Xu, Yuanping, Xu, Zhijie, Jiang, Xiang and Paul, Scott

Developing a Knowledge-based System for Complex Geometrical Product Specification (GPS) Data Manipulation.

Original Citation

Xu, Yuanping, Xu, Zhijie, Jiang, Xiang and Paul, Scott (2010) Developing a Knowledge-based System for Complex Geometrical Product Specification (GPS) Data Manipulation. Knowledge Based Systems, 24 (1). pp. 10-22. ISSN 0950-7051

Knowledge-Based Systems 24 (2011) 10–22

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Developing a knowledge-based system for complex geometrical product specification (GPS) data manipulation

Yuanping Xu [a,*], Zhijie Xu [b], Xiangqian Jiang [a], Paul Scott [c]

[a] Centre for Precision Technologies, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, United Kingdom
[b] Informatics Department, School of Computing and Engineering, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, United Kingdom
[c] Taylor Hobson Ltd., P.O. Box 36, 2 New Star Road, Leicester LE4 9JQ, United Kingdom

ARTICLE INFO

Article history:
Received 12 October 2009
Received in revised form 3 May 2010
Accepted 3 May 2010
Available online 27 May 2010

Keywords:
Category theory
Categorical object model
GPS

ABSTRACT

Geometrical product specification and verification (GPS) matrix system is a universal tool for expressing geometrical requirements on product design drawings. It benefits product designers through providing detailed description of functional requirements for geometrical products, and through referring to corresponding manufacturing and verification processes. In order to overcome current implementation problems highlighted in this paper, a GPS knowledge base and a corresponding innovative inference mechanism have been researched, which led to the development of an integrated GPS knowledge-based system to facilitate rapid and flexible manufacturing requirements. This paper starts with a brief introduction of GPS, GPS application problems and the project background. It then moves on to demonstrate a unified knowledge acquisition and representation mechanism based on the category theory (CT) with five selected examples of this project. The paper concludes with a discussion on the future works for this project.

1. Introduction

In order to ensure quality of geometric products and to facilitate global outsourcing, almost all the so-called "world class" manufacturing companies nowadays are applying various tools and methods to maintain the consistency of product characteristics through out the manufacturing life cycle. Among these, for ensuring the consistency of the geometric characteristics, a tolerancing language – the geometrical product specification (GPS) has been widely adopted to precisely transform the functional requirements from customers into manufactured workpieces expressed as tolerance notes in technical drawings.

The initial GPS standards were set up by the International Organization for Standardization (ISO) to determine geometrical features of workpieces through the use of a symbolic language for expressing tolerances in technical drawings [1,2]. It has also been used to manage geometrical variations of workpieces according to their specifications as well as to suggest measuring instruments and their calibration methods [3]. The next generation of GPS standards (second generation) aim at integrating all the data concerning essential steps of a production cycle right down to the macro or nano scale, which cover the whole spectrum of manufacturing design and production stages through specifying and verifying parts' sizes and dimensions, geometrical tolerances, and surface properties [4–6]. They are ordered in a matrix in which all the rows constitute 18 chains of standards in total (size, distance, radius, angle, etc.) with each column defining various characteristics of geometrical features. Therefore, the whole GPS (ISO/CEN) is also called as the GPS matrix system [3]. Table 1 shows the chain of standards relating to profile surface texture that are grouped into six aspects: the product documentation indications, definition of tolerances, definitions of characteristics of actual feature, assessment of the workpiece deviations, measurement equipment, and calibration requirements/measurement standards. In an ideal case, based on each chain of standards, the process of manufacturing a geometric product can be clearly defined by taking into factors such as setting up unambiguous specifications, and interpreting manufacturing specifications and verification information.

The latest version of general GPS matrix system is composed of 108 cells (6 by 18) and each of them contains at least one standard. In the future, there will be more standards to be filled into those cells. Hence the GPS matrix system will become more complex and difficult to be handled.

Although commonly acknowledged by industrial users as one of the most successful efforts in integrating existing manufacturing life-cycle standards, current GPS implementations and software packages suffer from several drawbacks in their practical use, possibly the most significant, the difficulties in inferring the data for

* Corresponding author. Tel.: +44 07922423188.
E-mail addresses: ghxpy@hotmail.com, y.xu@hud.ac.uk, creekline017@hotmail.com (Y. Xu).

0950-7051/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2010.05.002

Table 1
The chain of standards relating to the "profile surface texture".

Chain link number Geometrical characteristic of feature		1 Product documentation indication-Codification	2 Definition of tolerances – Theoretical definition and values	3 Definitions for actual feature – Characteristic or parameter	4 Assessment of the deviations of the workpiece – Comparison with tolerance limits	5 Measurement equipment requirements	6 Calibration requirements – Measurements standards
Roughness profile	M-system-Ra, Rz	1302	4287, 11562	4287, 11562	4288	3274, 11562	5436, 12179
	M-system-RSm	1302	4287, 11562	4287, 11562	4288	3274, 11562	5436, 12179
	Motif method-R, Rx, AR	1302	12085	12085	4288, 12085	3174	5436, 12179
	Rk family	1302	11562, 13565-1, 13565-2	11562, 13565-2	4288	3274, 11562	5436, 12179
	Rpq, Ruq, Rmq	1302	11562, 13565-1, 13565-3	11562, 13555-2	4288	3274, 11562	5436, 12179
Waviness profile	M-system-Wa, Wz,...	1302	4287, 11562	11562	4288, 12085	3274, 11562	5436, 12179
	Motif system-W, AW, Wx, Wte	1302	12085	12085		3274	5436, 12179
Primary profile	M-system-Pa, Pt,...	1302	4287, 11562	4288		3274, 11562	

the "best" solutions. The problem stemmed from the foundation of data structures and knowledge-based system design. This indicates that there need to be a "new" software system to facilitate GPS applications. The rationales for this decision can be summarized as:

- GPS expresses different product requirements in a language of geometry. Thus, the link between functional requirements and the desired GPS parameters is implicit. Designers need to have the skill to translate various mechanical requirements into the geometrical language of GPS or vise verse.
- The GPS standards are still growing and becoming more complex, sometime too complex, abstract, and theoretical for ordinary designers and engineers in the manufacturing industry to read, understand, remember and make full use of them [7].
- Data defined in a GPS system is difficult, if not possible, to be divided into the strict formats of flat tables applied in all Relational database models. For example, Fig. 1 shown below cannot be easily divided into indexed rows and columns.
- The inter relationships between different GPS files are vaguely defined. Only experienced GPS experts are capable of deciphering and cross-referencing them to achieve specific manufacturing tasks, which are often in short supply in most of the small medium sized enterprise (SMEs).
- Current GPS standards are often stored in text-based electronic file formats (e.g. PDF), which are difficult for users to search and access without knowing specific searching criteria. It is almost impossible to apply computerized automated querying process in their current forms.

Surface	Functions applied to the surface			Parameters				
				Roughness profile			Waviness profile	
	Designations		Symbol	R	Rx	AR	W	Wx
With relative displacement	Slipping(lubricated)		FG	•			≤0,8R	
	Dry friction		FS	•		∘		•
	Rolling		FR	•			≤0,8R	•
	Resistance to hammering		RM	∘		∘	∘	
	Fluid friction		FF	•		∘		
	Dynamic sealing	with gasket	ED	•	∘	∘	≤0,6R	•
		without gasket		∘	•		≤0,6R	

Fig. 1. Example of a classification of surface function for motif parameters ISO 12085 [3].

- There are no existing de facto knowledge systems to manage this large maze of GPS standards and to keep its data integrity and version consistency for GPS applications. Current research efforts and pilot systems to resolve those problems do not seem to provide mechanisms for GPS users to manage and query "intelligently" those information; never mention to customise or even add their own new knowledge relating to certain processes. Next paragraph will explain this point in detail.

The aforementioned six major shortcomings relating to applying ISO/CEN GPS standards indicate that the paper based GPS standards need to be computerized. However, the current computer aided design and manufacturing software systems are still struggling to meet the demands of the global and dynamic manufacturing environments and failed to stand for the complexity of the whole GPS world due to the following reasons:

- Almost all current GPS knowledge base systems are based on the first generation of GPS standards, which did not integrate the "Specification" part and the "Verification" part together. Thus, the integrity of a geometric product cannot be ensured.
- Most systems do not provide precise drawing indication.
- Current GPS knowledge base systems are week in GPS technical documentation. The lacks of effective communications have resulted in wide misunderstanding between the design concept and the real product. Experience has shown that the average costs resulting from such shortcomings of incomplete GPS technical documentation can amount to as much as 20% of the production turnover [5].
- Almost all current GPS knowledge base systems are only focuses on filter selection, filtering calculations and measured data analysis. An internet-based surface texture analysis and information system, developed by Center for Precision Metrology in the University of North Carolina, claimed to solve the problem that current surface texture analysis systems are weak in developing process knowledge or mapping the observed effect to causes [8]. For example, after taking several measurements on a workpiece, users can use traditional systems to filter the profile at a standard cut_off length and then get a table of calculated parameter values. However, these systems cannot provide a documentation mechanism to store process parameters with metrology data for observing how process parameters relate to variability in the surface parameters. This system focuses on filter selection, filtering calculations and measured data analysis.

