# A Comparative Study of Representation Methods for States of Information

Poongothai. P [#1], Devi. T [*2]
[#] Department of Computer Applications
Bharathiar University, Coimbatore, India
[1] poongothaimca39@gmail.com
[*] Department of Computer Applications
Bharathiar University, Coimbatore, India
[2] tdevi5@gmail.com

**Abstract—Concurrent engineering is a product development approach which requires collaborative multidisciplinary team environment with effective communication. The nature of approach determines the information needs of the project team. The flow of information can be of various types such as Sequential, Partial Overlap, Parallel, and Coupled. Among these types, Coupled is the more realistic process of simultaneous engineering, where information transfer is essential and involves iteration. Concurrent engineering involves free flow of information among team members and vague or partial information helps for follow up and overlapping tasks.The states of evolving information are: null, vague, partial and complete. As vague or partial information can represent valuable orientation for overlapping tasks, it is very essential to represent this state of information. None of the existing models have any room to represent the states of such information.This paper reports a detailed study on existing methodologies related to data quality and comparison of the methodologies.**

**Keyword-**Concurrent Engineering, States of Information.

## I. Introduction

Manufacturing industries are under constant pressure to increase the speed of product development and the product choices that are offered to customers, and product introduction is a vital process for most firms' growth and prosperity. A product is introduced by adding new functionalities or by improving the available functionalities of an existing product. Through the process of product introduction, ideas and needs are converted to the information from which technical systems and products can be made and sold. Concurrent engineering is practiced by industries to introduce new products to the market quickly. Concurrent Engineering involves interdependent activities and hence demands continuous availability of correct and complete information for efficient process management. Interacting activities are dependent on each other and they demand sharing of uncertain or vague information as they represent valuable orientation for overlapping tasks. This paper is concerned with a comparative study on completeness of information in order to find a novel mechanism to represent the states of information. As an idea is being analysed, a new product is being designed and developed, information is constantly accessed, interpreted, augmented, transformed and deployed. The basic elements of product introduction process are activities that require input information, take time to execute, and produce decisions or output information for input to other tasks [10]. After each new step, it may become necessary to upgrade or improve the results of the last, that is, to repeat it at a higher information level, and to reiterate until the necessary improvement has been made.

Concurrent Engineering is being practised by industries to bring new products quickly to the market. Interdependent tasks are involved in concurrent engineering, where information transfer is essential and iteration is typical. The basic effects of simultaneous work can be represented by a bi-directional information flow between tasks worked on in parallel. The first direction aims at early providing of results of the leading task. As this task is not yet completed, these results are just partial; containing some initial decisions already made, and is vague. Nevertheless uncertain or vague information can represent valuable orientation for following or overlapping tasks. Information created by individual gains value when it is shared with others. Operational with a growing number of content- generating devices, more information is being produced by individuals than by businesses. When created, information resides locally on product devices such as cell phones, cameras, and laptops. Data, whether structured or unstructured, it does not fulfil any purpose for individuals or businesses unless it is presented in a meaningful form. Businesses need to examine data for it to be of value.

## II. Concurrent Engineering

Concurrent Engineering (CE) has emerged as one of the most commonly referenced enablers for improved product introduction [14]. The definition of concurrent engineering proposed in the work of [8] is arrayed using the print format to reflect the arrangement of thoughts: "CE is a *systematic approach* to integrated product development that emphasizes *response to customer expectations* and embodies *team values* of co-operating, trust and sharing in such a manner that *decision making* proceeds with large intervals of*parallel working* by all life-cycle perspectives in the process, *synchronized* by comparatively *brief exchanges* to product *consensus*." A key term in a more precise definition of CE is to perform "interacting" activities at the same time. Thus parallel performance by itself is not a sufficient condition to characterize a process as concurrent. More managerial effort is needed in performing interacting activities rather than serial ones as they require considerable co-ordination effort to utilize resources effectively [7]. The effective deployment of concurrent engineering requires a smooth flow of information across the organization boundaries involved in the PI project. It is therefore expedient to carry out concurrent engineering with the aid of computer technology. However, computer-assisted concurrent engineering can only be implemented when concepts have been developed to fulfill the following requirements - ✶ uniform data models for displaying the product characteristics at the various phases of development, ✷ a system of information management to organize the information arising in the various phases and amendments to this information, and ✸ communication capabilities to support the simultaneous exchange of information between various disciplines [18]. In order to manage and control the PI process, it would be necessary to integrate the above mentioned data models and communication capabilities with the process models that represent the process and activities that generates and make use of the information. Thus, a marriage of process and information in the form of process-based information representation would seem inevitable in light of the growing interests in these two areas. The infrastructure required for realizing this identifies several tools and research issues. Many of these tools and issues deal with technology that is currently maturing, while others identify potential research initiatives.

