

PRODUCT STRUCTURE ONTOLOGY TO SUPPORT SEMANTIC SEARCH IN  
MANUFACTURING REQUIREMENTS MANAGEMENT

NOOR NADHIYA BINTI MOHAMMAD

UNIVERSITI TEKNOLOGI MALAYSIA

PRODUCT STRUCTURE ONTOLOGY TO SUPPORT SEMANTIC SEARCH IN  
MANUFACTURING REQUIREMENTS MANAGEMENT

NOOR NADHIYA BINTI MOHAMMAD

A thesis submitted in fulfilment of the  
requirements for the award of the degree of  
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems  
Universiti Teknologi Malaysia

MARCH 2010

## ABSTRACT

Automated information extraction of 2D CAD engineering drawing ensures more accurate extracted information of product manufacturing requirements. However, an occurring problem from the automation process is the existence of heterogeneous terms in the engineering drawing. The problem can be solved by formalizing the knowledge in this domain. Therefore, a dynamic ontology called the Product Structure Ontology (PSO) has been developed. The process of developing the PSO is extended from Noy and McGuiness's methodology. It consist of nine steps; determining ontology domain and scope, considering ontology reuse, enumerating important terms, defining classes and class hierarchies, creating instances of classes, designing anatomy and database schema, creating an evidence code, creating an annotation and developing the PSO artifacts. With the aim of enabling the PSO to be reused and extended, the PSO artifacts such as website, browser, database and documentations have been shared on the World Wide Web (WWW). In order to test the applicability and usage of the PSO in digital engineering drawing extraction, Semantic Ontology-based Searching Algorithm (SOBSA) has been developed. The SOBSA entails the use of PSO to overcome the limitation of keyword-based search by using information content approach and considering the three types of ontology relationships; subsumption, meronymy and association. The performance of SOBSA has been tested by using real digital engineering drawing and evaluated by using the standard information retrieval measures which are precision, recall and F1. The experimental evaluation demonstrates that the query search SOBSA improves the accuracy of the query retrieval results compared to conventional method. Besides that, the performance of SOBSA also depends on the type of relationship used and the completeness of the knowledge base.

## ABSTRAK

Pengekstrakan maklumat daripada lukisan kejuruteraan 2D CAD secara automatik dapat menjamin maklumat yang diekstrak lebih tepat untuk keperluan pembuatan produk. Namun, masalah yang timbul daripada proses automasi ini ialah wujudnya kepelbagaian istilah di dalam lukisan kejuruteraan. Ianya boleh diselesaikan dengan memformalkan pengetahuan yang terdapat di dalam bidang ini. Oleh itu, satu ontologi dinamik yang dinamakan Ontologi Struktur Produk (PSO) telah dibangunkan. Proses membangunkan PSO dibuat dengan menggunakan metodologi Noy dan McGuinness yang telah ditambahbaik. Methodologi ini mempunyai sembilan langkah; menentukan bidang ontologi dan ruang lingkungannya, mempertimbangkan penggunaan semula ontologi, menyenaraikan satu per satu istilah-istilah penting, mentakrifkan kelas-kelas dan hirarkinya, mewujudkan contoh-contoh kelas, mereka bentuk anatomi dan skema pangkalan data, mencipta satu kod bukti, mewujudkan anotasi dan membangunkan sumber-sumber PSO. Untuk membolehkan PSO digunakan semula dan diperkembangkan, satu rangka kerja pembangunan PSO berserta artifaknya seperti laman web, pelayar internet, pangkalan data dan dokumentasi telah dibangunkan dan diterbitkan supaya ianya boleh dikongsi di dalam Jaringan Sejagat Sedunia (WWW). Untuk menguji kebolegunaan PSO yang dibangunkan dalam pengekstrakan lukisan kejuruteraan, satu Algoritma Pencarian Berasaskan Ontologi (SOBSA) telah dibangunkan. SOBSA melibatkan penggunaan ontologi untuk mengatasi kekurangan pencarian berasaskan kata kunci dengan menggunakan pendekatan kandungan maklumat dan mempertimbangkan tiga jenis hubungan ontologi iaitu pengkelasan, hubungan bahagian kepada keseluruhan dan perkaitan. Prestasi SOBSA telah diuji menggunakan lukisan kejuruteraan berdigital yang diambil daripada sampel sebenar dan dinilai ketepatannya menggunakan ukuran seperti ketepatan, dapatan semula dan F1. Penilaian ujikaji menunjukkan bahawa hasil yang diperolehi bagi pertanyaan menggunakan SOBSA adalah lebih tepat berbanding dengan kaedah semasa. Selain itu, prestasi SOBSA juga bergantung kepada jenis hubungan yang digunakan dan kesempurnaan pangkalan pengetahuannya.

## TABLE OF CONTENTS

<b>CHAPTER</b>	<b>TITLE</b>	<b>PAGE</b>
	<b>DECLARATION</b>	ii
	<b>DEDICATION</b>	iii
	<b>ACKNOWLEDGEMENTS</b>	iv
	<b>ABSTRACT</b>	v
	<b>ABSTRAK</b>	vi
	<b>TABLE OF CONTENTS</b>	vii
	<b>LIST OF TABLES</b>	x
	<b>LIST OF FIGURES</b>	xii
	<b>LIST OF ABBREVIATIONS</b>	xiv
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Background	1
	1.2 Current Method in Formalizing Knowledge	3
	1.3 Challenges of Formalizing Knowledge	4
	1.4 Statement of Problem	6
	1.5 Objectives of the Study	7
	1.6 Significance and Scope of the Study	7
	1.7 Organization of the Thesis	9

<b>2</b>	<b>LITERATURE REVIEW</b>	<b>10</b>
2.1	Introduction	10
2.2	Product Manufacturing Requirements	11
2.3	2D CAD Engineering Drawing and Product Structure	14
2.4	Extraction Algorithm	18
2.5	Knowledge Formalization	20
2.6	Ontology	26
2.7	Semantic Similarity Measures	29
2.8	Trend and Tendencies	31
2.9	Summary	32
<b>3</b>	<b>RESEARCH METHODOLOGY</b>	<b>33</b>
3.1	Introduction	33
3.2	The Framework of the Study	34
3.3	Data Sources	39
3.4	Instrumentation and Result Analysis	41
3.4.1	Hardware and Software Requirements	41
3.4.2	Case Study and Implementation	41
3.4.3	Performance Measurement	42
3.5	Summary	43
<b>4</b>	<b>THE DEVELOPMENT AND DESIGN OF PRODUCT STRUCTURE ONTOLOGY</b>	<b>44</b>
4.1	Introduction	44
4.2	Product Structure Ontology (PSO)	48
4.3	Features of PSO	54
4.4	The Proposed PSO	56
4.5	PSO Anatomy and Schema	59
4.6	PSO Annotation	61
4.7	PSO Evidence	64