In general, the major software systems at present are still weak on functionality and relying on ambiguous dimensioning and tolerancing practices based on the nominal model methodology and geometry theory. Features such as product function, surface properties and the related verification principles, measuring equipment, calibration requirements, uncertainty and measurement traceability are often largely ignored. One of the major reasons for causing these drawbacks is that the traditional knowledge modelling mechanisms cannot efficiently support the modelling of the various complex GPS data structures in a unified mathematical foundation to reflect the complicated relationships among parts and GPS standards, especially for the multi-level mappings, which are essential for comprehensive analysis and data manipulations to solve practical production problems. Thus, a unified knowledge acquisition and knowledge representation mechanism to retrieve and organize knowledge from various GPS documents has been devised using Category Theory to solve this problem. Therefore, this paper names the knowledge modelling mechanism devised in this project as the "categorical" modelling mechanism.

2. Project background

The so-called virtualGPS project being carried out in the University of Huddersfield focuses on developing an integrated GPS knowledge platform with knowledge generation and accessing facility based on the GPS matrixes defined in GPS standards. The term "virtual" at here refers to the effort in removing the borders and integrating the GPS information (especially these specified in the CEN and ISO standard documentation [5]) and the corresponding GPS realization methodologies into a single framework regardless their physical storage locations. The system spans knowledge domains from dimensional and geometrical tolerancing, surface properties and related manufacturing processes/equipments, to verification principles and calibration requirements, as well as uncertainty and measurement traceability. This paper focuses on discussing the Category Theory modelling capability for GPS standards through five key examples (see Section 3.3, 3.4.1, 3.4.2 and 3.5–3.6) and demonstrating the implementations of these examples through showing the real software system snapshots. The envisaged potential benefits of the virtualGPS system can be summarized as:

- To provide a unified knowledge base for supporting engineering decisions in choosing appropriate GPS parameters according to the required functional performances.
- To enable an automated querying mechanism for guiding designers with relevant GPS specifications.
- To equip a rating and ranking inference engine for locating and retrieving GPS-recommended manufacture processes and equipments.
- To link similar functions for aiding decisions on measurement procedures and equipments.

To achieve the system functions, the proposed software specifications have following design features:

- Flexible data storage to enable data sharing and maintenance through representing GPS information in the form of knowledge objects in the object-oriented (OO) style, which can be readily adopted by other platforms and tools.
- Client/Server structure for data synergy and remote collaborations between geographically dispersed designers, production engineers and metrologists.
- User-friendly system interfaces for accessing system data and functions such as cross-referencing and updating.

It is anticipated that when completed, the virtualGPS system will enable non-GPS-expert to use GPS-matrixes in an efficient manner. It will also ensure that when product design changes, the relevant GPS specifications will be updated automatically to remain consistent with relevant GPS standards. Moreover, with the trend of globalisation in manufacture industries, the remote data access features and web-based user interfaces of the system will become a norm for future systems.

In this paper, a case study analyses and demonstrates the design process of a cylinder liner will be discussed with five examples in Section 3. A cylinder liner is one of the central working parts of a reciprocating engine, and it is the space in which a piston travels. The movement of a piston inside the cylinder can drive a vehicle moving (see Fig. 2). Normally, a piston moves inside each cylinder with several metal piston rings fitted around its outside surface in machined grooves – typically two for compressional sealing and one for oil sealing. They are commonly made of spring steel and having close contact with the hard walls of the cylinder bore, which rides on a thin layer of lubricating oil to keep the engine from seizing up. The contact between the cylinder liner and its counterpart piston rings requires the cylinder to have a good bearing surface but also retain a reservoir of oil for lubrication. Furthermore, the space surrounded by the cylinder bore and piston rings needs tight seal to contain the compression of fuel and air mixtures.

Among all the design features, the most important functional demands of the cylinder and piston rings are oil consumption, blow-by, and wear; especially at the top-dead centre (TDC). The surface texture parameters defined in the latest GPS standards have direct influences on the functional performance of the cylinder and piston. After performing a factorial designed experiment (FDE) where surface roughness was correlated to important functional performance indicators – oil consumption, wear, and blow-by, in a 10-liter truck engine, it proves that 'oil consumption' is strongly correlated to the Rz parameter measured on the cylinder liner. The biggest influence on 'blow-by' is the Ra parameter measured on the piston rings with a negative variation. The 'wear' is also strongly correlated to the Ra value measured on the piston rings followed by the Rz measured on the cylinder; both are having the same variation with the 'wear'.

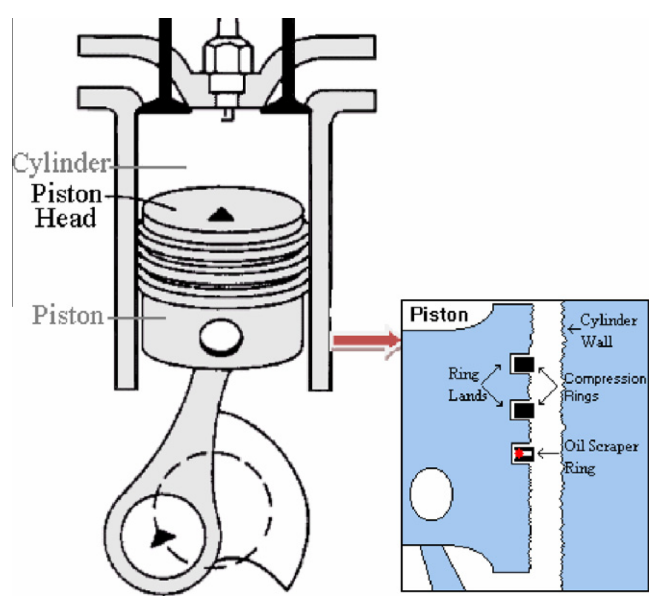

Fig. 2. Piston rings.

3. Category theory based implementation

3.1. Rationales

The so-called category theory (CT) originally rose in mathematics and was defined as an abstract way to deal with mathematical structures and relationships between them. It offers a formal basis and abstraction for handling the passage from one type of mathematical structure to another through mappings that preserve structures [9]. It is still a maturing mathematical subject, which firstly emerged in 1945 in Eilenberg and MacLane's paper entitled "General Theory of Natural Equivalences" [10]. In last three decades, category theory has found new applications in the theoretical computer science, algebra and database applications attributing to its firm mathematical roots, which contributed, among other things, to the development of semantic programming and new logical systems. The notions and constructions of CT include objects, arrows, categories, functors, natural transformations, coproducts, products, pullbacks, cones and so on. For an in-depth understanding of category theory, readers can refer books [9,11].

To development of a software system, a suitable system modelling strategy needs to be chose and clarified in advance for the whole system design process. As the VirtualGPS is a knowledge-based system, the system design should focus on the knowledge/application modelling and database modelling. To avoid the error-prone and misunderstanding process of mapping the data stored in a database into objects in the knowledge base of the VirtualGPS system or vice versa, researchers in this project devised a unified modelling mechanism which can be used both on the application side and on the database side. In this project, CT is used to solve five important issues relating to the design of the VirtualGPS:

(1) CT is employed to organize the whole system architecture to specify, visualize and document the entire software system (see an example in Section 3.3). Moreover, as mentioned earlier, VirtualGPS is a "growing" system which needs the addition and updating of new features for future usage, CT provides basis for supporting this incremental process (data refinement process) with its inherent multi-level features (subcategories, functors, natural transformations, fibration and adjointness).
(2) CT is adopted to describe and constrain the knowledge extracted from existing GPS matrixes. It guides knowledge-base designers to build a categorical object model that can clearly reflect knowledge-base structures with form mathematical formulizations (see an example in Section 3.4).
(3) Both dynamic aspect (e.g. methods) and static aspect (e.g. attributes, objects) of the VirtualGPS can be modelled uniformly using arrows. The type and definition of arrow will determine what it's actually roles. This is much better than Set Theory that uses two notions – set and function to represent static and dynamic aspects in separate way (see an example in Section 3.6).
(4) CT is applied to define a stable measurement procedure. As claimed by Kappel and Vieweg, the processing modelling step is of vital importance to manufacturing applications [12]. Within which, measurement procedures are key to the final quality of a manufactured product. CT serves greatly in terms of improving the stability of a selected measurement procedure (see an example in Section 3.5).
(5) The categorical object model is also used for the "categorical" database management system (DBMS) developed for the VirtualGPS system. In this project, system designers uses CT to model the complex relationships and constrains among GPS standards, so it is wise to use same model mechanism in the database side since there is no need to grogram any mapping codes between the data in the database and the data in the application. Thus an object-oriented DBMS fully supporting the categorical data model is required in this project. Comparing with conventional relational data model based on the set theory, this categorical DBMS relies on the CT to provide a rigorous mathematical foundation, which will support handling of complex data structures and manipulations (see an example in Section 3.6).