Teamworking - The birth of multidisciplinary teams in PI is a result of implementing concurrent engineering practices and the need for employing the best resources for effective decision-making. The success of the implementation of concurrent engineering methods, tools and technologies in any industry will depend on the success of their teamwork [14]. Cross-functional teams work co-operatively to address downstream development issues as early as possible to create a high-quality, low-cost product more quickly. However, in a number of organisations, physically co-locating a team of relevant experts is very difficult or even impossible. For a variety of reasons, the relevant people, their tools and computer equipment are in geographically dispersed locations, and the cost involved in co-locating the people and the appropriate tools may be too high. One of the key factors for teamwork to be effective is 'good communication' i.e. the ability of team members to build on each other's ideas. The importance of information sharing and information flow in teamworking has been discussed by many researchers [12][19]. Thus, the foundation for the team-based co-operation is sharing of information, and the teams would require an information infrastructure that supports the sharing of information.

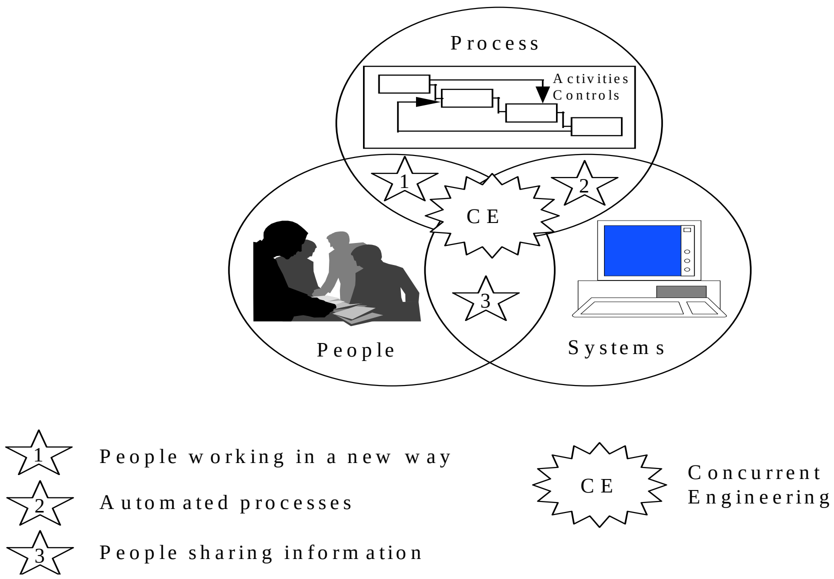



Fig.1. A structure of concurrent engineering [3]

In order to introduce high-quality, low-cost products quickly, now a days, concurrent engineering technique is practiced by industry. Concurrent Engineering (CE) has been loosely used to mean a host of different things. Both 'concurrent' and 'simultaneous' have the meaning 'happen at the same time', but the word 'concurrent' has the additional meaning 'agreeing; co-operating'[15]. Concurrent Engineering, with its emphasis on product teams made up of individuals from different departments and even different companies, and parallel working on processes that were previously carried out in series aims to overcome the disadvantages of the traditional method. Its purpose is to deliver the product in the shortest time possible and at the highest quality level [4] and it necessitates the focusing of the total effort towards achieving the common goal. Concurrency is present in three respects. At the highest level, different domains work on different perspectives of the product introduction process in parallel, if at all possible. Furthermore, several parts or components going into the final product may be worked on simultaneously; this is the second aspect of concurrency. And finally, within one domain, a group of designers guided by the group leader may be working on several different analyses of alternatives [8]. According to [6], concurrent engineering involves the integration of people, systems and information into a responsive, efficient system. According to [3], it involves the integration of process, people and systems (Fig. 1).

