success of the automated processes and to determine how to improve the success rate.

3.3.6.5 Create Attributes

Measuring the success of the automated processes can be accomplished by combining the reports provided by the web application and PIC charts. For PIC charts, the following attributes are required:

1. User ID
2. Edit Code(s)

3.3.6.6 Configure and Measure

To measure the success of the automated processes it is best to use the reports provided by the web application. Each web application is specific to the customer and environment. Refer to web application-specific documentation to determine how to create and read the reports provided by that web application.

PIC can be used to help determine how to improve the execution of the automated process. If, after reviewing the reports from the web application, the success rate for a particular edit code is low because there are other edit codes on the claim that prevent the claim from finalizing, use PIC to help.

First, learn what those other edit codes are. Then, determine if those edit codes can be automated. Then, determine what the success rate can be if they are automated. Finally, create the requirements documents so that they can be automated.

Described above is the improvement process outlined in this document. The difference is that it will be necessary to have the User ID and Edit Code filters defined and selected before starting the process.

3.3.7 Summary

This document described an overall process and a set of detailed steps that are flexible enough for each organization to fit it to their environment and situation as necessary. Each stage can be considered a separate sub-process and can be enhanced to provide results needed for any particular organization.

Experience shows that following the process leads to quantifiable, sustainable improvements to claims processing, including cost savings.

In this chapter a fairly complete example of a toolset that automates the major steps of the C-K-theory based process. The collectors and analytics components build the data in the required Event-Actor-Object format as input for Process Discovery. The analytics components provide the first clustering off coherent event streams into services, to simplify and reduce the input data. A major role is played by the Process Intelligence Cluster which automates the disjunction and the actual analysis of the process conceptual representation (in terms of Hidden Markov Models). An extension of PIC with Formal Concept Analysis is scheduled in subsequent versions of *Comprehend* in 2013 – 2014. Finally, *Comprehend* also provides support for conjunction and activation of insights in value leaks and process improvements, by means of its possibilities for simulation and user replay of processes.

Chapter 4

Case Studies I: Administrative Processes

4.0 Introduction

This section defines a series of case studies using the *Comprehend* product suite that has been developed by OpenConnect Systems, Dallas, Texas, USA, in conjunction with research performed at KU Leuven, Leuven, Belgium. The *Comprehend* Product suite was designed to perform automated business process discovery specifically focused on operational business processes (e.g., claims processing, call center productivity, etc…). It has been used extensively by healthcare insurers in the United States to reduce the administrative costs associated with processing healthcare claims. *Comprehend* has been used in two major approaches, first as an analytical tool to give insight into what is actually happening with an ICT-enabled service process, specifically, to discover and expose issues of concern regarding time and cost with both the process itself as well as the resources used in the process. This is what we have called earlier "value leaks". Second, if he process being investigated is a combination humans assisting automated ICT-enabled activities (e.g., claims processing or any high transaction activity), use the information form *Comprehend* as a targeting mechanism for further process automation. In this regard, *Comprehend's* automated discovery capability can also be used to document the "repetitive steps" used by the human to complete the task so as to produce the documentation for an automation script that can be used to generate software code. In this case, the level of automation and efficiency are increased and the overall labor component of the task is reduce thus increasing the overall efficiency of the process. In essence, this is accomplished by reducing or "plugging" value leaks.

4.1 Overview of healthcare claims payment process

The health insurance claims payment process is designed to ensure the correct payment is made to the appropriate healthcare provider (physician, laboratory, hospital, etc…) in the timeliest manner possible. In the U.S., over 6 billion claims are processed annually at an estimated cost of $1.23 billion (AHIP 2006). The process depicted below (Figure 4.1, T.M. Floyd &Company 2006) is representative of the overall claims review, validation and payment process, commonly called claims adjudication:

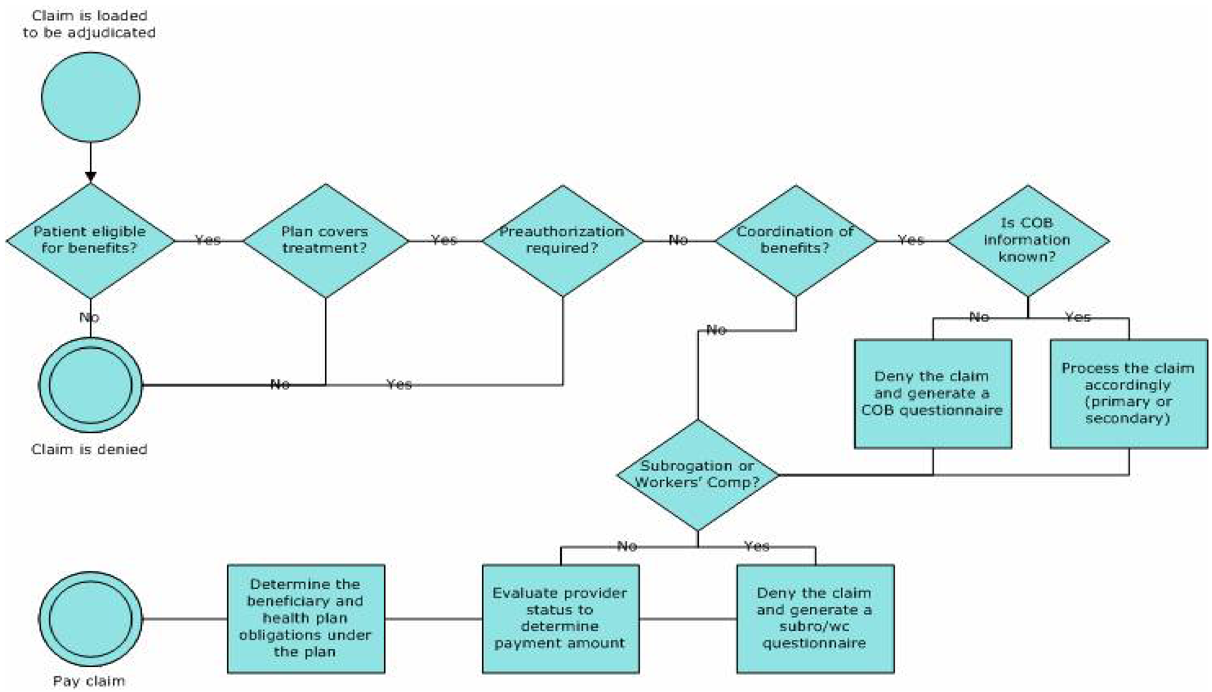


Figure 4.1 – Overall Claims Review

In its simplest form, a claim represents the documentation of an event where a person was treated by some provider for a specific healthcare situation (flu, discomfort, etc…) on a specific date and time. In essence, a claim is the record of a healthcare event for which a request for reimbursement for services is made to the health insurer. As the insured is a subscriber or member of the health plan, he is able to receive such payment. In the case of a provider, if they are a member of the plan network, they may be paid directly for providing such service at a specified rate outlined in the agreement. However, many types of agreements exist and there is much room for error in recording the event (claim) so as to qualify for payment. Any time a claim does not automatically pass the validation steps of the process above (e.g., patient/subscriber validation, treatment coverage validation as per contract, benefit coordination with other plans, etc…) the claim will be invalidated or denied. This process of validation and subsequent payment is termed adjudication. If a claim passes through the system automatically without human intervention, it is called auto-adjudication. If a claim does not pass through automatically when submitted, it can either be placed in to a holding file or "pended" to be manually edited by a human claims examiner and be re-entered into the system for processing. This step is an example of *muda* or a value leak since it is re-work caused by a defect somewhere in the system that is increasing the overall cost of processing the transaction. If the claim is "pended" but then is automatically re-worked by a software robot or virtual claims examiner, it will fall into a special category of *operational auto-adjudication* and therefore mitigate the labor costs of claims adjudication process through the substitution of technology cost for human labor cost. Therefore, the goal of healthcare claims payment is to have a high rate of auto-adjudication often called the *first pass rate* or *operational first pass rate*.

As computer hardware, software and telecommunications are all inherent in claims systems, the healthcare claims process is a classic example of an ICT-enabled service. Further, while the underlying healthcare delivery service possesses tangible actions (medical procedures) and is people focused, the claims process deals mainly with intangible actions (business rules and electronic money transfer) and enrollment (insurance policy) focused. Claims processing therefore have many of the characteristics of intangible service processes, including difficulty in measuring many varied workflows, operating on large scale. Large insurer's process tens of millions of claims annually and often employ hundreds of human examiners to handle claims involving very complex business rules that many times are not automated. The rules are complex as the insurance industry is highly competitive and it is often the case that an insurer allows the group representative (HR or finance person) to customize the terms of the policy, offering non-standard benefits, in order to obtain or renew a given customers business. In this manner, customers and insurance salesmen routinely create new services. ICT must then create new services and apply business rules which often vary by customer. There are thousands of business rules modeling the terms of all the group policies. While it takes some data entry on the main claims system to implement deductibles, co-payments and other terms of service, manual claims processors must be aware of the main rules for each group and have access to process documentation for the more arcane situations. Due to the complexity and large scale, manual claims processing operations executives generally organize their staff into a number of teams who routinely handle a few customers and claims types. Problems arise when the complexity and human involvement lead to variation and defects, resulting in a costly adjustment. However, as shown later, the majority of adjustments are caused by the providers as they document the services they deliver.

At the same time, manual claims processing has many of the characteristics of a manufacturing (goods-focused) process, as front line associates spend little time waiting for work. They generally have queues of claims and adjustments waiting for them, and they need only click the top one on the queue to begin the manual claim adjudication or adjustment work. Additional *muda* arises because examiners are measured by the number of claims and adjustments they process, with varying points awarded based on complexity. Experienced examiners often bypass the queuing system which prioritizes claim based on age, so that they can meet their quotas and to make up for having to work a number of complex claims.

4.1.1 Value Creation and Value Leaks in Health Insurance Claim Processing

Health insurance companies (payers) create value for insured members when they reimburse the healthcare provider or member for the cost of health care delivered to the member. The insurance service exists in the first place to spread risk among a pool of members, often in so called groups (employer-sponsored plans) such that no member has the burden of paying the entire cost for expensive health care services or medicines, while most of the time, no individual uses the insurance at all. In order to obtain that value, member pay and payers receive monthly premiums, then process claims for reimbursement. The payer's profits are the mainly the difference between the premiums and the reimbursements, after subtracting the costs of providing the service.

All health insurers have claims processing systems that take in claims through electronic means, apply business rules relating to covered services, usual and customary fee schedules, deductibles and co-pays (called auto-adjudication), then finalize the claim by issuing a check or denying the claim. The payer sends an explanation of benefits (EOB) to the member, notifying him of the results of the process. Most large insurers are able to perform this first-pass automation of claims processing successfully on about 70% - 85% of claims, using their main claims processing system. The rest of the claims were traditionally processed manually, by human claims examiners, though there are now software robots that emulate humans performing some manual claims processing work. Not all work can be automated, as in the case when an examiner must speak to a provider or member, or when human judgment is required. Over time, members or providers can challenge the EOB, creating a claim adjustment.

According to industry survey's (AHIP 2006), a claim that is received electronically and adjudicated by the main claim system or a software robot emulating a human costs the insurer $0.20, while one processed by a human costs about $2.05 (Figure 4.2).

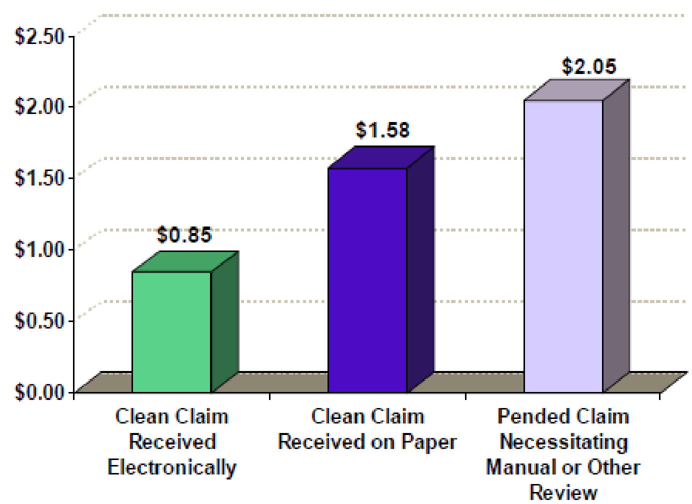


Figure 4.2 - Average cost to Process a claim (AHIP 2006)

On average, the healthcare claims processors Pend 14% of claims received. In the United States, the estimate of overall claims processed by the healthcare industry is estimated to be 5 billion claims per year (CMS). This yields the following calculation:

[5,000,000,000 claims * 0.14= 700,000,000 pended claims]
[700,000,000 pended claims * $2.05/per claim = $1,435,000,000/year in cost of pended claims]

This cost is a good description of a value leak that exists in the healthcare claims payment sector. As mentioned above, healthcare claims payers try to reduce this cost through the use of automation in the form of software robots to reduce the manual labor costs. The best cost observed by OpenConnect in practice among the healthcare payers is $.073 per claim. This particular payer also had a operational first pass rate of 95.4%, an industry best. If all pended claims could be paid at this rate, the total industry cost to process pended claims would be $511 million/year yielding a potential savings of $924 million/year to the healthcare industry with the closing of one value leak alone!

4.1.2 Healthcare Claims Adjustments

While there is no official industry estimate of claims adjustment volumes and costs, it has been determined through discussion with industry leaders that approximately 5% of all healthcare claims are subject to adjustment at a typical cost of $15/claim. Based on this information, the estimated cost of adjustments to the US healthcare payers is as follows:

> [700,000,00 pended claims * 5% adjustments to pended claims= 35,000,000 adjustments/year]
> [35,000,000 adjustments * $15/per claim = $525,000,000/year in costs of adjusted claims]

When this cost is added to the $1,435,000,000 per year in the cost of pended claims, it shows that the healthcare claims payment process in the United States leaks approximately $2 Billion/year between pended claims and adjusted claims alone. Also, it must be noted that for claims that are not processed in a timely manner (usually defined by a service-level agreement in a specific contract), interest and penalty charges can typically add an additional $25 to the cost of a claim: this has not been included in the estimated above. In the following case study, adjustments were costing the healthcare payer over $30 million per year.

4.2 Healthcare Administrative Processes Case One: Claims Adjustments

This case study was performed at a large not for profit healthcare claims payer in the southeastern part of the United States which shall be called *Atlantic Healthcare Company* (fictionalized name). The healthcare insurance company studied has over 8 million members and processes on average 5.45 million claims per month yielding approximately 65.34 million claims per year. The average operational first pass rate is at 90% with adjustments being approximately 5% of overall claims (3,267,000).

4.2.1 Context for this process discovery case study

As discussed previously, adjustments are major value leaks, or *muda* (waste), in the claims process. They represent expensive re-work and present a "lose-lose-lose" situation for the insurer, the provider and the member. Each party has to go through the adjustment process, creating extra work for the insurer, delays in reimbursement for the provider, and generally a number of follow up and tracking activity by the member, who is ultimately responsible for reimbursement to provider, whether insurance is available or not. As this insurer spends $30 million annually on adjustments ($9.18 per claim), much of it in manual labor costs, adjustments potentially represented a major opportunity to realize large savings

All insurers know the volumes of claim holds (e.g., pends), number of adjustments by reason code and have addressed some of the most costly value-leaks. A disproportionally large component of the adjustment cost is the manual labor, where abundant opportunities often exist to increase process flow by reducing waste (*muda*). The caveat in healthcare insurance claims service is the previously stated problems of large scale parallel execution, complexity of business rules and the difficulty in measuring precisely how much time is spent at each step. This healthcare insurer had achieved significant cost reductions through business rule changes, by automating a significant portion of new claims processing with software claims robots, incentives to providers, and other techniques. However, the additional human interaction involved with adjustments made it significantly more difficult to analyze and improve the adjustment process.

Many healthcare insurers have increasingly come to points of diminishing return in process improvement over the years because they did not have the time dimension available in their process maps. *Comprehend* collects and analyzes data from the customer's business processes as users interact with systems, automatically creating Markov Process diagrams, including the time dimension**,** as shown in Figure 4.3 below:

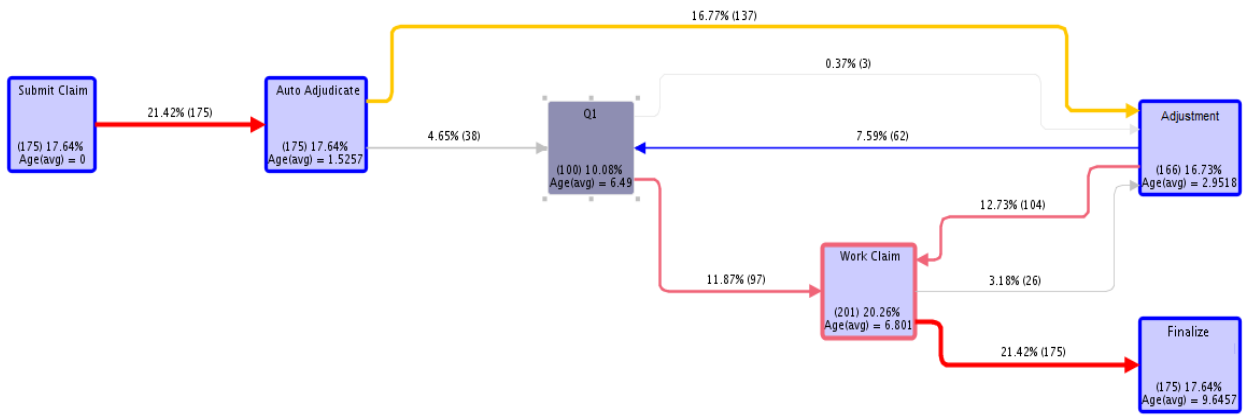


Figure 4.3 - Claims Adjustments

Observe how the above Hidden Markov Model is normalized, in the sense of the definition discussed in chapter 2. Not only are the Markov chain percentages documented, but the time spent at each state is calculated as well and stored in the process events. This additional dimension of time allows customers to focus improvement activities more accurately on costs to the business, in order to increase process flow and hence reduce cost per transaction.

It is a truism that to improve a process one must first understand it in detail: each of the steps, what contributes to a quality result, how workers and systems behave as they perform the process, and so on. A problem that arises repeatedly when trying to understand service-based processes is that many of the activities are performed by humans in non-standard ways, making service processes much more difficult to understand and improve than manufacturing processes. Process variation (value leak) arises as workers constantly adjust their actions when performing work activities to complete a task, so the process evolves over time without control or documentation.

There is significant inherent complexity involved with improving a claims system where the work is distributed between desktops and back-end servers and databases. Since not all the data traverses the network, no single PC or server-based technology can capture all the necessary content, nor can a passive network-based collector do so, especially since much of the claim activity occurred before the project started. In order to obtain necessary process history for a claim, database queries were built into *Comprehend*, augmenting the visible, real time process flow with historical information. Additionally, each of the activities in the process event generally consists of several user interactions and screens. As an example, within the adjustment activity, the user may navigate several claim history screens in order to process the adjustment in question.

4.2.2 Initial Process Analysis and Improvement

When an examiner brings up a claim screen, the claim number becomes the key for a predetermined database query run later in the collection process. The claim number and other information from the database, such as diagnosis code, line of business, provider ID, fee for the service, location of the service, etc… provide *Comprehend* with the history needed to understand each important event in a claim's life cycle. Since an examiner's time is a large portion of the overall cost, the *Comprehend* desktop agent software tracks how examiners spend their time: whether in production claims work, researching membership or benefit information or not performing production work at all. With this approach, analysts have the information they need to choose where to focus their improvement efforts.

After *Comprehend* discovers and maps the process, business analysts can perform their own *ad hoc* analysis in support of an improvement project. Figure 4.4 and Figure 4.5 are *Comprehend* analyses of the number of minutes spent by a subset of claims examiners, who worked on adjustments over a few months.

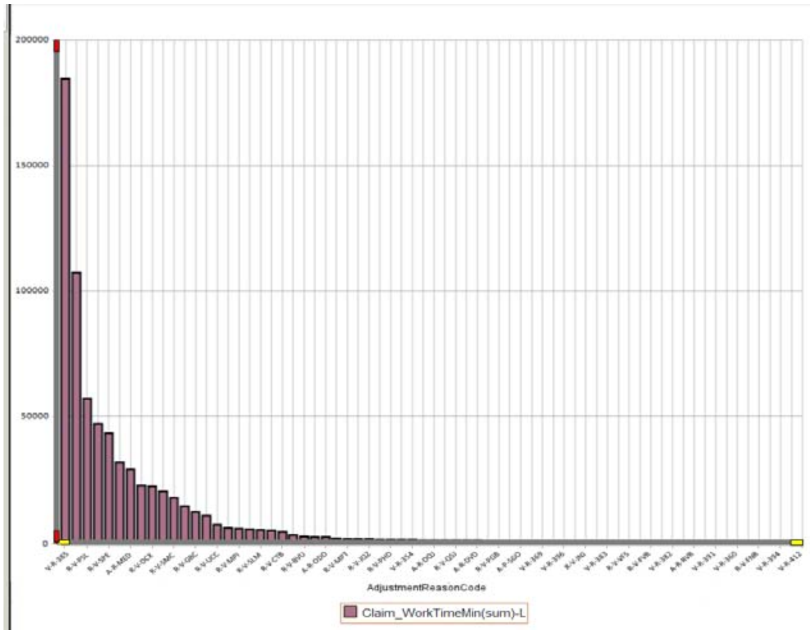


Figure 4.4 - All 69 adjustment codes

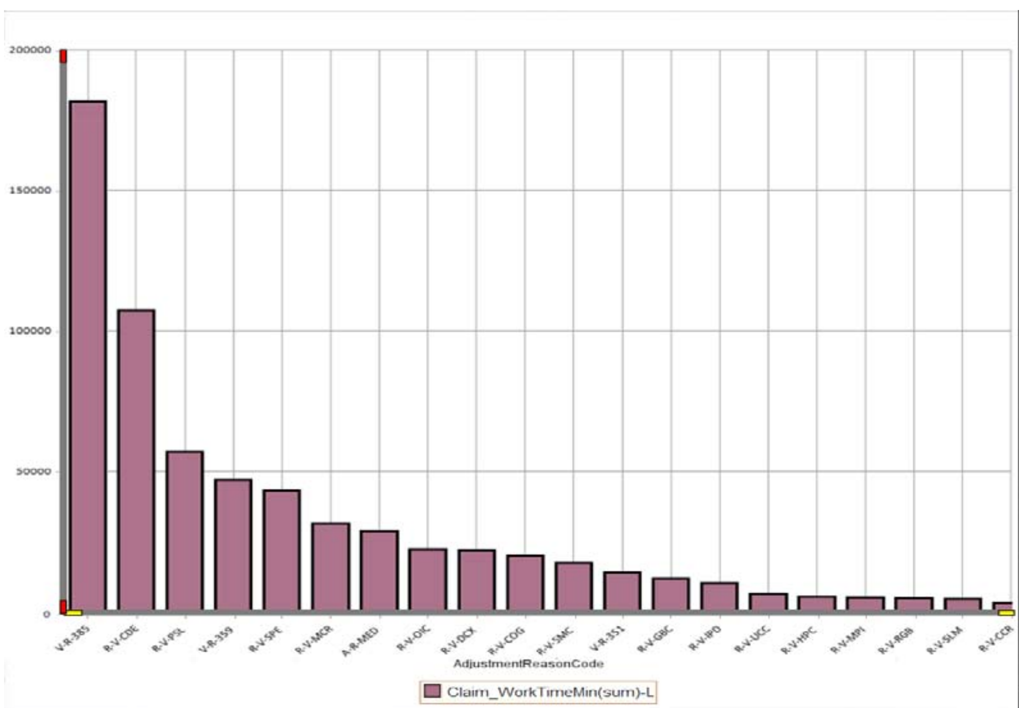


Figure 4.5 - Filtered View of Adjustment Codes

Since labor is a large component of adjustment costs, it is a useful indicator of potential improvement projects, a pointer to which projects to pursue, and a unique way to gather, analyze and report on the data needed to complete the project. With 69 adjustment reasons in use, it is necessary to narrow the focus, to keep the scope of the improvement project reasonable and to avoid improving adjustments with little impact.

In this case, the analyst exported to MS Excel all adjustment codes and the time spent by examiners working them to build a Pareto chart shown in Figure 4.6. This revealed that nine adjustment reasons out of 69 accounted for over 80% of the time spent on adjustments by this organization. With better than an 80-20 rule, there is a reasonable chance to make major improvements with a small number of process changes and projects. In order to optimize the chances for breakthrough improvement, the analyst concentrated on those top nine adjustment reasons, again using *Comprehend's* data export feature to build a filtered adjustment code analysis (see Figure 4.5)**.** Further analysis with *Comprehend* showed that approximately 20% of the providers accounted for over 75% of the adjustments, reinforcing the judgment about the chances to succeed in a reasonable time, with a small number of projects.

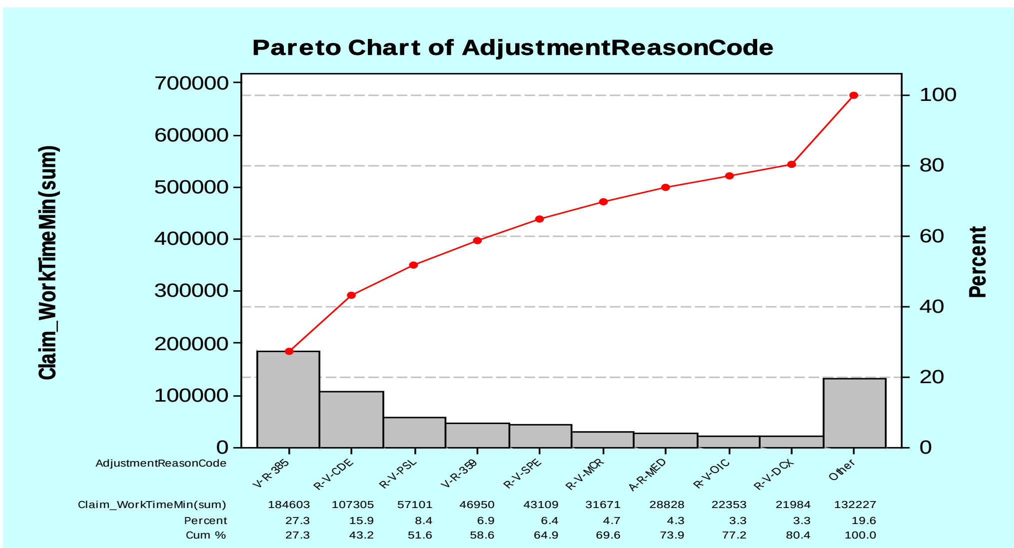


Figure 4.6 - Nine Critical Adjustment Codes

Reason codes, as shown in Table 4.6, were stratified into those caused by the provider and those caused by the payer. Providers caused the top three codes adjustment codes, accounting for the vast majority of time spent on adjustments. The top adjustment codes relate to providing timely and accurate billing information. Working with a few of the providers who were involved with the largest number of adjustments in a pilot project, the insurer and provider billing personnel analyzed the most common specific errors and built a checklist-based procedure to use every time the provider submitting a claim. The procedure was just common sense, but making it a part of the process to reduce errors and rework, focused providers on behaving in ways that were win-win. Both parties worked together, put some rigor into the error analysis and improvement of the claim submission process and used the proven technique of a checklist to dramatically reduce omissions, errors and late submissions. Of course, this type of project is only possible when one can measure time spent in detail.

For the pilot group, adjustments caused by the top three reason codes dropped by over 50% within two months and by 65% within four months.

Claim_WorkTimeMin (sum)	Adjustment Code	Adjustment Reason	Responsible Party
184,603	V-R-385	HVB Pricing Change	Provider
107,305	R-V-CDE	Reversed - Initial Claim Line Level Info Incorrect	Provider
57,101	R-V-PSL	Reversed - Provider Submitted Charges Late	Provider
46,950	V-R-359	Price Changed	Payer
43,109	R-V-SPE	Reversed - System Processing Error	Payer
31,671	R-V-MCR	Reversed - Insurer Paid Medicare as Primary Claim	Provider
28,828	A-R-MED	Adjusted - By Medicare	Payer
22,353	R-V-OIC	Reversed - Other Insurance Carrier Information Changed	Payer
21,984	R-V-DCX	Reversed - Diagnosis Code Changed	Provider

Table 4.1- Time Spent on Top Adjustment Reason Codes

Atlantic Healthcare Company used the pilot results to work with the rest of the 20% of the providers who were involved in a large number of adjustments and achieved similar results. Six months into the project, recurring annual savings were projected to be $2.1 million if implemented across all of business. The insurer then rolled out the new process into production, with minor revisions, and made it a standard part of training providers received for submitting claims. Annual savings actually topped $2.5 million (US) with the reduction/redeployment of 55 Full Time Equivalent (FTE) personnel in the examiner organization. The insurer also saw modest reductions in interest and penalties, training costs and facilities as a positive side effect of correcting these errors. The results were better than projected because the new procedure also reduced a number of errors with other claim submissions and focused the providers' attention more on the quality of their claims work.

4.2.3 Additional Value Leak Identification and Reduction

Essentially, all adjustments can be considered to be *muda*, since they are, by definition, rework of a previously completed claim. Therefore, in an effort to use “discovered” intelligence to drive as much waste from the process as possible, the project team decided to analyze adjustments in more depth. In parallel with the project focusing on improving providers' claims submissions, the project team looked into improving the adjustment process from an internal viewpoint.

As mentioned earlier, the *Comprehend* product suite electronically monitors user interactions with application systems at the user interface itself. This includes monitoring the user navigating among screens, changing fields, recording idle time and so on. Once configured with meaningful titles for the activities and the user data to be stored, *Comprehend* produces a process map, essentially providing the analyst with detailed process map of digitally observed activities and events detailing sequence and time.

In order to understand the adjustment process in more depth, *Comprehend* was configured to collect data on the adjustment process one layer deeper than before.