All in all, CT provided a good unified tool that enabled the system design from high-level system architecture down to the knowledge base, and from static aspects to dynamic aspects in same mathematic mechanism. Thus, different modelling powers from different modelling mechanisms can be unified in single mathematical foundation. Moreover, it provides good abstractions that give a deep insight into the essence of knowledge and knowledge processing, which cannot be obtained simply from a large number of details.

3.2. Fundamental categorical notations

The two fundamental concepts in CT are *arrows* and *internal objects*.

For example, a category $\boldsymbol{P}$ (Fig. 3) consists of three internal objects $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ together with arrows: $\boldsymbol{f}$, $\boldsymbol{g}$, an associated arrow $h = g \circ f$ and identity arrows $\boldsymbol{I_A}$, $\boldsymbol{I_B}$, $\boldsymbol{I_C}$, respectively. A composition operator on each pair of arrows $\boldsymbol{f}$ and $\boldsymbol{g}$ satisfies $cod(f) = dom(g)$ (a composite $g \circ f : dom(f) \to cod(g)$).

CT also provides various constructs (functors, natural transforms etc.) to build hierarchies of 'category of categories'. A general type of structure preserving mapping (arrow) between categories, called a functor, is considered an arrow in a 'category of categories' (here called 'a functorial category'). The formal definition of a functor is [9]:

"Let $\boldsymbol{C}$ and $\boldsymbol{D}$ be categories. A functor F: $\boldsymbol{C} \to \boldsymbol{D}$ is a map taking each object A in $\boldsymbol{C}$ to an object $F(A)$ in $\boldsymbol{D}$ and each arrow $f{:}A \to B$ in $\boldsymbol{C}$ to a arrow $F(f){:}F(A) \to F(B)$ in $\boldsymbol{D}$ while holding the following two properties:

- $F(id_A) = id_{F(A)}$
- $F(g \circ f) = F(g) \circ F(f)$ for all arrows $f{:}A \to B$ and $g{:}B \to C$."

3.3. Categorical representation for the VirtualGPS architecture

Because of the large size and high complexity of the VirtualGPS knowledge-based system, researchers in this project set up a unified CT based framework for representing the system through appropriate mathematical formalizations. The overall VirtualGPS system architecture contains two major parts: system modularized framework (system deployment graph) and system sequent diagram. The system modularized framework focuses on specifying the functions of all the modules, their mutual interactions and

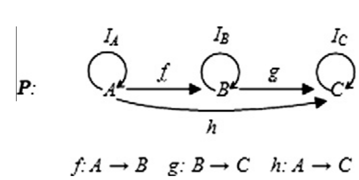

Fig. 3. A category ***P***.

transformations. The sequence diagram shows how messages are sent sequentially from module to module. This paper takes surface texture part in GPS as an example (see Figs. 4 and 5).

There are three categories in Fig. 4: '*Surface Texture Client*', '*Categorical DBMS*' and '*Surface Texture Module Server*', as well as two typed functors: '*Internet*' and '*Intranet*' are shown. The components in each module are modelled as internal objects of categories. For example, the '*Surface Texture Module Server*' contains four internal objects (lower level categories): '*Function*', '*Specification*', '*Manufacture*', and '*Verification*'.

In the Fig. 4, the client-side browser provides users with an interface to access the GPS knowledge organised by rules and standards devised in the knowledge base. The Surface Texture module server controls an inference engine to organise four lower level categories and generates an XML report for each of them. In order to form an accurate and comprehensible 'knowledge' from the maze of GPS-matrixes, a back-end database system that can directly store and manage GPS standards modelled using the devised model has also been developed.

The diagram highlights the perceived process flow for utilizing the surface texture module in a typical manufacturing cycle, which can be described as follows:

For designers: Product designers activate the VirtualGPS system; the "Function" component will search and advise users by translating functional performances (e.g. fluid friction or dry friction) into surface texture parameters defined in GPS-matrixes; and then generates a function analysis report using a so-called "pattern" language. Therefore, the function component is responsible for translating the design intent into requirements of GPS characteristics for designers.

For designers and manufacturing engineers: The generated "Specification" component produces the details of the GPS specification on the technical drawing in the form of complete 'callouts', based on the selected surface texture parameters.

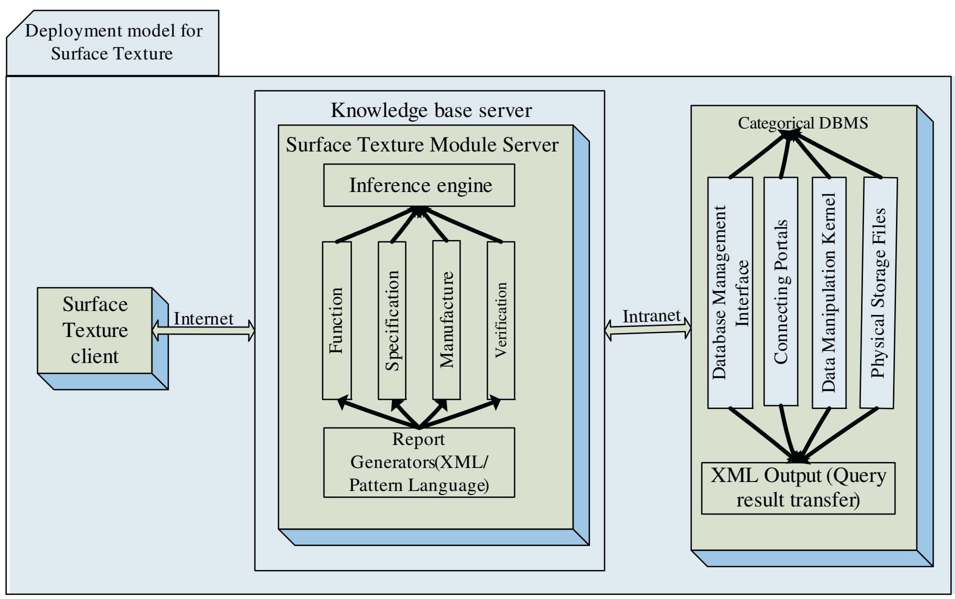

Fig. 4. Overall architecture of the VirtualGPS system.

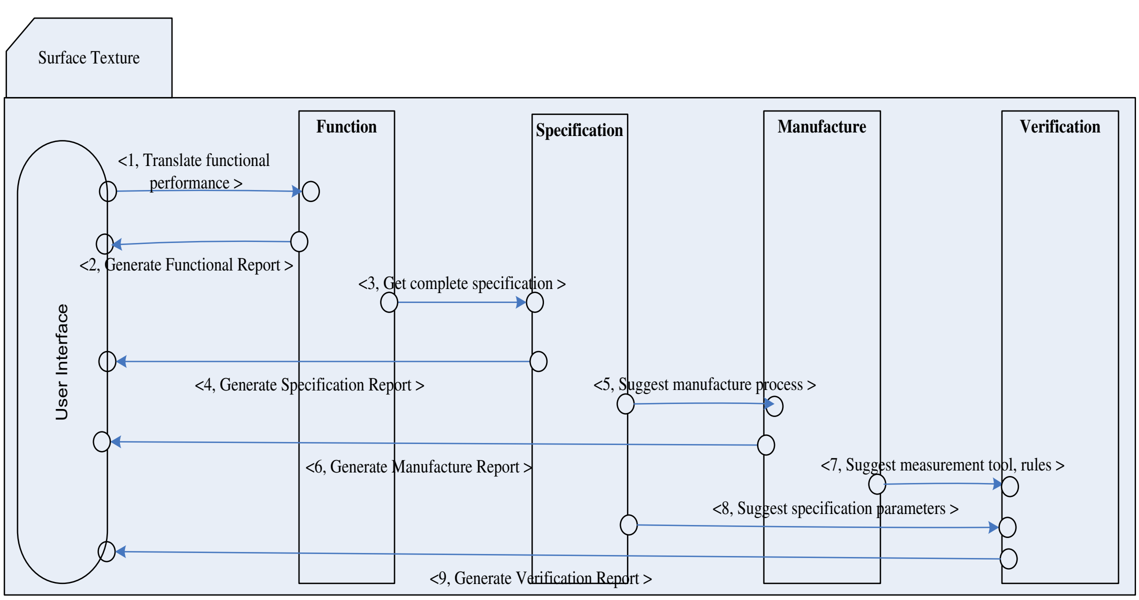

Fig. 5. The sequence diagram for the surface texture module.