4.8	The Shared PSO Resource	65
4.8.1	The PSO Website	65
4.8.2	The PSO Browsers	66
4.8.3	The PSO Database	67
4.8.4	The PSO Documentation	67
4.9	Summary	68
<b>5</b>	<b>SEMANTIC ONTOLOGY-BASED SEARCHING ALGORITHM</b>	<b>69</b>
5.1	Introduction	69
5.2	Semantic Similarity Searching Algorithm (SSSA)	71
5.2.1	Information Content Approach	71
5.2.2	Semantic Ontology-Based Searching Algorithm (SOBSA)	73
5.3	The PSO-EDEX Framework	76
5.4	Statistical Measurement	78
5.5	Results and Discussion	79
5.6	Summary	104
<b>6</b>	<b>CONCLUSION</b>	<b>105</b>
6.1	Concluding Remarks	105
6.2	Contributions	108
6.3	Future Works	109
6.4	Summary	110
	<b>REFERENCES</b>	<b>111</b>

## LIST OF TABLES

<b>TABLE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	A table of crosses represents a very simple and often used data type	21
3.1	The features comparison between PSO, MASON and PRONTO	40
4.1	Attributes of PSO terms	62
4.2	Size of PSO data	68
5.1	The comparison of the query results between semantic ontology-based and keyword-based searching using the standard information retrieval measures there are recall (R), precision (P) and F1	81
5.2	The evaluation of precision (P), recall (R) and F1 of the query results that have been retrieved using 50% PSO	84
5.3	The evaluation of precision (P), recall (R) and F1 of the query results that have been retrieved using 75% PSO	85
5.4	The evaluation of recall, precision and F1 of the query results that have been retrieved using 75% PSO	86
5.5	The precision values of the query results that have been retrieved using 50%, 75% and 100% of PSO at standard recall	91
5.6	The evaluation of recall, precision and F1 of the query results that have been retrieved using subsumption relationship	93



5.7	The evaluation of recall, precision and F1 of the query results that have been retrieved using meronymy relationship	93
5.8	The evaluation of recall, precision and F1 of the query results that have been retrieved using association relationship	94
5.9	The evaluation of recall, precision and F1 of the query results that have been retrieved using combination of subsumption and meronymy relationships	94
5.10	The evaluation of recall, precision and F1 of the query results that have been retrieved using combination of subsumption and association relationships	95
5.11	The evaluation of recall, precision and F1 for the query results that have been retrieved using combination of meronymy and association relationships	95
5.12	The evaluation of recall, precision and F1 for the query results that have been retrieved using combination of subsumption, meronymy and association relationships	96
5.13	Variance of recall for each types of relationship(s) used	98
5.14	Variance of precision for each types of relationship(s) used	99
5.15	Variance of F1 for each types of relationship(s) used	100
5.16	Statistical significance of differences between relationship(s) for recall	101
5.17	Statistical significance of differences between relationship(s) for precision	102
5.18	Statistical significance of differences between relationship(s) for F1	103

## LIST OF FIGURES

<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE</b>
2.1	Screenshot of the FuturePlan Requirements Management Software	13
2.2	2D CAD engineering drawing	15
2.3	AutoCAD software screenshot	17
2.4	Example of concept lattice. The information contained in the formal context is preserved. The concept lattice is the basic for further data analysis.	22
2.5	Example of semantic networks visualized	23
2.6	Example of ontology visualized	25
2.7	Protégé ontology editor screenshot	28
3.1	The overview of the study framework	37
4.1	The framework adopted to develop the PSO	50
4.2	PSO design in Protégé-owl editor	52
4.3	The PSO in DAG representation	53
4.4	The fragment of PSO classes and individuals	55
4.5	Overview of the PSO's classes, individuals and relationships	56
4.6	Part of "Product Specific Knowledge"	57
4.7	Part of "Industry Specific Knowledge"	58
4.8	Part of "Generic-term"	59
4.9	Anatomy of the PSO terms	60
4.10	Database structure scheme	63

4.11	Screenshot of the PSO web site	66
4.12	Screenshot of the PSO browser	67
5.1	The semantic ontology-based searching algorithm	74
5.2	An example of semantic similarity score results for the <i>end trim</i> query engineering term	75
5.3	The flowchart of the PSO-EDEx framework	77
5.4	The $A \cap (B \cap C)$ is the shaded region	79
5.5	The recall of the query results for the ontology-based semantic searching and keyword-based searching using 50 sheets digital engineering drawing	82
5.6	The precision of the query results for the ontology-based semantic searching and keyword-based searching using 50 sheets digital engineering drawing	82
5.7	The F1 of the query results for the ontology-based semantic searching and keyword-based searching using 50 sheets digital engineering drawing	83
5.8	Precision-recall values for representative the effectiveness of SOBSA using 50% of PSO	87
5.9	Precision-recall values for representative the effectiveness of SOBSA using 75% of PSO	87
5.10	Precision-recall values for representative the effectiveness of SOBSA using 100% of PSO	88
5.11	The average of interpolated precision-recall for different percentage of PSO	92
5.12	The average F1 for 5 query results that have been retrieved using different percentage of PSO	92
5.13	The average of the query results that have been retrieved using different types of relationship.	97

**LIST OF ABBREVIATIONS**

2D	-	Two-Dimensional
3D	-	Three-Dimensional
A&RM	-	Automated and Robotised Manufacturing
AI	-	Artificial Intelligence
ALPS	-	A Language for Process Specification
APMSSB	-	APM Springs Sdn Bhd
BOM	-	Bill-of-Material
CAD	-	Computer Aided Design
CAM	-	Computer-Aided Manufacturing
CCO	-	Cell-Cycle Ontology
CPU	-	Central Processing Unit
DAG	-	Directed Acyclic Graph
DAML+OIL	-	DARPA Agent Markup Language + Ontology Interchange Language
DOLCE	-	Descriptive Ontology for Linguistic and Cognitive Engineering
DXF	-	Drawing Exchange Format
EDEx	-	Engineering Drawing Extraction
EF	-	Experience Feedback
ERP	-	Enterprise Resource Planning
FCA	-	Formal Concept Analysis
FMA	-	Foundational Model of Anatomy
GO	-	Gene Ontology
GOC	-	Gene Ontology Consortium