## III.INFORMATION FLOW AMONG THE ACTIVITIES

Problem solving can be considered as a form of information conversion, and it demands a constant flow of information [17]. The nature of the task determines the information needs of the project team [11]. Information delivery to a decision maker is clearly a critical factor in the effectiveness of the decision making process [1]. It would be essential to synchronize and manage the flow of the information and the resources (capacity in time) available. The interdependent or coupled tasks model given in Fig. 2 is a more realistic diagram of simultaneous engineering, where information transfer is essential and iteration is typical.

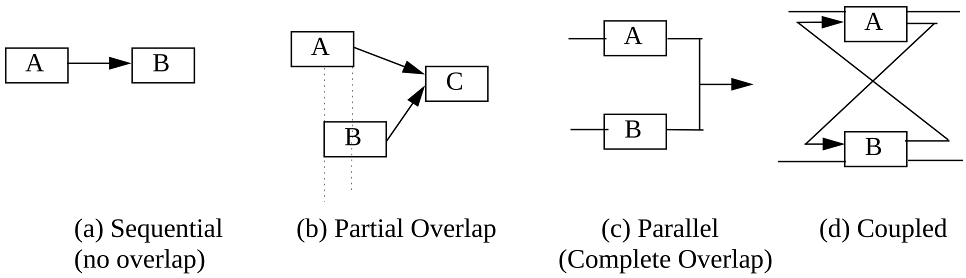



Fig. 2 Various types of tasks

The basic effects of simultaneous work can be represented by a bi-directional information flow between tasks worked on in parallel (Fig. 3). The first direction in Fig. 3 aims at early providing of results of the leading task 1. As this task is not yet completed, these results are just partial, containing some initial decisions already made, and are vague. Nevertheless uncertain or vague information can represent valuable orientation for following or overlapping tasks. The second direction is backwards from task 2 to task 1. In task 2 input information is analysed under task specific conditions. For example, the suitability of the design methods in achieving the critical parameter values and effects can be reported back early. This analysis provides basic information for a productive feedback and allows iteration loops [13].

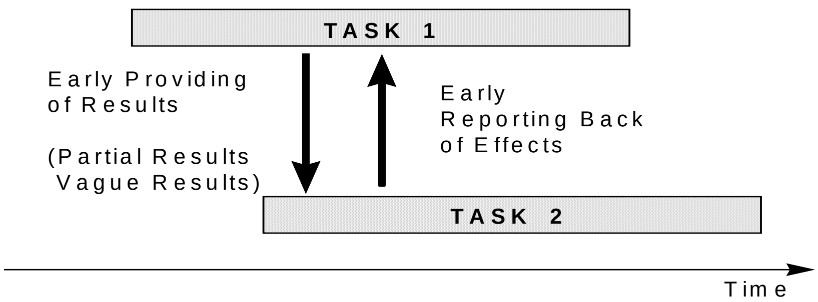



Fig. 3. Bi-directional information flow within simultaneous tasks[13]