For manufacturing engineers: In accordance with the deduced specification report and any extra criteria defined (such as material types and quantity), the "Manufacture" component can suggest appropriate manufacturing processes for the designers. In order to enable cross comparing among different processes, a manufacturing process report for each recommended process plan will be formed, which includes details such as process description, material suitability, process variations, costing issues and sample applications.

For metrologists: The final "Verification" component enables metrologists to choose from recommended measurement instruments and filtering techniques to formulate a measurement strategy.

Fig. 6 demonstrates the software implementation result which seriously follows the Figs. 4 and 5. According to Fig. 5, in this paper, the software system starts with "Function" component for designers to deal with the design process of a cylinder liner.

3.4. Categorical representation for knowledge-bases

This project provides an innovative way for manufacturing engineers to establish knowledge-bases derived from GPS raw standards without specialised expert computer skills. The knowledge-base of the pilot system contains four derived sub-components: function, specification, manufacture and verification. The following two sub-sections explain the designing processes of the function and specification components in CT terms. The other two – manufacture and verification – have adopted similar designing process.

3.4.1. Function component

In the function component, CT is applied through describing function requirements in a so-called pattern language, which guides the inference engine to generate a function performance report to highlight the suggested specific surface roughness parameters according to the inputted function performance requirements [13]. The pattern language is used at here for facilitating function decomposition and to structure connection process. This paper gives an example on using the partial order set and the product order of CT to represent and record decomposition alternatives of the pattern language [14]. A partial order is a binary relation R over a set S, which is reflexive, antisymmetric and transitivity. The set S with a partial order is called a partially ordered set (*poset*). The function performance report generated from the Function component contains six patterns specified in the explained pattern language format. These patterns are connected with each other by the context of each pattern, and ordered by the design sequence:

(1) Pattern 1 specifies the surface requirements.
(2) Pattern 2 analyses the functional performances according to the output of Pattern 1.
(3) Pattern 3 selects suitable specification to ensure the surface functions correctly.
(4) Pattern 4 suggests a function correlation approach between surface texture parameters and the functional performances.
(5) Pattern 5 provides an alternative route through the surface change monitoring approach to find the relations between functional performances and surface parameters.
(6) Pattern 6 specifies the limit values for the parameter selected from Patterns 4 and 5.

In this project, every pattern in the Function component is represented as a "category" which contains six internal objects: name, problem, solution, forces, examples, and next pattern. All of them are represented in *posets*. As illustrated in Fig. 7, all patterns are connected with each other, which form also an integrated *poset*.

Actually, Fig. 7 is a product order which is a cartesian product of two *posets*: namely, patterns collection *poset* and internal object poset. As the transitivity definition of *poset*, all arrows among internal objects must commutes (e.g. if f: context1 → name, g: name → problem, so $f \circ g$ must equal to k: context1 → problem). This is used to ensure the consistency of a functional report. The Patterns 4 and 5 are two optional approaches to find the relations between function performance and surface parameters, which uses injection functors to form Pattern 6 together.

CT gives an implementable representation for function knowledge-base with an open platform for GPS experts add more knowledge conforming to the Fig. 7 in the future. Fig. 8 gives an example for a function report generated for designing the cylinder liner.

The Fig. 8 demonstrates how knowledge inferred by the "Function" component is organized by using the pattern language. In this case, it provides options for engineering designers with suitable surface parameters that match the required functional performances in the predefined patterns as shown below:

Pattern 1 – surface requirements
For a cylinder liner on an engine block, the counterpart is the piston ring; the surface requirement is to maintain a good bearing surface while retaining a reservoir of oil for lubrication.
Pattern 2 – functional performance
The most important functional demands in this case are correct oil consumption, blow-by, and wear especially at the top-dead centre (TDC).
Pattern 3 – surface parameters selection
The texture parameters Rk and Rz have been shown to have a functional correlation with the desired surface tasks. One

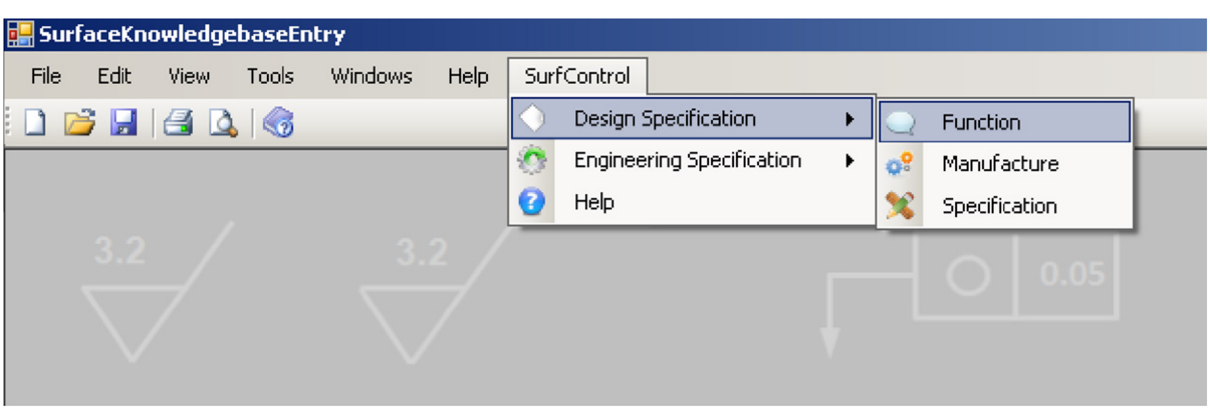

Fig. 6. The main user entry interface.

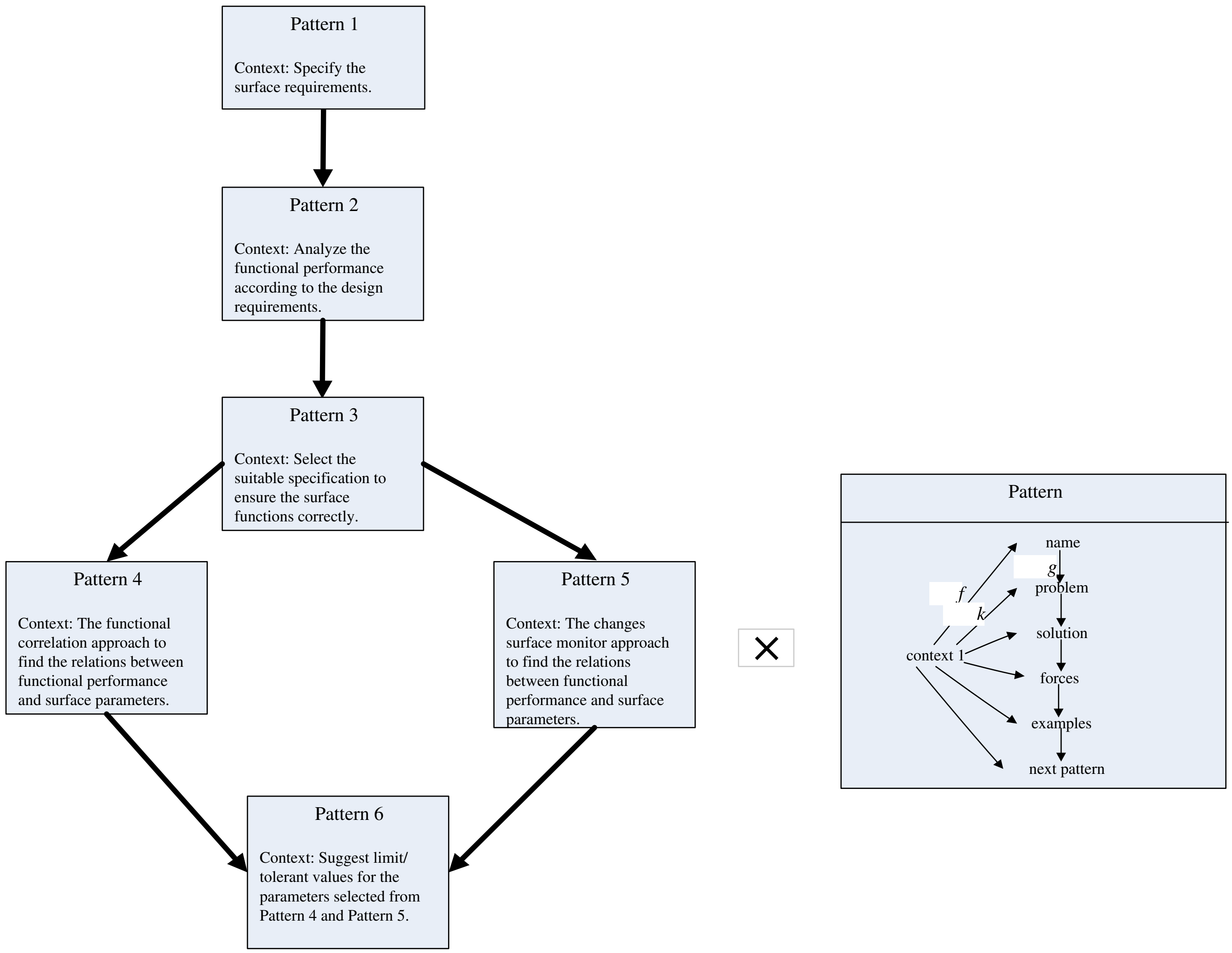

Fig. 7. Product order of function knowledge-base.