GOLD	-	General Ontology for Linguistic Description
I/O	-	Input/Output
IDE	-	Inferred By Domain Expert Statement
IFRM	-	Intelligent Feature Recognition Methodology
IGES	-	Initial Graphics Exchange Specification
ITD	-	Inferred From Technical Documentation
JSP	-	Java Server Pages
KIF	-	Knowledge Interchange Format
KL-ONE	-	Knowledge Language One
MASON	-	Manufacturing's Semantic Ontology
MGI	-	Mouse Genome Informatics
MLE	-	Maximum Likelihood Estimation
MPS	-	Master Production Schedule
MRD	-	Market Requirements Document
MRMS	-	Manufacturing Requirement Management System
MRP	-	Material Requirements Planning
NAS	-	Non-Traceable Author Statement
ND	-	No Engineering Data Available
NIC	-	Network Interface Card
NIST	-	National Institute of Standards and Technology
OCML	-	Operational Conceptual Modelling Language
OFS	-	Ontological Filtering System
OWL	-	Web Ontology Language
PC	-	Personal Computer
PDF	-	Portable Document Format
PRONTO	-	Product Ontology
PSL	-	Process Specification Language
PSO	-	Product Structure Ontology
RAM	-	Random Access Memory
RDBMS	-	Relational Database Management System
RDFS	-	Resource Description Framework Schemas
SGD	-	Saccharomyces Genome Database
SOBSA	-	Semantic Ontology-based searching Algorithm

SSSA	-	Semantic Similarity Searching Algorithm
SWEET	-	Semantic Web for Earth and Environmental Terminology
TAS	-	Traceable Author Statement
UOM	-	Unit of Measures
WAVE	-	Web Analysis and Visualization Environment (WAVE)
WIP	-	Work in Process
WWW	-	World Wide Web
XML	-	Extensible Markup Language

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 Background**

Product structure provides important information that has become a part of product manufacturing requirements in a real manufacturing field. Product structure is a hierarchical classification of items that form a product, and also shows the product complexity. It illustrates the material, components, parts, sub-assemblies and other items in a hierarchical structure that represents the grouping of items on an assembly drawing or the grouping of items that come together at a stage in the manufacturing process. With product structure, the understanding of the components which compose a product as well as their attributes can be represented. Product structure is normally presented on a Bill-of-Material (BOM) or a parts list or a variant specification. A BOM is a formally structured list for an object (semi-finished or finished product) which lists all the component parts of the object with the name, reference number, quantity and unit of measure of each component in any product instance that may be manufactured. The engineering BOM normally lists items according to their relationships with parent product as represented on assembly drawings. This study attempts to formalize knowledge related to product structure in engineering manufacturing area.

Knowledge is often held as free text and a set of vocabulary definitions that capture a community's knowledge of a domain. Ontology is now in widespread use as a means of formalizing domain knowledge. In manufacturing domain, there has been several works that used ontology to formalize knowledge of manufacturing requirements especially those which involve product structure. For example, Product Ontology (PRONTO: Vegetti *et al.*, 2005) defines concept, relation among them and axioms to be applied in the complex product modeling domain. It is primarily related with complex product structure which involves associated hybrid structures (combining composition and decomposition types of operation) to end product like in some food (milk and meat) and petrochemical industries. A study which engages ontology in order to represent product has also been done by Lee *et al.* (2006). The study reports the effort to build an operational product ontology system for government procurement services which is designed to serve as a product ontology knowledge base. Thevenot *et al.* (2006) presents their approach to retrieve and reuse relevant information when redesigning product. They used heuristics and shared ontological component information and proposed a framework to capture, store, retrieve, reuse and represent information for product family redesign using ontology, graph query, formal concept analysis, commonality assessment and genetic algorithm-based optimizer. As a reference to utilize ontology in manufacturing field, Manufacturing's Semantic Ontology (MASON: Lemaignan *et al.*, 2006) aimed to draft a common semantic net in this domain. Their paper discusses the use of ontology for data formalization and sharing, and also shows how adequate ontologies are especially in manufacturing environment. Meanwhile, an innovative ontology-based approach called Ontological Filtering System (OFS: Lepratti, 2006) proposes to formalize natural language contents in a systematic way. They present a concept for improving the human-machine interaction by means of an innovative procedure for the processing of natural language instruction.

Ontology is famous as means of formalizing domain knowledge. One of the most important roles of ontology is that it provides a vocabulary of terms and some specification of their meaning. This includes definitions and an indication of how concepts are inter-related which collectively impose a structure on the domain and constrain the possible interpretations of terms. The following sections in this chapter



discuss the motivation and challenges involved in formalizing knowledge. Then, the current methods in formalizing knowledge will be presented. This is followed by the problem to be solved in this study. The next section will describe the aim of this research as well as the summary of the research objectives. Further in this chapter, the scope of this study will be presented and finally the overview of the organization of this thesis will be illustrated.

## 1.2 Current Methods in Formalizing Knowledge

Generally, there are three basic groups of methods to formalize knowledge:

- (i) Formal Concept Analysis (FCA) is a theory of concept formation and conceptual classification. It is based on the mathematical order theory and the theory of complete lattices (Ganter and Wille, 1999; Wille, 1992). Several researchers have further elaborated FCA in different domains like geographical domain (Kavouras and Kokla, 2002), medicine domain (Spangenberg and Wolff, 1999), biology domain (Ganter and Wille 1989), sociology domain (Ganter and Wille 1989) and information and computer science (Priss, 1999; Deogun *et al.*, 1998; Faid *et al.*, 1997; Schmitt and Saake, 1997; Kent and Neuss, 1995).
- (ii) A semantic network or net is a graphic notation for representing knowledge in patterns of interconnected nodes and arcs. Computer implementations of semantic networks were first developed for artificial intelligence and machine translation, but earlier versions have long been used in philosophy, psychology and linguistics. Example of works are in philosophy domain (Ceccato, 1961; Margaret, 1961), psychology domain (Hendler, 1992; Shastri, 1992; Fahlman, 1979; Newell and Simon, 1972; Quillian, 1966; Groot, 1965)

and linguistics domain (Brachman, 1979; Woods, 1975; Klein and Simmons, 1963; Tesnière, 1959).