## IV. METHODS FOR REPRESENTING INFORMATION COMPLETENESS AND COMPARISONS

The various methods for representing information completeness have been studied in detail. The comparative details are given in Table I. The methods are compared for Approaches followed, Existing Models / Tools, Functions used, Usage, Supporting level such as Conceptual, Measurement and Prototype, Data Quality Parameters such as Coverage, Density, Consistency and Completeness, Tradeoff Data Completeness and Data Consistency, Supporting level of completeness, and applications of the approach. The methods proposed by F. Naumann and J. C. Freytag – NFA [2000], Donald P. Ballou and Harold L. Pazer – BPF [2002], Yan Zou and Zhi Xiao – YZA [2008], Wenfei Fan and FlorisGeerts – FGM [2009], Shashi K. Gadia et. al., - SGM [1992], CinziaCappielloet. al., - CCM and Devi Thirupathi – DTM [1998] are considered for this research. Out of seven methods, three of them follow data analysis approach-, they are FGM, SGM, and CCM. First two methods NFA and BPF follow formal approach to the measurement of information and the combination of such measurement for information integration. Data analysis is a process of cleaning, transforming, and modeling data with the aim of discovering useful information and supporting decision making. Data analysis has multiple facets that encompass diverse techniques under a variety of names, in business, and science domains. Apart from this, data query analysis approach is proposed by FGM, data analysis is needed in connection with query processing, to produce data summary information in form of rules that allow semantic query optimization or direct query answering without connecting the data itself. The methods use some existing models / tools such as Value model, Bayesian model, Relational model, Storage model, incomplete model, and Information system architecture model. NFA approach is based on value model that incorporates both coverage (extensional value) and density (intensional value) of information. YZA approach is based on Bayesian model, it is predicted by experts experience and statistics and probability distribution form the basis.

Relational model is followed by both FGM and SGM. The relational model has been considerable in the area of incomplete information. The storage model must be considered in two points for incomplete information in databases. First, a given objects may know the values for a given attribute at some points but the values at other points may be unknown. Second, at some points sure that the object must exist in the relation but other points the existence of the object in the relation is not certain. Apart from the storage model, to maintain partial information so using a powerful algebra to query the incomplete information. CCM is tested on information system models for validation and evaluation of information in multi-channel financial institutions. An integrated information model is followed by DTM.

By looking at the functions used in the seven approaches, each of them follow different functions such as Merge function, Utility function, Mapping function, Mathematical (algebraic) function and Information mapping function. NSA follows Merge function; i.e. by comparing two databases, merge the changes into the database. The merge function is provided for adequately scoring integrated results and also combined the two criteria to an overall completeness criterion that formalizes the intuitive notion of completeness of query results. The completeness measure tool is used to assess source size and to predict result of queries in integrated information. This measure is an important step towards the usage of information quality for source selection, query planning, query optimization and quality feedback to users. Applications can be found in web, such as meta search engines, stock information systems and commercial applications such as eBusiness applications - marketplaces and eProcurement systems for inter – or intra organizational projects. BPF follows Utility function; to introduce a utility function for optimizing the completeness – consistency tradeoff. A tradeoff is a situation that involves losing one quality of something in return of gaining another quality. A tradeoff is commonly expressed as opportunity cost which is preferred alternative when taking an economic decision. The utility of various combinations of completeness and consistency for fixed and variable budgets provides guidance as to the appropriate tradeoff for specific decision contexts. BPF is applied in the Human Resource data system wherein different evaluators generate information regarding a particular employee.

YZA follows Mapping function; in view of the particularity of the value of domain of mapping function in soft sets. For standard soft sets, the decision value of an object, and the weight of each possible choice values is decided by the distribution of other objects and for fuzzy soft sets based on the method of average – probability. In fuzzy soft sets, the data analysis approach is different from the approach of standard soft sets, because in fuzzy soft sets each object belongs to the mapping function with a specific probability. YZA is applied in analysis, data mining and fuzzy control particularly in data deletion and data filling in databases. FGM is used to investigate the question of whether a partially closed database has complete information to answer a query. There has been a host of work on missing values, under the Closed World Assumption and equally important to data quality is how to handle missing tuples under the Open World Assumptions. FGM has a uniform framework that can deal with both relative information completeness and consistency. It is applied in Master data management of departments and projects. Both SGM and CCM apply mathematical functions such as algebraic expressions and operations. The set of algebraic expressions can be divided into three mutually exclusive groups: temporal expressions, Boolean expressions and relational expressions. The definitions of these expressions and operators were strengthened to give more information, but the results obtained are not reliable. CCM shows how different distributed architectures impact on the level of completeness of data and CCM is applied in multi-channel financial services. DTM is used to represent the product data and process data. DTM methods are used in new product introduction and product development. As it can be seen from Table I the first two approaches provide conceptual models and relevant measurements, Whereas YZA model provides only measurements. The last four models FGM, SGM, CCM and DTM provide prototype. The various parameters for data quality are CV, DN, CS and CP. NFA represents CV, DN and CP. BPF and FGM represent CS and CP. YZA represents CS. SGM and CCM represent CP. DTM represents CV and CS. The tradeoff between CP and CS is dealt in BPF. It can be clearly observed that none of the approaches represent the states of the evolving information from vague to completeness.