option for the manufacture process is to adopt a plateau-honed surface. Rq & Rsk can be used to monitor for surface changes.
Pattern 4 – functional correlation
The surface texture parameters Rk and Rz have been shown to have a functional correlation with the desired surface tasks.
Pattern 6 – suggestion of limit values
According to the factorial designed experiment (FDE), when Rz of the cylinder increased, oil consumption, blow-up and wear all increased since they have the same variation. Therefore, the limit value of Rz is suggested at 4 μm in this case.

3.4.2. Specification component

The Specification component provides detailed geometrical specifications for the selected surface parameters including information obtained from partition, extraction and filtration processes. For example, to satisfy the functional requirements of a cylinder liner, the VirtualGPS system suggests using the surface texture parameter Rz with a toleranced value at 4 μm. The Specification component in turn recommends a complete set of operational procedures such as evaluation length for extraction, and the bandwidth for filtration. Due to the complexities and intertwined attribute relationships and constraints among all viable operational procedures, CT is used to build a categorical object model. Fig. 9 gives an example of the model when defining a default constraint between the operations of extraction and partition.

As shown in Fig. 9, extraction and filtration are modelled as categories. The constraint r demonstrates a relationship between extraction and partition, which is structured by the construct of a "pullback" of CT. The expression "equals:: sampling_length × up_limit" is the name and type of the product, where "*Extraction*$_r$ × *Filtration*" is the restricted product over r (r is restricted condition – "equal" here). The notations $\pi_1 r$ and $\pi_2 r$ are projections of the product into the initial instance of the "*Extraction*" and "*Filtration*" categories, respectively with $\lambda_1 r$, $\lambda_2 r$ represented as arrows injecting the initial objects into the pool of instances of this constraint. The detail explanations on the construct of "pullback" and how it can be used in representing constraints among entities can be located in a paper published by Nelson etc. in 1994 [15]. Thus, in this project, by representing surface texture operational procedures as categories, attributes of them as internal objects, and the corresponding relationships and constraints as pullbacks between categories, the whole Specification component can be logically and structurally expressed. All arrows in Fig. 9 must commute in a manner to ensure consistency. Based on a set of constraints and inter-relationships modelled in the categorical way as Fig. 9, the completed specifications extracted from GPS standards can be generated for *Rz* with limited value 4 μm of a cylinder liner (see Fig. 10).

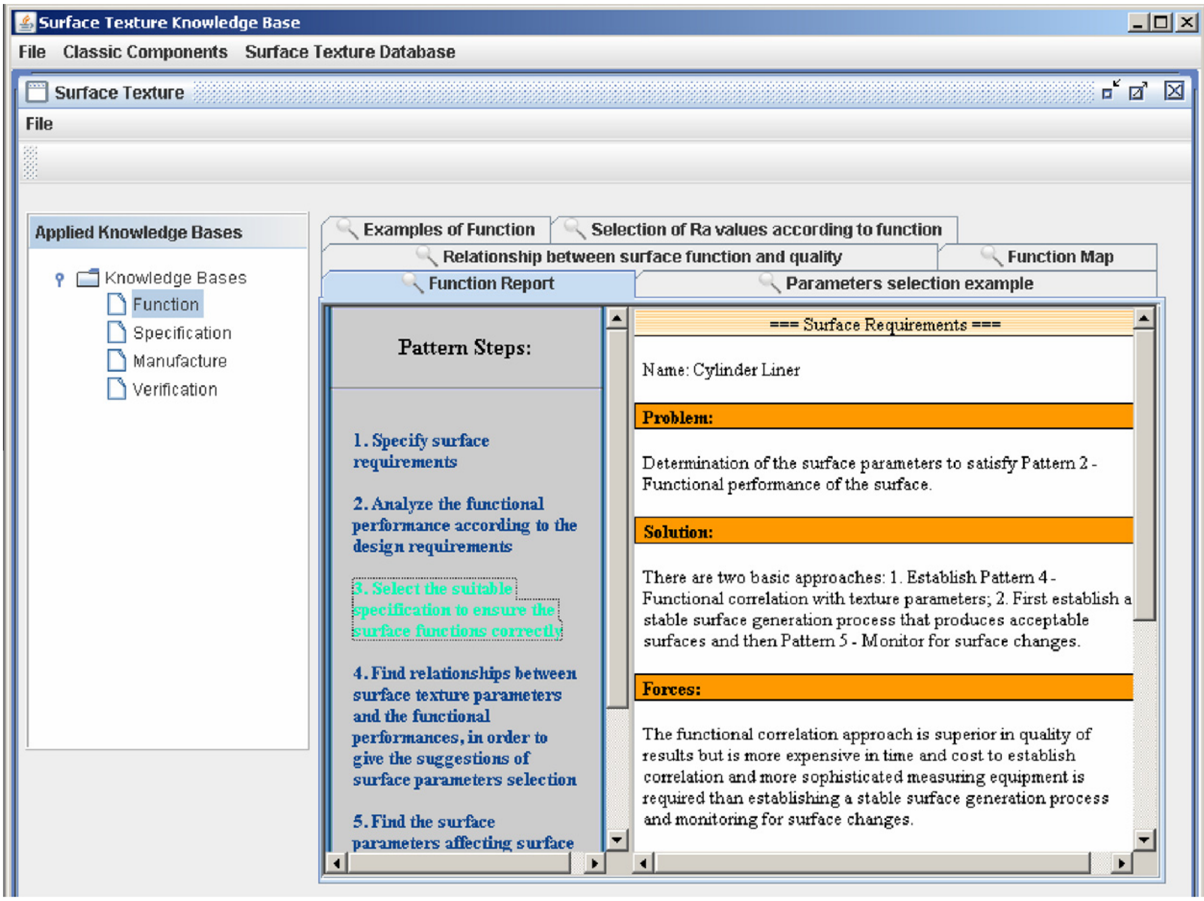

Fig. 8. An example of a function performance report.

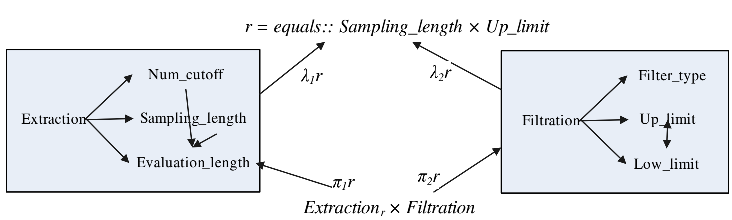

Fig. 9. Pullback representation of the constraint "equals".

3.5. Categorical representation for measurement theory

One of the main attractions of CT in this project is that it provides a rigorous mathematical foundation to define the measurement theory. As being successfully proven in the past, the representational measurement theory can be used to define the stability of the measurement procedure [16,17]. The measurement procedures relating to this project contains three key points in terms of the applied representational measurement theory:

(a) An empirical relational system (ERS); consisting of a set of objects on which a measurand is defined together with the relations between other relevant measurands.
(b) A numerical relational system (NRS); comprising numbers (derived values) and the relationships between them.
(c) A set of mappings; referred as the measurement procedures, map from ERS to NRS, in such a way that the relationships between measurands are matched by relationships between numbers.