- (iii) Ontology-based approach is a hierarchically structured set of concepts describing a specific domain of knowledge that can be used to create a knowledge base. Several works in different domain used ontology-based approach for knowledge formalization like in bioinformatics domain (Aranguren *et al.*, 2008; Gary *et al.*, 2005; The Gene Ontology Consortium, 2001), linguistic domain (Ferrario and Guarino, 2008), business domain (Mustafa *et al.*, 2003; Gangemi, 2001) and manufacturing domain (Foguem *et al.*, 2008; Ashin and Yoshizu, 2001). The interest for the use of ontology is that it is not only consistent but is also upward compatible and easily modifiable.

### **1.3 Challenges of Formalizing Knowledge**

There are several challenges in formalizing knowledge specifically in manufacturing requirements domain. Designing the ontology, performing technical analysis about domain concept and creating annotation are several of the main challenges. There are not many ontology design released by researcher especially in product manufacturing requirement knowledge. However, several methodologies for ontology building have been proposed and can be used as a guide to create ontology (Darlington and Culley, 2008; Noy and McGuinness, 2001; Uschold and Gruninger, 1996). Performing technical analysis of what exactly these concepts are need a careful analysis and is very time consuming. Although there are several works that involved ontology in manufacturing field, it seems that there is no suitable ontology that exists in this domain to be reused or extended. PRONTO present concepts related with the complex product structure which focuses on non-atomic raw materials that are used to make certain final products in meat industry. Thus, PRONTO is not suitable to be used because this study considers discrete

manufacturing industries where products are always fabricated by putting parts together (composition) in assembly process. The knowledge provided by MASON is only more suitable to be used for manufacturing product's cost estimation during design phase. Product information defined by its manufacturing concept can only be used to determine the cost of producing the order and the price to be quoted. Thus, MASON ontology is not suitable to be used because the aim of this research is to complete the creation of product structure which is the first step on the back of the cost estimation during design phase. Moreover, these works does not cover the heterogeneity of terms problem in manufacturing requirements domain. Unlike bioinformatics field, research for annotations in product manufacturing requirement fields had not been done and computational tool for this work has not yet been developed. Therefore, the work to identify any exact matches and relationships between vocabulary of terms and engineering products in order to create annotation is not an easy task.

The second challenge in this study is to apply ontology since there are not many researches that especially apply ontology in manufacturing requirements domain. For example, Two-Dimensional (2D) Computer Aided Design (CAD) engineering drawing information extraction involves table extraction and process identification. The challenges are to standardize the terms used in product structure and to identify the synonymous term from the information extraction process. This process has a high possibility to produce synonymous term which involved product manufacturing requirements information contained in the 2D CAD engineering drawing. An identification of these terms is very important in order to create accurate and precise product structure. However, to identify these synonymous terms using ontology, an appropriate searching technique is needed to find similar terms. The study of searching technique is also a challenge since the search result seriously relate to semantic issue. This issue must be considered to ensure the effectiveness of the information extraction process that used ontology as the main component.

#### **1.4 Statement of Problem**

The problem on formalizing knowledge specifically in manufacturing requirements domain to be studied can be described as follows:

“Given heterogeneous products and terms in manufacturing requirements domain, the challenge is to formalize the knowledge using ontology and create a simple development methodology and design which will enable the ontology to be reused and extended and at the same time, applies the ontology in 2D CAD engineering information extraction in order to show the applicability and usage of the developed ontology”.

This study assists in solving the problem of formalizing knowledge in manufacturing requirements domain focus in product structure. The first factor that has to be considered to produce the best solution for this problem is by developing ontology to standardize the engineering terms in this domain and annotation to analyze the usage of shared terms between products. The second factor that has to be considered in order to enable the ontology to be reused and extended is by developing ontology development methodology and prepares ontology artifacts that are shareable and publicly available on the World Wide Web (WWW). Finally, the third factor that has to be considered is by developing a searching algorithm in order to apply the ontology in 2D CAD engineering information extraction. This searching algorithm can be used to identify the synonymous term and subsequently helps in complementation and creation of product structure. Consideration of these factors is expected to formalize manufacturing requirements knowledge in this domain with feature to be reused and extended and furthermore lead to correctness and comprehensiveness of the ontology in this domain.

## **1.5 Objectives of the Study**

The goal of this study is to formalize knowledge in manufacturing requirements domain with reused and extended features using ontology. In order to realize this goal, several objectives must be achieved:

- (i) To formalize manufacturing requirements knowledge by developing Product Structure Ontology (PSO) and annotation in order to standardize the engineering terms and analyze the usage of shared terms between products in this domain.
- (ii) To share the ontology in manufacturing requirements domain by developing PSO development framework and artifact such as website, browser, database and documentations in order to enable the ontology to be reused and extended.
- (iii) To test the applicability and usage of the proposed ontology by developing Semantic Ontology-Based Searching Algorithm (SOBSA) with the aim of applying the ontology in 2D CAD engineering information extraction.

## **1.6 Significance and Scope of the Study**

In this study, ontology will be used to formalize manufacturing requirements domain specifically involving terms in product structure. The reason to use ontology is that the ontology can give formal representation of a set of concepts within a domain and the relationships between those concepts. This formalization of knowledge is needed to solve the problems of heterogeneity of product and terms in product design. Therefore, we come out with a dynamic ontology called PSO in which the vocabulary can be controlled and standardized, making it suitable to be

applied to all automotive spring products. Then, annotation was designed to solve the heterogeneity of product in product configuration. The proposed ontology will be applied in automated extraction of 2D CAD engineering drawing information to enable the manufacturing requirements to be identified effortlessly and accurately. Thus, it can reduce the implementation cost and time, reduce human error and standardize the engineering terms. Furthermore, product structure can be created in an easier manner. The product structure or also known as Bill-of-Material (BOM) contains product attributes, materials, components, unit of measures (UOM) and relationship between them. The product structure is needed to support manufacturing lifecycle such as production, testing and maintenance.