## V. Representation of States of Information

As vague or partial information can represent valuable orientation for overlapping tasks, it is very very essential to represent this state of information. The states of evolving information are: null, vague, partial and complete. None of the existing models have any room to represent the states of information. When a representation mechanism omits significant effects, then important system behaviour remains unexplained.

Hence the paper towards representation of states of information needs to be initiated in a serious manner. In existing databases, mostly data represented are of complete nature. To represent and store the data, in databases, tables are used as shown in Table II. A table represents an entity. Columns represent the attributes and rows represent the records i.e. instance of an entity. When the states of information need to be represented, every data item may have four possible values i.e. null, vague, partial and complete in Fig. 4. Hence, with every cell of the table given in Table II, an array of maximum size four has to be attached as shown in Table III. Thus a three dimensional representation mechanism is essential to represent the states of information.

Table I. Comparative Study on Representing States of Information

| | **Approaches Followed** | **Existing Models / Tools** | **Function** | **Usage** | **Applications** | **Supporting Level** | | | **Data Quality Parameter** | | | | **State of information** | **Trade off Data CP and Data CS** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | **CPL** | **MSRT** | **PRT** | **CV** | **DN** | **CS** | **CP** | | |
| **NFA** | Formal Approach | Value Model and Complete measure tools | Merge functions | Source selection, query planning, query optimization and quality feedback from users | Meta search engines, Integrated stock information systems, or bibliographic services, eBusiness and eProcurement systems | ✓ | ✓ | × | ✓ | ✓ | × | ✓ | × | × |
| **BPF** | Systematic Approach (Frame work) | | Utility function | To evaluate the completeness and consistency tradeoff for both fixed and variable budgets for decision makers | Industries (Human resource data) different evaluators generate information regarding a particular employee | ✓ | ✓ | × | × | × | ✓ | ✓ | × | ✓ |
| **YZA** | Data analysis approach | Bayesian model | Value domain of mapping function | Data analysis for standard and fuzzy soft sets weights of incomplete information | Decision analysis, Data mining, fuzzy control. (Data deletion and data filling in databases) | × | ✓ | × | × | × | ✓ | × | × | × |
| **FGM** | Data query analysis approach | Relational model | | To investigate the question of whether a partially closed database has complete information to answer a query under CWA and OWA | Master data management (departments and projects) | × | × | ✓ | × | × | ✓ | ✓ | × | × |
| **SGM** | Data analysis approach | Relational model, storage model, incomplete model | Algebraic expressions, operations used | Querying the incomplete temporal information in databases | Employee database in an organization | × | × | ✓ | × | × | × | ✓ | × | × |
| **CCM** | | Information System architecture model | Mathematical | How different distributed architectures impact of the level of accuracy and completeness of financial data | Multi-channel financial services | × | ✓ | ✓ | × | × | × | ✓ | × | × |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **DTM** | Integrated Information Model | Methodology Modelling | Information Mapping | Representing Process data and Product data | New Product Introduction and Product Development | ✓ | × | ✓ | ✓ | × | ✓ | × | × | × |