A measuring procedure is regarded as mathematically stable, when a "small" difference in the derived values can imply a "small" difference in the measurand. Relationships between measurement values should reflect functional significant properties between the measurands; if not, the measurement is rendered unusable [16]. Since in Topology open set can be used to define "small" differences between points, the stability condition of measuring procedure can be described using the Topology and Set Theory. In 2004, Scott [17] devised stability corollaries that can be used to justify when a measurement procedure is stable or not using following rules:

Corollary 1. *"Finite sets of measurands and derived values with partial pre-orders and increasing mappings map one-to-one onto finite topologies with continuous functions."*

Stability corollary: *"If for a measurement procedure, the relational structures of the measurand and the derived values are both partial pre-orders and the mapping between them are also increasing mappings then the measurement procedure is stable."*

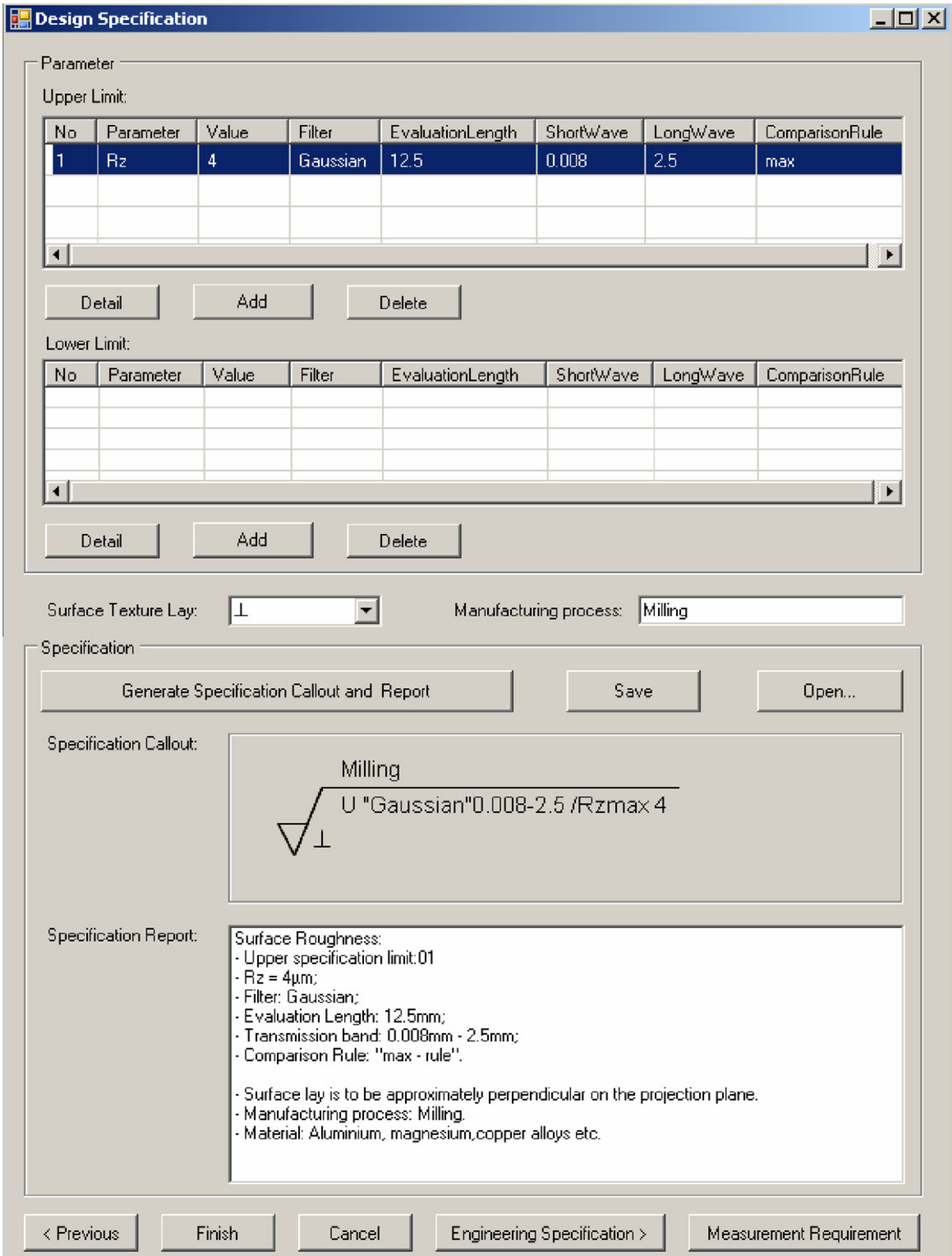

Fig. 10. Output of a specification report.

Based on the above rules, if define topologies on the space of measurands and the space of derived values, the stability condition is just a continuous mapping from the measurands to the derived values (if the inverse image of every open set on the topological space of the derived values is an open set on the topological space of the measurands, this is a topological definition of a continuous mapping). Researchers in this project found CT can provide a visual framework to vigorously represent the corollay1 and stability corollary using notions and constructions defined in CT. The following points give a short explanation on how CT represents the stability corollary:

(1) In order to satisfy the stability corollary, both ERS and NRS for a measurement procedure should be partial pre-orders with properties of reflexive and transitive, so categories are used to represent ERS and NRS while arrows inside the

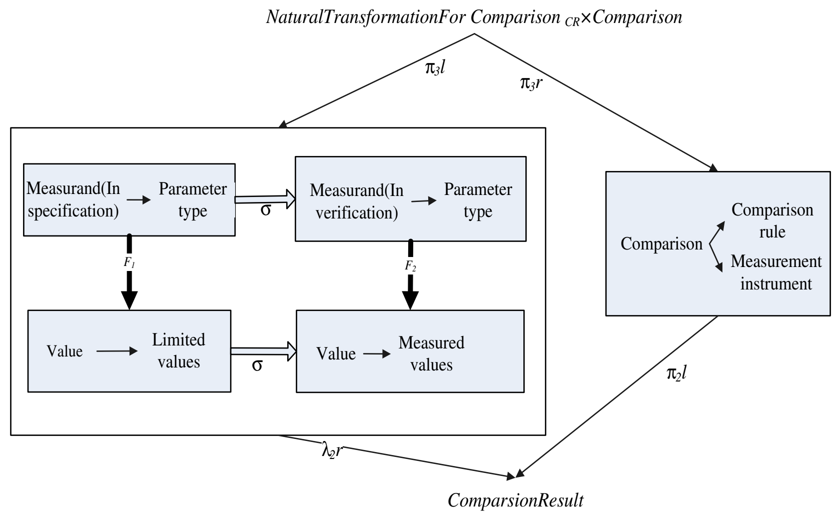

Fig. 11. Comparison between specification and verification.

category are used to represent partial pre-order. Moreover, objects in ERS or derived values in NRS are represented as internal objects of category. In Fig. 3 for transitive property, if arrow f: $A \rightarrow B$ and g: $B \rightarrow C$ represent binary relations in ERS, so $g \circ f$: $A \rightarrow C$ is categorical representation of transitive property in partial per-order.

(2) According to the definition of functor, functor must preserve identity arrow and the compositions of arrows inside categories. Therefore, functor can gracefully represent increasing mappings between ERSs or NRSs with partial pre-orders defined.

Based on the above two definitions, the categorical way of defining a stability corollary can be restated as: "If for a measurement procedure, the relational structures of the measurands and the derived values are both partial pre-order categories and the mapping between them is functor then the measurement procedure is stable". The development of the stability corollary in categorical way is beneficial for retrieving useful features from the observable data relating to this project and ensuring consistency of the knowledge acquisition for this knowledge-based system. Moreover, by adding the "natural transformation" notion of CT, the whole verification process can be refined as Fig. 11.