Specifically, instead of the mapping the life cycle of a claim, the product looked at the life cycle of an adjustment. Instead of entire logical units of work among various applications, activities were now typically one or a few screens within a single application. The resulting session measurements are shown in Table 4.8 below:

Adjustment Step	Occurrence Count	Step Time (seconds)	Total Time Spent (seconds)
System Login	0.032	4.100	0.131
Check E-mail	0.031	31.200	0.967
Navigate Queue	0.985	4.132	4.070
Open Adjustment	0.978	6.974	6.821
Return To Queue	0.071	4.140	0.294
Review History	0.975	280.303	273.295
Open Base Claim	0.846	7.280	6.159
Review Rules	0.247	96.491	23.833
Chk Member Data	0.213	67.338	14.343
Route	0.039	4.290	0.167
Reroute	0.012	4.205	0.050
Call Provider	0.279	191.639	53.467
Call Member	0.273	270.430	73.827
Pend	0.086	6.910	0.594
Modify Header	0.367	48.954	17.972
Modify Line	6.975	29.603	206.495
Finalize	0.792	5.871	4.650
End	1.000	0.000	0.000
Average Adjustment Time (seconds)			687.137

Table 4.2

It was helpful to summarize the workflow in the manner below however, in order to find and eliminate waste, a graphical representation (Figure 4.7) of the process, automatically generated by *Comprehend*, is a useful step.

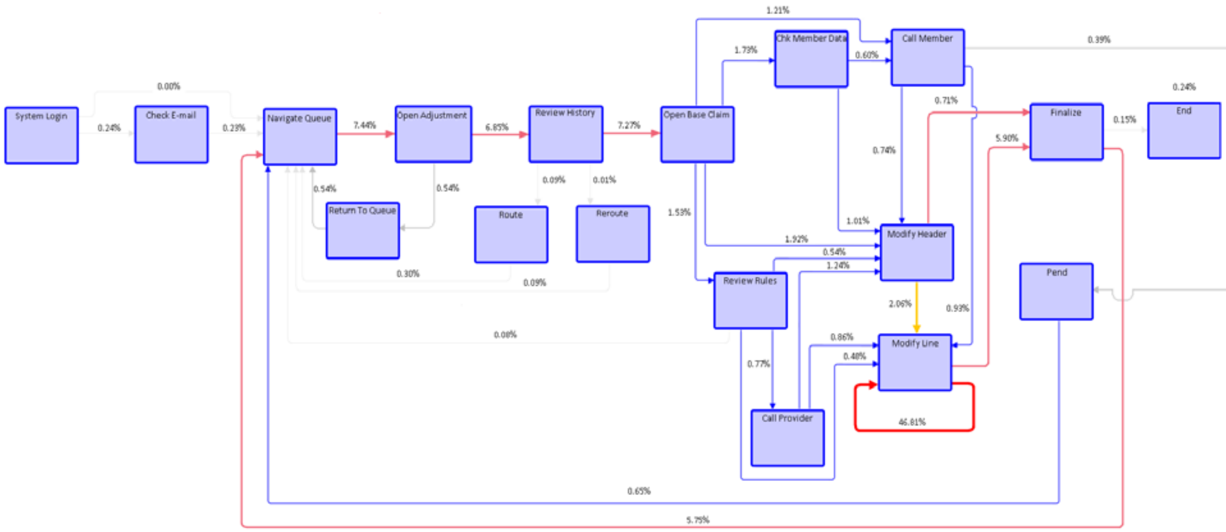


Figure 4.7 - Adjustment Markov Diagram

Observe how “Modify Header” and “Modify Line” are the two places where the Markov Model is not normalized. It is natural to focus on them as potential value leaks. By noting that adjusters repeatedly modify lines on a claim (and spend a good bit of time doing so), the team chose that as a focus. Identifying the fields an adjuster modified in this repetitive activity, it became clear that they often modified the same field on each line. Analysis showed that hospital claims averaged 19 lines per adjustment. It was determined that approximately one-half of a percent of the entire adjustment process, or $150,000, could be saved by reducing user keystroke- based navigation. A relatively simple desktop augmentation script was written in three days to implement the change. In this case, the adjuster used a PC shortcut key, invoking a "macro". The macro automatically moved the cursor to the same field on each line of the claim after the user pressed the Enter key, rather than using a mouse or repeated tab and page forward keys. Training was minimal and results, while incremental, were immediate. *Comprehend* analysis confirmed the time and dollar savings.

Additional analysis showed that claims examiners created a large amount of adjustments when manually pricing claims' services. Manual pricing involves a number of steps to lookup usual and customary service rates, group benefits, state mandates and other information. Using well accepted techniques during the original manual pricing process, the team avoided adjustments downstream.

An additional follow on project focused on analyzing other forms of waste using *Comprehend* as an automatic process diagram generator. Because there are a number of ways to successfully navigate the adjustment process, it was useful to identify exactly how the faster adjusters worked on the various types of adjustments. Once that was complete, the operations manager incorporated those workflows into training. In addition to saving approximately $300,000 annually through improved flow, the approach helped new employees flattened the learning curve. New employees then learned the knack of quickly completing adjustments with high quality, more quickly acquiring the skills of seasoned adjusters without nearly so much time spent in trial and error approaches.

The improvement team is currently involved with projects to identify additional automation improvements, improving the audit process, and in sales and marketing, based on the ability to measure service workflows that were not possible before acquiring the collection and analysis product.

4.2.4 Value Leak Recapture Summary

As discussed earlier, adjustments, in and of themselves are value leaks. To reduce the impact, the improvement team implemented the following value recapture projects (Table 4.3):

Value Leak	Type of Imperfection	Recapture Action	Estimated Savings
Improve provider data	Incompleteness of Information	Joint payer/provider improvement project	$2.5 million
Reduce data entry	Asymmetry of meaning	Desktop script	$150,000
Improve overall adjustment workflows	Asymmetry of meaning	Workflow analysis, training	$300,000
More intelligent audit	Asymmetry of Information	TBD – Identify best audit indicators, less en random choice	TBD

Table 4.3

4.3. Healthcare Administrative Processes Case Two: Determining the real costs of claims processing with a Business Process Outsourcer (BPO).

In order to maximize core organizational competencies and reduce costs, some firms have chosen to outsource functions where they believe they cannot offer differentiated value. This case study addresses issues associated with this action in a healthcare claims environment.

4.3.1 Context of this Case Study

Lonestar Healthcare Company (*fictionalized name of a healthcare insurer located in the southwest region of the US*) had been working on an automatically renewing contract with a major business process outsourcing (BPO) provider. Originally, this effort was planned to be temporary and, as such, there were few documented service- lifecycle agreement (SLA) requirements. Also, the initial BPO processing price of $2.40/claim had never been reviewed with the outsourcer in question. *Lonestar Healthcare Company* looked to *Comprehend* to help them document and create SLA's that were previously missing and hereto for, were not possible to measure. Additionally, *Lonestar Healthcare Company* needed to understand the breakeven price point at which it was more cost effective to pay a business process outsourcer (BPO) to process claims rather than use in-house examiners. Therefore, the challenge for *Lonestar Healthcare Company* was to appropriately measure the administrative costs of in-house examiners to be able to determine the most cost effective price for an outsourcer to do the same work.

To accomplish this, *Lonestar Healthcare Company* initially used *Comprehend* to determine the cost of each claim (by claim type) that was worked by in-house examiners. *Comprehend* enabled *Lonestar Healthcare Company* to measure the time taken for each individual claim and filter out complicated claims (normally performed by in-house examiners) versus less complicated claims similar to the type routinely sent to the BPO provider.

4.3.2 Business Process Outsourcing: Contract renewal

Lonestar Healthcare Company had been working on an automatic renewal (temporary) contract for the last few years with their BPO (Business Process Outsourcing) provider. Since this contract was meant to be temporary, few SLA requirements were documented and the price had never been reviewed. The Company looked to *Comprehend* to help them create SLA's that previously were difficult to measure. More importantly, the Company needed to understand the price point at which it was more cost effective to pay a BPO rather than perform the activity themselves.

The challenge for *Lonestar Healthcare Company* was to appropriately measure the administrative costs of "in-house examiners" to be able to determine the most cost effective price for an outsourcer to do the same work. Using *Comprehend* to analyze the cost of each group, filtered by the type of claims worked, it was determined that "at-home" workforce was less expensive than the BPO contract in place. Having this analysis at hand during negotionations allowed the company to negotiate a reduced price of their BPO activities from $2.40 per claim to $.90 per claim, a significant decrease in their adminsitrative costs (Figures 4.8).

Type of Examiner	**Average Film (minutes)**	**Time per Number**	**Average Cost Per Claim**
BPO Examiners	3.94		$2.40
At-Home Examiners	4.01		$1.35
In-Building Examiners	4.33		$1.46
Special/Risk Examiners - Home Based	8.27		$2.78
Special/Risk Examiners - In Building	13.99		$4.71

Figures 4.8 Time and Cost of Different Examiner Groups

4.3.3 Workforce Reporting

As previously mentioned, the company had several types of claims examiners; at-home examiners; in-house examiners and Business Process Outsourcing. As opposed to the last issue, here the initial step was to look at how time was being spent by each individual examiner during their time at work. The first set of reports focused on providing visibility into the productivity of each group and how they went about administrating claims. This included monitoring each examiner and team of examiners as to what applications they were logged into during the day and the duration of their usage as well measuring the type of claim worked and the duration of each activity. The *Comprehend Workforce Reporting and Analysis* reports provided visibility into the cost of administrating claims per workforce group as well as by claim type (Figure 4.9).

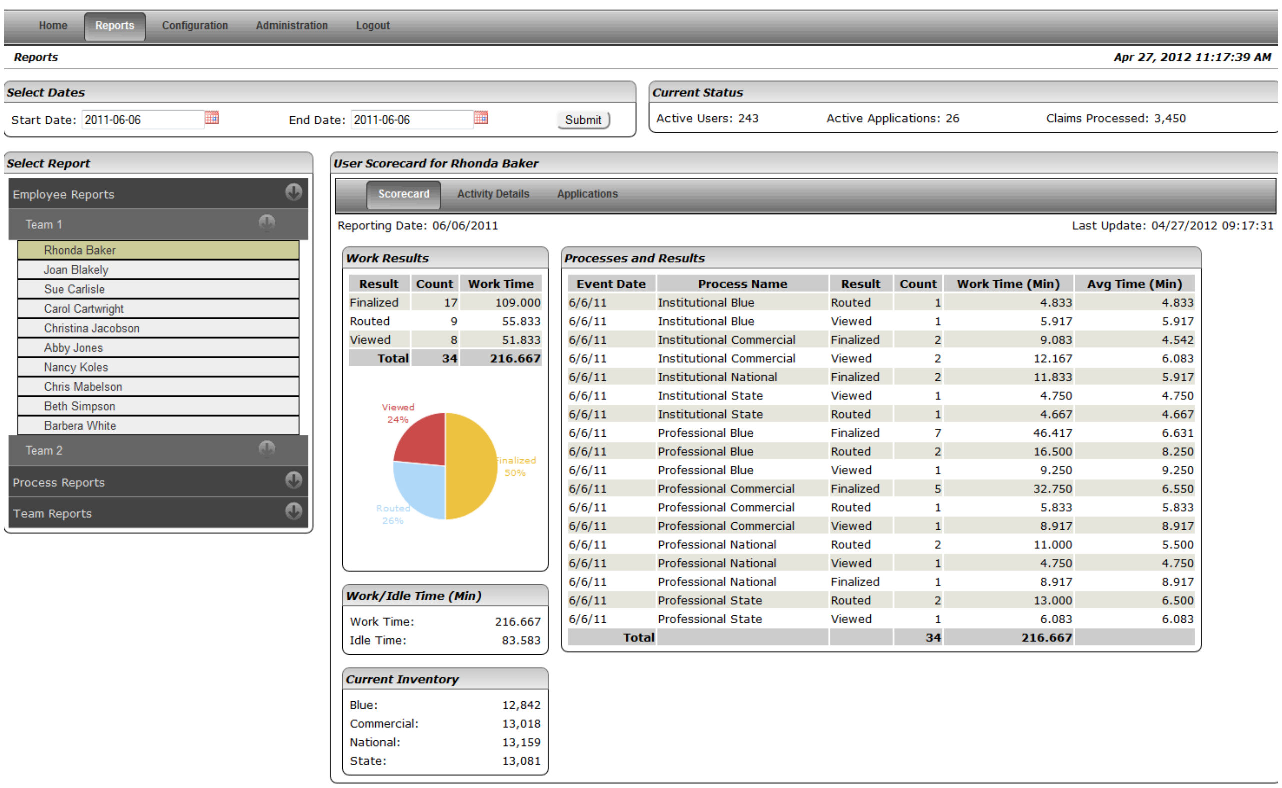


Figure 4.9 Costs by Workgroup

4.3.4 Examiner Productivity

After two weeks of reporting on the different workforce groups, a trend was discovered. Additional analysis with the *Comprehend* Process Intelligence Component (PIC) proved that the business was about to make a decision that was meant to make examiners more productive and save administrative costs but instead was moving their most cost effective examiners to the least productive area. The effort started with an analysis of the examiners designated as "power users" or those thought to be the best examiners on staff. Figure 4.10 shows an analysis of the volume of claims processed by user id with the blue line showing average time.

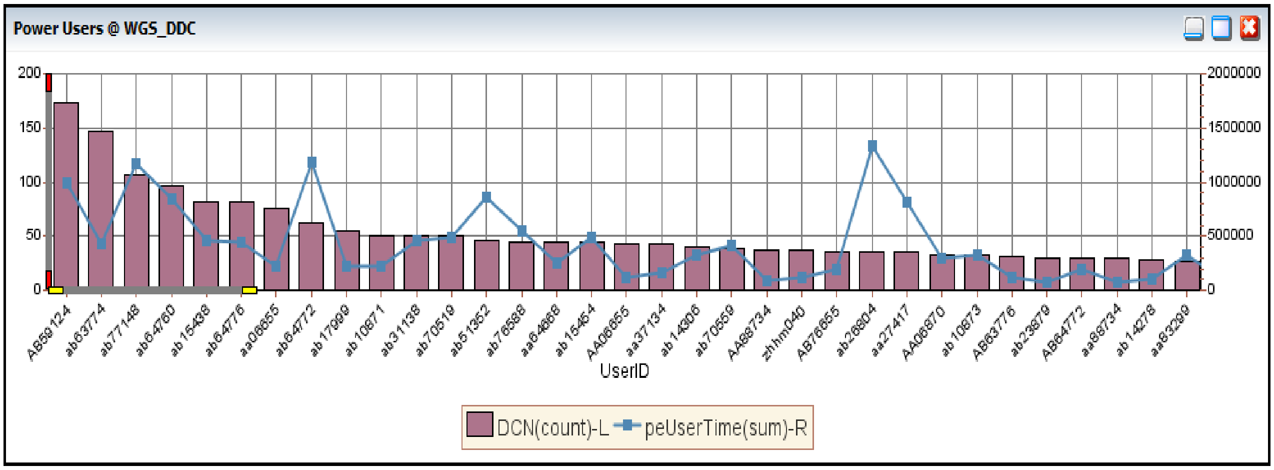


Figure 4.10 – Analysis of Volume Claims Processed by User ID

Using *Comprehend* to analyze each "at-home examiner" and comparing them to examiners working in the office showed that the most productive workforce was the at-home examiners. These examiners worked in longer blocks of time and finished claims quicker than their counterparts working in the office. The company was about to take their most productive users and put them in the least productive environment. Instead, with this analysis, they reversed their decision and gave incentives for the workforce to work from home which also allowed them to reduce facilities (real estate) costs.

4.3.5 Process Analysis ToS (Time of Service)

Atlantic Healthcare had been paying approximately $2 million/year in penalties and interest for paying claims late and incorrectly. Working with over 40 States, each with their own rules and regulations, made it difficult to have visibility into the different processes that incurred these costs. By using *Comprehend* to analyze the lifecycle of a claim, the company was able to filter down to the claims that were incurring these fines and gave the visibility into the process and the root cause of the late payments. *Comprehend* was able to analyze this by State with each of their separate rules. This root cause analysis showed several inefficiencies to multiple processes that led to late payments (Figure 4.11).

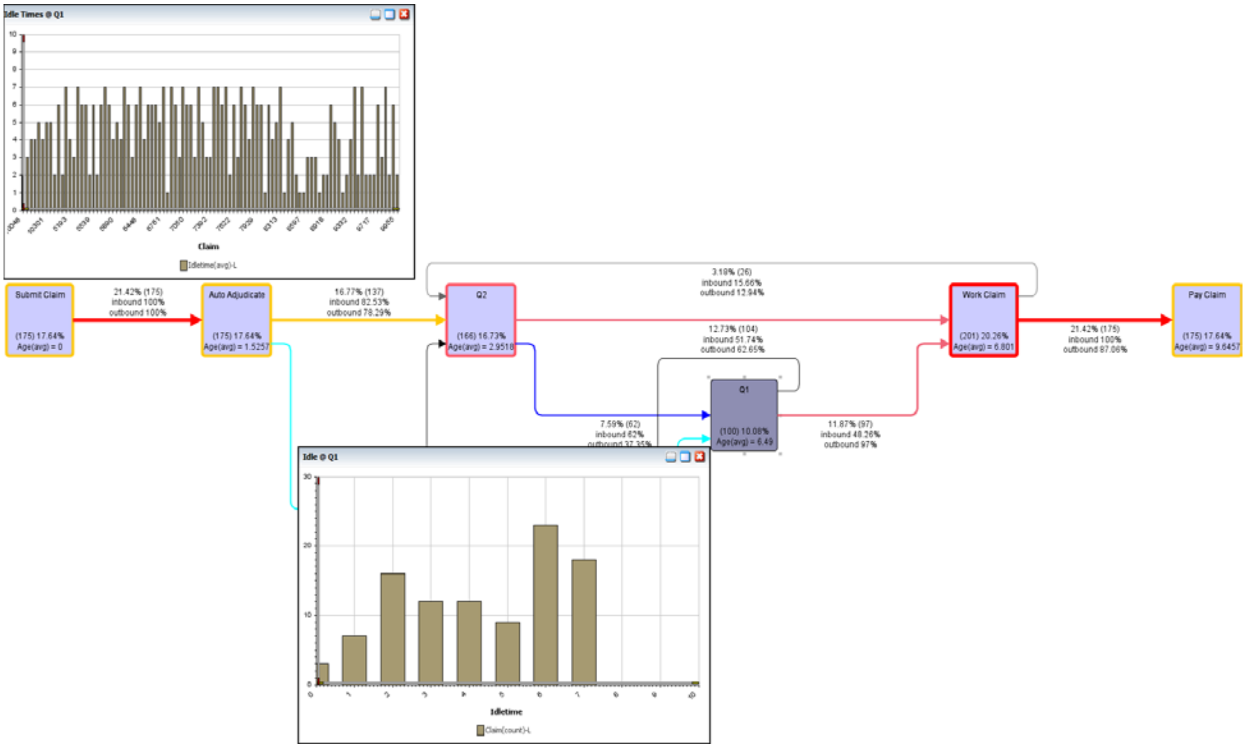


Figure 4.11 - Claims life-cycle process with inset of every claim by top in queue by time (top) and average claim idyll time (bottom)

Also included in this was an analysis of multiple touches to a claim by examiners and BPO's skirting around SLA's that beforehand were not visible to the company. This situation where a claim was "touched by a BPO examiner" multiple time resulted in the claim being paid at a higher cost. This analysis is shown in Table 4.4. As BPO examiners were doing this disproportionally to *Atlantic Healthcare* examiners, an investigation revealed that revealed that the BPO examiners were doing this purposely to get to the higher reimbursement level thus, abusing, if not perpetrating fraudulent activities.

User ID Processor 1	Examiner Type	Film Numbers Worked	Average Time per Film Number (minutes)
lxj	BPO	50	7.15
mdw	BPO	36	11.25
mxv	BPO	36	10.37
dzn	BPO	27	6.78
dyd	BPO	6	3.02
dnd	A HCO	2	0.61
eak	A HCO	1	0.20
mxz	A HCO	1	6.42
pau	A HCO	1	1.10
sgq	A HCO	1	24.79
		161	8.51

Table 4.4

Many of the issues were able to be fixed within the first week of analysis saving the company million's in late fees.

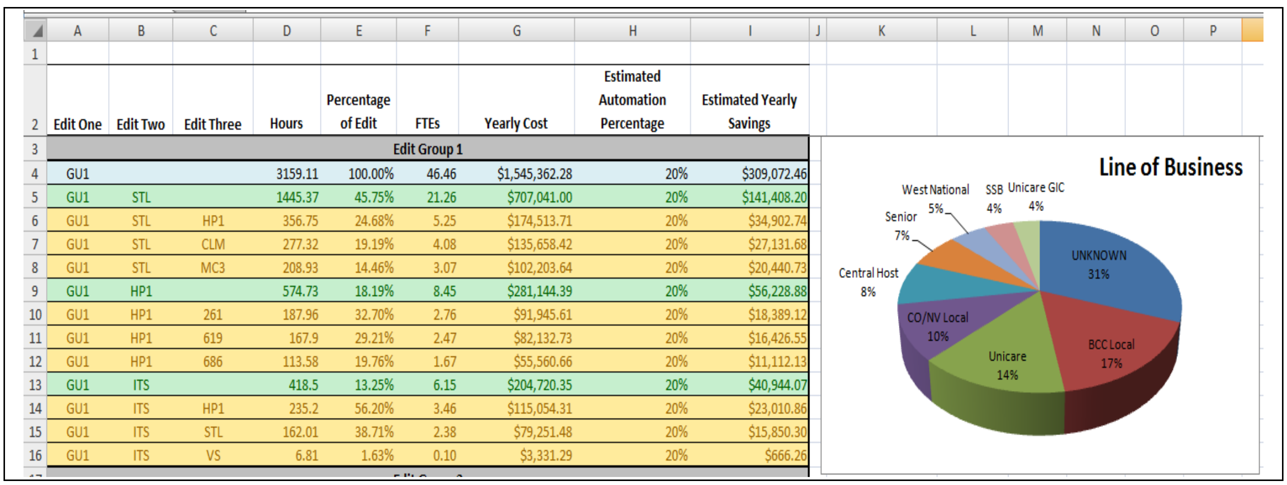
4.3.6 Process Analysis – Claims Impact to the Business

One of the most cost inefficiencies in the healthcare system is the manual adjudication of claims. Business Impact analysis can help determine which edit codes to automate or improve based on the actual cost to the business. The business impact is determined by the total time spent by users to work specific edit codes. The actual impact to the business is calculated using the formula:

[total time in hours / hours of work each day) * hourly rate]

The CAR (*Comprehend* Analysis Report) (Figure 4.12) displays the top ten edit codes in position one based on the impact to the business. It also display the top three edit codes in position two for each edit code in position one, and the top three edit codes in position three for each code in position two.

The purpose of the CAR is to show the top edit codes for potential automation and to highlight edit code patterns. Many edit codes fall into patterns, so that working one may fix others. The Analysis Report allows the business to see the relationship between edit codes and perhaps to see certain edit codes appear in the list more often than others.



Edit One	Edit Two	Edit Three	Hours	Percentage of Edit	FTEs	Yearly Cost	Estimated Automation Percentage	Estimated Yearly Savings
Edit Group 1								
GU1			3159.11	100.00%	46.46	$1,545,362.28	20%	$309,072.46
GU1	STL		1445.37	45.75%	21.26	$707,041.00	20%	$141,408.20
GU1	STL	HP1	356.75	24.68%	5.25	$174,513.71	20%	$34,902.74
GU1	STL	CLM	277.32	19.19%	4.08	$135,658.42	20%	$27,131.68
GU1	STL	MC3	208.93	14.46%	3.07	$102,203.64	20%	$20,440.73
GU1	HP1		574.73	18.19%	8.45	$281,144.39	20%	$56,228.88
GU1	HP1	261	187.96	32.70%	2.76	$91,945.61	20%	$18,389.12
GU1	HP1	619	167.9	29.21%	2.47	$82,132.73	20%	$16,426.55
GU1	HP1	686	113.58	19.76%	1.67	$55,560.66	20%	$11,112.13
GU1	ITS		418.5	13.25%	6.15	$204,720.35	20%	$40,944.07
GU1	ITS	HP1	235.2	56.20%	3.46	$115,054.31	20%	$23,010.86
GU1	ITS	STL	162.01	38.71%	2.38	$79,251.48	20%	$15,850.30
GU1	ITS	VS	6.81	1.63%	0.10	$3,331.29	20%	$666.26

Figure- 4.12 - *Comprehend* Analysis Report (CAR)

Comprehend collects the data for the CAR from identified events in the process map as shown in Figure 4.13.

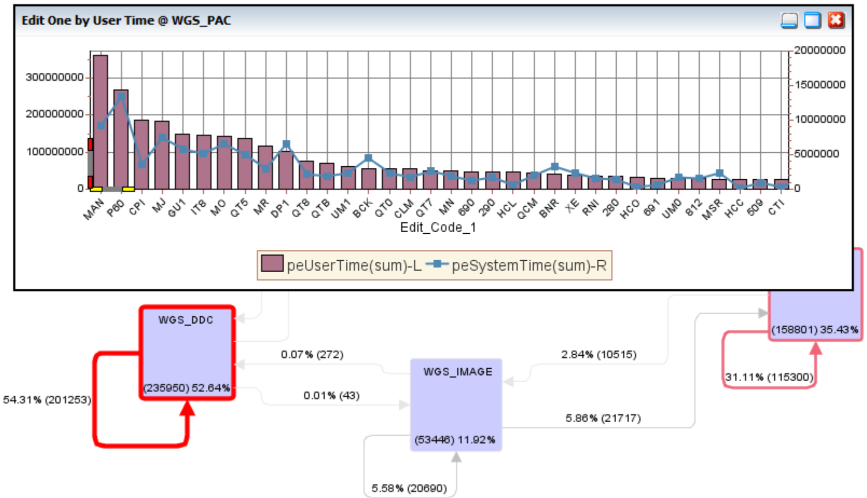


Figure 4.13 - Process Event- edit claims showing edit by user time and system time

4.3.7 Process Reporting- Claims Impact to the Business

The biggest cost impact to a healthcare insurance payer's business is the administrative costs of adjudicating claims. As previously described, most of these costs come from manually adjudicating claims (e.g., human labor costs) that fail to process through automated means. *Comprehend Reporting* provides an accurate cost of adjudicating claims and allows an organization to breakdown the costs by type of claim processed. As with most claims payers, an average cost per claim is used when analyzing impacts to the business. In practice, without the ability to understand the true nature of the process, most organizations attempt to automate claims by volume, looking at Operational First Pass Rates (OFPR) instead of having a true visibility into the real costs of the different type of claims.

Comprehend's reporting of stop or edit codes shows where the examiners are spending their time and allows the improvement team to pinpoint where additional analysis is needed to get understand the root cause of the inefficiencies. The organization can use this information to develop a road map of auto-adjudication possibilities that accurately forecasts how the suggested changes will affect the process.

This is accomplished by determining which edit codes have the highest cost followed by a detailed analysis of the actual work (keystrokes) the examiners perform. This information can be used as input by the *Comprehend* auto-adjudication solution to create "virtual examiners" (e.g., software robots) to work the claims hence offloading the work from costly human examiners.

4.3.8 *Comprehend* Automation

Comprehend collects all of the users sessions and can be used to create "virtual examiners" to automate what the examiners do to work specific claims. All of the business logic is collected and reused in the "virtual examiner" through the *Comprehend* Controller Application(Figure 4.14).

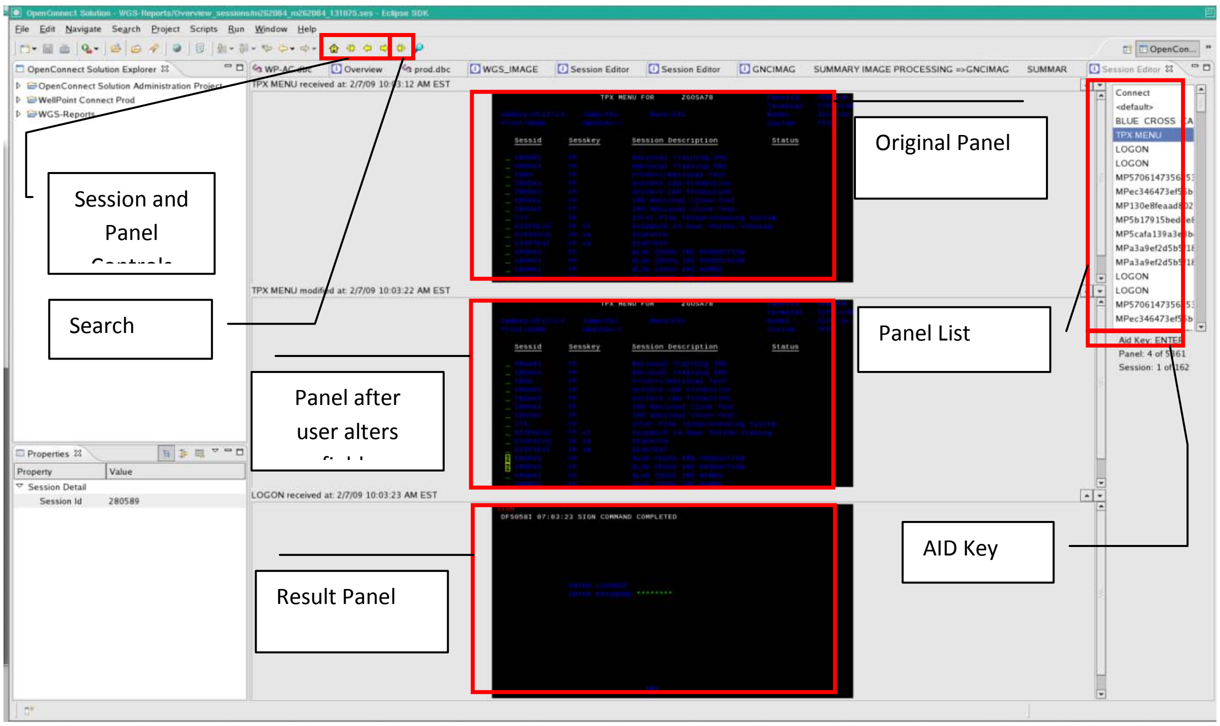


Figure 4.14 - *Comprehend*Automation Controller

The *Comprehend* Auto-Adjudication Solution allows the healthcare payer to reallocate or use fewer resources to manually adjudicate claims saving millions of dollars ($US) in a matter of months.