2D CAD engineering drawing and technical reports from APM Springs Sdn Bhd (APMSSB: <http://www.apm-automotive.com>) are used to collect the dataset for this research. Besides that, information from engineering manufacturing books and websites related to this domain are also used. Then the data is analyzed to form formal knowledge in manufacturing requirements domain called PSO. The proposed PSO methodology is introduced to explain the detail steps involved in PSO development. The PSO artifacts called PSO category and PSO anatomy are also presented to make sure PSO can be extended and reused, duplicated into other products as well as its annotation being made in a comprehensive and complete manner. In addition, the PSO resources are provided which include the PSO website, PSO browser, PSO database and PSO documentation. PSO website provides information about PSO, downloadable versions of the ontology and hyperlink to PSO browser. The Java-based PSO browser provides interface to the ontology that allows user to use PSO tree to query the database. It displays the hierarchy of product structure, information of terms and the annotation of the terms. PSO database give information about PSO database design and schema while PSO documentation provide general information to guide users to understand and using PSO. Lastly, the PSO will be applied in Engineering Drawing Extraction (EDEX) to test the correctness and comprehensiveness of the ontology in this domain. EDEX is a tool that has been developed to extract elements specifically tables and labels in digital engineering drawing (Ahamad et al., 2009).

## 1.7 Organization of the Thesis

A general content description of the subsequent chapters in this thesis is given as follows:

- (i) Chapter 1 describes the challenges, current methods, problems, objectives, scope and significance of the study.
- (ii) Chapter 2 begins with the introduction on product manufacturing requirements followed by 2D CAD engineering drawing and product structure. Reviews on extractor algorithm, knowledge formalization, ontology and semantic similarity measure are also described. The last sections of this chapter will present the trend and tendencies related to this study.
- (iii) Chapter 3 begins with a brief review of the proposed ontology development framework, followed by detailed descriptions of hardware and software requirements, data sources, testing and analysis procedures and performance measurement used.
- (iv) Chapter 4 gives a brief overview on PSO design and implementation including PSO category, anatomy, schema, annotation and its evidence and shared resources.
- (v) Chapter 5 describe the Semantic Ontology-Based Searching Algorithm (SOBSA) that is developed to applying the PSO in digital 2D CD engineering drawing extraction.
- (vi) In Chapter 6, the conclusion of the study and the achieved results to date is presented. The contributions and future works of the study are also described.

## REFERENCES

- Advanced Computer Solution (2008). Caddie. <http://www.caddiesoftware.com/>.
- Ahamad M. K., Tap M. A., Majid A. A., Samsudin, Z. F., Othman M. R, and Rahim M. S. (2009). Patent submitted to MyIPO - Automatic Extraction of Tables and Labels from 2D CAD Engineering Drawing. Malaysia: MyIPO.
- Ahmad, N. and Haque, A. F. M. A. (2001). Manufacturing Feature Recognition of Parts using DXF Files. *Proceedings of the 4th International Conference on Mechanical Engineering*. 26-28 December. Dhaka, Bangladesh: Unpublished. 111-115.
- APM Springs Sdn. Bhd. (2006a). Manufacturing Process Standard. Technical Report. APM Spring Sdn. Bhd., Malaysia.
- APM Springs Sdn. Bhd. (2006b). Machine Capacity and Capability. Technical Report. APM Spring Sdn. Bhd., Malaysia.
- Aranguren, M. E., Antezana, E., Kuiper, M. and Stevens, R. (2008). Ontology Design Patterns for Bio-ontologies: a case study on the Cell Cycle Ontology. *BMC Bioinformatics*. 9(5), 1-13.
- Ashin, T. and Yoshizu, H. (2006). Formalization of Material Property Data Analysis with Web Ontology. *Proceedings of the PV 2005, Ensuring Long-term Preservation and Adding Value to Scientific and Technical Data*. 21-23 November. Edinburgh, UK: UKOLN. <http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-papers/031.pdf>.
- Autodesk (2009). AutoCAD. <http://usa.autodesk.com/>.
- Bard, J., Rhee, S. Y. and Ashburner, M. (2005). An Ontology for Cell Types. *Genome Biology*. 6(2), R21.



- Batres, R., West, M., Leal, D., Price, D., Masaki, K., Shimada, Y., Fuchino, T. and Naka, Y. (2007). An Upper Ontology based on ISO 15926. *Computers and Chemical Engineering*. 31(5-6), 519-534.
- Bioinformatics Core Facility (2003). Disease Ontology. <http://diseaseontology.sourceforge.net/>.
- BioPathways Consortium (2005). BioPAX. <http://www.biopax.org/>.
- Brachman, J. R. (1979). *On the Epistemological Status of Semantic Networks*. London, UK: Academic Press.
- Brandt, S. C., Morbach, J., Miatidis, M., Theißen, M., Jarke, M. and Marquardt, W. (2008). An Ontology-based Approach to Knowledge Management in Design Processes. *Computers and Chemical Engineering*. 32(1-2), 320-342.
- California Institute of Technology (2008). SWEET. <http://sweet.jpl.nasa.gov/>.
- Catron, B. and S. Ray, S. (1991). ALPS - a Language for Process Specification, *International Journal of Computer. Integrated. Manufacturing*. 4(1), 105-133.
- Ceccato, S. and Maretta, E. (1961). *Linguistic Analysis and Programming for Mechanical Translation*. New York, USA: Gordon and Breach.
- Chen, R. -S., Lu, K. -Y., Yu, S. -C., Tzeng, H. -W. and Chang, C. C. (2003). A Case Study in the Design of BTO/CTO Shop Floor Control System. *Information and Management*. 41(1), 25-37.
- Cicek, A. and Gulesin, M. (2007). A Part Recognition based Computer Aided Assembly System. *Computer in Industry*. 58(1), 733-746.
- Crosby, M. A., Goodman, J. L., Strelets, V. B., Zhang, P., Gelbart, W. M. and The FlyBase, Consortium (2007). FlyBase: genomes by the dozen. *Nucleic Acids Research*. 35(suppl\_1), D486-491.
- Cuffaro, D. (2006). *Process, Material and Measurement: All the Details Industrial Designers Need to Know But Can Never Find It*. Massachusetts, Sydney: Rockport Publisher.
- Darlington, M. J. and Culley, S. J. (2008). Investigating Ontology Development for Engineering Design Support. *Advanced Engineering Informatics*. 22(1), 112-134.
- Davis Business Systems (1997). BS1 Manufacturing Software. <http://www.dbsonline.com/>.