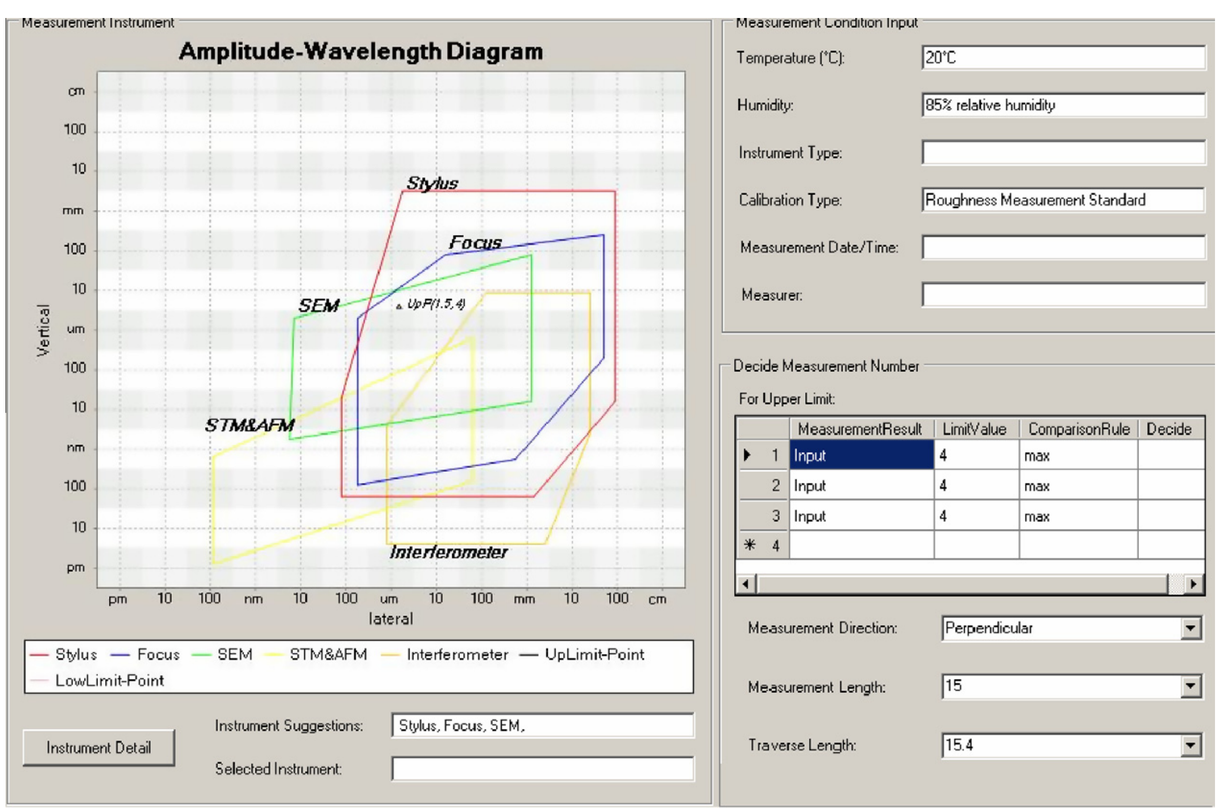

Fig. 12. Output of a verification report.

As a natural transformation provides a feasible way for transforming between functors while respecting the internal structure of the categories involved, the final comparison between the suggested GPS parameters with toleranced values in specification knowledge-base and the measurands with measured values in verification knowledge-base is represented as natural transformations. In Fig. 11, F_1 and F_2 are functors mapping from partial pre-order category "*Measurand*" to partial pre-order category "*Value*". The σ is natural transformation mapping from F_1 to F_2. The F_1, F_2 and σ form a natural transformation square. Fig. 11 also shows a 2-ary pullback relationship structure between a natural transformation square and a class category "*Comparison*" for storing the final comparison results in the database.

The final software implementation for "Verification" component is shown in Fig. 12.

This final design step also uses the verification component to find out suitable measurement information for the cylinder liner, which can include traverse length, sampling space, measuring instruments. The virtualGPS can automatically provide the measurement results after inputting the measurement data by users. All comparison constraints and multi-level mappings will be automatically applied and preserved in the system based on the Fig. 11.

3.6. Categorical object model for DBMS

After describing the knowledge-base in CT terminology, this project moves on to the next phase of developing an innovative DBMS with the ability of fully supporting the devised categorical object model. The first step in developing this categorical DBMS is to do further refinements on the categorical object model. Once the refined object model is established, the DBMS will have a sound mathematical foundation to ensure the integrity of databases when applying operations such as addition, deletion, and modification. The object model in categorical DBMS refines object model in knowledge-base by letting it more computing focused from following aspects:

- This project has used Java to implement the system and Java is a strongly typing language, the categorical object model adds the typing mechanism. In Category Theory, one discrete item is identified by the single category **1** [18]. Hence, typing can be added to show the types upon which the item is taken from in form of $\mathbf{1}_{TYP}$, where ***TYP*** can be the base types in OO programming language (e.g. String), other class categories, or other defined complexity such as arrays or lists. When **1** is other class categories or arrows, the values of **1** are class category or arrow names.

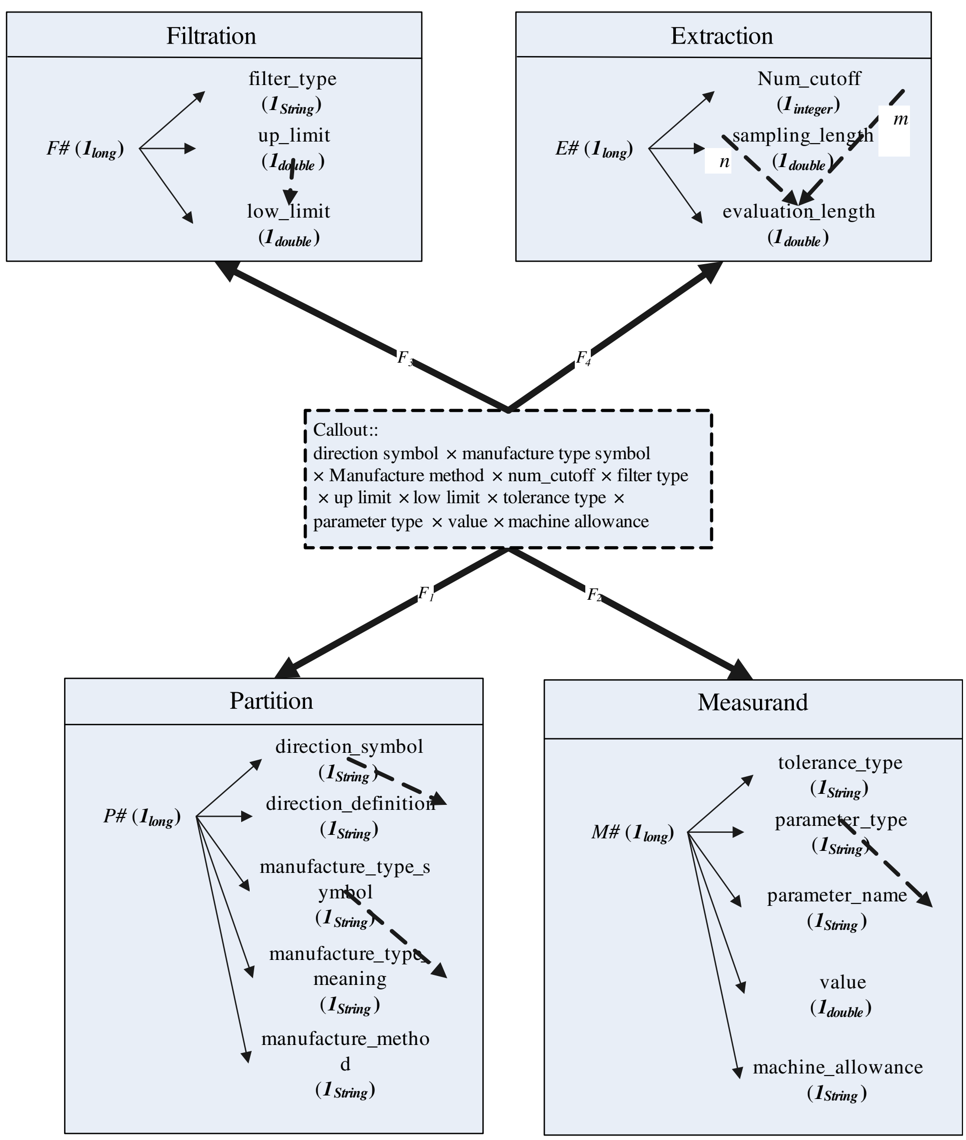

Fig. 13. Categorical representation of "Callout".

- The "pullback" structure for relationships or constraints in modelling of knowledge-base is represented in a universal product. Comparing with relational algebra that defines relationship as projection or Cartesian product, the categorical representation of relationships in the categorical DBMS is the universal product. At category level, the product is formed by categories and functors instead of internal objects and arrows with the vertex of the product become a category – relationship category. The relationship categories are stored and managed in the DBMS as class categories and instance categories.
- Because the modern programming languages such as Java can fully support objects embedding and other complex data structures such as list, array, tree, hashtable, hashmap and so on, some low-level categories represented in GPS knowledge-base can be embedded into other categories as internal objects. In addition, the constraints between two internal objects that belong to different categories often occur during the process of embedding some low-level categories into internal objects of high-level categories, so the categorical data model uses the functor notation of CT with unique name to model this situation.
- As an object-oriented DBMS assigns a unique identifier to every instance of a database entity, the vertex of the universal cone (limit) can be used to model the unique identifiers (see Fig. 13) [18]. If view the universal cone as a category, the vertex of the universal cone is actually the initial object in this category with an arrow from itself to every other objects (attributes) in the same category, which stores a unique automatically generating identifier values. These ID values cannot be modified by applications at run time and they are independent of how objects are created and manipulated. By modelling the database in this style, users have been spared the task of defining keys (primary keys or candidate keys).