4.3.9 Value Leak Re-Capture Summary

A summary of the discovered value leaks, the type of service imperfection they represent, the re-capture action taken and estimated resulting value re-captured is described in Table 4.6 below:

Value Leak	*Type of Imperfection*	*Recapture Action*	*Estimated Savings*
Outsourcing Accountability	Asymmetry of Information	A negotiated rate of $.90 with BPO will save HealthMarkets in-home examiners. This is based on Film number counts. BPO is reimbursed by sub-claim number	$550K
Increase Examiners Productivity	Incompleteness of Information	Increasing the average time per day examiners spend processing claims. Raising the average time spent processing claims by an examiner to 75% nets a savings of $495K annually. Analyze top producers, as identified by management. FTE savings = 11	$495K
Identify Best Practices	Asymmetry of Meaning	Process Improvement based on implementation of "Best Practices". Based on current customer base 20% of claims operating costs. 20% savings from 3.5 million is administrative costs - $700k savings. FTE Savings = 16	$700K
Reduce Compliance Penalties	Asymmetry of Information	Time of Service Improvements. Penalties – 10% Improvement delays that cause penalties. $180K savings	$180K
Increase Discounts Received	Asymmetry of Information	Reducing penalties will increase discounts received. Ongoing use of *Comprehend* to continually monitor and analyze claims processing environment. Sustainable claims improvement process provides additional annual savings.	$100K

Table 4.6

4.4 Summary

In this chapter the first set of case studies, those that are concerned with administrative process, was introduced. As both of the case studies are in the domain of healthcare insurance, an overview of the healthcare claims review and payment process was described. Specifically, it was noted that "pended" claims and "adjusted" claims are the largest source of variable cost in the claims processing environment. Both are a type of re-work and incur added labor costs. The former has been estimated to be applicable to 14% of all healthcare claims in the US at a cost of $1.45 Billion (US) per year. The latter is estimated to add $545 MM (US) per year bringing the total cost of this value leak to over $2 B (US) per year.

The first case study addressed the issue of healthcare claims adjustment costs. The costs of adjustments to the company in question, called *Atlantic Healthcare* (fictionalized name), is

$30mn (US) per year. As much of the cost is in manual labor expended by knowledge workers during the review process, the potential savings are significant. *Comprehends* ability to record the time associated with each step of the event associated with a claim (e.g., researching subscriber history, benefit information, etc…).

The first value leak was associated with improving the completeness and quality of the data sent to *Atlantic Healthcare* by the various healthcare providers. The amount of time spent by claims analysts working on population of adjustment reason codes was subjected to a Pareto analysis. This revealed that the top 9 out of 69 codes represented 87% of the time spent in the process. Further analysis showed that 20% of the providers accounted for 75% of all adjusted claims. This insight led to a pilot project with the top three providers who accounted for most of the adjustment issues. The solution implemented a “checklist” procedure to be followed by the providers before submitting a claim. This improvement corrected the incompleteness of information issue (service imperfection) and yielded a 65% improvement over a four month period. This allowed *Atlantic Healthcare* to reduce the examiner workload costs for adjusted claims by $2.5 MM (US) per year.

In addition to the reduction in examiner workload, a further analysis of the process and worker repetitive actions uncovered two additional value leaks, both resulting from an asymmetry of meaning, (service imperfection) yielding an additional $450,000 (US) savings per year.

Lastly, a fourth value leak, stemming from the current process of randomly identifying claims for audit, could add an undetermined amount to the payment recovery effort. This issue stemmed from the inability to have reliable audit identifiers to trigger likely recovery candidates. This asymmetry of information (service imperfection) between the provider organization(s) and *Atlantic Healthcare* could be corrected through the development of said indicators. This would result in a more intelligent and effective audit process.

Overall, *Atlantic Healthcare* was able to discover value leaks and service imperfections that were draining approximately $3MM (US) per year from the organization.

The second administrative case study focused on the cost issues associated with the claims processing environment when working with a Business Process Outsourcer (BPO). *Lonestar Healthcare Company* (*fictionalized name of a healthcare insurer located in the southwest region of the US*) had been working on an automatically renewing contract with a major business process outsourcing (BPO) provider. Originally, this effort was planned to be temporary and, as such, there were few documented service- lifecycle agreement (SLA) requirements. Also, the initial BPO processing price of $2.40/claim had never been reviewed with the outsourcer in question.

In this case study, five value leaks resulting from four service imperfections were discovered. Specifically, the accountability for properly managing the BPO claims processing costs were addressed. This required an investigation of the work being done by in-house and outsourced examiners. It was determined that the outsourced examiners were being overpaid for work performed. This information was not previously available to *Lonestar Healthcare* but was available to the BPO firm. This asymmetry of information (service imperfection) allowed the BPO to realize higher rents than were justified, resulting in a contract re-negotiation saving *Lonestar Healthcare* $550K (US) per year.

The second value leak concerned examiner productivity. *Lonestar Healthcare* held a false assumption as to the productivity of its claims examiners. It believed that at-home examiners were less efficient than those on premises and was about to eliminate the at-home workers. After investigation with *Comprehend* it was determined that the at-home group was actually the most productive. Through rectifying the incompleteness of information (service imperfection) regarding examiner productivity, estimated saving increased by $495K (US).

The third value leak was in the area of best practices. Through the use of *Comprehend*, it was discovered that examiners used different process paths to work claims. As each person developed their own way to perform the work, variances existed in time among the multiple paths. This asymmetry of meaning (service imperfection), when addressed revealed the potential of an additional $700K (US) per year savings.

The fourth value leak focused on reducing compliance penalties which were costing the Company $2MM (US) per year. Performing a root cause analysis with *Comprehend* showed several inefficiencies to multiple processes that led to late payments. This was compounded by different rules that were established by 40 different (US) States resulting in an asymmetry of information environment (service imperfection). Through correcting this imperfection it was estimated that at approximately 10% of penalty cost could be saved resulting in an additional $420K (US) per year.

Lastly, associated with the issue mentioned above, it was determined that by reducing compliance penalties (e.g., correcting the imperfection of asymmetry of information), *Lonestar Healthcare* could avail itself to discounts available to prompt payers. This had the potential of adding an incremental $100K (US) to the projected savings.

Overall, *Lonestar Healthcare* was able to discover value leaks and service imperfections that were draining an estimated $2MM (US) per year from the organization.

CHAPTER 5

Case Studies II: Care Processes

5.0 Introduction

This section discusses the discovery of value leaks that occur in care centered business processes. Whereas healthcare claims processing value leaks were mostly repaired by replacing human labor with software robots, in care processes, the value and quality of human-human contact is often are area of focus. Hence, this interaction, while a focus for efficiency, is moderated within an element of quality. Specifically, what good is it if the process is simultaneously fully automated and fully hated by the users of the process. This level of dissatisfaction can show up as low customer satisfaction rates and possibly declining enrollments and revenues. The three case studies look at this issue from a variety of perspectives. First, the contact center for a large multi-national bank is discussed. This case looked at the costs associated with customer churn due to re-work (*muda*) and errors in the system. The second case looks at customer contact in the case of a municipality and its tax service. Especially interesting is that it is staffed with handicapped employees so quality and efficiency, while an objective is moderated by the desire to give employment to a select group of the population. In this case, the discovery of value leaked concerning both time inefficiencies as well as quality is addressed. In the third case, a hospital post-surgical case ward is examined. Here the process inefficiencies were about the quality of care as it addressed the patient experience. The value leaks in case concerned how to stop value in quality for draining from the system. Each case shows the unique differences when both discovering and plugging value leaks in caring environments.

5.1 Customer Contact Center Case One: Multi-National Bank

This case study describes a business process improvement engagement that was executed at a world leading European financial services institution. The engagement was with their retail banking division, which was seeking targeted initiatives to reduce costs and improve customer satisfaction within call center, back office and self-service operations. The bank has over 600 individual processes and over 5,000 employees involved in these areas.

5.1.1 Goals

The goal of the effort was to identify ways to reduce costs, and increase revenue across the retail operations. Costs were driven based on resource requirements to provide the retail services across both front office (call center) and back office processing. Cost reduction could come from one of two areas: improve the speed (productivity) of service; or improve the quality of service, and thus reduce failure demand. Revenue increases are driven directly through improvements in customer retention.

5.1.2 Speed

Small improvements to the time taken to execute services, either in back-office processing or in the call center, produce large savings in staffing costs. As such, this organization had a large staff of trained business process improvement professionals that had executed many studies of the operations looking for improvements. The challenge faced, however, was that most of these studies required manual data collection. This created two issues: first; worker behavior changes under observation and second, manual data collection is expensive and often error prone [Hubbard, 2007].

5.1.3 Quality

Quality in a retail service organization is difficult to precisely define, and for this organization impossible to measure. Anecdotal evidence suggested that they were facing severe quality issues; however, they had no way to measure quality on a consistent basis.

5.1.4 Revenue (Customer Churn)

Customer churn is always a top issue. There is a large cost to customer acquisition and the main revenue driver is the profit driven from each account. Tracking customer satisfaction through survey was the best proxy for predicting customer churn for this organization. However, examining customer churn, and trying to identify specific causes or leading indicators was extremely difficult on an ad-hoc manual analysis basis.

5.1.5 Approach

The approach taken was to start with a limited proof of value. The goal of this proof of value was to show that *Comprehend* could discover the root cause of business process inefficiencies and produce quantifiable benefits as a result of improving speed, quality or revenue.

The engagement was embedded into a Business Process Improvement department of this bank responsible for the complete Retail Services Organization.

There was a dedicated project manager and analyst of the bank for the whole engagement. The project also included following personnel:

- Project Manager
- *Comprehend* Technical Architect
- *Comprehend* Analyst

The bank provided business SME's for the Direct Banking and Back Office applications to support the daily analysis work of the whole project team.

5.1.6 Project Methodology and Timeframe

The following simplified project plan was applied to this engagement.

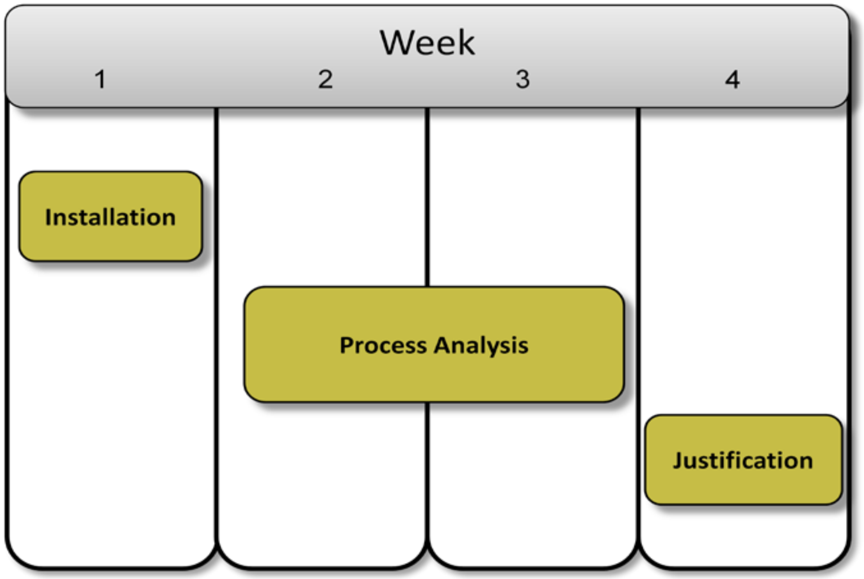


Figure 5.1 - Project Methodology & Timeframe

- **Week 1: Installation**
  - Installation of *Comprehend* software component
  - Definition of process metrics and analytics
  - Aggregation of collected session data and basic analytics
- **Week 2 + 3: Process Analysis**
  - Deep analytics on collected sessions
  - Meeting with Business SME's to verify analytics results
  - Document Process Improvement opportunities
- **Week 4: Justification**
  - Verification of Process Improvement opportunities with Business SME's
  - Evaluation of cost of Process Improvement with Business and IT areas
  - Creation of final results document and business case
  - Final results presentation to Executive Management

5.1.7 Analysis Approach

During the engagement, *Comprehend* collected and analyzed user activity for one week. The analytics phase of this engagement included daily meetings with Business SME's of these two applications to verify and evaluate analytics results.

On a weekly basis the steering committee, including Process Owners and Business Sponsors of this engagement, reviewed and evaluated the analytics results.

The final presentation was prepared and conducted by the customer's project manager to ensure an objective presentation of the engagement results to the Executive Management.

5.1.8 Technical Environment

Installation of *Comprehend* consisted of configuring two span ports:

1. a span port on the TCP/IP connection to collect all TCP/IP traffic between the call center & back office processors and the servers used for the applications supporting their respective business processes
2. a span port to collect all HTTP traffic between the customer using the self-service web application and the web server.

Comprehend was installed on an appropriate set of hardware, and configured to pick up the network traffic from the span ports. Additionally, the log files created by the IVR system were imported into the Process Intelligence Cluster to enable analysis on customer behavior using the IVR system.

The entire installation and initial configuration took 4 hours.

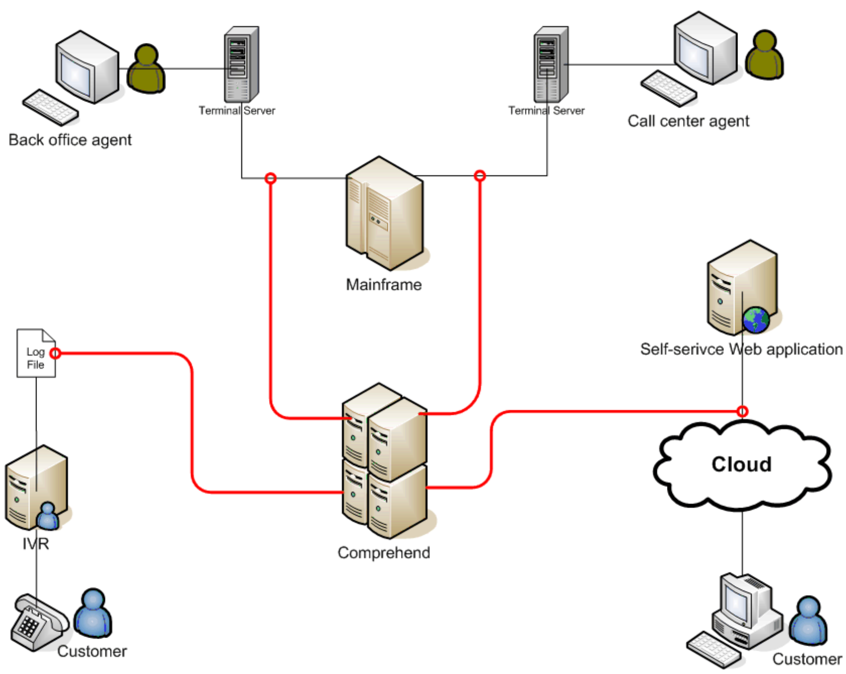


Figure 5.2 - Technical Environment

Comprehend Analytics Cluster was then used to configure the set of application screens used to support the ½ dozen processes both in back office and the call center. This configuration and validation for the processes took the rest of the first week.

The Process Intelligence Cluster was used to analyze the business processes to explore process variations. This analysis of the processes was an iterative analysis through the course of the second week.

5.1.9 Example Results

Comprehend Process Intelligence Cluster provides an interactive analytical workflow that allows for the discovery of root causes of process variations. Through the analysis of the six processes, a set of specific issues were uncovered. *Comprehend* provided the insights into the issues and root causes, as well as the metrics quantifying the value of addressing each issue.

5.1.9.1 Speed

Speed improvements were specifically targeted at finding productivity improvements in the back office and call center operations. By examining the day to day processes of the agents in these environments, an average of a 20% improvement in productivity was discovered across the six processes.

5.1.9.2 Account Balance Link (Figure 5.3)

Customer service agents were examined to understand a number of behavior patterns. A great deal of focus was placed on understanding the link between different service requests placed on the agents, including handoffs, process variations and typical call structures.

While the overall process of a call in general is structured solely based on the customers' requests, the process is always initiated with an authentication of the customer. Upon examining the process, immediately an interesting finding was made: 88% of all customer calls proceeded directly from authentication to a current account balance. The call center system required several 'clicks' for the agent to pull up a current account balance for the customer, which required on average 12 seconds.

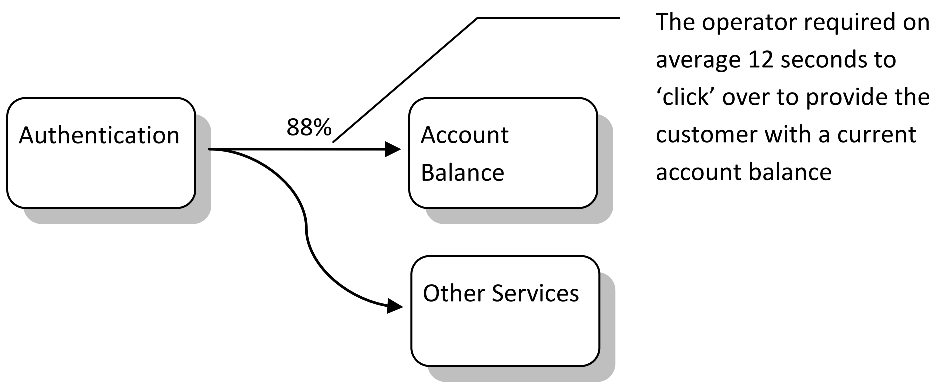


Figure 5.3 - Account Balance Link

5.1.10 Benefit

Working with the application team supporting the call center application, a detailed ROI was constructed based on the costs (application change, agent re-training, etc.) as well as the opportunity (*Comprehend* provided data on the frequency, and time required). The result was a measurable improvement in call times by immediately providing account balance at the end of the authentication process. The result was an overall 8% decrease in call time, which provided significant staffing reductions in the call center.

5.1.10.1 Agent Improvement (Figure 5.4)

Comprehend provided detailed by Agent by Service productivity data. This allowed for the analysis of Agent performance at a level of detail that was previously impossible. This analysis looked at the average time to process service requests by Agent. The results of this analysis found two immediate areas for improvement:

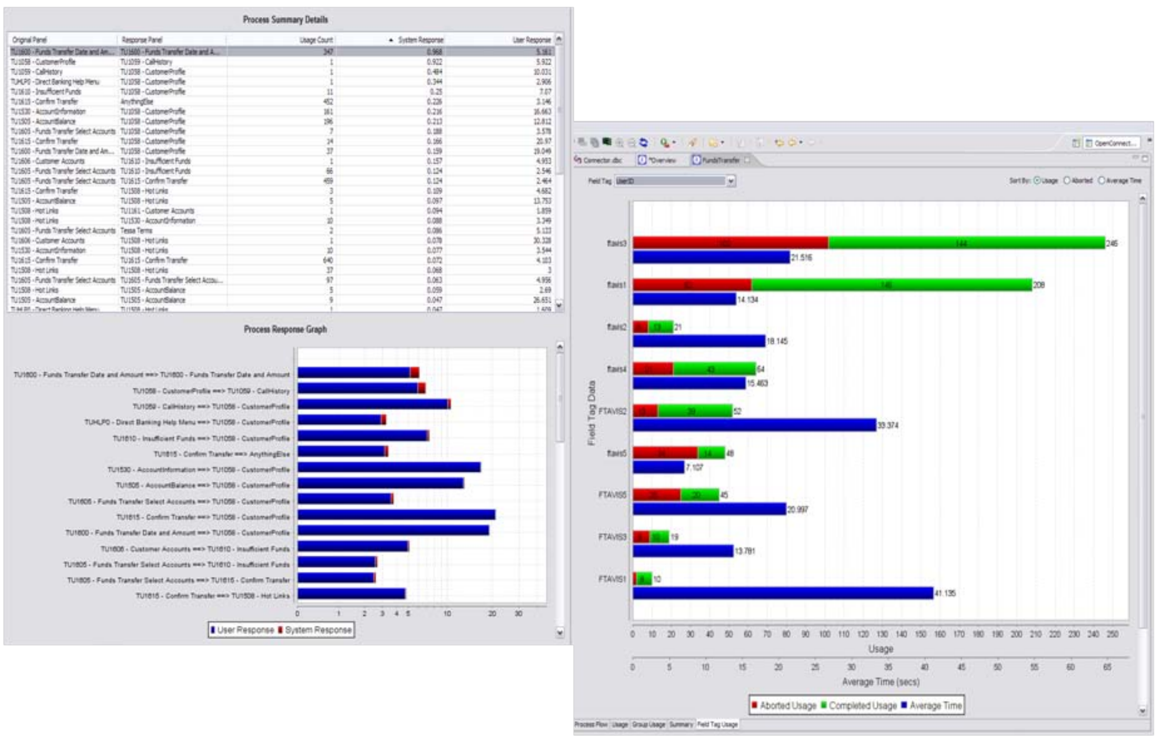


Figure 5.4 - Agent Improvement

1. One set of agents had acceptable productivity on the majority of the services they provided, however, were found struggling on several services. For these agents, a targeted training program was put in place, where with the identification through *Comprehend*, the agents were pulled out of the line, and provided a short targeted training session on the specific services of issue.
2. Another set of agents were found to have lower than acceptable productivity over the entire range of services offered. These agents were placed into a broader employee improvement program that included evaluation to overall fit with the job function.

Implementing the targeted agent improvement program resulted in an overall reduction in average call time of 10%, creating a significant reduction in agent staffing and resulting cost reduction.

5.1.10.2 Quality

Quality was defined on a per-service basis, but generally represented the goal of first contact resolution. While conceptually simple, creating a metric for quality was beyond the capability for this client. For example, consider measuring first contact resolution of a change of address request. The request could be originated via a letter/fax sent into back office operations. However, the customer, seeing either no results or erroneous results may subsequently repeat the attempt via any channel, back-office again or a self-service or call center channel.

5.1.10.3 Change of Address (Figure 5.5)

The change of address service was offered via every channel, and the concern for quality was that the scenario where the customer perceived the service to have been rendered and over the course of time realizes that the address change was not performed correctly.

Comprehend captured every change of address event across all channels and provided the analysis of those events where the same customer had executed a change of address request within two weeks.

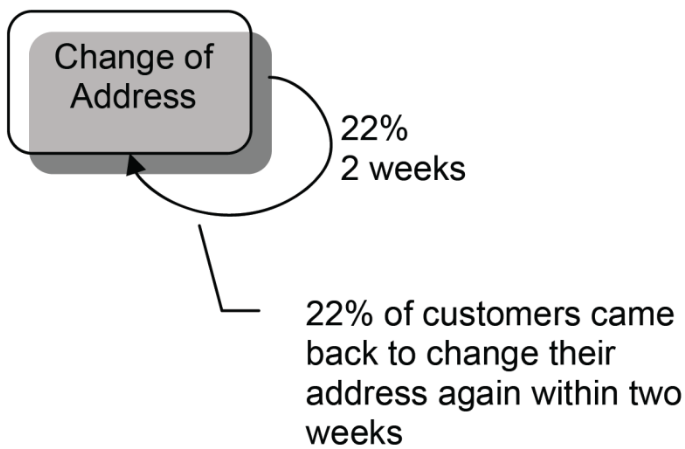


Figure 5.5

Comprehend indicated that first contact resolution for change of address was 78%, providing the metric for the first time. More importantly, *Comprehend* provided the analysis of the underlying events to gain insights into the root causes of the change of address failure. By examining the events related to the failed change of address requests it was discovered that 5-line addresses were disproportionally represented. Drilling deeper into the process, it was found that the entry of 5-line addresses into the underlying systems was a confusing, poorly documented activity.

By simply adding some better documentation to the systems and brief training to the operations staff, quality was improved over 10% resulting in a 6% reduction in change of address work, and a measurable improvement in customer satisfaction.

5.1.10.4 Self-Service Site Improvement

With over one million unique customer logins to their self-service site a day, this organization knew through survey research that 60% of their over 100,000 calls a day were from customers that were attempting self-service before their call. Having 60,000 customers a day failing on self-service was not only keeping customer service agents busy, but also impacting customer satisfaction. Surveys showed that every failure led to a 10% or better drop in customer satisfaction.

Comprehend connected the dots between the customers' web experience, the ensuing voice recognition session and call center service to provide invaluable insights into what the customers were really trying to accomplish on the self-service site. *Comprehend* provided the detail on the exact services those customers, who were just on the web, requested of the call center. This analysis provided the a segmented view of the web experience for just those customers that requested a specific service from the call center after being on the web. A detailed analysis of the click stream behavior of this segment revealed significant improvements to the online experience to improve the first contact resolution rate.

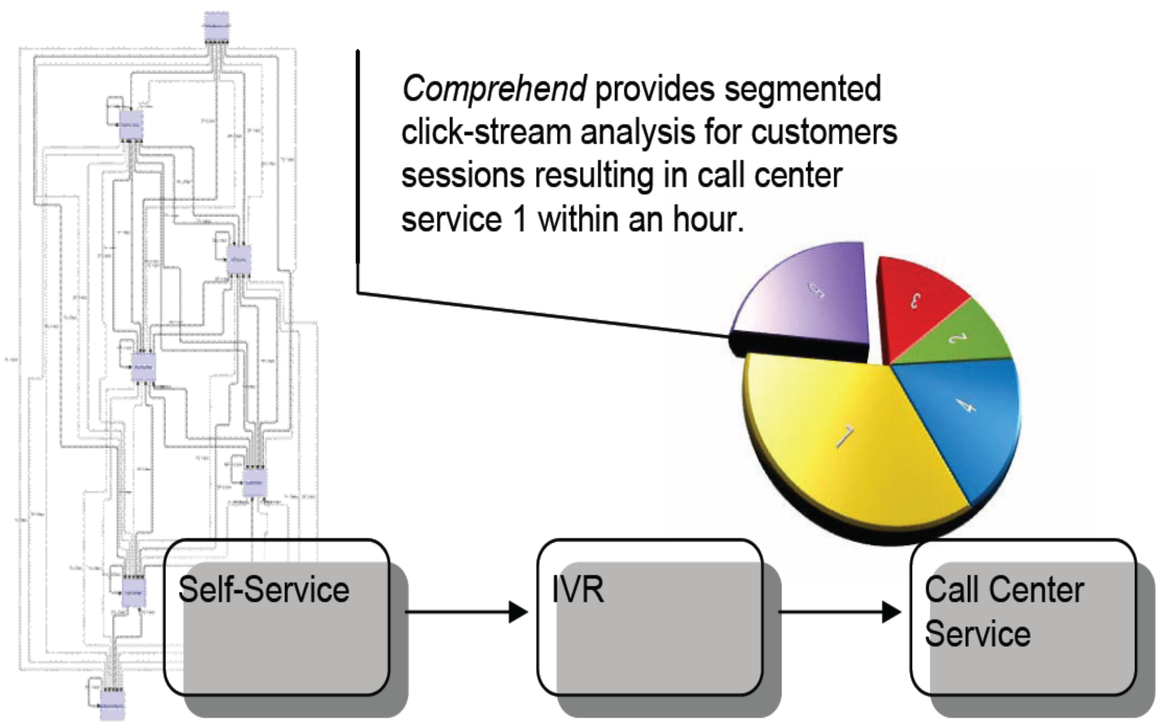


Figure 5.6 - Self-Service Site Analysis

Within the first month of analysis, several significant challenge areas were identified, and improvements were designed and tested, resulting in a 5% reduction in the calls for those services originating from online customers.

5.1.11 Revenue

Revenue improvement was examined from two perspectives, first, customer churn and second, cross-sell opportunities. Customer attrition reduction provides not only revenue increases (due to continued account profitability), but also cost savings when considering the cost of customer acquisition. Marketing additional account services through cross-sell opportunities provides an obvious revenue lift of additional services.

5.1.11.1 Leading Indicator of Churn

Through the course of the analysis, the effects of improved speed and quality were expected to have a measurable improvement in customer satisfaction, and thus a carry-on affect on customer churn. However, a more direct analysis of customer churn was also provided to discern patterns of customer behavior preceding the closing of the account.

Comprehend provided the ability to filter customer activity across channels to only look at the interactions of customers that had closed their accounts. By comparing process deviations between the closed account customers and the general population, interesting observations on precursors to account closing were discovered. For example, one such perhaps obvious observation was that customers that ran their account to a zero balance for a billing cycle tended to close their accounts when calling in subsequently. In this case, customers that met this condition were routed immediately to a customer retention specialist instead of the general agent pool.

The results of this analysis provided several observations that led to a measurable reduction in customer churn, resulting in a net 3% decrease in closed accounts.

5.1.11.2 Cross-Sell (Figure 5.7)

Agents were all trained and measured on cross-selling activities to appropriate target groups of callers. *Comprehend* provided the details on the process followed by successful agents.