- Dean, P. R., Xue, D. and Tu, Y. L. (2009). Prediction of Manufacturing Resource Requirements from Customer Demands in Mass-customisation Production. *International Journal of Production Research*. 47(5), 1245-1268.
- Deogun, J. S., Raghavan, V. V. and Sever, H. (1998). Association Queries and Formal Concept Analysis. *Proceedings of the 6th International Workshop on Rough Sets, Data Mining and Granular Computing*. 23-28 October. North Carolina, USA: Elsevier Science Publishing. <http://yunus.hacettepe.edu.tr/~sever/rncar.ps>.
- Elgh, F. (2007). Modelling and Management of Manufacturing Requirements in Design Automation System. In Loureiro, G., and Curran, R. eds. *Complex System Concurrent Engineering*. London, UK: Springer. 321-328.
- Elgh, F. (2008). Supporting Management and Maintenance of Manufacturing Knowledge in Design Automation Systems. *Advance Engineering Informatics*. 22(4), 445-456.
- Fahlman, S. E. (1979). NETL: A System for Representing and Using Real-World Knowledge. Massachusetts, USA: MIT Press.
- Faid, M., Missaoui, R. and Godin, R. (1997). Mining Complex Structures using Context Concatenation in Formal Concept Analysis. *Proceeding of the 2nd International KRUSE Symposium*. August 11-13. Vancouver, Canada: John Wiley and Son. 45-59.
- Ferrario, R. and Guarino, N. (2008). Towards an Ontological Foundations for Services Science. In: Fensel, D. and Traverso, P. *Proceedings of Future Internet Symposium*. Vienna, Austria: Springer Verlag. 152-169.
- Foguem, B. K., Coudert, T., Be'ler, C. and Geneste, L. (2008). Knowledge Formalization in Experience Feedback Processes: An Ontology-based Approach. *Computers in Industry*. 59(1), 694-710.
- Ganter, B. and Wille, R. (1989). Conceptual Scaling. In Roberts, F. ed. *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences*. New York, USA: Springer-Verlag. 139-167.
- Ganter, B. and Wille, R. (1999). *Formal Concept Analysis*. Berlin, Germany: Springer.
- GOLD Community. (2008). GOLD. <http://www.linguistics-ontology.org>.

- Grenon, P., Smith, B. and Goldberg, L. (2003). Biodynamic Ontology: Applying BFO in the Biomedical Domain. *Proceedings of the Workshop on Medical Ontologies*. 8-9 October. Amsterdam, Netherlands: IOS Press. 20-38.
- Gruber, T. (2008). Ontology. In: Liu, L. and Özsu, M. T. *Encyclopedia of Database Systems*. Heidelberg, Germany: Springer-Verlag.
- Halbleib, H. (2004). Requirements Management. *IS Management*. 21(1), 8-14.
- Hendler, J. A. (1992). Massively-parallel Marker-passing in Semantic Networks. *Computers Mathematics Application*. 23(2-5), 277-291.
- Hirst, G. and Onge, D. S. (1998). Lexical Chains as Representations of Context for the Detection and Correction of Malapropisms. In: Fellbaum. *WordNet*. Massachusetts, USA: MIT Press. 305-332.
- Hu, J., Kashi, R., Lopresti, D. and Wilfong, G. (2000). Medium-independent Table Detection. *Proceedings of the SPIE Document Recognition and Retrieval VII*. 26-27 January. San Jose, California: SPIE. 291-302.
- Huang, H. C., Lo, S. M., Zhi, G. S. and Yuen, R. K. K., (2008). Graph theory-based approach for automatic recognition of CAD data. *Engineering Application of Artificial Intelligent*. 21(7), 1073-1079.
- Institute of Cognitive Science and Technology. DOLCE. <http://www.loa-cnr.it/DOLCE.html>.
- Janardanan, V. K., Adithan, M. and Radhakrishnan, P. (2008). Collaborative Product Structure Management for Assembly Modeling. *Computers in Industry*. 59(8), 820-832.
- Jiang, J.J. and Conrath, D. W. (1998). Semantic Similarity based on Corpus Statistics and Lexical Taxonomy. *Proceedings of the International Conference Research on Computational Linguistics*. 22-24 August. Taipei, Taiwan: Morgan Kaufmann. 19-33.
- Kalpakjian, S. (1995). *Manufacturing Engineering and Technology*. (3rd. ed.). Massachusetts, Sydney: Addison Wesley.
- Kavouras, M. and Kokla, M., (2002). A Method for the Formalization and Integration of Geographical Categorizations. *Geographical Information Science*. 16(5), 439-453.
- Kent, R. E. and Neuss, C. (1995). Creating a Web Analysis and Visualization Environment. *Computer Networks and ISDN Systems*. 28(1), 109-117.

- Klein, S. and Simmons, R. F. (1963). Syntactic Dependence and the Computer Generation of Coherent Discourse. *Mechanical Translation*. 7(2), 50-61.
- Knowledge Media Institute of Open University (2004). Apollo. <http://apollo.open.ac.uk/>.
- Lagos, N. and Setchi, R. M. (2007). A Manufacturing Ontology for e-Learning. *Proceedings of the 3rd Intelligent Production Machines and Systems Conference*. July 2-13, 2007. Cardiff, UK: PROMS Network of Excellence. 552-55.
- Leacock C. and Chodorow M. (1998). Combining Local Context and WordNet Similarity for Word Sense Identification. In: Fellbaum, C. *WordNet: A Lexical Reference System and its Application*. Massachusetts, USA: MIT Press. 265-283.
- Lee, J. H., Kim, M. H. and Le, Y. J. (1993). Information Retrieval based on Conceptual Distance in is-a Hierarchies. *Journal of Documentation*. 49(2), 188-207.
- Lee, T., Lee, I., Lee, S., Lee, S., Kim, D., Chun, J., Lee, H. and Shim, J. (2006) Building an Operational Product Ontology System. *Electronic Commerce Research and Applications*. 5(1), 16-28.
- Lefteri, C. (2007). *Making It: Manufacturing Techniques for Product Design* (1st. ed.). Southampton Row, London: Laurence King Publishing Ltd.
- Lemaignan S., Siadat A., Dantan, J. and Semenenko, A. (2006). MASON: A Proposal for an Ontology of Manufacturing Domain. *Proceedings of the Distributed Intelligent Systems: Collective Intelligence and Its Applications*. 15-16 June. Prague, Czech Republic: IEEE. 195-200.
- Lepczyk, C. A., Lortie, C. J. and Anderson, L. J. (2008). An Ontology for Landscapes. *Ecological Complexity*. 5(3), 272-279.
- Lepratti, R. (2006). Advanced Human-Machine System for Intelligent Manufacturing. *Journal of Intelligent Manufacturing*. 17(6), 653-666.
- Lim, T. (2005). Manufacturing Standard Quality Control. Technical Report. APM Spring Sdn. Bhd., Malaysia.
- Lin, D. (1998). An information-theoretic definition of similarity. *Proceedings of the 15th International Conference on Machine Learning*. 24-27 July. California, USA: Morgan Kaufmann. 296-304.