Fig. 13 gives an example on modelling the Specification using the categorical object model for database side, which is actually a 4-ary pullback relationships – "Callout:: direction symbol × manufacture type symbol × manufacture method × num_cutoff × filter type × up limit × low limit × tolerance type × parameter type - value × machine allowance". The *P#*, *E#*, *F#*, and *M#* in the diagram are unique identifiers for Partition, Extraction, Filtration and Measurand, respectively. The F_1, F_2, F_3, F_4 are functors that project from relationship category "Callout" to the four class categories. In the "*Extraction*" category, the "evaluation length = num_cutoff × sampling_length" clause indicates the two arrows (m and n) are method arrows and the rest arrows are dependency arrows.

Further theoretical explanation on the categorical object model can be referred in a paper published by the team in 2008 [19]. Fig. 14 shows that the categorical DBMS can directly storage all diagrams conforming to the categorical object model and keep their constraints.

The database management system is developed based on a P/FDM system with following extensions [20,21]:

- Strictly defined object-orientation.
- Consistent data management and efficient querying mechanism and interfaces.
- Supporting complex data structures such as trees and lists.

Those features empowered users when updating existing categories (classes) and their instances, adding new patterns, querying data information (rendered in the tree structure) and generating

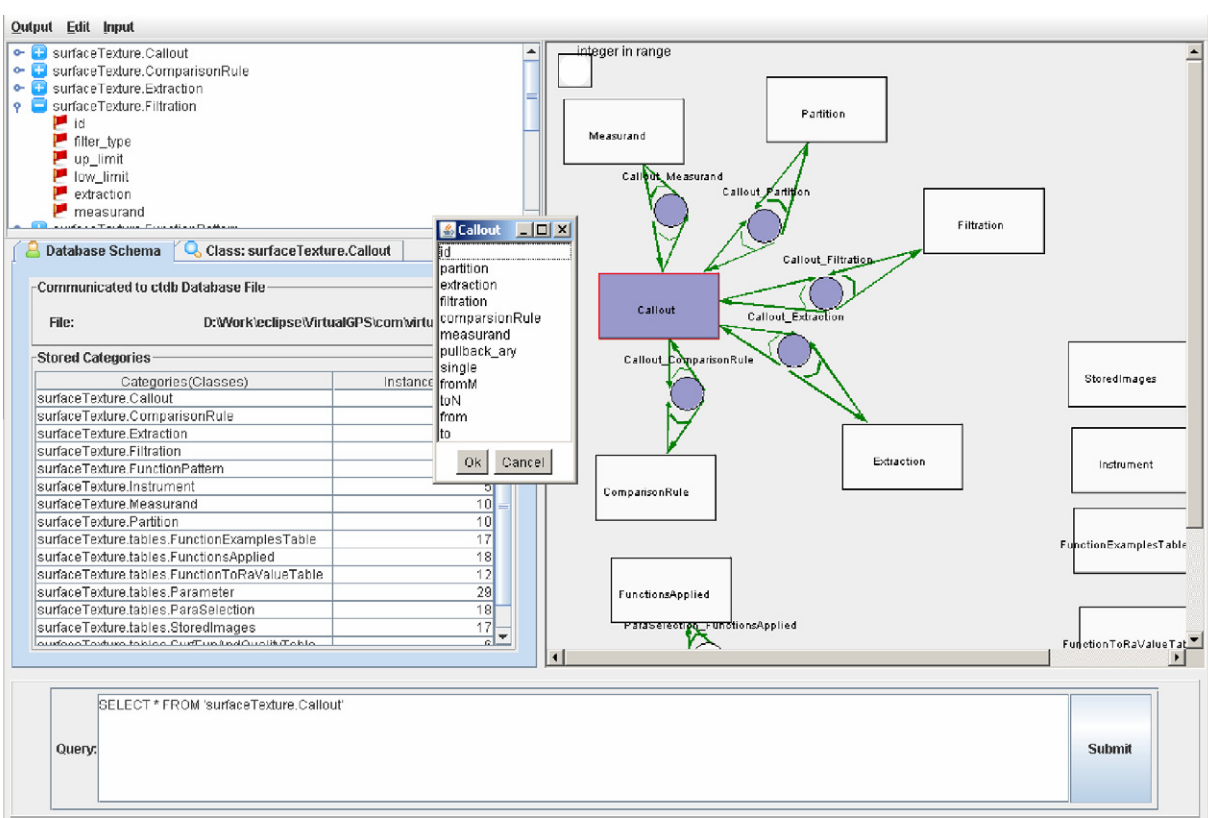

Fig. 14. Main interface for the categorical database management system.

XML files to transfer query results crossing networks or contact with CAD software.

4. Conclusions and future work

As demonstrated in this paper, CT has sufficient capability when defining and implementing most aspects of the proposed Virtual-GPS system including measurement procedure formation, knowledge generation, and data modelling. Manufacturing engineers and system designers can be benefited by the "virtual" GPS expert on improved communication and understanding. Moreover, the VirtualGPS database developers only need a few simple changes on the categorical object model devised by knowledge-base designers to form a refined categorical data model for database implementation.

It is hoped that the platform will be developed further to handle more complex GPS knowledge inferences to link closer between the real world applications and the ISO GPS standards. Future work of this project aims at adding more task specific features to help GPS users to improve the design and manufacture geometrical products. Fuzzy neutral network-based inference engine has also been planned to improve the "intelligence" of the VirtualGPS.

References

[1] A. Ballu, L. Mathieu, 1995. Univocal expression of functional and geometrical tolerances for design, Manufacturing and Inspection, 4th CIRP Seminar on Computer Aided Tolerancing, Tokyo, Japan, pp. 31–46.

[2] N.M. Durakbasa, P.H. Osanna, A general approach to workpiece characterization in the frame of GPS (geometrical product specification and verification), International Journal of Machine Tools and Manufacture 41 (13–14) (2001) 2147–2151.

[3] Z. Humienny et al., Geometrical Product Specifications Course for Technical Universities, Warsaw Bielsko-Biala Erlangen Tallinn Vienna, Huddersfield, 2001.

[4] ISO/TR 14638, 1995. Geometrical Product Specifications (GPS) –Masterplan, International Organisation for Standardisation, Geneva, pp. 1–23.

[5] ISO TC/213, 2001. Vision Statement [online], Available at http://isotc213.ds.dk.

[6] Y. Wang, et al., 2004.The structure of surface texture knowledge, [online], Available at http://www.iop.org/EJ/abstract/1742-6596/13/1/001.

[7] X. Jiang, A knowledge-based intelligent system for engineering and bio-medical engineering surface texture (VirtualSurf), Huddersfield: Research Project Report for University of Huddersfield, 2004.

[8] S.H. Bui et al., A framework for internet-based surface texture analysis and information system, Journal of Precision Engineering 29 (3) (2005) 298–306.

[9] M. Barr, C. Wells, Category Theory for Computing Science, Prentice-Hall International Series in Computer Science, London, 1996.

[10] S. Eilenberg, S. MacLane, General Theory of Natural Equivalences, Transactions of the American Mathematical Society 58 (2) (1945) 231–294.

[11] B.C. Pierce, Basic Category Theory for Computer Scientists, MIT Press, 1991.

[12] G. Kappel, S. Vieweg, Database requirements for CIM applications, Information Management in Computer Integrated Manufacturing 973 (1995) (1994) 136–164.

[13] L. Rising, The patterns handbook: techniques, strategies, and applications, Cambridge University Press, 1998.

[14] J. Neggess, H.S. Kim, Basic Posets, World Scientific, 1998.

[15] D.A. Nelson, B.N. Rossiter, M. Heather, The functorial data model – an extension to functional database, University of Newcastle Upon Tyne, Technical Report Series No. 488, 1994.

[16] P.J. Scott, The case of surface texture parameter RSm, Measurement Science and Technology 17 (2006) 559–564.

[17] P.J. Scott, Pattern analysis and metrology: the extraction of stable features from observable measurements, Proc. R. Soc. Lond. A 460 (2004) 2845–2864.

[18] D.A. Nelson, B.N. Rossiter, Prototyping a categorical database in P/FDM, in: Proceedings of the Second International Workshop on ADBIS, Springer, 1995, pp. 247–258.

[19] Y. Xu, Z. Xu, X. Jiang, Category theory-based object-oriented data management for web-based virtual manufacturing, International Journal of Internet Manufacturing and Services 1 (2) (2008) 136–159.

[20] S. Embury, User's manual for P/FDM V9. Object database group, Department of Computing Science University of Aberdeen (1995) 1–87.

[21] Y. Xu, Z. Xu, X. Jiang, Machining surface texture knowledge management using a category theory-based object-oriented database, System and Information Science Notes 1 (1) (2007) 83–88.