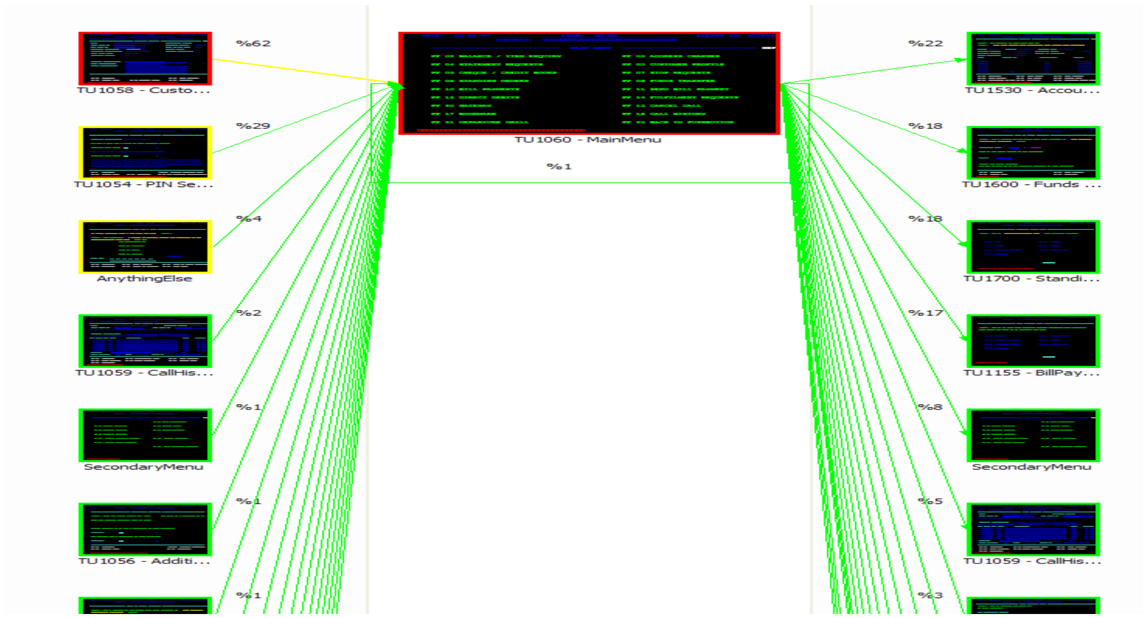


Figure 5.7 - Cross Sell Best Practice

Best practice was discerned from an analysis of the process details for the best performing agents. This combined with the training and improvement program for agents not following the best-practice policy resulted in a measurable 4% improvement in cross-sell affectivity.

5.1.12 Justification and Rollout

With over 9,000 individual processes across the organization, the implementation of *Comprehend* followed a roll-out program. This roll-out program involved a series of projects to leverage the technology against specific sets of processes based on an identification of which processes might leverage the best returns. This provided a roll-out plan where there were immediate returns for the analysis of these processes, and an ongoing extension of the analysis into new process areas and a continued acceleration of the benefit as the scope of deployment grew.

5.1.13 Value Summary

Comprehend, found business process improvement opportunities with an improvement value of more than £9MM per year by analyzing core processes across back office and call center operations.

The main benefit streams were found in quality improvement and back office telephony process improvement.

- Quality Improvements (RFT): *Comprehend* provided the data required to increase Right First Time (RFT) from 75% to 85% (worst case) or 90% (best case) yielding a NPV £12.0M to £15.9M of benefit from reduced re-work labor costs over a 4 year period.

- <u>Cycle Time (Speed)</u>: *Comprehend* provided the detailed steps performed by every operator for every process yielding the insights to improve process end-end times by 10-25% yielding a NPV £8.2M to £16.6M over a 4 year period.

As a result of a reduction in errors and an improved understanding of the processing work, it can be assumed that there are additional opportunities for improvement in customer attrition.

- <u>Customer Attrition</u>: Identification of individual customer activity, highlighting preferences and allowing for the tailoring of service provision and resource allocation to suit actual customer behavior yielding a reduction in attrition from 1-2% providing a benefit of £4.7M to £9.4M in MTA customer retention over a 4 year period.

In addition to the hard benefits identified during the Proof of Concept, *Comprehend* provided additional soft benefits associated with:

- <u>Staff Analysis:</u> Detailed monitoring of operator activity showing actual transaction completion times per staff member, allowing identification of best practice and under-performance, highlighting targeted training opportunities.

Value Leak	**Type of Imperfection**	**Recapture Action**	**Estimated Savings**
Improve Quality	Asymmetry of Information	Reduced re-work costs	£12.0MM to £15.9MM
Cycle Time	Asymmetry of Information	Reduced processing time	£8.2MM to £16MM
Reduce Customer Attrition	Asymmetry of Information	Increased customer retention	£4.7MM to £9.4MM
Staff Analysis	Asymmetry of Information	TBD – Identify and train Staff on best practices	TBD

Table 5.1 - Value Summary

5.2 Customer Contact Center Case Two: Municipality Contact Center

The municipality audit service of a major Dutch city was looking for a new approach to and the development of new techniques for auditing the effectiveness of their supporting processes, in particular for their call center processes. Severe need for cost reduction was the principal driver for this case study. Standard audits in the past had not identified any hard root causes of noncompliance or major possibilities for cost reduction as they started auditing bases on the a priori existing process- and quality system. Previous benchmarking on costing and quality (being done by means of a quick scan comparison with other municipalities) had indicated that in terms of quality, this call center was compliant and the quick scan had showed that in this call center cost issues were below average.

The municipality contact center emerged about one decade ago. It was majorly a “self-made” organization that never invested heavily in dedicated call center tools, so most of the tooling was in-house development. They acquired, at some point in time, a partial license for professional call center software, which we call for the sake of this paper ContactRight. The contact center employed a fixed number of contact agents in-house but also used a variable workforce, which could, by teleworking, be deployed in peak moments. One such moment is typically when municipality tax declarations are sent out and a lot of questions emerge from those declarations.

A particular issue not to be neglected in this case study is the fact that there was a policy in place for the preferred employment of handicapped people. As such, the contact center had also a social employment function in the municipality.

The yearly budget of the contact center was around 9.5 million EUR, and one of the purposes of the study was to find “smart savings” in this budget, preferably without reducing full-time employees for reasons just described. This case study was performed in collaboration with DENION, an innovative network structured company providing organizational advice to mainly governmental customers.

5.2.1 Customer Value in Contact Centers

As in the other case studies, the primary question remains: what constituted the *perceived value* of a contact center for the customers of that type of support and guidance interacting service.

It is well-known in the literature that how most contact centers optimize on *customer contact handling time*, which becomes the primary parameter to evaluate the efficiency of the contact agents. The basic idea seems (again): *the faster the customer is answered the less costly the customer handling*. Very often the length of the customer contact is subsequently translated in the “cost per contact”, which is considered as a number 2 performance indicator…

As pointed out previously, in many processes, the first-time pass rate is a primary process performance indicator. The same turns out to be true for contact centers. As [Zbikowski 2007] found, the First Customer Contact Resolution Rate (which is in this case the equivalent of the first time pass rate) correlates strongly with the customer satisfaction. He reports the case where a call center was struggling with low levels of customer satisfaction until they found out that the percentage of contacts resolved on initial contact (irrespective whether the customer got the final answer or not. But, in case of no answer, it became clear to the customer why there were no direct answer) with the customers was low, about 70%. The organization initiated a number of initiatives to improve the knowledge level and handling skills of the customer contact agents, which resulted in a substantial improvement in both first contact resolution rate as well as customer satisfaction. The following shows the strong correlation between first customer contact resolution rate and customer satisfaction:

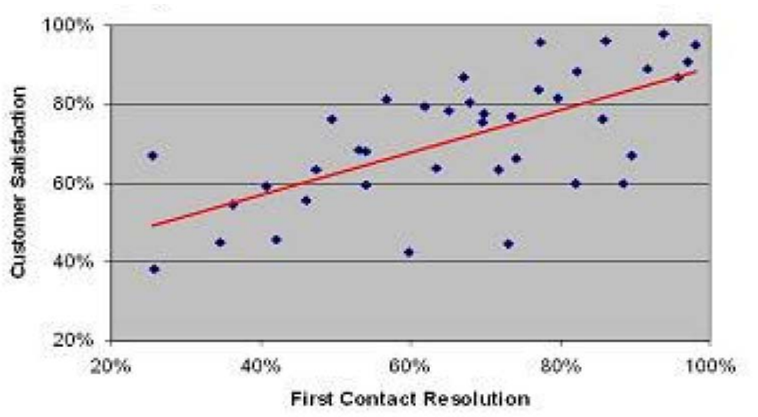


Figure 5.9 - First Contact Resolution

Observing this strong correlation, it is clear that detecting and understanding the value leaks in the first contact resolution rate are crucial in improving customer contact center operations.

The check-list of *muda* that was developed earlier in this thesis was used again to provide guidance for potential value leaks. Perceived customer value is eroded whenever:

- The customer experiences (for him apparent unnecessary) *repetitions*, such as “Can you call back later and try to contact this person…?”.

- The customer experiences unnecessary “lost” waiting times (before the call really starts, or during the call). Typical waiting times that are invisible (and hence difficult to understand for the customer) are *search time* on behalf of the customer agent, or *access time to ICT-services* that might be needed to answer some questions (for example, if a citizen has questions out his/her tax calculation, the customer agent might need to have access to the tax sheet in the municipality tax handling information system).

- *Connect through* times, when the contact agent is trying to get in touch with second line support, for example for more specialized questions.

- *Unnecessary steps* are present in the process. Agents know when this happens: when the customer asks the agent “why do you ask this?”; it is for them an indication that – at least from a customer perspective – this seems like unnecessary bureaucracy. Although for some questions (for example, tax related) it is evident that the customer must provide a proof of identity. Such a proof however is overkill for simple questions (such as asking for the opening hours of some municipality front office desks).

- *Lengthy explanation times.* Whereby it seems, for the customer, as if the agent does not well understand the answer to the question, and – in the worst case – is just reading some piece of text found in a Frequently Asked Questions list…

- Agents are *overstressed*. A typical norm for this is 80%: whenever contact agents have a higher utilization, they become nervous, less friendly, more snobbish, less caring, and so on, as a consequence of the higher utilization. Clearly, above 80% utilization, contact agents show underperformance.

  The utilization is calculated as follows:

  ((Avg #calls handled) x (Avg call handling time length)) / (Available Agent Time)

  It is obvious that shortening contact handling times is ONE way to reduce the agent utilization, but then the length of the contact is not the norm: the agent utilization is the norm!

Seeing this, it is clear that the above list gives a number of guidance points in auditing a customer contact handling center. Process discovery was used as the methodology to perform this audit.

5.2.2 Understanding the current process

As mentioned before, the ICT-infrastructure for this customer contact center was rather limited. Basically, a Windows-front-end application was in-house developed in VISUAL BASIC, with the following facilities:

- Registering the different steps in a customer contact, although without time information, except for the beginning and the end. This was needed because the contact center used a cost recovery schema that was based on the contact cost, which was simply:

  (contact time in seconds) x (rate per contact second)

Hence this information was needed for the monthly invoicing.

- Giving access to a FAQ (Frequently Asked/Answered Questions) database. This was merely a text database, with simple keyword searches. Boolean combinations and sub-searches are not available (although under development, since this was an in-house system). The content of the FAQ database was also copied regularly to the municipality web portal: they were published as text pages. (*The smart reader sees already something very ironical: once published on the Web, using Google as a search engine provides already the more advanced search facilities that could be needed…).*

- Provide access to the ICT-systems that give specialized support (for example: reservation systems for permissions, appointments… tax information system; etc…).

- Provide the possibility to switch the contact session to second line support, having a more experienced contact agent in case of fairly deep questions (for example, environmental pollution issues).

The system had basic registrations of all these activities, however, without detailed timing information on the sub steps.

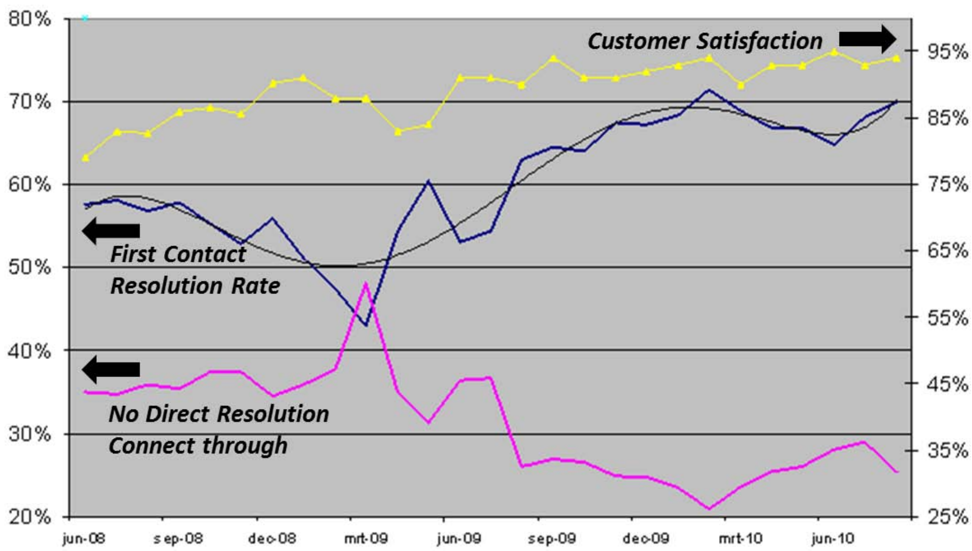


Figure 5.8 - First Contact Resolution Rate

This picture was taken from the municipality web portal. It showed, in a report to the municipal board, the contact center performance. The yellow line is the customer satisfaction, which is actually very good, and increasing. Nevertheless the First Contact Resolution Rate is relatively low, with (in some months) a high level of connect through and recall actions. *It is clear that this aspect requires further investigation. It turns out later in this chapter that the first contact handling rate was underestimated in this contact center, due to a misunderstanding of the actual process model.*

Apart from the front-end system, there was a manual, describing the “standard” customer handling process. Ignoring the details (for reasons that become clear in the next paragraphs: the real process turned out to be very different) this was the “official” Process model, as it was designed in VISIO.

At first sight, this process seemed to be clear and simple, perhaps too simple. So, before transforming this representation into a Hidden Markov Model and doing the quantitative analysis in looking for value leaks, a "manual" reality check was needed.

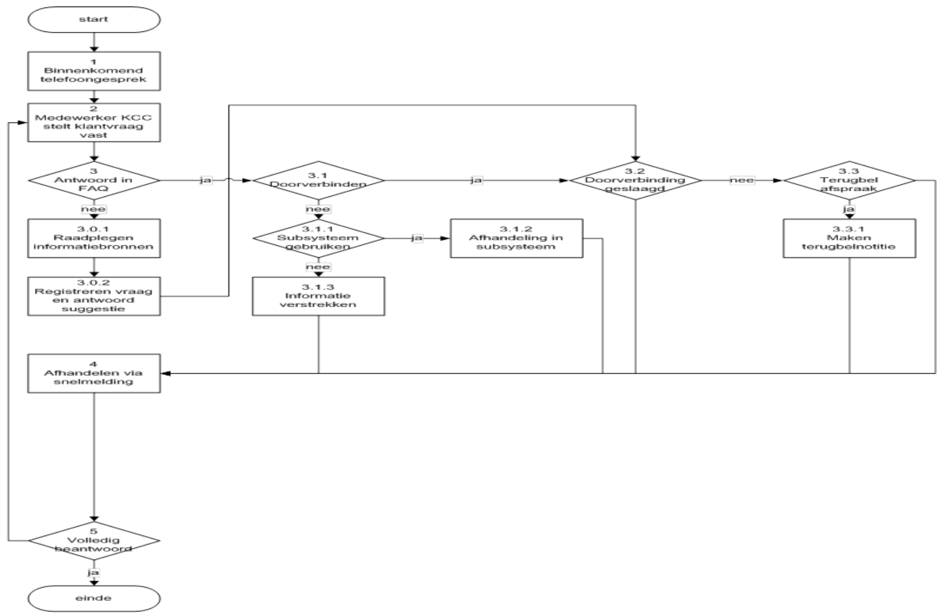


Figure 5.9 - The a priori assumed Process Model

This manual process audit was done by means of interviews with the following "stakeholders" in the process:

- A representative of the "experienced", say "senior" contact agents
- A representative of "junior" contact agents, typically "in training"
- A representative of the municipality cost administration, being the main "client" of this contact center, and being responsible for the cost handling
- The contact center manager

All of them were confronted with the above process drawing, to validate how close it actually comes to the reality on the work floor of this contact center. The interviews were not monitored by the contact center manager, who allowed the contact agents to speak freely about their perceptions of the contact handling process, and their frustrations. Just as in the case with Taylor, we made it very clear to the contact agents that our main goal was not to "control" them, but rather to understand what was really happening in the contact processes.

The major points discussed in the interviews were the following:

- The actual process on the work floor was reviewed, step by step, to understand the mapping to the assumed process model.

- There is a fairly high turnover of young, inexperienced contact agents. This puts double pressure on the senior agents: they are constantly challenged for the advanced questions (which in reality are sometimes not so advanced) and they are supposed to coach and train the junior contact agents.

- The access to the supporting information systems in various process steps was discussed in detail.

- The experiences with second line support (back-offices).

- The Service Levels as experienced on the work floor.

- The availability of measurement data on the Contact Process steps, and the collaboration to “time & motion” experiments (which actually turned out to be necessary in this case).

Another important category of potential value leaks in processes is related to availability and capacity. In the case of a customer contact center, the standard model to calculate the required capacity is the Erlang-C model. For this highly non-linear model, several open source calculators are available, such as:

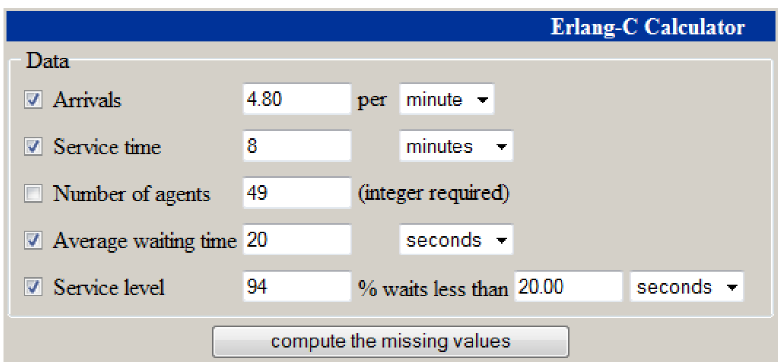


Figure 5.9 - Erlang-C Calculator

This model enables one to calculate the required number of agents based on the arrival rate of the contact requests, the average session length (which is here the service time) and the average waiting time before some free agent starts the contact and the service level for this waiting time. The demand for this contact center was 5 contacts per minute. The assumed customer contact time for capacity calculations was 5 minutes, with at time, another 10% final contact handling time (administrational finalization of the contact).

It is obvious that the time interval during which this model is used is quite crucial for the accuracy of the capacity. As in most services, excess capacity is a significant value leak (Ng et al., 1999). At the time of the study, the municipality administration was assisting the contact center by providing call forecasts on a **weekly** basis. A typical example: the municipality administration knows exactly when they send out the tax bills for the citizens, and translate this into an appropriate forecast for the contact center. *The question remains whether the weekly frequency was the most appropriate forecasting interval for this particular type of contact center?*

Another important element in the understanding of the Customer Contact process is the impact of the supporting ICT-systems. The following list shows the various ICT-components:

- The frontend-application is an in-house developed Windows/VisualBasic application, which allows the customer contact agents to:
    - Register the customer contact (including the citizen identification)
    - Search the Frequently-Asked-Question (FAQ) backend database (which is a ContactRight systems application)
    - Connect-through to the (mostly in-house developed) Municipality Backend systems
    - Connect-through to second line support
    - Formally close the customer contact (to register the handling time and details)
- A FAQ-database, which is automated in a specific contact center software technology, ContactRight. This software product has the capabilities to support most of the functionality for a contact center. For budget reasons, and because of the non-trivial connectivity to the other Backend systems, ContactRight was not chosen as an integral contact center system. Of course, the ContactRight system logs all activity on the FAQ-database. It should be noted that the FAQ-database was hierarchically organized by Municipality Back-Office service, to facilitate the connect-through.
- The various Municipality backend systems (such as tax administration, parking allowance and permit systems, environmental issue enforcing systems, and so on…), several of them realized in ORACLE.

From the above description it becomes clear that measurement data on the various customer contact process steps are scattered over several systems, and for many steps there were simply NO automated measurement data. There is no systematic logging of the contact events; only aggregated information in view of enforcing the service levels was available.

To apply Process Discovery for this case, there turned out to be no alternative but to utilize a time & motion series of experiments. First of all, it was explained to the contact agents that the aim of the experiments was not to “control” them, but to gain a better understanding of the actual process. A mixture of senior as well as junior agents would be measured and the team was urged to “act as normal” during the experiments. This manual “time and motion” logging (using stopwatches and the maps of the “official” process) was done for one week and resulted in several hundreds of logged customer contacts. This is the basis for this Process Discovery. Of course during the manual logging, attention was given to remain conformant to the Event-Object-Actor format in the recorded data.

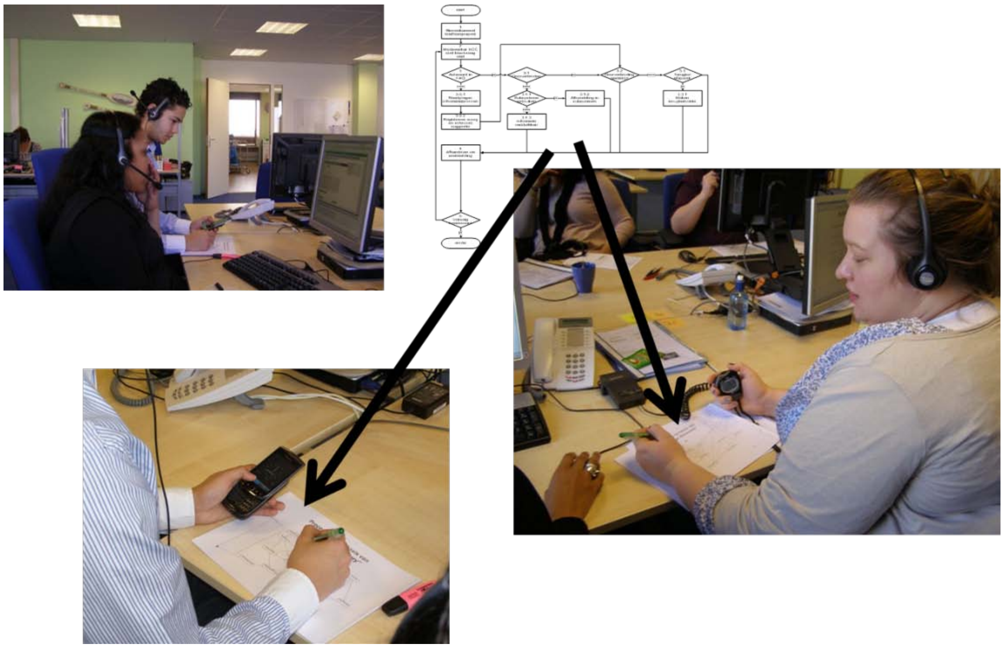


Figure 5.10 – Call Center Workflow

5.2.3 Process Discovery of the actual contact process

The manually obtained logging data were the starting point for discovering the actual Hidden Markov Model (HMM) for the Customer Contact process and the search for Value Leaks in the process. This was basically done by feeding the loggings records from the Time & Motion experiments as XML input records for the Process Discovery algorithms. The observations in the Time & Motion experiments made it evident that the actual process was *really different* from the assumed, documented process.

First of all, a new step was introduced at the beginning of the process: the verification of the customer identification. This step was not formally documented, but was introduced to be able later to have a better understanding of the types of questions typically asked by some types of customers. Customers are typically citizens or local enterprises, but could also be visitors, inhabitants of other municipalities or even tourists. As will become clear later, this mandatory "customer identification" was experienced as extremely tedious and unnecessary. The irritation becomes so high that at some point a local newspaper article was published, in-which the journalist bundled all customer complaints. Of course, this was another clear signal to the Municipality Board members that the customer contact process had some problems…

Another undocumented step was "immediate connect-through" which basically means a highly specialized question is immediately connected through to another experienced senior agent, who is supposed to handle the question without any access to systems.

Furthermore, it turned out that several decision points were actually not documented in the process model, such as the conditions for connecting through as well as the circumstances under which the customer is asked to call back (and for which a call-back ticket is issued). This is the contact process model recovered from the interviews and the Time & Motion experiments:

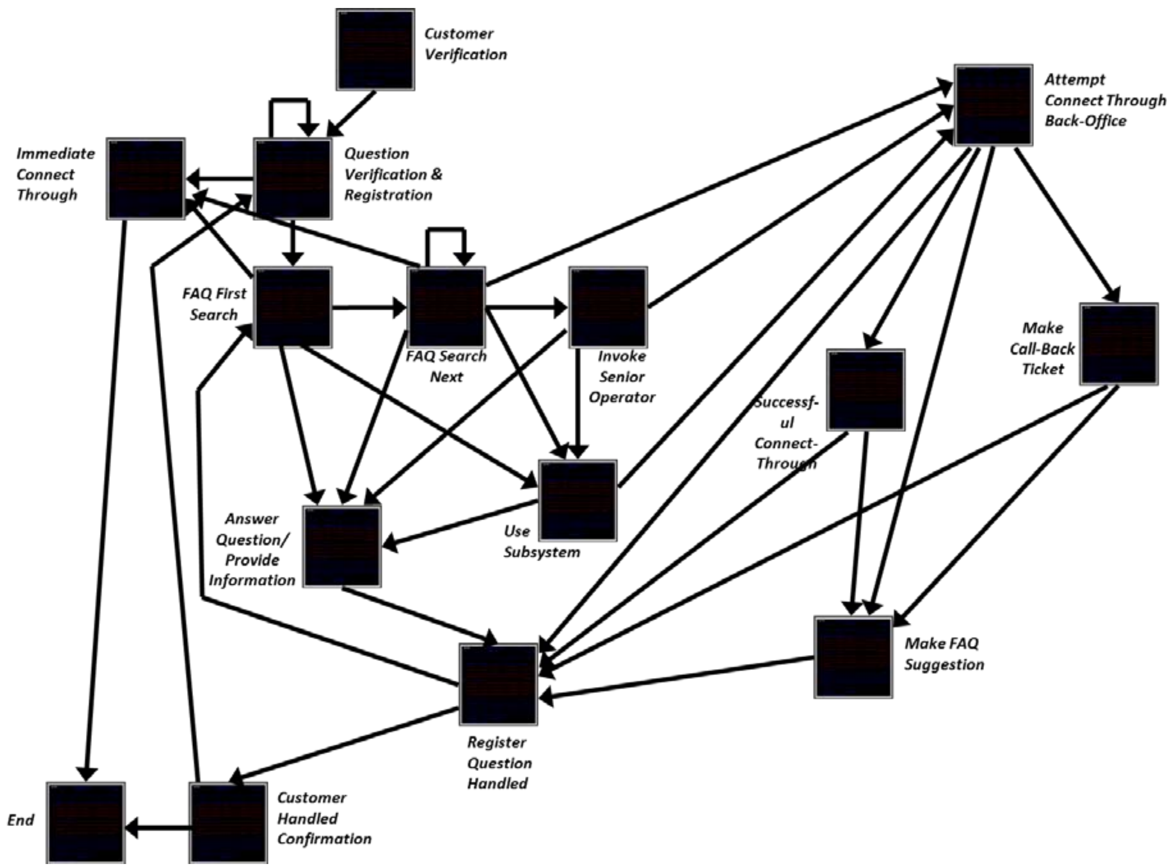


Figure 5.11 - The actual discovered Contact Process Model

In short, this was the story that the agents described:

1) The first thing that must be done when a call arrives is the verification of the customer, and the registration/verification of the question asked by the customer. In fact, the agent is not only obliged to register the question, but also to repeat the question to the customer for verification.

2) Next the agent searches the FAQ database. Possible answers in the FAQ may directly provide information for the answer, or the link to a subsystem or the possibility to connect-through to the back-office. There may be multiple searches and the agent may invoke a senior (experienced) agent.

3) The ultimate goal is to answer the question and provide the right information, sometimes by invoking subsystems or connect-through to a back-office.

4) After providing the information, the agent must register that the question was handled (which is just pressing a “confirmation” button in the front-end software program). Next, this the agent must formally ask the customer if the question was indeed sufficiently handled from the customer perspective. If not, the question is refined as a new question, and the process basically starts over again.

5) When a customer is connected-through, the connect-through may be successful or not successful. In the case of not successful, the customer may request that a call-back ticket is registered or the customer may judge that this is not needed. The agent has in the front-end application a quick editor to make a FAQ suggestion (suggesting to the back-office to add an appropriate FAQ entry). Even if the connect-through is successful (the back-office is able to take the call) the agent must register the handling and customer confirmation…In principle the customer can still choose to omit the connect-through and formulate another question.

The transition probabilities as well as the durations of the process steps were separated for “average” contact agents and new, inexperienced agents. This is the HMM for the average agent.

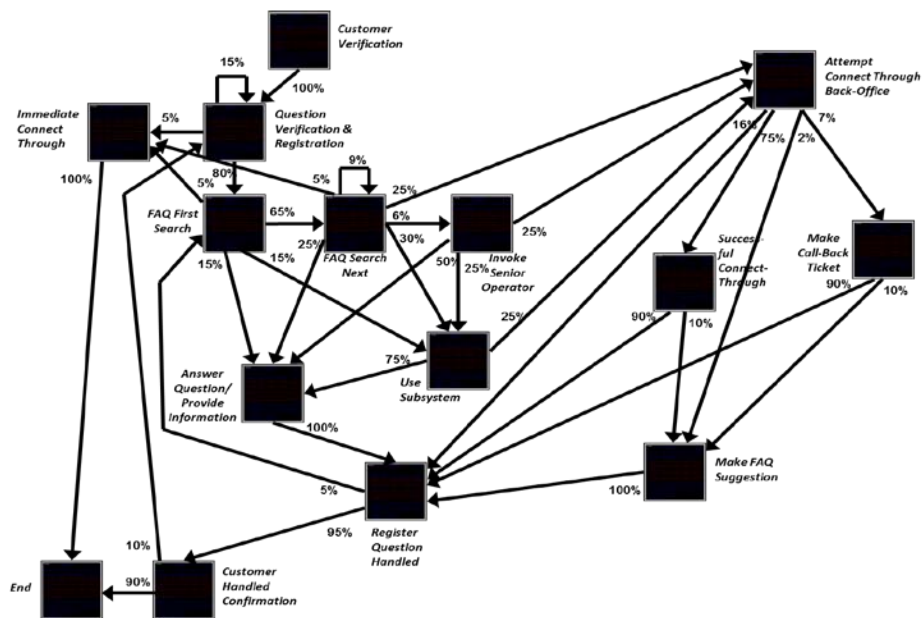


Figure 5.12 - The process model as an HMM (including the probabilities for an average contact agent)

Observe how the discovered first-order HMM contains many un-normalized places (busy many-to-many transitions). As will become clear further on in the results, most of these places contain value leaks in the contact process.