- Lord, P. W., Stevens, R. D., Brass, A. and Goble, C. A. (2003). Investigating Semantic Similarity Measures Across the Gene Ontology: the Relationship between Sequence and Annotation. *Bioinformatics*. 19(10), 1275-1283.
- Lu, T., Yang, Y., Yang, R. and Cai, S. (2008). Knowledge Extraction from Structured Engineering Drawings. *Proceedings of the 5th International Conference on Fuzzy Systems and Knowledge Discovery*. 25-27 August. Shandong, China: IEEE. 415-419.
- Margaret, M. (1961). Semantic Message Detection for Machine Translation, using an Interlingua. *Proceedings of the 1st International Conference on Machine Translation of Languages and Applied Language Analysis*. 5-8 September. London, UK: HMSO. 438-475.
- Masterman, M. (1961). Semantic Message Detection for Machine Translation, using an Interlingua. *Proceedings of the 1961 International Conference on Machine Translation of Languages and Applied Language Analysis*. 5-8 September. London, UK: Springer. 483-475.
- Morbach, J., Yang, A. and Marquardt, W. (2007). OntoCAPE - a Large-scale Ontology for Chemical Process Engineering. *Engineering Applications of Artificial Intelligence*. 20(2), 147-161.
- Mula, J., Poler, R. and Garcia-Sabater, J.P. (2007). Material Requirement Planning with Fuzzy Constraints and Fuzzy Coefficients. *Fuzzy Sets and Systems*. 158(7), 783-793.
- Mustafa, J., Verlinden R. and Meersman R. (2003). Ontology-based Customer Complaint Management. *Proceedings of the 1st International Workshop on Regulatory Ontologies and the Modelling of Complaint Regulation*. 4 November. Sicily, Italy: Springer. 594-606.
- Naphade, M., Smith, J. R., Tesic, J., Chang, S. -F., Hsu, W., Kennedy, L., Hauptmann, A. and Curtis, J. (2006). Large-scale Concept Ontology for Multimedia. *IEEE Multimedia*. 13(3), 86-91.
- Nasr, E. S. A. and Kamrani, A. K. (2006). A New Methodology for Extracting Manufacturing Feature from CAD system. *Computer and Industrial Engineering*. 51(1). 389-415.
- Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. New Jersey, USA: Prentice-Hall.

- Nilsson, P. and Andersson, F. (2004). Process-driven Product Development - Managing Manufacturing Requirements. *Proceedings of the Fifth International Symposium on Tools and Methods of Competitive Engineering*. 13-17 April. Lausanne, Switzerland: Chalmers Publication Library. 395-404.
- Noy, N. F. and McGuinness, D. L. (2001). *Ontology Development 101: a Guide to Creating your First Ontology*. Technical Report, Stanford Knowledge Systems Laboratory, Stanford, USA.
- Othman, R. M., Deris, S. and Illias, R. M. (2008). A Genetic Similarity Algorithm for Searching the Gene Ontology Terms and Annotating Anonymous Protein Sequences. *Journal of Biomedical Informatics*. 41(1), 65-81.
- Petridis, K., Precioso, F., Athanasiadis, T., Avrithis, Y. and Kompatsiaris, Y. (2005). Combined Domain Specific and Multimedia Ontologies for Image Understanding. *Proceedings of the 28th German Conference on Artificial Intelligence*. 11-14 September. Koblenz, Germany: ACM. 1-7.
- Plant Ontology Consortium (2003). Plant Ontology. <http://www.plantontology.org/>.
- Pongcharoen, P., Hicks, C. and Braiden, P.M. (2004). The Development of Genetic Algorithms for the Finite Capacity Scheduling of Complex Products, with Multiple Levels of Product Structure. *European Journal of Operational Research*. 152(1), 215-225.
- Prabhu, B. S., Biswas, S. and Pande, S. S. (2001). Intelligent System for Extraction of Product Data from CADD models. *Computers in Industry*. 44(1), 79-95.
- Priss, U. (1999). Efficient Implementation of Sematic Relations in Lexical Databases. *Computational Intelligence*. 15(1), 79-87.
- Rada, R., Mili, H., Bicknell, E. and Blettner, M. (1989). Development and Application of a Metric on Semantic Nets. *IEEE Transactions on Systems, Man, and Cybernetics*. 19(1), 17-30.
- Ratsch, E., Schultz, J., Saric, J., Lavin, P. C., Wittig, U. and Reyle, U. (2003). Protein-interaction Ontology. *Comparative and Functional Genomics*. 4(1), 85-89.
- Resnik, P. (1992). Wordnet and Distributional Analysis: A Class-based Approach to Lexical Discovery. *Proceedings of the AAAI Symposium on Probabilistic Approaches to Natural Language*. 5-8 September. San Jose, California: AAAI Press. 56-64.