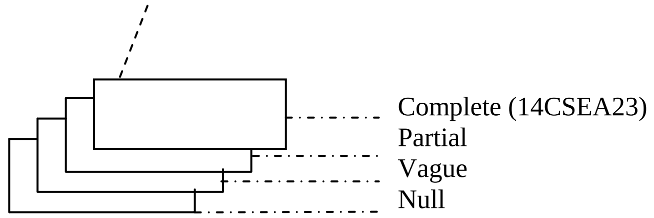
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CPL | → | Conceptual | CP | → | Completeness | NFA | → | F. Neumann and C. Freytag 2000 | CCM | → | CinziaCappiello et al., |
| MSRT | → | Measurement | CWA | → | Closed World Assumptions | BPF | → | P. Ballou and L. Pazer 2002 | DTM | → | Devi Thirupathi 1998 |
| PRT | → | Prototype | OWA | → | Open World Assumption | YZA | → | Yan Zou and Zhi Xiao 2008 | | | |
| CV | → | Coverage | CS | → | Consistency | FGM | → | Wenfei Fan and FlorisGeerts 2009 | | | |
| DN | → | Density | | | | SGM | → | Shashi K. Gadia et al., 1992 | | | |

Table II. Student Data Table

| Record no | Student ID | Name of student | City | % marks |
|---|---|---|---|---|
| 1 | Null | XXX | CBE | 75% |
| 2 | 23 | YYY | CHE | 80% |
| 3 | Null | ZZZ | CHE | 90% |
| 4 | Null | SSS | CBE | 60% |

Table III. Student Data Table with States of Information

| Record no | Student ID | Name of student | City | % marks |
|---|---|---|---|---|
| 1 | Null | XXX | CBE | 75% |
| 2 | 23 | YYY | CHE | 80% |
| 3 | Null | ZZZ | CHE | 90% |
| 4 | Null | SSS | CBE | 60% |




States of Information

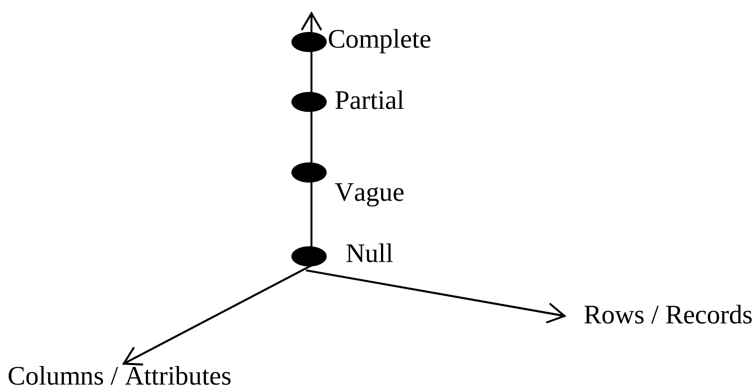



Fig. 4. Three Dimensional Representations

## VI. CONCLUSION

This paper reports an investigation into the information required for Concurrent Engineering, methodologies related to representing data quality and concepts to represent the states of information. The outcomes of this paper are: ①A review on Concurrent Engineering, Information flow in Concurrent Engineering, Information Completeness; ②A detailed study on methodologies related to representing data quality; ③A comparative study of the above methodologies and ④Concepts to represent various states of evolving information in databases. None of the existing methodologies address the representation techniques of states of information. Concurrent engineering involves free flow of information among team members and vague or partial information helps for follow up and overlapping tasks. Concepts for representing the states of information are proposed in this paper. The authors are developing these concepts further.

## ACKNOWLEDGEMENTS

The authors thank the University Grant Commission for providing Rajiv Gandhi National Fellowship to support this research by research fellowship.