The standard calculation of the steady-state probabilities results in the following average number of “visits” to each of the processing steps. The Service Times for each processing step were derived from the Time & Motion experiments. Multiplication of the number of visits with the service times results in the average time demanded for each processing step. Adding up all these times gives the total session length for a customer contact. This allows a calibration of the model with the observed measurements.

These are the detailed results for an “average” contact agent (and the summary for a junior agent):

Customer Contact Processing Step	# Visits	Service Time	Service Demand	Service Demand
		Average contact agent		Junior Agent
Customer Verification	1.000	41	41.0	41.0
Question Verification & Registration	1.287	23	29.6	(*)48.3
Immediate Connect-through	0.157	30	4.7	5.4
FAQ First Search	1.079	16	17.3	17.0
FAQ Search Next	0.770	15	11.6	(*)24.9
Answer Question/Provide Information	0.681	90	61.3	58.5
Use Subsystem	0.404	145	58.7	54.0
Invoke Senior Operator	0.046	75	3.5	(*)9.0
Attempt Connect-through Back-Office	0.305	68	20.8	21.0
Successful Connect-through	0.229	41	9.4	9.5
Make Call-back Ticket	0.021	197	4.2	4.3
Make FAQ suggestion	0.031	15	0.5	0.5
Register Question Handled	0.986	11	10.8	10.5
Customer Handled Confirmation	0.937	6	5.6	(*)7.3
End	1.000	0	0.0	0.0
TOTAL SESSION LENGTH			278.8	311.1

Table 5.2 - Average Contact Agent Session Results

Similar calculations for junior, less experienced agents result in an average session length of 311.1 seconds. The process steps where junior agents typically spent more time are indicated by (*).
These measurements are fairly consistent with the 5 minutes average session time that had been measured by the Municipality services.

It is already obvious from the calculations that the most time consuming steps in the process are:

- The verification of the Customer: 41 seconds
- The registration of the Question
- The providence of Information: 60+ seconds
- The use of the subsystems: 54+ seconds
- The FAQ searching: 30+ seconds
- The connect-through: 20+ seconds

In fact, in each of these steps, value leaks will be identified.

5.2.4 Value Leaks in the discovered contact process

The discovered process is only the initial departure for the quest for value leaks in the process, whereby un-normalized places in the Markov model as well as excessive (*muda*) time are both pointers to potential value leaks. Let it be clear, again that the aim is not to reduce the contact time as such; the purpose is to look for *muda*, which represents meaningless loss of time, in the customer contact session duration.

d) The customer verification

The very first step in the discovered process is the mandatory verification of the customer, accounting for 41 seconds session time. Requiring a citizen to disclose his/her civil ID before the contact agent can continue the session is a heavy requirement. A more reasonable assumption would be to require this verification for access to customer-sensitive information (such as financial or medical information). Asking this verification even for simple questions (such as opening times for front offices) seems overacting indeed.

Further investigation revealed that the intention of this customer verification was to develop more consistent customer history profiles and customer intelligence. However, at this moment, without any such facilities, this process step merely appears to be *muda.*

e) *Providing the information to the customer*

When the contact agent found a seemingly close question in the FAQ-database, the actual answer in presented as a plain text, which may be longer than one A4-page. One major reason for doing so is the fact that the same FAQ-database is used to feed the FAQ-section in the Municipality Portal.

However, it is well-known (Santos-Gomez, 1992) that alternative representations of plain text may significantly speed up the providing of information. In particular decision tables typically speed up information providing by a factor 2 to 3.

f) *Access to the subsystems*

The access to the Municipality subsystems involves two factors: logging-on to the systems, and the actual use of the systems. It turned out that there was no single log-in for the subsystems. For security reasons, sessions inactive for more than one hour were automatically canceled, which requires the contact agents to frequently log-in again to the subsystems, with different userid/password combinations for each system.

Furthermore there were no proper service level agreements for the subsystems, except regarding the availability of the systems. Days with major system unavailability could be exempted as exceptional days, which were not included in the service level measurements for the contact center. However, no service levels have been agreed on the actual usage of the subsystems, regarding response times, for example.

g) *Connect-through activities*

The municipality desired to have strict control on connecting through to second-line support. Ideally 70% of the contacts should be resolved without connecting through. In case the agents cannot find the answers themselves nor with the help of senior agents, they want to help the customer anyhow and try to connect through.

The way the municipality decided to control the connect-through facilities was by building in the connect-through possibilities only in specific questions in the FAQ database. This database was ordered in the first place according to the Back-offices (which were actually responsible for feeding the FAQ with the appropriate content).

So what happened in reality, if an agent was on a question for which there was no direct answer, he/she would go on and search for another FAQ entry which had the connect-through facility included. It is clear that as a result, any logical connection between the questions and the FAQ entries in practice was lost, and Business Intelligence is meaningless on this customer contact database.

h) *Search facilities*

The FAQ-database was organized in ContactRight, which – being a contact center software – has advanced Boolean search facilities. For unclear reasons however, the front-end VisualBasic application did not take over those search facilities and provides only simple searches in the FAQ-database to the contact agents. Already in the interviews with the (senior) agents it became clear that this was a major frustration.

Ironically enough, since the FAQ database was also available on the Municipality portal, some experienced agents had parallel computer sessions where they also searched via standard search engines in the Municipality websites to find the answer. Of course, such sessions were not appropriately logged in the systems…

i) *Contact agent utilization*

In the contact value discussion it was mentioned how the contact agent utilization should not exceed 80% on a daily basis. Given the profile of the specific employment in this case study, the actual stress level is even lower.

Of course, everyone likes to be helped by the senior agents. To make the problem even worse, some Back-Offices “outsourced” their second-line support to the senior agents. As a result, our calculations showed that in peak-moments, the utilization of the agents was in general more than 85% and for the senior agents higher than 95%. This partially explains the turnover of agents in this center…

j) *Lack of knowledge transfer to the junior agents*

Related to the previous Value Leak is the next one: since the experienced senior agents are over-utilized and do not have sufficient time to train the junior agents, transfer best practices, tips and tricks, and so on. Additionally, they are discouraged to do so, because of the turnover of the junior agents.

k) *Performance of the Back-Offices*

Second-line support is costly because it goes directly into the Back-Offices. Nevertheless, for specific questions there is no alternative, bearing in mind that the first contact resolution rate is the primary contributor to customer satisfaction.

Unfortunately, just as in the case of the subsystems, no service level agreements had been obtained with the Back-Offices. Ironically, in the absence of such agreements, they could optimize sometimes their costs by not being available.

l) *Imprecise understanding of customer contact handling time*

In the discovered process there are two steps at the end of the process before actually potentially ending the process. First of all, the customer contact session is “registered”, meaning a record with the customer contact length is effectively created (also including some other details, such as the access to subsystems). After this step, the agent has a formal step in which the customer must acknowledge that the contact session was effectively addressing his/her problem. If not, a new search in the FAQ is initiated and the process restarts almost entirely again, actually initiating a second contact session.

m) *Cost accounting versus Performance measurement*

Currently the duration of the contact sessions is used as a basis to charge the customer contact center costs to the Back Offices. They pay a rate per second in the session length. This has unwanted side-effects. First of all it puts pressure on the contact agents, especially for the administrational contact handling times. Longer times are discouraged or even penalized.

A better alternative would be to isolate cost accounting from process performance management, for example, by introducing Activity Based Costing with a fixed rate for the majority of the contact sessions.

n) *Granularity of the contact demand forecasting activity*

Changing the WEEKLY frequency into a HALF-DAY frequency can significantly increase the accuracy of the forecasts. It was found that in the weekly forecasting, the peak-to-average ratio could go up to 25%. This creates significant leaks due to overcapacity at some moments and lack of capacity (resulting in the over-utilization discussed earlier) at other moments. Some of these deviations are simply due to the fact that the incoming stream of customer calls is in reality not completely “random” but may contain bursting (grouped around events, such as sending out the municipality tax bills). However, the Erlang-C models assume random arrival streams…

The following table presents a classification of the types of Value Leaks in this case study:

Value Leak	Type
Customer Verification	Asymmetry of Information/Meaning (Customer History/Needed?)
Providing Information to the Customer	Information Incompleteness
Access to the Subsystems	Asymmetry of Meaning (Security versus Single Login)
Connect-through Activities	Asymmetry of Information
Search Facilities	Information Incompleteness
Contact Agent Utilization	Asymmetry of Meaning (Meaning of Agent Idle Time)
Lack of Knowledge Transfer to Junior Agents	Asymmetry of Information
Performance of the Back-Offices	Asymmetry of Information
Customer Handling Time	Asymmetry of Meaning (Customer Handling Time)
Cost Accounting vs. Performance Mgmt.	Asymmetry of Meaning (Conflicting goals)
Granularity of Forecasting	Asymmetry of Information

Table 5.3 - Value Leak Types

A final Asymmetry of Meaning was detected during the discussion of the results. Remember how the First Contact Resolution Rate was reported to be low, whereas the perceived user satisfaction was fairly high. The discovered process model provided some evidence for this anomaly:

a) In the reports to the Municipality, only the directly answered contacts, where the information was provided were considered as the First Contact Resolution Rate. The discovered Process Model shows a visit ratio of 0.681, 95% of which are considered as actually handling the question properly. This results in a rate of 64.7 %. This also configures to the numbers found the Municipality reports.

b) However, successful connect through sessions should also be considered as contribution to the First Contact Resolution Rate. In the discovered Process Model the visit ration for this is 0.229. Adding this to the above visit ratio and applying the 95% results in an actual First Contact Resolution Rate of 86.6 %, which is much more conformal to the standard Contact center benchmark data.

5.2.5 Estimated Value contribution of Process Discovery in this case study

Because a sufficiently precise process model was discovered, it was possible to simulate the effects of solving Value Leaks in this process. The experienced agents also gave their views on realistic improvement and of course they knew very well how frequently some Value Leaks actually occurred.

The following Value contributions could be estimated:

Value Leak in Contact Process	Estimated % reduction of Contact Handling Time after fixing the Value Leak	Estimated % increase to First Contact Handling Rate after fixing the Value Leak
Mandatory Customer Verification	2%	3% (e.g. anonymous questions)
Information Providing Time	1-4%	-
Performance Subsystems	1-5%	1%
Connect Through Limitations	3-6%	-
Availability Back-Offices	1-3%	2-3%
Improved Session Registration	1%	1%
Utilization of the Agents	-	1-3%
Granularity of the Forecast	-	-
Overall	9-21%	8-11%

Figure 5.4 - Estimate Value Contributions

Of course, the Municipality board wanted to see these numbers translated into tangible savings. A key element in the translation is the agent capacity utilization. The majority of the contact agents are full-time employee contracted by the contact center. To deal with variations in the call workload, the contact center can fall back to a network of part-time agents. These agents are only engaged on demand, and run their contact center activities by means of a teleworking infrastructure from home. Once this is understood, it becomes clear that the two major elements contributing to reductions in the Total Cost of Ownership are the following:

- The *improved granularity of the forecasting* leads to less excess capacity. Also the cases of shortage of capacity (with unnecessary high stress levels, mainly for senior agents) are better handled, resulting in a reduction of absence due to stress-related sickness.

- The *reduction of muda in the customer contact handling time* results in a lower agent utilization, actually allowing them to do slightly more work with the full-time workers, reducing the need for additional capacity.

All in all, ***the overall reduction in Total Cost of Ownership is around 15%***. As a result, after resolving the value leaks, ***this Contact Center can do the same work with a budget of 8 Million EUR*** (instead of 9.5 Million EUR), without firing any of the full-time contracted people. We think that this case study is an example of “smart saving”.

It is interesting to compare these results with some findings of McKinsey Research on applying Innovative ICT in contact centers (Pietraszek & Ramchandran, 2006). They found an average Return on Investment of 15% on investments in improved usage of ICT, with the following ranking:

Type of ICT Investment	Estimated ROI%
1. Portfolio management *Improvements to the Customer self-service facilities, such that less (trivial) calls come to the contact center*	37.5%
2. Call balancing *Improvements to have less second-line support Connect-through*	31.25%
3. Workforce management *Improvements to workforce forecasting and scheduling*	25%
4. Knowledge database *Improvements to the FAQ search and content management*	5%
5. ICT Architecture *Improvements to the Contact center software and its relations to specific supporting ICT-systems*	marginal
Total for 1 – 2 – 3	32%
Overall total	17%

Table 5.5 - Estimated Value Contributions

In this municipality contact center, the last 4 elements of this overview have clearly been addressed. In the McKinsey study the ROI for the last 4 elements is 13.4 %.

At the same time this case study reveals that perhaps even more value can be generated by focusing on a next element: Portfolio management. Portfolio management should optimize the “right” communication channel for the “right” service, including self-service facilities. This one is already partially enabled because the same content that is used for the FAQ-database is also used in the Municipality web portals. However, the current hierarchical organization of the information and the text-only representation of the information are the major inhibitors for a better portfolio management in the calls.

5.3 Clinical Pathways for Breast Cancer

To date, in a competitive health-care market, hospitals have to focus on ways to streamline their processes in order to deliver high quality care while at the same time reducing costs. Additionally, from the governmental side and the side of health insurance companies, more and more pressure is put on hospitals to work in the most efficient ways possible, whereas in the future, an increase in the demand for care is expected (Mans et al. 2008a).

Chou et al. (1999) declared that clinical pathways are being promoted as instruments that on the one hand can reduce costs and on the other hand can increase or maintain the quality-of-care and the optimization of health outcomes. These benefits derived by the use of clinical pathways, are recognized by numerous researchers (e.g., Lemmens et al. 2008; Skalkidis et al. 1996; Bryson et al. 1999). These researchers stated that the utilization of clinical pathways led to a reduction of the costs of patient care, to an increased nurse satisfaction, and to an improvement in the education of professionals.

Also Skinner et al. (2008) and Mans et al. (2008b) refer to the huge value that information and communication technologies can provide to the healthcare system. Skinner et al. (2008) declare that the solution to poor quality is not to increase the supply of physicians or specialists or hospital beds, but instead to improve health care systems and incentives to ensure that existing physicians and hospitals provide the best possible quality at the lowest cost. More specifically, they argue that one of the items that often lead to better health quality is the availability of information technology. Furthermore, Mans et al. (2008b) declares that "*nowadays it is of the utmost importance to evaluate existing infrastructures in Health-Care Organizations, and the services offered by these organizations*". The researchers state that therefore it is of crucial importance to explore and process the data collected by the Health-Care Organizations systems. These data can be process logs from a process management system, or databases from the electronic clinical chart system, or unstructured data, such as patient blogs.

5.3.1 Context for this process discovery case study

This case study was performed at a specialized Breast Cancer unit of an association of hospitals in Belgium, and is focused on Primary Operable Breast Cancer (POBC) patients. The number of patients for this pathology evolved as follows:

2002: 194
2003: 256
2004: 245
2005: 385
2006: 391
2007: 379

This care unit consists of a multidisciplinary team of 34 specialist doctors, 52 nurses and 14 paramedical staff members. The context for this case study is provided by the situation of the patients and the team in 2007, as reflected in (Verheyden et al., 2007). By using continuous improvement techniques, this team succeeded in improving dramatically the patient satisfaction over the years, as shown in the following results:

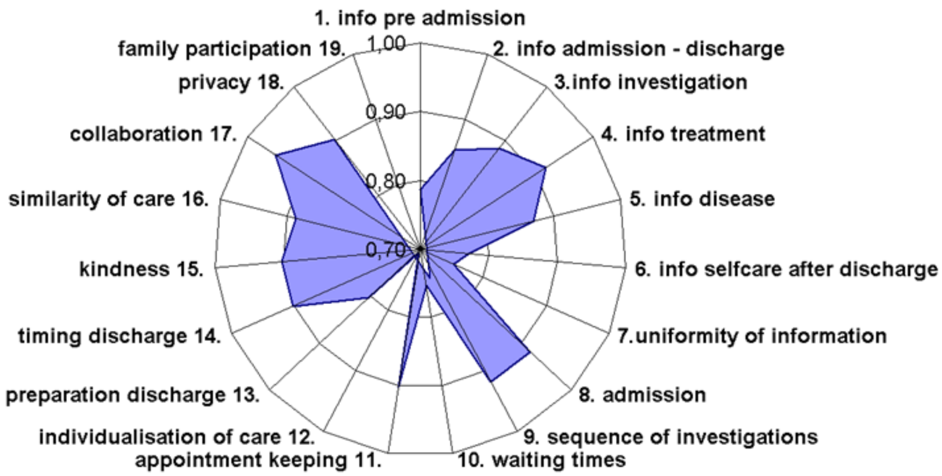


Figure 5.13- The Patient Satisfaction in 2003

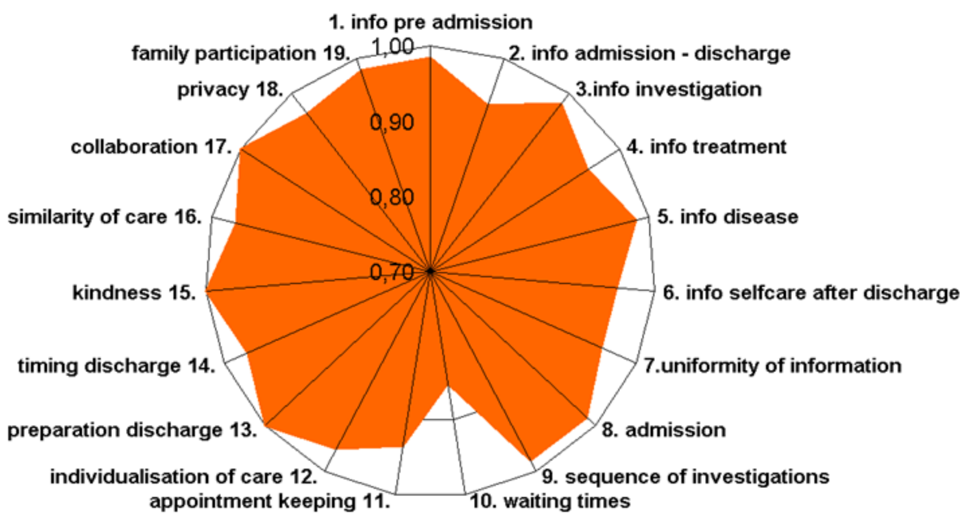


Figure 5.14 - The Patient Satisfaction in 2007

In the Belgian Healthcare system, hospitals are strongly encouraged to minimize the length of stay (LOS) in the hospital. It is one of major governmental Key Performance Indicators.

This is the evolution of the average LOS, measured in days:

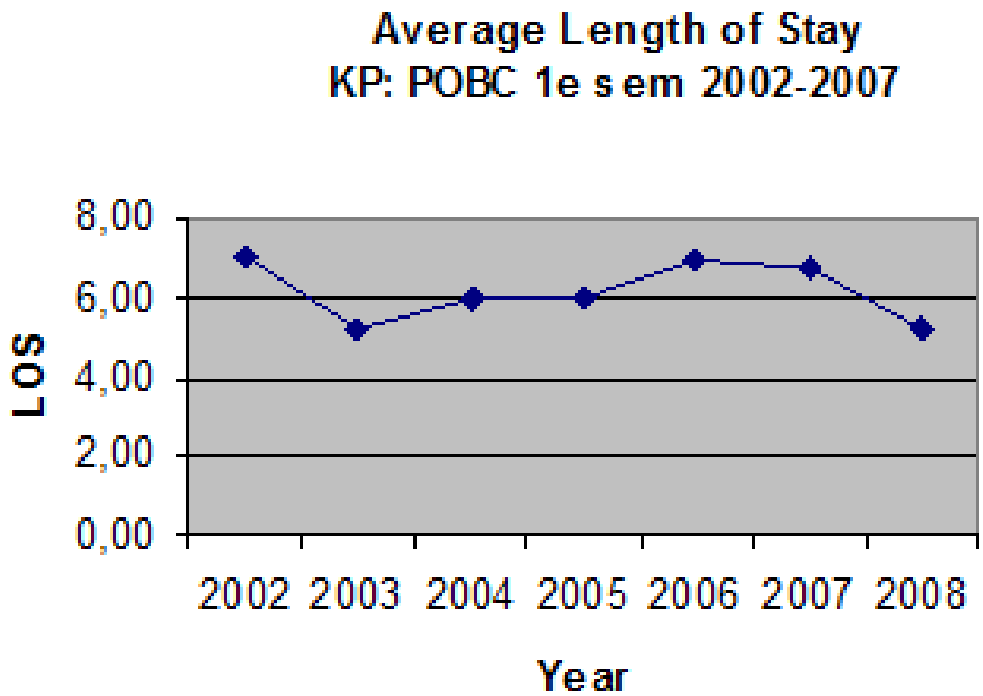


Figure 5.15 - The evolution of the Length of Stay

Despite these encouraging results, the quality team still noticed important *value leaks* in the POBC Clinical Pathway:

a) Clinical pathways are integrated in the ICT-systems of the breast cancer clinic. So, based on the actual pathway that is present in the ICT-systems, the ICT-system may indicate in which phase the patient is in the (presumed) pathway. Of course, also the nurse knows where the patient, based on her execution of the care process, actually is. Potential value leaks might emerge if (and of course this is as such a hypothetical if) variations on the pathways would emerge on the actual work floor. In that case, the nurse knows exactly what she is doing in "her" variation of the care process, whereas the actual tracking may follow another process (variation). Of course this is not a problem if the clinical pathways are enforced by the ICT-systems. It is, as we will see later in this case, the challenge of Process Discovery to find if there are indeed relevant variations in the clinical pathways. This is potentially an *asymmetry of information*, since in reality the patient and the staff knows where he/she is, but the ICT-system might follow a variation of the process.

b) The actual tracking of the pathway. If some steps are inadequately performed and require rework, potential value leaks emerge. Fortunately, if the ICT-systems enforce the clinical pathways, rework is largely prevented. However, also the skipping of some steps (to shorten the length of the stay of the patient, for example) could also result in value leaks. If the skipped steps are not traced, an *incompleteness of information* emerges.

c) At some points there was a unintentional lack of information, meaning *incompleteness of information*. One typical example was trying to understand when and why patients ask for pain killers.

d) Finally, sometimes inconsistent information emerged, such as the proper impact of the Length of Stay on the overall process/pathway. Governmental enforcement rules in the Belgian Healthcare system stimulate minimal Length of Stay for patients. An obvious question is whether the ***minimal length of stay*** is also the ***optimal length of stay***. This is an *asymmetry of meaning.* Observe that this potential value leak is not bound to this case study only, and arises in general for the Belgian Healthcare enforcement system.

This case study (Poelmans et al., 2010), (Poelmans, 2010) was a first attempt to get a better understanding on these value leaks by using Process Discovery. The dataset consists of 148 breast cancer patients that were hospitalized during the period from January 2008 till June 2008. The treatment of breast cancer consists of 4 phases. Before the patient is hospitalized, she/he ambulatory receives a number of pre-operative investigative tests. During the surgery support phase she is prepared for the surgery she will receive, while being in the hospital. After surgery she remains hospitalized for a couple of days until she can safely go home. The post-operative activities are also performed in an ambulatory fashion. Every activity or treatment step performed to a patient is logged in a patient results database and in the dataset all the activities performed during the surgery support phase to each of these patients were included.

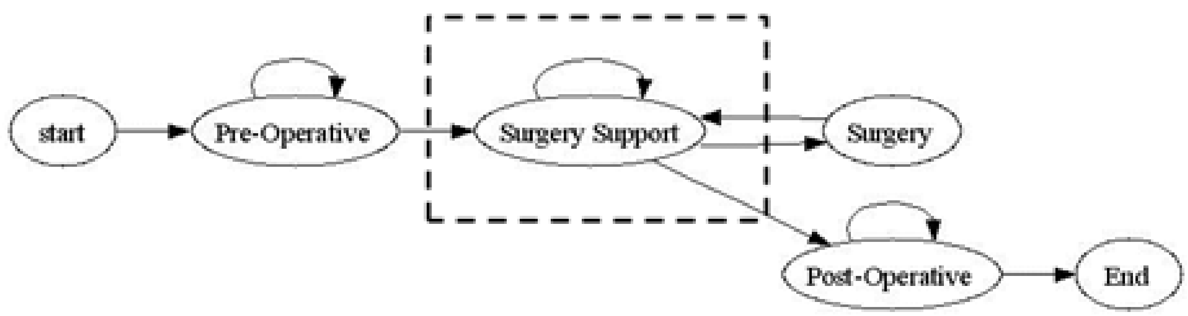


Figure 5.16 - The high-level overview of the patient care activities

Each activity has a unique identifier: 469 identifiers in total were introduced for the clinical pathway POBC. Using the timestamps assigned to the performed activities, we turned the data for each patient into a sequence of events. These sequences of events, conform to the Event – Actor – Object model, were used as input for the process discovery methods described in Chapter 2.

5.3.2 A first discovery analysis examining quality of care

The dataset was initially submitted to the *Comprehend* toolset, using the facilities for manual XML-input for Event – Actor – Object compliant data. Since the Length of Stay (LOS) is an official governmental Key Performance Indicator, it is obvious to observe in a first step the discovered Process models by LOS. In fact this led to a first remarkable observation:

a) We observed a relatively linear process for the group of patients with a length of stay in the hospital less than 10 days.

b) However, there were 12 patients for which the process model was very complex. They all had in common that their length of stay in the hospital was longer than 9 days.

The picture below shows screenshots from the output produced by the PIC in the *Comprehend* toolset (see Chapter 3). The upper part displays the obtained process map on the set of patients with a length of stay lower than or equal to 10 days in the hospital and the lower part displays the obtained map for the patients with a length of stay lower than 10 days.

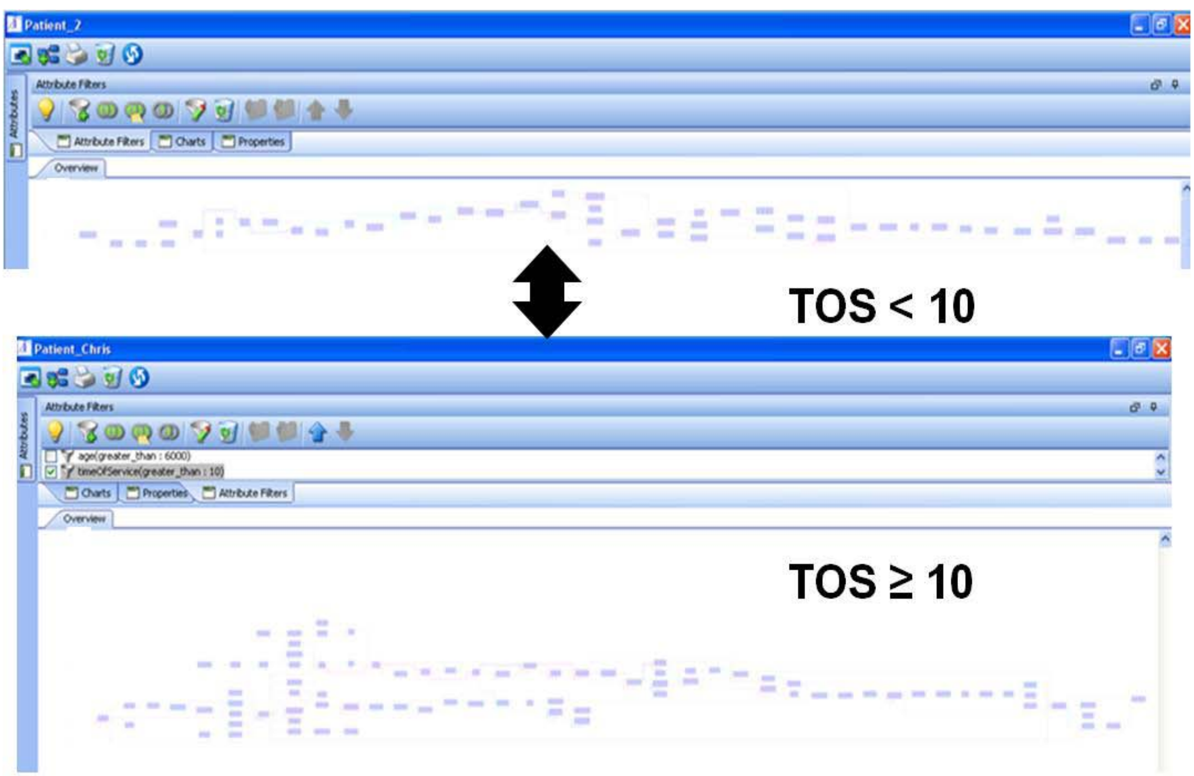


Figure 5.17 - A first Length of Stay analysis in *Comprehend*

To gain a better understanding of what happened to the 12 patients with an extended stay, added was a Data Discovery technique qualitative, capable of revealing potential relations between different steps in the clinical pathways. Formal Concept Analysis (FCA) was used, using the 12 patient records as “objects” and the pathway step indicators as “attributes”. FCA automatically discovers relationships between the objects and attributes into concepts, and organizes these concepts into a mathematical order structure, known as a lattice.

This resulted in the following observations.

1) Pain comfort is one of the very tangible perceived values for a patient in hospital care. One of our clinical indicators is the pain score which tells us at which days the pain experienced by patients reaches its highest level. We always saw peaks on 1 and day 4 of hospitalization however until now we had no idea why. The FCA lattice gave us an interesting suggestion that this might be due to an overlooked connection between removal of the wound drains and insufficient pain medication. We were able to find that wound drains is probably the most contributing factor to an increased pain score experience by patients and that pain medication should be administered (or – at least – presented to the patient) before removing the drains. It is clear that this was a significant *value leak*, in particular showing *asymmetry of information*: the patients knew very well what is painful or not, but until this analysis this was overlooked by the team. Needless to say that this observation was a plain eye-opener for the staff.

2) Another small data quality leak emerged in the analysis. For 1 patient, perhaps due to an enforcement leak in the ICT-systems, a psychosocial activity was not recorded in the systems. Later it turned out how the patient actually received the psychosocial caring, as it was reflected in the medical patient file. So there was no issue of lack of quality in the actual nursing of the patient.