- Resnik, P. (1995). Using Information Content to Evaluate Semantic Similarity in a Taxonomy. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*. 22-25 August. Montreal, Canada: Morgan Kaufmann. 448-453.
- RibbonSoft (2008). QCAD 2D CAD Software. <http://www.qcad.org/>.
- Richardson, R. and Smeaton, A. (1995). *Using WordNet in a Knowledge-Based Approach to Information Retrieval*. Technical Report, Dublin City University, Dublin, Ireland.
- Romanov, L. (2005). Notes Re: Personal Knowledge Formalization and Modeling. *Journal of Knowledge Management Practice*. 6(1), <http://www.tlinc.com/articl94.htm>.
- Ryma Technology Solutions (2006). FeaturePlan Requirement Management Software. <http://www.featureplan.com/>.
- Schlenoff, C., Ivester, R. and Knutilla, A. (1998). A Robust Process Ontology for Manufacturing Systems Integration. *Proceedings of 2nd International Conference on Engineering Design and Automation*. Maui, USA: National Institute of Standards and Technology. [http://www.mel.nist.gov/publications/get\\_pdf.cgi?pub\\_id=821235](http://www.mel.nist.gov/publications/get_pdf.cgi?pub_id=821235).
- Schmitt, I. and Saake, G. (1997). Merging Inheritance Hierarchies for Schema Integration Based on Concept Lattices. Technical Report, University of Magdeburg, Magdeburg, Germany.
- Selamat, A. and Selamat, M. H. (2004). Routing Algorithm of Mobile Agents for Query Retrieval using Genetic Algorithm. *Malaysian Journal of Computer Science*. 17(2), 1-10.
- Seremeti, L. and Kameas, A. (2007). Multimedia Ontologies. *Proceedings of the 3rd International Conference on Mobile Multimedia Communications*. 27-29 August. Nafpaktos, Greece: ICST. 1-7.
- Spangenberg, N., and Wolff, K. E. (1999). Concept Lattices as Indicators of change in the Therapeutic Process: does Formal Concept Analysis of Repertory Grids Represent a Paradigm Change of Data Evaluation?. In: Spangenberg, N. and Wolff, K. E. *Psycho Analytic Research by means of Formal Concept Analysis*. Munster: Lit Verlag.
- Stanford University School of Medicine (2009). Protégé 2000 Ontology Editor and Knowledge Acquisition Tool. <http://protege.stanford.edu/>.

- Sussna, M. (1993). Word Sense Disambiguation for Free-text Indexing using a Massive Semantic Network. *Proceedings of the 2nd International Conference on Information and Knowledge Management*. 1-5 November. Washington, USA: ACM. 67-74.
- TechTarget (2009). SYSPRO Requirement Planning. <http://www.2020software.com>.
- Tesnière, L. (1959). *Elements de Syntaxe Structurale*. Paris, France: Klincksieck.
- The Gene Ontology Consortium (1999). Gene Ontology. <http://www.geneontology.org/>.
- The Gene Ontology Consortium. (2000). Gene Ontology: Tool for the Unification of Biology. *Nature Genetic*. 25(1), 25-29.
- The Gene Ontology Consortium. (2001). Creating the Gene Ontology Resource: Design and Implementation. *Genome Research*. 11(8), 1425-1433.
- The University of Washington (2006). Foundational Model of Anatomy. <http://sig.biostr.washington.edu/projects/fm/>.
- Thevenot, H. J., Nanda, J. and Simpson, T. W. (2006). Redesigning Product Families Using Heuristics and Shared Ontological Component Information. *Proceedings of the IEEE International Conference on Information Reuse and Integration*. 16-18 September. Waikoloa, USA: IEEE. 330-335.
- University Bremen (2008). Generalized Upper Model. <http://www.fb10.uni-bremen.de/anglistik/langpro/webpace/jb/gum/>.
- University of Cyprus (2005). CContology. <http://www.cellcycleontology.org/>.
- University of Maryland (2006). Swoop. <http://www.mindswap.org/2004/SWOOP/>.
- University of Montpellier (2008). CoGITaNT. <http://cogitant.sourceforge.net/>.
- USC Information Sciences Institute (1998). Ontosaurus. <http://www.isi.edu/isd/ontosaurus.html>.
- Uschold, M. and Gruninger, M. (1996). ONTOLOGIES: Principles, Method and Application. *Knowledge Engineering Review*. 11(2), 93-155.
- VariCAD. (2007). VariCAD. <http://www.varicad.com/>.
- Vegetti, M., Henning, G. P. and Leone, H. P. (2005). Product Ontology. Definition of an Ontology for the Complex Product Modeling Domain. *Proceedings of the 2nd Mercosur Congress on Chemical Engineering*. 14-18 August. Rio de Janeiro, Brazil: Elsevier Science Publishers. [http://www.enpromer2005.eq.ufjf.br/nukleo/pdfs/1016\\_productontology.pdf](http://www.enpromer2005.eq.ufjf.br/nukleo/pdfs/1016_productontology.pdf).



- Villa, F. (2007). A Semantic Framework and Software Design to Enable the Transparent Integration, Reorganization and Discovery of Natural Systems Knowledge. *Journal of Intelligent Information Systems*. 29(1), 79-96.
- Wang, Y., Philips, I. T. and Haralick, R. M. (2004). Table Structure Understanding and its Performance Evaluation. *Pattern Recognition*. 37(7), 1479-1497.
- Waters, T. F. (1996). *Fundamentals of Manufacturing for Engineers*. (1st. ed.). London, UK: UCL Press.
- Wilks, Y. and Fass, D. (1992). The Preference Semantics Family. *International Series in Modern Applied Mathematics and Computer Science*. 24(1), 205-221.
- Wille, R. (1992). Concept Lattices and Conceptual Knowledge Systems. *Computers and Mathematics with Applications*. 23(1), 493-515.
- Williams, R. J., Martinez, N. D. and Golbeck, J. (2006). Ontologies for Ecoinformatics. *Web Semantics: Science, Services and Agents on the World Wide Web*. 4(4), 237-242.
- Woods and William, A. (1975). What's in a Link: Foundations for Semantic Networks. In Bobrow, D. G., and Collins, A. M. eds. *Representation and Understanding*. London, UK: Academic Press. 35-82.
- Wu, Z. and Palmer, M. (1994). Verbs Semantics and Lexical Selection. *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*. 27-30 June. Las Cruces, New Mexico: Association for Computational Linguistics. 133- 138.
- Xiong, M., Tor, S. B., Khoo, L. P., and Chen, C. (2003). A Web-enhanced Dynamic BOM-based Available-to-promise System. *International Journal of Production Economics*. 84(2), 133-147.
- Zuyev, K. (1997). Table Image Segmentation. *Proceedings of the 4th International Conference Document Analysis and Recognition*. 18-20 August. Ulm, Germany: IEEE. 705-708.