## REFERENCES

[1] ANGUS J., and MURDOCH T., 1993, Supporting Product Definition Process, Proceedings of the International Conference on Engineering Design (ICED'93, The Hague, 17-19 August), pp.1725-1732.
[2] BALLOU D.P. and PAZER H.L., 2002, Modeling Completeness versus Consistency Tradeoffs in Information Decision Contexts, IEEE Transactions on Knowledge and Data Engineering, Vol.15,No. 1.
[3] BROOKS B., 1995, Time Equals Money - But where does it all go?, IEE Colloquium on Concurrent Engineering - Getting it Right First time, (London, 8 June), pp. 4/1 - 4/23.
[4] BUDILL E.J., 1989, How Process Logistics Planning Can Enhance the Effectiveness of Simultaneous Engineering, Proceedings of the Simultaneous Conference, pp. 73-81.
[5] CAPPIELLO C., BATINI C., FRANCALANCI .C, and MAURINO A., 2009, Methodologies for data quality assessment and improvement. ACM Computing Survey - vol.41, No.3.
[6] CARVER G.P., and BLOOM H.M., 1991, Concurrent Engineering through Product Data Standards, (U.S.: Department of Commerce, May).
[7] CERC, 1992, Minutes - CERC's First Workshop on Product Development Process Modeling and Characterization, Research Note, CERC Technical Report Series, (West Virginia : West Virginia University, CERC-TR-RN-92-006), pp. 14.
[8] CLEETUS K.J., 1992, Definition of Concurrent Engineering, CERC Technical Report Series, (West Virginia: West Virginia University).
[9] DEVI THIRUPATHI., 1998, Integrated Information Model for Managing the Product Introduction Process, University of Warwick, UK.
[10] EPPINGER S.D., WHITNEY D.E., SMITH R.P., and GEBALA D.A., 1994, A Model-Based Method for Organizing Tasks in Product Development, Research in Engineering Design, pp. 1-13.
[11] HAUPTMAN O., 1986, Influence of task types on the relationship between communication and performance: the case of software development, R & D Management, 16(2), 127-139
[12] KARINTHI R., JAGANNATHAN V., MONTAN V., PETRO J., RAMAN R., and TRAPP G., 1992, Promoting Concurrent Engineering Through Information Sharing, CERC Technical Report Series, (West Virginia : West Virginia University, Nov, CERC-TR-TM-92-10).
[13] KRAUSE F.L., and OCHS B., 1992, Potentials of Advanced Concurrent Engineering Methods, Manufacturing in the Era of Concurrent Engineering, Edited by Halevi G., Weill R.D., IFIP, (North-Holland: Elsevier Science), pp. 15-27.
[14] KUSIAK A., 1993, Concurrent Engineering - Automation, tools and Techniques, (John Wiley & Sons).
[15] LINDBERG L., 1993, Notes On Concurrent Engineering, Annals of the CIRP, 42(1), pp. 159-162.
[16] NAUMANN F. and FREYTAG J.C., 2000, Completeness of Information Sources, Humboldt-UniversitatZu Berlin, Institut fur Informatik, Tech. Rep. 135.
[17] PAHL G., and BEITZ W., 1984, Engineering Design, Edited by Wallace K., (London:Design Council).
[18] PFEIFER T., EVERSHEIM W., KONG W., and WECK M. (Ed.), 1994, Manufacturing Excellence, the competitive edge, (London : Chapman & Hall).
[19] RANKY P.G., 1994, Concurrent/Simultaneous Engineering (Methods, Tools and Case studies), (Guildford: CIM ware).

## AUTHOR PROFILE

Ms. P. Poongothai received her Bachelor of Computer Science from Velalar College for Women, Erode in 2008, Master of Computer Applications from Vellalar College of Engineering and Technology, Erode in 2012 and Master of Philosophy in Computer Science from Bharathiar University, Coimbatore in 2014. Her area of interest is Concurrent Engineering.

Dr. T. Devi received Master of Computer Applications from P.S.G. College of Technology, Coimbatore in 1987 and PhD from the University of Warwick, United Kingdom in 1998. She is presently heading the Department of Computer Applications, Bharathiar University, Coimbatore. Prior to joining Bharathiar University, she was an Associate Professor in Indian Institute of Foreign Trade, New Delhi. Her current research includes Software Engineering, Product Introduction, Technical Process Management and Concurrent Engineering. She has contributed more than 74 papers in various National / International conference / journals.