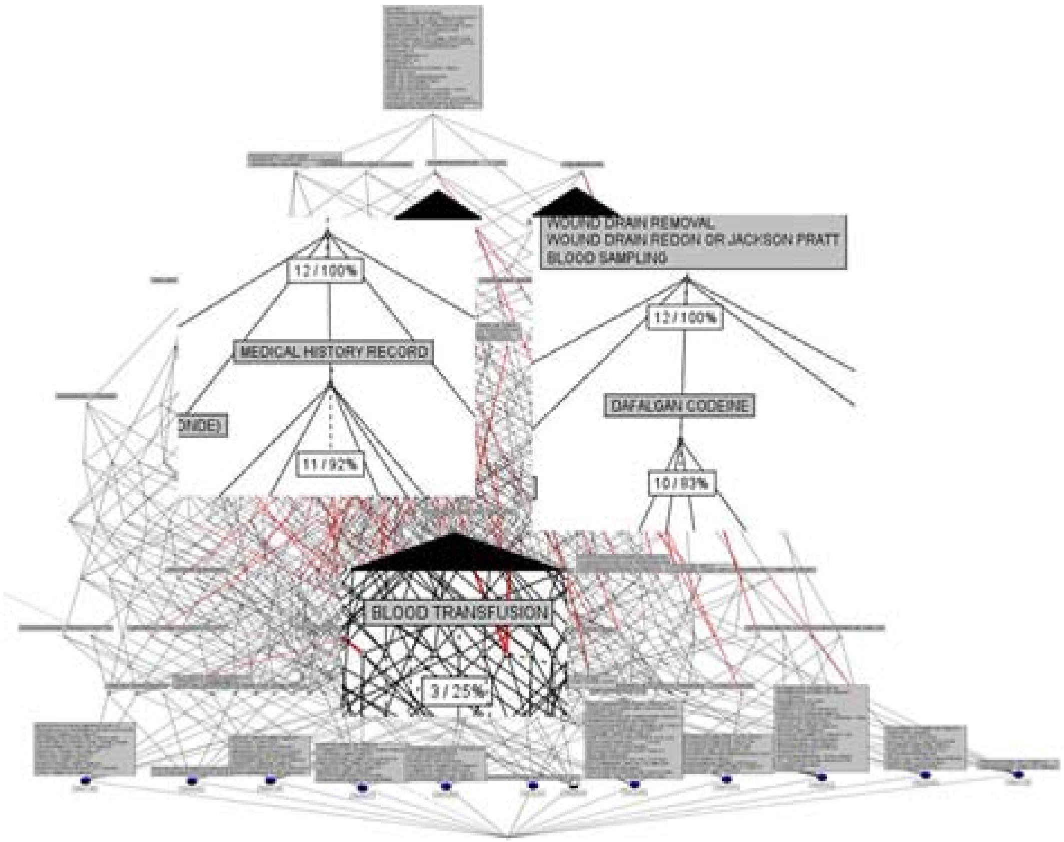


Figure 5.18 - The FCA-analysis showing the event lattice

3) Finally 3 main indicators for an increased length of stay were found to be the following:

   - wound infection
   - subsequent bleeding
   - neurological/psychiatric problems

The presence of these factors makes the care process more complex and results in more investigative tests. Since these additional morbidities are probably one of the root causes for this increased length of stay, there treatment should be anticipated on and hopefully be optimized during the preoperative phase.

5.3.3 A second discovery analysis in search of process variations

There are five types of breast cancer surgery:

- mastectomy
- breast conserving surgery
- lymph node removal
- the combination of mastectomy with lymph node removal
- the combination of breast conserving surgery with lymph node removal

For each of these surgery types, we extracted the corresponding patients in the dataset and constructed a process model and an FCA lattice for in-depth analysis of the characteristics of these groups.

Mastectomy surgery consists of completely removing the breast whereas during breast conserving surgery only the tumor is removed. The process models showed that the complexity of the care process is much higher for the mastectomy patients. Since mastectomy is a more complex surgery type, we expected that the FCA lattices would also be more complex than for breast conserving surgery. Surprisingly we found out that this was not true. The complexity of the lattice was larger for the breast conserving surgery patients and we found that this was due to the less uniform structure of this care process, in which for many patients, some essential care interventions were missing. This is the lattice showing the interventions performed to the 60 patients receiving breast-conserving surgery with lymph node removal (Figure. 5.19)

The lattice shows that 3 of these patients did not receive a consultation from the social support service. 15 patients did not have an appointment with a physiotherapist and did not receive revalidation therapy. 1 patient did not receive a pre-operative preparation and 2 patients were missing emotional support before and after surgery. All of these are again clear indications of *value leaks* in the care processes, mainly due to *incompleteness of information* about the actual care processes.

Of course, the next question is; Why? One potential root cause that was given by the team was the fact that the originally developed breast-conserving surgery care pathway was written for a certain length of stay for the patients in the hospital. As seen in the first paragraph of this case study, this length of stay was significantly reduced over the past years without modifying the care process model. As a consequence, it became impossible to execute the prescribed process model correctly in practice, and subsequently some patients are receiving suboptimal care. The activities performed on the patients should be reorganized and a new care pathway, taking into account this time restriction, should be developed.

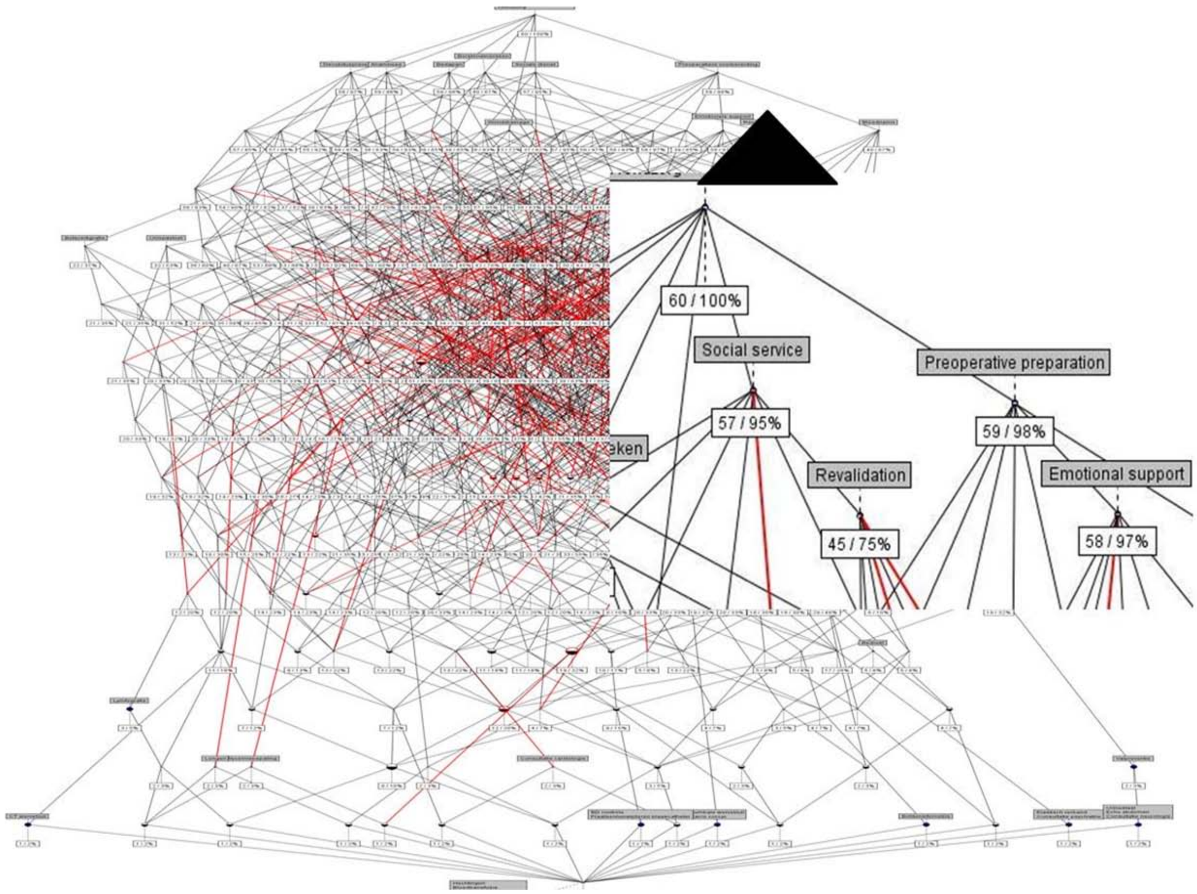


Figure 5.19 - The FCA analysis for a process variation with a shorter Length of Stay

This shows the lattice for the 37 patients receiving mastectomy surgery with lymph node removal.

First of all, this lattice has a much less complex structure than the lattice for the breast conserving surgery with lymph node removal. For the mastectomy patients, we found that most patients received all of the key interventions prescribed in the clinical pathway. Only for two patients there was a potential quality of care issue, namely 1 patient did not receive emotional support and 1 patient did not receive breast prosthesis. These shortcomings in the provided care however may have serious consequences for the psychological well-being of the patient.

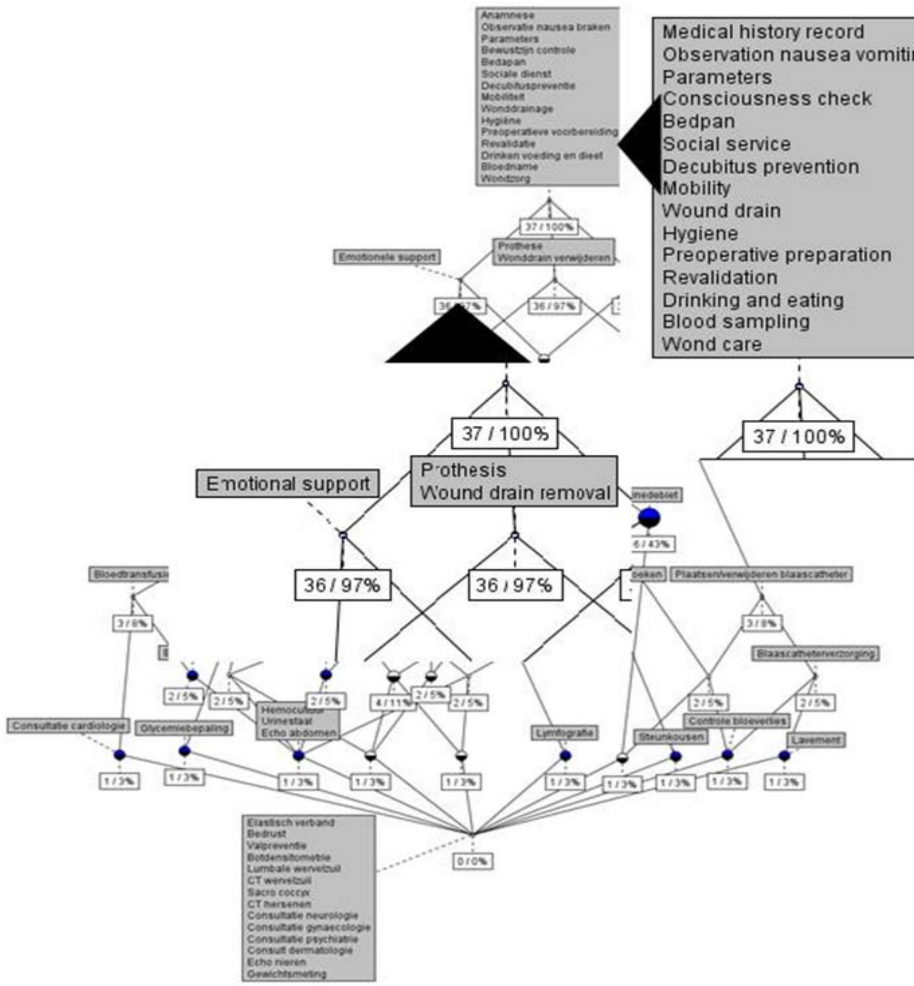


Figure 5.20 - Another process variation FCA analysis

5.3.4 Discovering the "natural" pathways.

Length of Stay, which is the official KPI for Hospital performance, can often be a major root cause of potential value leaks in care processes. To discover this, we constructed a lattice that contains indications on the LOS, and tried to find the relations with responsible team members (in particular, specialist doctors). This is the lattice for the 60 patients receiving breast conserving surgery with lymph node removal, with as attributes the length of stay and the doctor performing the operation (the "actor").

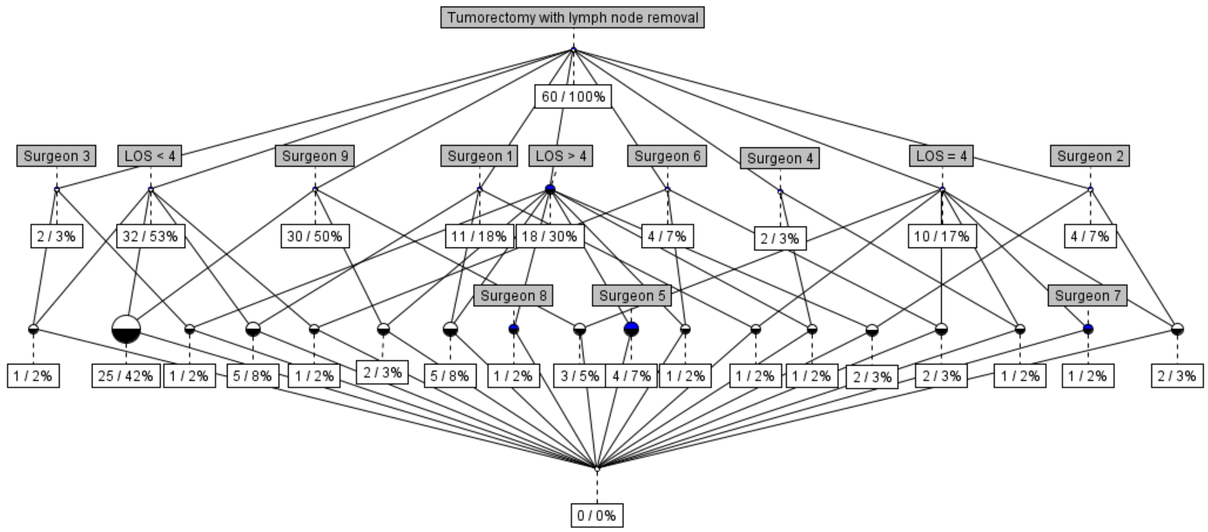


Figure 5.21 - An Actor lattice

In the lattice we see that 25 patients with a length of stay smaller than 4 days were treated by "surgeon 9", whereas almost all patients treated by the other doctors had a longer length of stay.

We extracted these subsets of patients and constructed a process model for the groups of patients with a length of stay smaller than 4 days, equal to four days and larger than 4 days. This shows the HMM process model extracted from the datasets with the 10 breast-conserving surgery patients with a length of stay in the hospital of 4 days (the average length of stay).

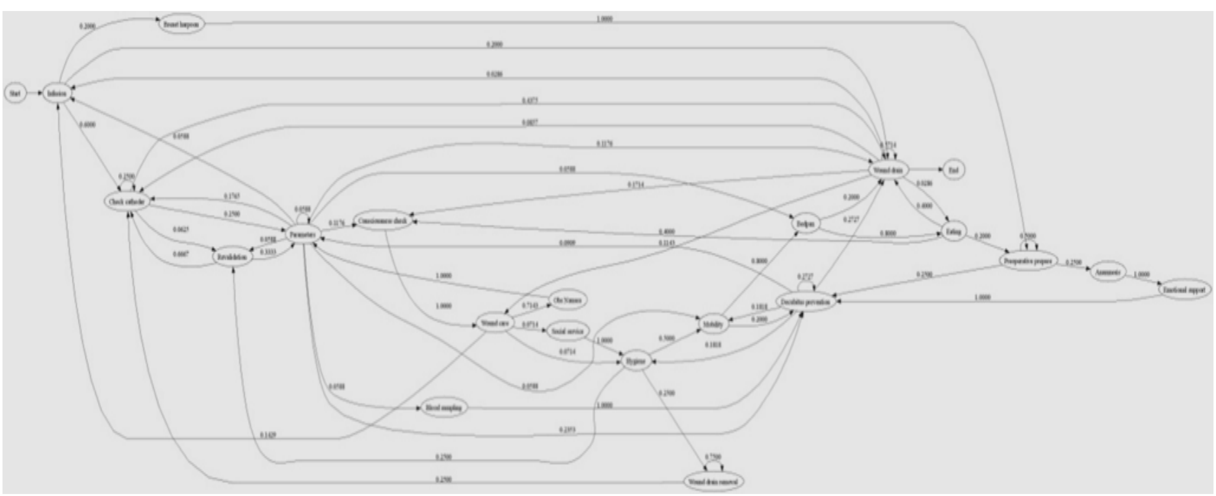


Figure 5.22 - The HMM for a Length of Stay of 4 days

This "natural" process model was chosen because of its simplicity in comparison with the other models and since it most closely resembles the standard care process as perceived by the domain experts.

To verify the degree of variation in the processes, some complexity measures for these process variations have been calculated, as follows:

SURGERY/LENGTH OF STAY		LOW	AVG	HIGH
Breast Conserving Therapy with Lymph Node Removal	Length of stay	< 4 days	= 4 days	> 4 days
	# patients	32	10	18
	Avg. # activities	97	146	184
	# unique activities	32	23	35
	# unique act filtered	24	22	22
	# connections filtered	98	80	92
	Structural Complexity filtered	122	102	114
Mastectomy with Lymph Node Removal	Length of stay	< 7 days	= 7 days	> 7 days
	# patients	17	4	16
	Avg. # activities	187	206	268
	# unique activities	27	21	36
	# unique act filtered	19	20	24
	# connections filtered	83	78	100
	Struct. Complex. filtered	102	98	124

Table 5.6 - Process characteristics

For each surgery type and length of stay, the number of patients, the average number of activities and the number of unique activities performed on these patients is given. For visualizing the process maps, we set a cutoff point at 5%, (i.e. all transitions with a lower probability of occurrence were removed from the process representation). The table also contains the number of remaining unique activities and the number of connections after filtering. The structural complexity measure after filtering is the sum of these two measures.

Similar results on the process variations were obtained for the patients receiving a Mastectomy with lymph node removal. The average LOS process model (for 7 days) is by far the simplest – almost linear – process:

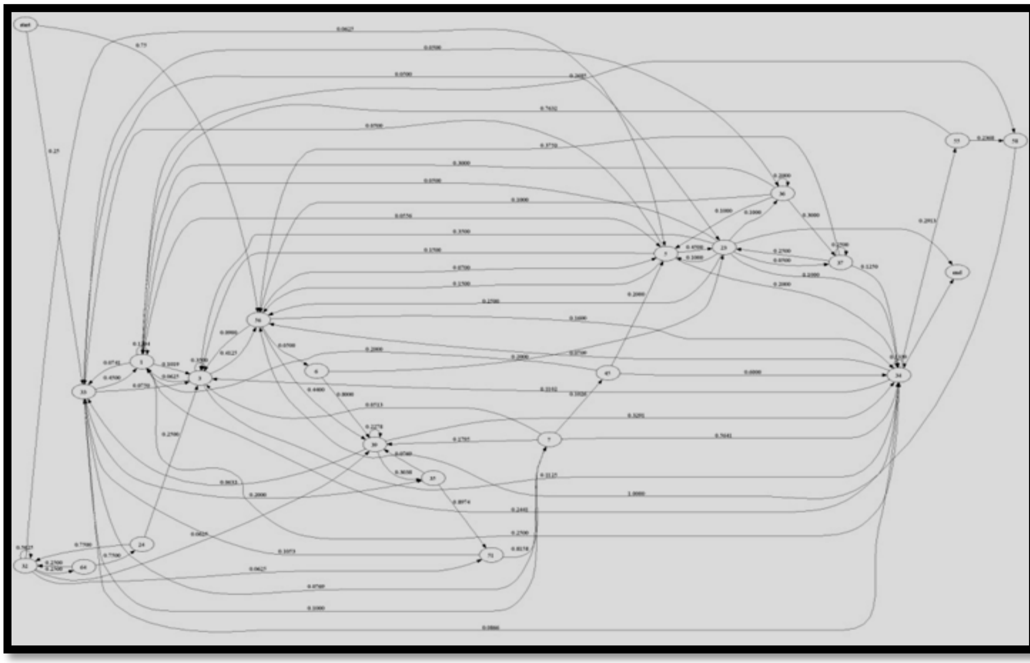


Figure 5.23 - The HMM for a Length of Stay of 7 days

The model for patients with a shorter length of stay is more complex, and more risky, as explained before. Visually it is also clear that this process model contains more places that are not normalized. This is again striking: more places that are not normalized means more potential value leaks…

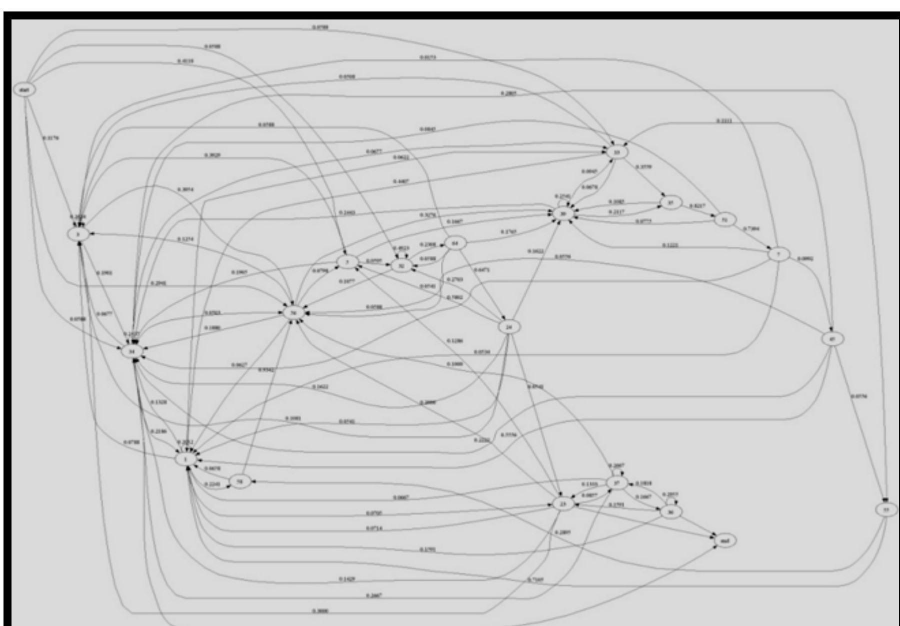


Figure 5.24 - The HMM for a Length of Stay smaller than 7 days

The process model for patients with a longer LOS is also more complex, but in this case this is almost naturally due to the complications that give rise to the longer LOS. Nevertheless, the complications, which in themselves are also potential value leaks (increased risks for patients) are reflected by a higher number of places that are not normalized in the process Markov Model.

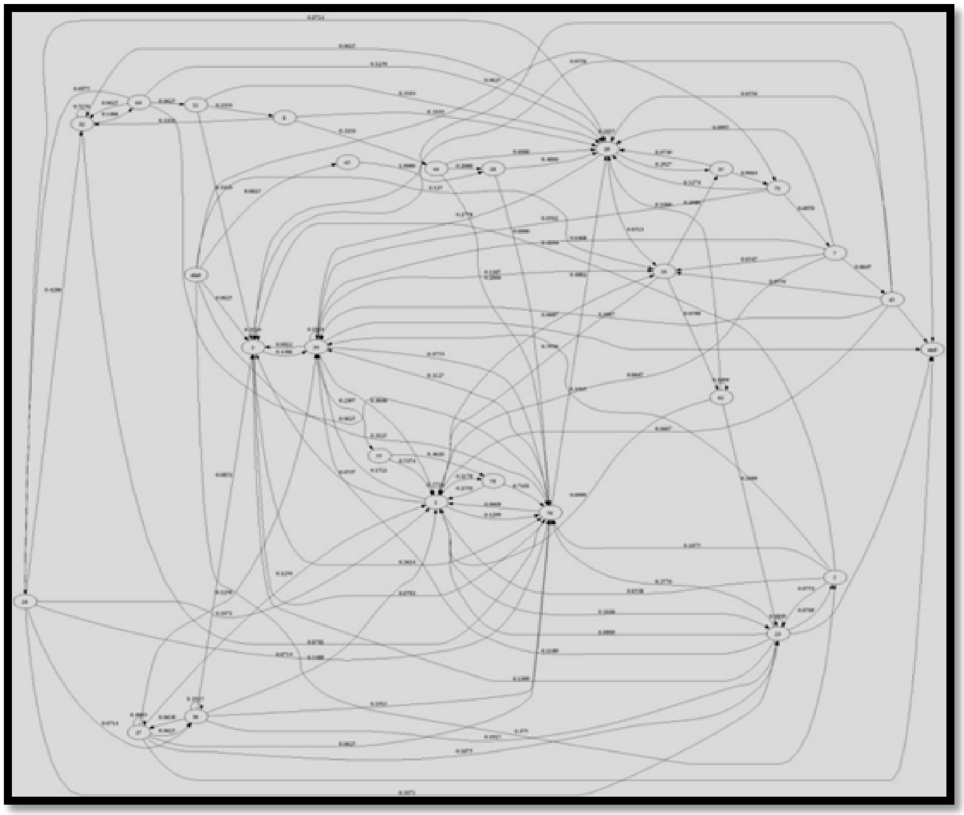


Figure 5.25- The HMM for a length of Stay longer than 7 days

5.3.5 Discussion

We learned a lot about Hospital Information Systems and, in particular about patient result servers. An important point is the time-stamping of the events that are registered. Typically the systems contain time-stamps not only for the moment when the event actually happened, but also for the moment when the event was registered, which may be significantly different. Much is related to the data entry and data enforcement policies of the ICT-systems for the responsible nursing staff. Sometimes the workflow procedures for data entry may also influence the time-stamps.

A major lesson learned from this last case study is the deep synergy between Process Discovery and Data Discovery that became evident. Although FCA has no proper process representation, it has proven to be very useful for process analysis, (i.e. to analyze anomalies and exceptions in detail).

In particular, concept lattices were used to improve the detection and understanding of outliers in the data. These exceptions are not noise, but are the activities performed on human beings, so every exception counts and must be understood. Concept lattices were also used to reduce the workspace, to cluster related events together in an objective manner.

Using this combination of techniques, we exposed multiple quality of care issues. We gained a better understanding of the process variations and better understood where we should take action to improve our healthcare processes. The impact of co-morbidities of patients on the overall care process was found to be of importance and offers some opportunities for improving quality and efficiency of care. Further, reducing the length of stay of breast-conserving therapy patients was discovered to be the root cause for a suboptimal care, missing some key interventions, provided to patients. Finally, we found the length of stay for patients receiving breast-conserving surgery was significantly different for different surgeons. This situation may be improved by making the discharge criteria more uniform.

5.4 Summary

Chapter 5 described three caring case studies as opposed to the administrative cases described above. The first two studies dealt with customer contact center operations and described activities that were both highly interactive as well as highly transactional. The third case study described activities in a post-surgical breast cancer ward that were highly interactive while low transactional. In each instance the concepts and techniques of process discovery were applied in order to discover, categorize and, where appropriate, provide a financial value to the discovered leaks.

The first case analyzed operations at a major international bank. In that instance, *Comprehend* discovered business process value leaks from core back office processes with an improvement value of more than £9MM per year. The main benefit streams were found in quality improvement and back office telephony process improvement. These included:

- Quality Improvements (RFT): *Comprehend* provided the data required to increase Right First Time (RFT) from 75% to 85% (worst case) or 90% (best case) yielding a NPV £12.0M to £15.9M of benefit from reduced re-work labor costs over a 4 year period.

- Cycle Time (Speed): *Comprehend* provided the detailed steps performed by every operator for every process yielding the insights to improve process end-end times by 10-25% yielding a NPV £8.2M to £16.6M over a 4 year period.

As a result of a reduction in errors and an improved understanding of the processing work, it can be assumed that there are additional opportunities for improvement in customer attrition.

- Customer Attrition: Identification of individual customer activity, highlighting preferences and allowing for the tailoring of service provision and resource allocation to suit actual customer behavior yielding a reduction in attrition from 1-2% providing a benefit of £4.7M to £9.4M in MTA customer retention over a 4 year period.

In addition to the hard benefits identified during the Proof of Concept, *Comprehend* provided additional soft benefits associated with:

- Staff Analysis: Detailed monitoring of operator activity showing actual transaction completion times per staff member, allowing identification of best practice and under-performance, highlighting targeted training opportunities.

In the second case study, the call center operations for a municipal tax office were analyzed. As opposed to the last example, the municipal tax authority operated in a "not for profit" environment and, while improved productivity was expected, it was not expected to come in the form of a reduction in full time positions (FTE). This point was further stressed as many of the personnel were physically handicapped. Additionally, in contradistinction to the first case, the *Comprehend* product was not able to be used since it was not compatible with the software deployed in the municipal tax office. However, because a sufficiently precise process model was discovered through manual time and motion study techniques, it was possible to simulate the effects of discovering Value Leaks in this process.

Overall, nine (9) value leaks were discovered in the contact process with a projected reduction in call handling time ranging from 9% to 21%. Additionally, by correcting the discovered process value leaks, a commensurate increase of 8% -11% in call handling volume was expected. All in all, the overall reduction in Total Cost of Ownership was projected to be 15%. As a result, after resolving the value leaks, this Contact Center can do the same work with a budget of 8 Million EUR (instead of 9.5 Million EUR), without any reduction in FTE. This was achieved through the reduction overflow capacity that was required to fill the consequences of the process value leaks. This is a wonderful example of increasing productivity through the elimination of *muda* in an ICT-enabled service process.

In the third case study, we applied the techniques of process discovery in a hospital, post-surgical setting. In this case, there was a lot to be learned about Hospital Information Systems and, in particular, about patient result servers. An important point in such systems is the time-stamping of the events that are registered. Typically the systems contain time-stamps not only for the moment when the event actually happened, but also for the moment when the event was registered, which may be significantly different. Therefore, the veracity of the data is related to the data entry procedures and data quality enforcement policies of the ICT-systems for the responsible nursing staff. Additionally, the workflow procedures for data entry may also influence the time-stamps.

A major lesson learned from this last case study is the deep synergy between Process Discovery and Data Discovery that became evident. Although FCA has no proper process representation, it has proven to be very useful for process analysis, i.e. to analyze anomalies and exceptions in detail. .

Using this combination of techniques, we exposed multiple quality of care issues. We gained a better understanding of the process variations and better understood where we should take action to improve our healthcare processes. The impact of co-morbidities of patients on the overall care process was found to be of importance and offers some opportunities for improving quality and efficiency of care. Further, reducing the length of stay of breast-conserving therapy patients was discovered to be the root cause for a suboptimal care, missing some key interventions, provided to patients. Finally, we found the length of stay for patients receiving breast-conserving surgery was significantly different for different surgeons. This situation may be improved by making the discharge criteria more uniform.

Interactive, caring, ICT-based processes are by nature more difficult and complex. Caring processes contain relatively more decision points and potential repetitions (loops) when compared to pure administrative processes. This is natural, since the "pre" and "post" conditions and invariants for the various processing steps (e.g. the services) are less strict in caring environment. Caring means: giving ones best effort, even being confronted with situations that go beyond the predefined service contracts. Interactive processes hence contain more “unknown unknowns” situations; we can try to anticipate all of the potential complications for patients, but they can never be totally excluded.

Therefore, even more than in administrative processes, markers for potential value leaks are extremely useful. Process Discovery with first order transitions are readily capable to show the major places that are un-normalized in the Markov Chains. These are the "hotspots" for further investigation and, speaking in C-K-terms, analysis and conjunction. Of course, additional statistical investigation of relevant parameters, where time – and in particular “lost time”, *muda* – is a prominent one in the case studies of this chapter. But not time alone; the first two case studies revealed the root causes of two very complex value leaks, re-work and customer turnover. The results of the third caring case study, clinical pathways, are equally relevant to avoid patient re-admission (e.g., re-work) for the breast cancer clinic.

CHAPTER 6

6.0 Thesis Conclusion

The objective of this research is to develop an appropriate set of methods, tools and techniques that can be used to automatically discover the root cause of value leaks and service imperfections in ICT-enabled business processes. A second objective is for these methods, tools and techniques to provide accurate and reliable guidance information for ICT-enabled business process improvement as defined as increases in productivity and/or quality**.**

In chapter 1, this PhD thesis has explored both the supply and the demand side of ICT-based services. The demand side was described as the desire for some event which could best be achieved through the clear description of expected "pre" and "post" conditions. These conditions in effect represent a service level agreement (SLA) between two parties. The supply side was described as a bundle of technologies, resources and assets that are deployed to meet the "pre" and "post" conditions of the SLA. ICT-based services then were seen as the building blocks of business processes. It was also shown how, in their attempt for simplification, ICT organizations have standard service bundles that cannot deal with exceptional situations and thus destroy value. The concept of value creation and value leaks was explored and it was shown how value leaks were a major issue in business process performance.

Chapter 2 addressed the research question; ***"How can process models be improved so as to more accurately reflect the underlying reality of the event being represented"?***

The main original contributions to Process Discovery research in this chapter are the following:

- the specialization of the C-K-theory to the problem of Process Discovery. This thesis work is the first that relates C-K-theory in detail to Process Discovery, after the successful applications of C-K-theory to Data Discovery.

- although there have been indications in other thesis works that the Petri-net representation is less suited for Process Discovery and Process Analytics in general, this work illustrated how several problems (such as matching an actual event sequence with a process representation) is easier to deal with when using the Hidden Markov Model representation.

- the notion of Process Model Normalization, stressing the potential ambiguities induced by many-to-many transitions in Process representations. It is demonstrated how several complicated extensions to the Petri-net representation (such as in [Greco et al., 2006]) are merely needed because of models that are not normalized. Moreover, it is clearly shown how models that are not normalized can lead to erroneous analytical results.

Based upon the mathematical techniques described above, Chapter 3 addressed the main research question; *"How, through the appropriate framing of methods, techniques and tools, can the root cause of value leaks and service imperfections in ICT-enabled service processes can be discovered so as to give guidance to management on increasing service productivity"?*

It described the methodology, tooling, process and a set of detailed steps that are required for an organization to use *Comprehend*. This chapter also detailed the experiences and processes through which *Comprehend* was developed and how OpenConnect product offerings transformed from exclusively middleware infrastructure into data collection, analytics and automation products and services.

In chapter 4, the first set of case studies, those that are concerned with administrative process, was introduced.

The first case study addressed the issue of healthcare claims adjustment costs. The costs of adjustments to the company in question, called *Atlantic Healthcare* (fictionalized name) are $30mn (US) per year. As much of the cost is in manual labor expended by knowledge workers during the review process, the potential savings are significant. *Comprehends* ability to record the time associated with each step of the event associated with a claim (e.g., researching subscriber history, benefit information, etc.…).

The first value leak was associated with improving the completeness and quality of the data sent to *Atlantic Healthcare* by the various healthcare providers. The amount of time spent by claims analysts working on population of adjustment reason codes was subjected to a Pareto analysis. This revealed that the top 9 out of 69 codes represented 87% of the time spent in the process. Further analysis showed that 20% of the providers accounted for 75% of all adjusted claims. This insight led to a pilot project with the top three providers who accounted for most of the adjustment issues. The solution implemented a "checklist" procedure to be followed by the providers before submitting a claim. This improvement corrected the incompleteness of information issue (service imperfection) and yielded a 65% improvement over a four-month period. This allowed *Atlantic Healthcare* to reduce the examiner workload costs for adjusted claims by $2.5 MM (US) per year.

In addition to the reduction in examiner workload, a further analysis of the process and worker repetitive actions uncovered two additional value leaks, both resulting from an asymmetry of meaning, (service imperfection) yielding an additional $450,000 (US) savings per year.

Lastly, a fourth value leak, stemming from the current process of randomly identifying claims for audit, could add an undetermined amount to the payment recovery effort. This issue stemmed from the inability to have reliable audit identifiers to trigger likely recovery candidates. This asymmetry of information (service imperfection) between the provider organization(s) and *Atlantic Healthcare* could be corrected through the development of said indicators. This would result in a more intelligent and effective audit process.

Overall, *Atlantic Healthcare* was able to discover value leaks and service imperfections that were draining approximately $3MM (US) per year from the organization.

The second administrative case study focused on the cost issues associated with the claims processing environment when working with a Business Process Outsourcer (BPO). *Lonestar Healthcare Company* (*fictionalized name of a healthcare insurer located in the southwest region of the US*) had been working on an automatically renewing contract with a major business process outsourcing (BPO) provider. Originally, this effort was planned to be temporary and, as such, there were few documented service- lifecycle agreement (SLA) requirements. Also, the initial BPO processing price of $2.40/claim had never been reviewed with the outsourcer in question.

In this case study, five value leaks resulting from four service imperfections were discovered. Specifically, the accountability for properly managing the BPO claims processing costs were addressed. This required an investigation of the work being done by in-house and outsourced examiners. It was determined that the outsourced examiners were being overpaid for work performed. This information was not previously available to *Lonestar Healthcare* but was available to the BPO firm. This asymmetry of information (service imperfection) allowed the BPO to realize higher rents than were justified, resulting in a contract re-negotiation saving *Lonestar Healthcare* $550K (US) per year.

The second value leak concerned examiner productivity. *Lonestar Healthcare* held a false assumption as to the productivity of its claims examiners. It believed that at-home examiners were less efficient than those on premises and was about to eliminate the at-home workers. After investigation with *Comprehend* it was determined that the at-home group was actually the most productive. Through rectifying the incompleteness of information (service imperfection) regarding examiner productivity, estimated saving increased by $495K (US).

The third value leak was in the area of best practices. Through the use of *Comprehend*, it was discovered that examiners used different process paths to work claims. As each person developed their own way to perform the work, variances existed in time among the multiple paths. This asymmetry of meaning (service imperfection), when addressed revealed the potential of an additional $700K (US) per year savings.

The fourth value leak focused on reducing compliance penalties which were costing the Company $2MM (US) per year. Performing a root cause analysis with *Comprehend* showed several inefficiencies to multiple processes that led to late payments. This was compounded by different rules that were established by 40 different (US) States resulting in an asymmetry of information environment (service imperfection). Through correcting this imperfection it was estimated that at approximately 10% of penalty cost could be saved resulting in an additional $420K (US) per year.

Lastly, associated with the issue mentioned above, it was determined that by reducing compliance penalties (e.g., correcting the imperfection of asymmetry of information), *Lonestar*

Healthcare could avail itself to discounts available to prompt payers. This had the potential of adding an incremental $100K (US) to the projected savings.

Overall, *Lonestar Healthcare* was able to discover value leaks and service imperfections that were draining an estimated $2MM (US) per year from the organization.

Chapter 5 described three caring case studies as opposed to the administrative cases described above. The first two studies dealt with customer contact center operations and described activities that were both highly interactive as well as highly transactional. The third case study described activities in a post-surgical breast cancer ward that were highly interactive while low transactional. In each instance the concepts and techniques of process discovery were applied in order to discover, categorize and, where appropriate, provide a financial value to the discovered leaks.

The first case analyzed operations at a major international bank. In that instance, *Comprehend* discovered business process value leaks from core back office processes with an improvement value of more than £9MM per year. The main benefit streams were found in quality improvement and back office telephony process improvement. These included:

- Quality Improvements (RFT): *Comprehend* provided the data required to increase Right First Time (RFT) from 75% to 85% (worst case) or 90% (best case) yielding a NPV £12.0M to £15.9M of benefit from reduced re-work labor costs over a 4 year period.

- Cycle Time (Speed): *Comprehend* provided the detailed steps performed by every operator for every process yielding the insights to improve process end-end times by 10-25% yielding a NPV £8.2M to £16.6M over a 4 year period.

As a result of a reduction in errors and an improved understanding of the processing work, it can be assumed that there are additional opportunities for improvement in customer attrition.

- Customer Attrition: Identification of individual customer activity, highlighting preferences and allowing for the tailoring of service provision and resource allocation to suit actual customer behavior yielding a reduction in attrition from 1-2% providing a benefit of £4.7M to £9.4M in MTA customer retention over a 4 year period.

In addition to the hard benefits identified during the Proof of Concept, *Comprehend* provided additional soft benefits associated with:

- Staff Analysis: Detailed monitoring of operator activity showing actual transaction completion times per staff member, allowing identification of best practice and under-performance, highlighting targeted training opportunities.

In the second case study, the call center operations for a municipal tax office were analyzed. As opposed to the last example, the municipal tax authority operated in a “not for profit” environment and, while improved productivity was expected, it was not expected to come in the form of a reduction in full time positions (FTE). This point was further stressed as many of the personnel were physically handicapped. Additionally, in contradistinction to the first case, the *Comprehend* product was not able to be used since it was not compatible with the software deployed in the municipal tax office. However, because a sufficiently precise process model was discovered through manual time and motion study techniques, it was possible to simulate the effects of discovering Value Leaks in this process.

Overall, nine (9) value leaks were discovered in the contact process with a projected reduction in call handling time ranging from 9% to 21%. Additionally, by correcting the discovered process value leaks, a commensurate increase of 8% -11% in call handling volume was expected. All in all, the overall reduction in Total Cost of Ownership was projected to be 15%. As a result, after resolving the value leaks, this Contact Center can do the same work with a budget of 8 Million EUR (instead of 9.5 Million EUR), without any reduction in FTE. This was achieved through the reduction overflow capacity that was required to fill the consequences of the process value leaks. This is a wonderful example of increasing productivity through the elimination of *muda* in an ICT-enabled service process.

In the third case study, we applied the techniques of process discovery in a hospital, post-surgical setting. In this case, there was a lot to be learned about Hospital Information Systems and, in particular, about patient result servers. An important point in such systems is the time-stamping of the events that are registered. Typically the systems contain time-stamps not only for the moment when the event actually happened, but also for the moment when the event was registered, which may be significantly different. Therefore, the veracity of the data is related to the data entry procedures and data quality enforcement policies of the ICT-systems for the responsible nursing staff. Additionally, the workflow procedures for data entry may also influence the time-stamps.

A major lesson learned from this last case study is the deep synergy between Process Discovery and Data Discovery that became evident. Although FCA has no proper process representation, it has proven to be very useful for process analysis, i.e. to analyze anomalies and exceptions in detail. .

Using this combination of techniques, we exposed multiple quality of care issues. We gained a better understanding of the process variations and better understood where we should take action to improve our healthcare processes. The impact of co-morbidities of patients on the overall care process was found to be of importance and offers some opportunities for improving quality and efficiency of care. Further, reducing the length of stay of breast-conserving therapy patients was discovered to be the root cause for a suboptimal care, missing some key interventions, provided to patients. Finally, we found the length of stay for patients receiving breast-conserving surgery was significantly different for different surgeons. This situation may be improved by making the discharge criteria more uniform.

Interactive, caring, ICT-based processes are by nature more difficult and complex. Caring processes contain relatively more decision points and potential repetitions (loops) when compared to pure administrative processes. This is natural, since the "pre" and "post" conditions and invariants for the various processing steps (e.g. the services) are less strict in caring environment. Caring means: giving ones best effort, even being confronted with situations that go beyond the predefined service contracts. Interactive processes hence contain more “unknown unknown” situations; we can try to anticipate all of the potential complications for patients, but they can never be totally excluded.

Therefore, even more than in administrative processes, markers for potential value leaks are extremely useful. Process Discovery with first order transitions are readily capable to show the major places that are not-normalized in the Markov Chains. These are the "hotspots" for further investigation and, speaking in C-K-terms, analysis and conjunction. Of course, additional statistical investigation of relevant parameters, where time – and in particular “lost time”, *muda* – is a prominent one in the case studies of this chapter. But not time alone; the first case studies revealed the root causes of two very complex value leaks, re-work and customer turnover. The results of the third caring case study, clinical pathways, are equally relevant to avoid patient re-admission (e.g., re-work) for the breast cancer clinic.

6.1 Research Approach Debriefing

At the end of this thesis, it is appropriate to debrief the Research approach that was employed in this exploration of Process Discovery for Value Leaks. As pointed out in chapter 1, the approach that is used is a holistic multiple case design approach. The approach is not focusing on one particular improvement parameter, such as cost, but tried to consider “value” and in particular value-leaks in a holistic way. Furthermore, Discovery has no pre-assumed experimentation model, in contrast to mining that often relies on pure statistical techniques (regression, correlation…), that are very well known in the Lean Six Sigma approaches to industrial processes. The approach included multiple cases, obviously to meet the desire to understand various ICT-based process types, in particular the transaction versus interaction type of processes.

The research model was followed during this thesis work. The fundamental theories for Process Mining and Discovery (D3), as well as Process Representation/Normalization (D4) have been explored in Chapter 2. They have been confronted with the semantics of Process Representation in view of Process Discovery (D3, C4) under the assumption of a particular structuring of the input data sources (C3) under the Event – Object – Actor format.

In chapter 1 The fundamentals of Service-Oriented Architecture was defined by putting ICT-based Services as the granular Process steps (C1). The notion of value leaks was introduced and linked to service imperfections (B1) bearing in mind that intelligence approaches, augmenting Human intuition and intelligence are required to deal with the many “unknown unknowns” in Service Imperfections. The C-K-theory was involved to guide the intelligence augmentation process (C3) to come to a fundamental understanding of ICT-enabled processes (B2), meaning the understanding of potential value leaks. In chapter 2 an intimate relation with Process Normalization was put forward (referring back to D4).

In chapter 3 it was demonstrated how to achieve and deploy a scalable toolset for automating the C-K-process, with disjunction in the first place. The experiences of developing such a toolset – *Comprehend* – in practice, large- scale problems were guiding this effort (D1).

The experiences of the case studies (D2) have mostly been explored and described in chapters 4 and 5. By the end of the thesis, bringing all elements together results in a practical, and – based on the case studies – proven approach for Discovering Value leaks and Service imperfections in ICT-enabled Business Processes (A1).

During the research journey, step by step the answers to the research questions emerged:

- Examination of un-normalized places in Hidden Markov Model representations of processes are providing the appropriate framing of methods, techniques and tools, that point to the root causes of value leaks and service imperfections in ICT-enabled service processes. The case studies give overwhelming evidence in this fact.

- The popular mainstream process models (BPMN, Petri-nets) are improved so as to more accurately reflect the underlying reality of the events being represented. The improvements include Hidden Markov Models, Normalization and the techniques to achieve normalization (clustering and cloning of un-normalized places).

- The trade-offs between volume of data and mathematical accuracy also became apparent: processing intensive algorithms, such as expected maximization (EM) are clearly overkill for a first analysis and a detection of un-normalized places. Furthermore, in chapter 2, the question on how many clones/clusters are minimally needed to arrive at precise models was addressed. FCA-analysis of the Event – Actor – Object data enables further detection and understanding of process exceptions and variations. This was illustrated in particular in Chapter 5.

6.2 Future Research

This thesis is one, but only one discovery journey on “Process Discovery”. Many more research issues remain to be explored in potential future research efforts. Three major research lines can emerge subsequent to this thesis work.

First there is an extension to the theoretical work of chapter 2. At the end of the keynote for the International Conference on FCA in Leuven, 2012, I already pointed to the underlying algebraic "language" for Probabilistic Finite State Machines and Hidden Markov Models, namely Idempotent Semi-rings. These semi-rings contributed to the research of so-called process algebra's, which are mainly used to analyze computer program implementations for processes. Recent developments in FCA, pioneered by Valverde, extended FCA into idempotent semi-rings. His examples focus on Boolean algebra as applications. However, idempotent semi-rings are the "most liberal" algebraic structures that allows for one single partial ordering relation (lattice) that is consistent with the axioms of the semi-rings. This allows one to derive concepts that are not only individual events or actors, but also consistent process parts that are expressible as formulas in semi-rings. This extension can give a further theoretical underpinning of the so-called dynamic diagrams in MERODE, a methodology once developed by Verhelst, Dedene and Snoeck. Part of this work was already developed by Poelmans for the static diagrams in MERODE. Another theoretical extension is the application of probabilistic FCA (using belief-functions) as an additional dimension for FCA. Probabilistic FCA may provide probabilistic extensions for Data Discovery as well as Process Discovery. One obvious element is the fact that the probabilities make the exceptions more evident.

A second line of research would be a further focus on a more systematic classification of processes, value leaks and the remediation for value leaks. The described case study research has revealed patterns for value leaks, extending what was already achieved for un-normalized places in this thesis, as well as best practices for the remediation. Such patterns may come from mapping process representations on categorized mathematical representations. Fortunately, HMMs indeed allow for such mappings, as well as a systematic exploration of "complexity" measures that might further indicate potential value leaks. It is clear that an un-normalized place denotes a higher complexity than the normalized process equivalents. This information, when associated with events can provide useful information such as alerts and the beginnings of self-correcting actions.

Finally, while a relationship was made among value leaks, service imperfections and financial value, the question of its connection to corporate strategy remains unanswered. Clearly, as more of our economic activity is performed by ICT-enabled service process, there will be a commensurate need to increase the competence of organizations to manage and perform in such an environment. The effects on jobs, the relationship of automation versus human augmentation in the delivery of services and the associated role of service quality will need to be explored.

These are just three examples, each of which in turn, could constitute another PhD thesis. Hopefully, each of them will inspire others and collectively, shine a small light into a corner of darkness, reducing the number of *"unknown, unknowns"*.

BIBLIOGRAPHY

Arthur, W. (2011). The Second Economoy. *McKinsey Quarterly*. Retrieved May 17, 2012, from http://www.mckinseyquarterly.com/The_second_economy_2853

Aberdeen Group. (2010). *Operational Intelligence: Boosting Performance with "Real-Time" Business Insight*. In M. Lock. Boston, MA: Author.

Agrawal, R., Gunopulos, D., & Leymann, F. (1998). Mining Process Models from Workflow Logs , report available at http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.8660, pp. 1-21.

Agrawal, A., E. Hamon, M. Viertler (2009) *Cutting Sales Costs, not Revenue,* Mc. Kinsey Quarterly, 2009, no. 2, pp. 77 – 85.

Alexander Ch. (1966), A city is not a tree, *Design, London Council of Industrial Design, No. 206, 1966.*

Aires da Silva G., R. Ferreira (2009). *Applying Hidden Markov Models to Process Mining,* Sistemas e Tecnologias de Informação Actas da 4 Conferência Ibérica de Sistemas e Tecnologias de Informação AISTIFEUPUPF, 2009.

Aljazzar, H., K. Sprague, and L. Yee. "What Matters Using Technology to Improve Workforce Collaboration." *McKinsey* (2009). Print

Allanach, J. (2005), *Detecting and Tracking Terrorist Activities with Hidden Markov Models,* University of Connecticut, 2005, 108 pp.

Alter, S., (2008), *Service System Fundamentals: Work System, Value Change and Life Cycle,* IBM Systems Journal, Vol. 47, No. 1.

Anderson, E., W. C. Fornell, and R. T. Rust. "Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services." (2001). Print.

Anderson, C., (2006) The Long Tail: Why the Future of Business is Selling Less of More, Hyperion, 2006.

Anupam, Kumar Nath Rahul Singh, Lakshmi S Iyer,(2010), Web 2 . 0 : Capabilities , Business Value and Strategic Practice, *Proceedings of the 15th Americas Conference on Information Systems AMCIS 2009, San Francisco,*

Apte , U.M Karmakar., U.S. & nath H. L, (2008), Information Services in the US Economy: Value , Jobs, and Management Implications. *California Management Review.*

Ashby, W. Ross., (2011). Variety, Constraint, and The Law Of Requisite Variety, Emergence: Complexity & Organization 13, pp. 190-207.

Awad, E. M. & Ghaziri, H. M.,(2004). *Knowledge Management.* North Garden, VA: International Technology Group.

Bhardwaj, D., (2007), Relationship Marketing in Context to the IT Industry, *The Journal of Business Perspectives, Vol. 11 Issue 2*

Beisiegel, M., Booz, D., Edwards M., Herness, E., Kinder S., (2007), Software components: Coarse-grained versus fine-grained. IBM Developer Works, Retrieved from http://www.ibm.com/developerworks/webservices/library/ws-soa-granularity/

Berkowitz, B. (2003). *The New Face of War: How War Will Be Fought in the 21st Century*. New York: Simon & Schuster.

Bicheno, John, and Matthias Holweg. *The Lean Toolbox: The Essential Guide to Lean Transformation*. 4thth ed. Buckingham: PICSIE Books, 2009. Print.

Bonchi, F., F. Giannotti, G. Mainetto, D. Pedreschi (1999). Using Data Mining Techniques in Fiscal Fraud Detection, DaWaK '99 Proceeding of the First International Conference on Data Warehousing and Knowledge Discovery, Springer Verlag, 1999, pp. 369 – 376.

Borges, J., M. Levene (2005). Generating Dynamic Higher-Order Markov Models in Web Usage Mining, *Proceeding of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 05), LNAI 3721, Porto Portugal, Outubro, 2005, pp. 34 – 45.*

Borges, J., M. Levene (2000). A Heuristic to Capture Longer Web Navigation Patterns, Proceedings of the 1st International Conference on Electronic Commerce and Web Technologies (EC-WEB 2000), September 2000 London U.K., pp. 155 – 164.

Borges, J., M. Levene (2005). A Clustering-Based Approach for Modeling User Navigation with Increased Accuracy, *Proceeding of the Second International Workshop on Knowledge Discovery in Databases WKKDDS, In conjunction with PKDD 2005, Porto, Portugal, Outubro, 2005.*

Borges, J., M. Levene (2004). A Dynamic Clustering-Based Markov Model for Web Usage Mining, *Technical Report cs. IR/0406032*.

Borges, J., M. Levene (2007). Evaluating Variable Length Markov Chain Models for Analysis of User Web Navigation Sessions, *IEEE Transactions on Knowledge and Data Engineering, Vol 19 (8), June 2007, pp. 717 – 727.*

Bradford, J. W., Knott, D. G., Levine, E. H., & Zemmel, R. W. (2011, December). Accounting for the cost of U.S. health care. *McKinsey Center for U.S. Health System Reform*, 1-37.

Brynjolfsson, E., L.M. Hitt (1998). Beyond the Productivity Paradox: Computers are the Catalysis for Bigger Changes, *Communications of the ACM, Vol 41(4), pp. 49 – 55.*

Buckow, H., S. Rey (2010). McKinsey on Business Technology: Why Business Needs should shape IT Architecture, nr 19, Spring 2010.

Bilmes, J.A., (1998). A General Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussain Mixture and Hidden Markov Models, International Computer Science Institute, Berkeley, 1998.

Brue, G. (2002). *Six Sigma for Managers*. Madison: CWL Publishing Enterprises Book.

Budalakoti, S., A.N. Srivatava, M.E. Otey (2008). Anomaly Detection and Diagnosis Algorithms for Discrete Symbol Sequences with Applications to Airline Safety, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol 39 (1), January 2009, pp. 101-113

Cabral de Moura Borges, Jose Luis (2000). A Data Mining Model to Capture user Web Navigation Patterns, PhD at Department of Computer Science, University College London, London University, 2000. Supervised by Mark Levene

Cai, J. (2007). Designing the right IT services for the bottom of the pyramid. *Communications of the Association of Information system, 19, 513-52.*

Callut, J., and P. Dupont. (2005a)Inducing Hidden Markov Models to model longterm dependencies. In *16th European Conference on Machine Learning (ECML)*, number 3720 in Lecture Notes in Artificial Intelligence, Porto, Portugal, October 2005. Springer-Verlag pp. 513 – 521.

Callut, J., and P. Dupont. "Technical Report." *Learning Hidden Markov Models to fit long-term dependencies*. 2005-9th ed. Louvain: Universite catholique de Louvain, 2005. Print.

Chesbrough H. & Spohrer J. (2006), A Research Manifesto for Service Science, *Communications of the ACM, Vol. 49, no 7, pp. 35-40.*

Christen, P., D. Belacicl (2005). Automatic Probabilistic Address Standardization and Verification, *Proceedings of the Australasian Data Mining Conference AusDM05.*

Codd, E.F., (1970). A Relational Model of Data for Large Shared Data Banks, *Communications of the ACM, Vol 13, No 6, pp. 377 – 382.*

Cook, J. E., & Wolf, A. L. (1996). Discovering Models of Software Processes from Event-Based Data, *ACM Transactions on Software Engineering and Methodology, Vol 7, pp. 215 – 249.*

Cook, J. E., & Wolf, A. L. (1999). Software Process Validation: Quantitatively Measuring the Correspondence of a Process Model, *ACM Transactions on Software Engineering and Methodology, Vol 8, pp. 147 – 176.*

Cox, K., Perneel, G., & Wellens, L. (2010). Clinical pathways and the value of process discovery. *Technical Report Faculty of Business and Economics, KU Leuven 2010. 39 pp.*

Chan, P., Coles, R., Dodds, L., Grantham, A., Kaplinsky, R., Mohamed, S., & Mynors, D. (n.d.). Understanding and Assessing the Concept of Knowledge Leakage. In *www.cs.nott.ac.uk*. Retrieved May 18, 2012, from http://www.cs.nott.ac.uk/~dap/network/previous%20workshops%20etc/literature%20review/closing_the_gap_joint_review.pdf#page=162

Coles, R.L., Walsh, K., Mynors, D., Grantham, A., Chan, P. & Mohamed S. (2007). Developing the Concept of Knowledge Leakage in Relation to UK Productivity, *the International Journal of Knowledge, Culture and Change Management, Vol 6, pp. 7 – 14 2007.*

Collins, M. (2005). Notes on the EM Algorithm, echnical Report 2005.

Datta, A. (2000). Automating the Discovery of AS-IS Business Process Models: Probabilistic and Algorithmic Approaches. *Information Systems Research*, *9*(3), 275-300. Retrieved May 18, 2012

David, S., M.A. Ferrer, C. M. Travieso, J. B. Alonso (2004). gpdsHMM: A Hidden Markov Model Toolbox in the Matlab Environment, *International Conference on Complex System Intelligence and Modern Technology Application, France, 2004.*

De Hertogh S., Viaene S., Dedene G., (2011), Governing Web 2.0, *Communications of the ACM, vol. 54(3), 2011, pp. 124 - 130.*

Drucker, P. (1999). Knowledge-Worker Productivity: The Biggest Challenge. *California management review*, *41*(2), 79-94.

Dupont, J., F. Denish, Y. Esposito (2005). Links between Probabilistic Automata and Hidden Markov Models: Probability Distributions, Learning Models, and Induction Algorithms, Pattern Recognition, Vol 38, 2005, pp. 1349 – 1371.

Dustdar, S., T. Hoffmann, and Van de Aalst. *Mining of ad-hoc business processes with TeamLog, Technical Report; Technical University of Russia*. N.p.: n.p., 2004. Print.

Elzinga, Paul. *Formalizing the Concepts of Crimes and Criminals*. Amsterdam: University of Amsterdam, 2001. Print.

Engelbart, Douglas C. (1962). Augmenting Human Intellect: A Conceptual Framework, Technical Report Stanford Research Institute

Ezell, S. J., & Atkinson, R. D. (2011). The Case for a National Manufacturing Strategy. In *The Information Technology & Innovation Foundation*. Washington, DC: ITIF.

Fleming, M., J. Silverstein, and B McDonough. "IDC Whitepaper." *OpenConnect Comprehend Work Event Monitoring and Analytics Product End-To-End Process Transparency*. N.p.: n.p., 2009. Print.

Fleming, M., and J. Silverstein. "IDC Whitepaper." *Healthcare Insurer Uses Comprehend to Improve Claims Processing Productivity*. N.p.: n.p., 2009. Print

Flesia, A. G., R. Fraiman, and F. G. Leonardi. "Technical Report." *Pattern Recognition on Random Tress Associated to Functionality Families of Proteins*. New York: Cornell University, 2007. Print.

Francalanci, C., & Galal, H. (1998). Information Technology and Worker Composition: Determinants of Productivy in the Life Insurance Industry [Electronic version]. *MIS Quarterly*, *22*(2), 227-241.

Gao, J., A.G. Hauptmann, A. Bharucha, H.D. Wactlar (2004). Dining Activity Analysis Using a Hidden Markov Model, Carnegie Mellon University Research Showcase Computer Science Department, 2004.

Galassi, U. "Tesi de Dottarato." *Structured Hidden Markov Model: A General Tool for Modeling Process Behavior*. Turni: Universita degli Studi de Torino, 2008. Print.

George, M.L. (2002), Lean Six Sigma: Combining Six Sigma Quality with Lean Speed, Mc Graw Hill Professional, 200, . pp. 6-8.

Gladyshev, P. (2005). Finite State Machine Analysis of Blackmail Investigation, International Journal of Digital Evidence, Vol 4(2), 2005.

Goetz, T. (2011). Harnessing the Power of Feedback Loops. In *www.wired.com*. Retrieved August 23, 2011, from http://wired.com/magazine/2011/06/ff_feedbackloop/all/1

Ghanayem, D, Peters, E. (2012), Effective Strategies for Streamlining Administrative Processes to Meet MLR Requirements, AHIP Institute Conference, Salt Lake City, Utah, June, 2012.

Grant R. M. (1991), The Resource Based Theory of Competitive Advantage. Implications for Strategy Formulation. *Californina Management Review, Vol. 33, pp. 114 – 136.*

Greco G., Guzzo, A., Pontieri, L., Sacca, D., (2006), Discovering Expressive Process Models by Clustering Log Traces, *IEEE Transactions on Knowledge and Data Engineering, Vol 18(8), 2006, PP. 1010 – 1027.*

Gronroos, C.(1990). *Service management and marketing.* Lexington, MA: Lexington Books

Gronroos, C., & Ojasalo, K. (2004). Service productivity towards a conceptualization of the transformation of inputs into economic results in services. *Journal of Business Research, 57*, 414-423.

Gummesson, E. (1998). Productivity, Quality and Relationship Marketing in Service Operations. *International Journal of Contemporary Hospitality Management, 10*(1), 4-15.

Gupta, M., V. Nagadevara (2007). Audit Selection Strategy for Improving Tax Compliance – Application of Data Mining Techniques, *Foundations of E-Government, Proceedings of the 11th International Conference on e-Governance, India, 2007.*

Gupta, S. W.M. P (2007). Workflow and Process Mining in Healthcare, *Master's Thesis. Technische Universiteit Eindhoven, Eindhoven, The Netherlands, 2007.*

Haplern, J., (2008), Beyond Nash Equilibrium: Solution Concepts for the 21st Century, Reprinted from the Proceedings of the Twenty-Seventh Annual ACM Symposium on Principles of Distributed Computing, 2008.

Hatchuel, A., Le Masson, P., et Weil, B. (2006), The design of science based-products: an interpretation and modelling with C-K theory." *9th International Design Conference, Dubrovnik, 15th-18th May 2004*, pp. 33-44.

Heckermann, D., C. Meek, D. Koller (2004). Probabilistic Models for Relational Data, *Technical Report MSR-TR-2004-30, Microsoft Research, March, 2004.*

Herbst, J., D. Karagiannis (1998). Integrating Machine Learning and Workflow Management to Support Acquisition and Adaptation of Workflow Models, *International DEXA (Database and Expert System Applications) Conference, DEXA 98, Wien 1998, pp.745-752*

Hopp, W. J., Iravani, S. R., & Liu, F. (2009). Managing White-Collar Work: An Operations-Oriented Survey. *Production and Operations Management*, *18*(1), 2009, pp. 1-32.

Houck, Christopher. "Business Intelligence Journal." *Breaking though with Business Discovery*. Vol. 4. Dallas: n.p., 2007. 9. 12 vols. Print

Hubbard, D. W. (2007). How to Measure Anything: Finding the Value of Intangibles in Business. *New Jersey: John Wiley & Sons, Inc., 2007.*

Hunt, D., Manyika, J., & Remes, J. (2011). Why US productivity can grow without killing jobs. *McKinsey Quarterly*, 2011, pp. 1-6.

Hunt, D., Manyika, J., Remes, J., Nyquist, S., Malhotra, V., Mendonca, L., & Auguste, B. (2011). Growth and renewal in the United States: Retooling America's economic engine. *McKinsey Quarterly*, 2011, pp. 1-67.

Jans, M., M. Alles, M. Vasrhelyi (2011). Processing Mining of Event Logs in Internal Auditing: A Case Study, *Proceedings of the International Symposium on Accounting Information Systems, Rome, 2011*.

Jansen, D. N., Hermanns H. , J-P Katoen (2002). A Probabilistic Extension of UML Statecharts. Specification and Verification, In *Formal Techniques for Real-Time and Fault-Tolerant systems (FTRTFT).*. Vol 2469 of LNCS. Springer-Verlag, 2002, pp. 355–374.

Jansen, W., Steenbakkers W., Jägers, H., (2007), New Business Models for the Knowledge Economy, *Gower Publishing*, 2007, 160 pp.

Joshi, S. S., V.V. Phoha (2005). Investigating Hidden Markov Models Capabilities in Anomaly Detection, *Proceedings of the 43th ACM Southeast Conference, Kennesay U.S., 2005, pp. 18 – 20.*

Kanigel, Robert. *The One Best Way: Frederick Winslow Taylor and the Enigma of Efficiency*. Cambridge: Viking Penguin, 1997. Print.

Kannan, P.K. and Proenca, J.F., (2010), Design of Service Systems under variability: research issues, *Information Systems and E-Business Management, Vol. 8, No. 1, pp. 1-11*

Karlof, C., D. Wagner (2009). Hidden Markov Model Cryptanalysis, Report UCB/CSD-03-1244, University of California, Berkeley, 21 pp.

Karmakar, U. (2005). *Service industrialization and service productivity. eBRC workshop on service innovations and new service models.* Retrieved from www.smeal.psu.edu/cdt/ ebrcpubs/papers/05wshop/karmark05.pdf

Kelly, D. J. (2004, November). Marxist Manager amidst the Progressives: Walter N Polakov and the Taylor Society. *Journal of Industrial History, 6*(2), pp. 61-75.

Kelly, Robert. "Where can $700 billion in waste be cut annually from the U.S. healthcare system?." *http://info.thomsonhealthcare.com*. Thomson Reuters, Oct. 2009. Web. 10 Apr. 2013. Path: http://img.en25.com/Web/ThomsonReuters/WASTEWHITEPAPER_FINAL11_3_09.pdf

Keynes, J. M. (1963). Economic Possibilities for our Grandchildren. In *www.econ.yale.edu*. Retrieved May 17, 2012, from http://www.econ.yale.edu/smith/econ116a/keynes1.pdfhttp://www.econ.yale.edu/smith/econ116a/keynes1.pdf

Klein B., Crawford R. G., Alchian A. A., 1978, Vertical Integration, Appropriable Rents and the Competitive Process. *Journal of Law and Economics Vol 21(2), pp. 297 – 326.*

Knab, B., A. Schliep, B. Steckemetz, B. Wichern (2003). Model-based Clustering with Hidden Markov Models and its application to financial time-series data, *Proceedings of the GfKI 2002 Between Data Science and Applied Data Analysis, Springer 2003, pp. 561 – 569.*

Kooper, M., Maes, R., Lindgreen, E. R., 2011, On the governance of information: introducing a new concept of governance to support the management of information, *International Journal of Information Management, vol 31, pp. 195 – 200.*

Kotsiantis, S., E. Koumanakos, D. Tzelepis, V. Tampakas (2006). Predicting Fraudulent Financial Statements with Machine Learning Techniques, *Lecture Notes in Artificial Intelligence, SETN 2006, Springer Verlag, Vol. 3955, pp. 538 – 542.*

Lankhorst, Marc, M Iacob, H Jonkers, L. van der Torre, H. Proper, F. Arbab, and W. Janssen. *Enterprise Architecture at Work: Modeling, Communication and Analysis*. Berlin: Springer Verlag, 2005. Print.

Levine, R., Locke, C., Searls, D., Weinberger, D., 2001, The Cluetrain Manifesto: The End of Business As Usual, Perseus Books Group, 190 pp.

Lewis, G. J., Steward, N., (2003) The measurement of environmental performance: an application of Ashby's law. *Systems Research & Behavioral Science, vol 20, no. 1, pp. 31-52*.

Li, C., G. Biswas (1999). Clustering Sequence Data using Hidden Markov Model Representation, *Proceedings of the SPIE'99 Conference on Data Mining and Knowledge Discovery Theory, pp. 14 – 21*.

Liker, J. K. (2004). *The Toyota Way*. Madison, WI: CWL Publishing Enterprises, Inc.

Lindley D., 2007,Imperfect Information for Consumers. Consumer Policy Review, vol. 17, no. 3.

Liu, Y. (2004). Markov Model-based Methods for Web User Clustering and Surfing Recommendation, *Master Thesis Toronto University, 2004.*

Liu, Y., X. Huang, A. An, G. Promhouse (2004). Clustering: Web Surfers with Probabilistic Models in a Real Application, *Proceedings of International Conference on WEB INTELLIGENCE WI 2004, pp. 761 – 765.*

Liu, Y., A. An, X Huang (2007). Personalized Recommendation with Adaptive Mixture of Markov Models, *Journal of the American Society for Information Science and Technology, Vol 58(12), 2007, pp. 1851 – 1870.*

Lock, Michael. *Operational Intelligence: Boosting Performance with "Right-Time" Business Insight*. Boston: Aberdeen Group, 2010. 1-22. Web. 10 Apr. 2013. <http://resources.idgenterprise.com/original/AST-0009180_0172-6528-RA-OperationalBI-MDL-NSP-03.pdf>.

Lohr, S. (2011). More Jobs Predicted for Machines, Not People. In *www.nytimes.com*. Retrieved May 17, 2012, from http://www.nytimes.com/2011/10/24/technology/economists-see-more-jobs-for-machines-not-people.html?_r=1

Mahoney J. 1995. The Management of Resources and the Resource of Management. *Journal of Business Research* 33, pp. 91-101

Mans, R. S., Schonenberg, M. H., Song, M., & van der Aalst, W. P. (2008). Process Mining in Healthcare – A Case Study, *Proceedings of the First International Conference on Health Informatics (HEALTHINF 2008), Portugal 2008, pp. 118 – 125..*

Mans, R., H. Schonenberg, G. Leonardi, S. Panzarsa, A. Cavallini, S. Quaglini, W. van der Aalst (2008). Process Mining Techniques: An Application to Stroke Care, *Proceedings of MIE2008 the 21st International Congress of the European Federation for Medical Informatics, Studies in Health Technology and Informatics, Vol 136.*

Manyika, J., K. Sprague, L. Yee (2009). *McKinsey: What Matters Using Technology to Improve Workforce Collaboration,* http://whatmatters.mckinseydigital.com/author/James+Manyika,+Kara+Sprague+and+Lareina+Yee/

Maruster, L., Weijters, A.J.M.M., Aalst, W.M.P. van der & Bosch, A. van den (2006). A rule-based approach for process discovery: dealing with noise and imbalance in process logs, *Data Mining and Knowledge Discovery, 13(1),2006, pp. 67-87.*

Materialise, (2011), Leuven, see http://www.materialise.com/

Matthews, M. (2006). Medicare's Hidden Adminstrative Costs: A Comparison of Medicare and the Private Sector. *The Council for Affordable Health Insurance, 1*, 1-17.

Mercado, C., Dedene, G., Peters, E., Maes, R. (2012), Towards a Dynamic, Systemic and Holistic Theory for Strategic Value Creation in ICT-based Services, in “ A Focussed

Issue on Competence Perspectives on New Industry Dynamics", ed. R. Sanchez & A. Heene, Research in Competence-Based Management, Vol 6, pp. 153 – 207, Emerald Group Publishing Limited, 2012.

Mishra, D., Heid, J.B. & Cort, S.G. (1998).l Information asymmetry and levels of agency relationships, *Journal of Marketing Research,(JMR)M 35(3), 277-295.*

Mosakowski, E. & McKelvey, B., 1997, Predicting Rent Generation in Competence-based Competition in *Competence-based Strategic Management*, Edited by A. Heene & R. Sanchez, John Wiley & Sons Ltd., 1997.

McAfee, A. (2006), Enterprise 2.0, The Dawn of Emergent Collaboration, *MIT Sloan Management Review, Vol 47(3), pp. 21 - 28*

McAfee, A. (2009). *Enterprise 2.0*. Boston, MA: Harvard Business School Publishing.

Ng, I. C. L., Wirtz, J., & Lee, K. S. (1999). The strategic role of unused service capacity. *International Journal of Service Industry Management, 10(2), 211-238.*

Ostby, K. (2005). Manufacturing Competitiveness through Robotics and Automation. In *www.robotics.org*. from http://www.robotics.org/content-detail.cfm/Industrial-Robotics-News/Manufacturing-Competitiveness-through-Robotics-and-Automation/content_id/1076

Page, R. (2002). *Hope Is Not A Strategy*. New York, NY: McGraw-Hill.

Pande, Peter S., Robert P. Neuman, and Roland R. Cavanagh. *The Six Sigma Way*. New York: McGraw-Hill, 2000. Print.

Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing, 69, pp. 167–176,*

Panuccio, A., M. Biecego, V. Murino (2002). A Hidden Markov Model-based approach to sequential data clustering, Proceedings of SSPR/SPR 2002, pp. 734 – 742.

Penrose, Edith. *The Theory of the Growth of the Firm*. New York: John Wiley & Sons, 1959. 272. Print

Peters, E. M., & Dedene, G. (2011). Business Process Discovery & Workforce Intelligence Techniques in Healthcare, *International Journal of Health Management and Information,* vol. 2, no. 1, pp. 25 - 39.

Peters, E.M.L., G. Dedene, C. Houck (2009). Business Process Discovery and Workforce Intelligence Techniques with Healthcare Applications, Proceedings of the INFORMS ORAHS 2009 Conference, Leuven Belgium 2009, http://www.econ.kuleuven.be/eng/tew/academic/prodbel/ORAHS2009/page5.htm

Peters, E.M.L., G. Dedene, C. Houck (2007). Hybrid Techniques for Business Process Discovery, *Proceedings of the INFORMS Annual Conference 2007.*

Peters, E.M.L., (2011), Workforce Productivity in Today's Service-Based Economy, Innovation Nation at Risk, Tech America Foundation, September, 2011, (http://www.techamericafoundation.org/content/wp-content/uploads/2011/09/OpenConnect090111.pdf

Peters, E.M.L., Cicchino, John (2011), Driving Administrative Cost Reductions by Improving Service Processes and Workforce Productivity, AHIP Fall Forum, November 15, 2011 Chicago, IL.

Peters, E.M.L., Payack, P.J.J. (2010), *The Paid for Option, Using Process Intelligence to Close the Healthcare Knowledge Gap, 2010, Tower Oaks Press, Cedar Park, TX.*

Peters, E. (2012). Processes are Concepts, aren't they ? *Invited Keynote for the 10th Internationl Conference on Formal Concept Analysis ICFCA 2012, Leuven.*

Peters, E.M.L.,(2012), How to Slow the Healthcare Value Leak, Forbes, May 7, 2012, (http://www.forbes.com/sites/realspin/2012/05/07/how-to-slow-the-health-care-value-leak).

Peters, E.M.L., (20102), Greater Healthcare Payer Efficiency Through Process Intelligence and Workforce Analytics, Insurance and Technology, January 20102,

(http://www.insurancetech.com/business-intelligence/greater-healthcare-payer-efficiency-thro/232301494

Pfeffer, Jeffery, and Robert I. Sutton. Hard Facts, Dangerous Half-Truths & Total Nonsense: Profiting from Evidence-Based Management. Boston: Harvard School Publishing, 2006. pp. 5-43.

Phua, C., V. Lee, K. Smith, R. Gayler (2010). A Comprehensive Survey of Data Mining-based Fraud Detection Research, *Computing Research Repository abs/1009.6119 , 2010.*

Poelmans J, Elzinga P, Viaene S, Dedene G, 2010, Curbing domestic violence: instantiating C-K theory with formal concept analysis and emergent self organizing maps, *Intelligent Systems in Accounting, Finance & Management, vol. 17, pp. 167 - 191.*

Poelmans J, Dedene G, Verheyden G, Van der Mussele H, Viaene S, Peters E, 2010, Combining business process and data discovery techniques for analyzing and improving integrated care pathways, in *Lecture Notes in Computer Science, Advances in Data Mining. Applications and Theoretical Aspects*, vol. 6171, Industrial Conference (ICDM) (Leipzig (Germany)), pp. 505 - 517,

Poelmans J., (2010), Essays on using formal concept analysis in Information Engineering, KU Leuven PhD-thesis, 2010.

Poelmans, J., Dedene, G., Verheyden, G., Peters, E., Van Der Mussele, H., Viaene, S. (2010). Concept mining and discovery in clinical pathways. INFORMS Data mining and health informatics (DM-HI) Workshop. Austin, Texas (US), 7-10 November 2010.

Poelmans J., Elzinga P., Viaene S., Dedene G., (2011), Formally analyzing the concepts of domestic violence, Expert Systems with Applications, vol. 38, no 4, pp. 3116 – 3130.

Poelmans J, Dedene G, Snoeck M, Viaene S, (2012), An iterative requirements engineering framework based on Formal Concept Analysis and C–K theory, *Expert Systems with Applications, vol. 39 (9), pp. 8115 - 8135.*

Porter, Michael E., and Elizabeth O. Teisber. *Redefining Health Care: Creating Value-based Competition on Results*. Boston: Harvard Business School Publishing, 2006. Print.

Put, F. (1989), The ER approach extended with the action concept as a conceptual modelling tool, *in Entity-Relationship Approach: a bridge to the user*, C. BATINI (ed.), Elsevier, Amsterdam, 1989, pp. 423-440

Rabiner L. (1989), A tutorial on hidden markov models and selected applications in speech recognition, *Proc. IEEE 77 (2) (1989*), pp. 257–286.

Reinhart, C. M., & Rogoff, K. S. (2011). This Time is Different: A Panoramic View of Eight Centuries of Financial Golof, *Princeton University Press, 512 pp.*

Remenyi, D. , M. Sherwood-Smith (2001). Outcomes and Benefits Modeling for Information Systems Investment, *International Journal of Flexible Manufacturing Systems, Vol 13 (2), 2001, pp. 105 – 129.*

Rhodes, Geoff. *Health Claims Processing in the US*. Colombia: TM Floyd & Company, 2006. Print.

Rifkin, J. (2005). The End of Work . In *www.spiegel.de/international*. Retrieved May 17, 2012, from http://www.spiegel.de/international/0,1518,368155,00.html

Roberts, Roger. "How Web 2.0 is changing the way we work: An interview with MIT's Andrew McAfee." *McKinsey Quarterly* (2009). Web. 10 Apr. 2013. <http://mkqpreview1.qdweb.net/How_Web_2_0_is_changing_the_way_we_work_An_interview_with_MITs_Andrew_McAfee_2468>.

Ruffa, Stephen A.(2008) Going Lean: How the Best Companies Apply Lean Manufacturing Principles to Shatter Uncertainty, Drive Innovation, and Maximize Profits. New York: AMACOM, 2008.

Rumsfeld, Donald H. (Feb. 12, 2002), US Department of Defense, News Briefing

Ruiz-Castro, J.E., R. Perez-Ocon, D. Montoro-Cazorla (2002). Algorithmic and Computational Procedures for a Markov Model in Survival Analysis, *Proceedings of Compstat 2002, Physica-Verlag, pp. 147 – 152.*

Rust, M. & Teisberg, E. O. (2006). *Redefining health care: Creating value-based competition on results.* Boston, MA: Harvard Business School Presss. ISBN: 1591397782.

Sahebi, S., F. Oroumchain, R. Khosravi (2008). Applying and Comparing Hidden Markov Model and Fuzzy Clustering Algorithms to Web Usage Data for Recommender Systems, *ADIS European Conference on Data Mining, 2008. Proceedings of, IADIS Press, Amsterdam, Netherlands, pp. 179-181.*

Sanchez, R., Heene, A. & Thomas H., 1996, *Dynamics of Competence-based competition: theory and practice in the New Strategic Management*, London, Elsevier, 1996.

SAP (2005), Adaptive Manufacturing: enabling the lean six sigma enterprise, *SAP White Paper, 22 pp.8-10, http://www.sap.com/solutions/manufacturing/pdf/WP_Adaptive_Manufactuing_Lean_Six_Sigma.pdf*

Schliep, A. B. Georgi, W. Rungsarityotin, I.G. Costa, A. Schonhuth (2004). The General Hidden Markov Model Library: Analyzing Systems with Unobservable States, *Proceedings of the Heinz-Billing-Price (2004)*

Schumpeter J. A., 1934, *The Theory of Economic Development*. Harvard University Press: Cambridge, MA.

Sinur J., 2013, *Using Process Mining for Increased Profitability on Insurance Claims: Success Snippet*, Gartner Research Blog Network, March 2013 http://blogs.gartner.com/jim_sinur/2013/03/11/using-process-mining-for-increased-profitability-on-insurance-claims-success-snippet/

Sink, D. S. *Productivity Management: Planning, Measurement and Evaluations, Control and Improvement*. Hoboken: John Wiley & Sons, 1985. Print.

Smyth, P. (1999). Clustering Sequences with Hidden Markov Models, *Advances in Neural Information Processing Systems.*

Snoeck M, Dedene G, (1998), Existence dependency: The key to semantic integrity between structural and behavioral aspects of object type, *IEEE Transactions on software engineering, vol. 24, no. 4, pp. 233 - 251.*

Spanos Y. E., Lioukas S., (2001), An examination into the causal logic of rent generation: contrasting Porter's competitive strategy framework and the resource-based perspective, *Strategic Management Journal Vol 22 pp. 907 – 934.*

Spohrer, J., Maglio, P.P., Bailey, J., gruhl, D.,(2007) Steps toward a science of service systems, *IEEE Computer, vol. 40 (1), 2007, pp. 71–78.*

Sproull, B., (2009), The Ultimate Improvement Cycle: Maximizing Profits through the Integration of Lean, Six Sigma and the theory of constraints, Productivity Press, 2009.

Srivastava, A. (2005). Discovering System Health Anomalies Using Data Mining Techniques, *Proceedings of the Joint Army Navy NASA Air Force Conference on Propulsion, 2005, 11 pp.*

Stiglitz J. E.,(1985), Information and Economic Analysis: A Perspective. *The Economic Journal, Vol 95 supplement, pp. 21 – 41.*

Stiglitz, J. E., (2000), The Contributions of the Economics of Information to twentieth Century Economics, The Quarterly Journal of Economics, Vol 115, 1441 – 1478.

Stiglitz, J. E., (2002), Information and the Change in the Paradigm in Economics, American Economic Review, Vol. 92, no. 3, pp. 460-501.

Stiglitz, J. E., (2003), Information and the Change in the Paradigm in Economics, Part I, *American Economist, Vol. 47(2), pp. 6-26.*

Sundaravaradan, N., Ramakrishnan, N. & Hanauer D. A., (2012), Factorizing Event Sequences, *IEEE Computer, Vol 45(12)*, 2012, pp. 73 – 75.

Swank, C.K. (2003), The Lean Service Machine, *Harvard Business Review, Vol.81, No.10, pp.123-129.*

Taylor, Fredrick W. *The Principles of Scientific*. Charleston, SC: BiblioBazaar, 1999. Print.

Thompson, M.D, (2011). *e-Marketing Fundamentals: the Customer Relationship Curve,* http://www.scribd.com/doc/67480019/Relationship-Curve

Turner, M. A. (2006). *Why Secret Intelligence Fails*. Dulles, VA: Potomac Books, Inc.

Teicher, S. A. (2003, July 29). White-collar jobs moving abroad. In *www.csmonitor.com*. Retrieved May 17, 2012, from http://www.csmonitor.com/2003/0729/p01s03-usgn.html

Teisberg, E., Porter, M., & Brown, G., (1994), Making competition in health care work, *Harvard Business Review*, vol 72, no. 4, pp. 131-141.

Triplett, J. E., B.P. Bostworth (2000). Productivity in the Services Sector, *American Economic Society*, *26 pp.*

Truijens, O., (2004), Towards a Theory of Information Technology. Exploiting Market Opaquness in search for Inforent. Primavera Working Series van Universiteit of Amsterdam

Tsoi, A.C, S.Zang, M. Hagenbuchner (2005). Pattern Discovery on Australian Medical Claims Data –a Systematic Approach, *IEEE Transactions on Australian Medical Claims Data – A Systematic Approach, Vol 17 (10), 2005, pp. 1420 – 1435.*

USACERL. (1994). *Evaluating Knowledge Worker Productivity: Literature Review*. In B. E. Thomas, J. P. Baron, & . (Eds.). Springfield, VA: National Technical Information Service.

Van Ark. B., R. Inklaar, R. McGuckin (2002). *"Changing Gear" Productivity, ICT and Services Industries: Europe and the United States."* Paper presented to the Brookings Workshop on Services Industry Productivity, May 17.

Van der Aalst, W., & Weijters, A. (2004). Process Mining: A Research Agenda, *Computers in Industry*, *53*(3), pp. 231-244.

Van der Aalst, W., K. van Hee, J. M. van der Werf, Kumar A., Verdonk M.C., (2009). Conceptual Model for Online Auditing, *Decision Support Systems, Vol 50(3), pp. 636 – 647.*

Van der Aalst, W. (2006). Matching observed behavior and modeled behavior : an approach based on Petri nets and integer programming, *Decision Support Systems, 42(3), 2006, pp. 1843-1859.*

Van der Aalst, W., Adriansyah, A., Alves De Medeiros, A.K., Arcieri, F., Baier, T., Blickle, T., Bose, J.C., Brand, P.C.W. van den, Brandtjen, R., Buijs, J.C.A.M., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., Leoni, M. de, Delias, P., Dongen, B.F. van, Dumas, M., Dustdar, S., Fahland, D., Ferreira, D.R., Gaaloul, W., Geffen, F. van, Goel, S., Gunther, C.W., Guzzo, A., Harmon, P., Hofstede, A.H.M. ter, Hoogland, J., Ingvaldsen, J.E., Kato, K., Kuhn, R., Kumar, A., La Rosa, M., Maggi, F.M., Malerba, D., Mans, R.S., Manuel, A., McCreesh, M., Mello, P., Mendling, J., Montali, M., Motahari Nezhad, H., Muehlen, M. zur, Munoz-Gama, J., Pontieri, L., Ribeiro, J.T.S., Rozinat, A., Seguel Peréz, H., Seguel Pérez, R.E., Sepúlveda, M., Sinur, J., Soffer, P., Song, M.S., Sperduti, A., Stilo, G., Stoel, C., Swenson, K., Talamo, M., Tan, W., Turner, C., Vanthienen, J., Varvaressos, G., Verbeek, H.M.W., Verdonk, M.C., Vigo, R., Wang, J., Weber, B., Weidlich, M., Weijters, A.J.M.M., Wen, L., Westergaard, M. & Wynn, M.T. (2012). Process Mining Manifesto, *Business Process Management Workshops (BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I), (Lecture Notes in Business Information Processing, 99, pp. 169-194). Berlin: Springer*.

Vargo, S., Maglio P. P., Akaka M. A., (2008), On value and value co-creation: A service systems and service logic perspective, *European Management Journal Vol. 26, pp. 145-152*

Verheyden, G., Van der Mussele, H., Provè, A.M., Verkinderen, L., De Schepper, A., Vermeulen, P.B., Buelens, G., De Pooter, C., Dirix, L.Y. & Van Dam, P. (2007), Stepwise implementation of an evidence-based specialist breast care nurse model in a Belgian breast cancer cliinic: Impact on patient's satisfaction, a prospective study, *EJC Supplements, Vol 5 (4), pp. 416 - 448.*

Verner, L., (2004) The Challenge of Process Discovery, *BPRTrends White Paper, May 2004, also at http://www.businessprocesstrends.com/publicationfiles/05-04%20WP%20Process%20Discovery%20-%20Verner1.pdf*

Verschuren, P., and Doorewaard, H,. *Designing a Research Project*. 2nd ed. The Netherlands: Eleven International Publishing, 2010. 65 - 81,. Print.

Viaene S., Dedene G. & Derrig R.,(2005) Auto claim fraud detection using Bayesian learning neural networks, *Expert systems with applications, Vol. 29 (3), pp. 653 – 666*

Viaene, S., Van den Bunder, A.,(2011), The secrets to managing business analytics projects, *MIT Sloan Management Review, Vol. 53(1), pp 65 – 69*

Vidal, E., F. Thollard, C. de la Higuera, F. Casacuberta, R.C. Carrasco (2005). Probabilistic Finite-State Machines – Part 1, , *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 27(7), 2005, pp. 1013 – 1025*

Vidal, E., F. Thollard, C. de la Higuera, F. Casacuberta, R.C. Carrasco (2005). Probabilistic Finite-State Machines – Part 2, *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 27(7), 2005, pp. 1026 – 1039.*

Ware, J. P., & Grantham, C. E. (2007). Knowledge Work and Knowledge Workers. *WIRED West Michigan*.

Wille, R., 2005, *Formal Concept Analysis as Mathematical Theory of Concepts, inn Formal Concept Analysis,* Lecture Notes in Artificial Intelligence 3626, pp. 1 - 33, Springer-Verlag Berlin Heidelberg 2005.

Whitman, L., B. Huff (2001). On the Use of Enterprises Models, *International Journal of Flexible Manufacturing Systems, Vol 13(2), 2001, pp. 195 – 208.*

Womack, J. P. & Daniel T. J., (1996), *Lean Thinking*. New York: Simon & Schuster, 1996.

Wu F., Chui, I-Han (2005). Prediction of the intention of purchase of the user surfing on the Web using Hidden Markov Model, *Proceedings of ICSSSM International Conference on Services Systems and Services Management*, pp. 387 – 390.

Yin, Robert K. "Design and Methods." *Case Study Research*. 4th ed. Vol. 5. Los Angeles: SAGE Publications, 2009. Print.

Zeng, Y., J. Garcia-Frias (2005). A novel HMM-based clustering algorithm for the analysis of gene express time-course data, *Computational Statistics and Data Analysis,2006, pp. 2472-2494*