J. Basic. Appl. Sci. Res., 2(8)8344-8354, 2012
© 2012, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Meticulous analysis of Semantic Data Model An optimal approach for ERD

Muhammad Ishaq Raza[1,2], Qutab Jahan Zaib [3, 4] , Muhammad Shoaib Farooq[2], Adnan Abid[5], Sher Afzal Khan[1]

[1] Department of Computer Science, National University of Computer & Emerging Sciences, Lahore, Pakistan.
[2] Faculty of Information Technology, University of Central Punjab, Lahore, Pakistan.
[3] Department of Computer Science, Govt. College University, Lahore, Pakistan.
[4] IT/MIS, Sui Northern Gas Training Institute, Lahore, Pakistan.
[5] Department of Electronics & Information, Politocnico diMilano, Milan, Italy.

ABSTRACT

Medium to large scale modern information systems, process and manage their data through database system. In order to comprehend the requirements of a system, semantic models are required. For decades, semantic data model using Entity Relationship (ER) has been used as a powerful tool to understand the characteristics of different systems and model their schema. There are a variety of notations available for semantic modeling, but standardization does not exist due to a lot of variations in these notations. This causes a number of issues like maintenance, reusability and compatibility; hence a selection of an appropriate Entity Relationship Diagram (ERD) notation is very difficult. The purpose of this research is to contribute towards the standardization of these notations. For this purpose a meticulous analysis of the existing notations has been performed to devise an optimal ERD notation. The main contributions of our research are the following: i) Figure out the limitations in widely used data modeling notations ii) Provide an optimal ERD notation for data modeling. This will help the designer to model a real world scenario by choosing an appropriate notation which is supported by text books/CASE tools. Furthermore the proposed optimal notation will provide the liberty to model any real world scenario without limitations.

INDEX TERMS— ER: Entity Relationship, ERD: Entity Relationship Diagram, EER: Extended Entity Relationship, UML: Unified Modeling Language

1. INTRODUCTION

With the technology rush, every information system now processes and manages vast data in a database system. In order to comprehend the requirements of a system, semantic models are required. The Semantic Data Modeling (SDM) using ER provides a paradigm where a high-level conceptual schema can be developed without taking into account the internal-level issues such as physical data structures or the underlying database management system model (Date, 2004)**.** In 1976, Chen defined the ER model and ER Diagram, since then this model and the diagrams have been widely used for data modeling (Peter-Pin-Shan, 1976)**.** For decades, international conferences on ER modeling are being held, countless papers have been devoted to ER modeling that emphasize on the importance of the ER approach.

This model plays a major role in modeling schema for different real world scenarios. The understandability of ER and its power to model real world problems affirm that ER model provides a convenient illustrative procedure to logical database design. For effective correspondence between teams of people working on an application, intelligible symbolic notations are crucial. For this purpose, databases make use of entity-relationship diagrams (ERD) as it is an effective communications tool between database designers and end users (Andrea De Lucia, 2010)

There are several diagrammatic notations for ERD that have been constructed and are being used interchangeably in the referring material which may include published text, educational materials and diagrammatic modeling systems, so maintenance become very difficult. Re-usability principle in software industry is not efficiently implemented as a result of different notations. This is because the person reusing the model may not have proper understanding of that particular notation in which the model was made and hence

***Corresponding Author:** Sher Afzal Khan, Department of Computer Science, National University of Computer & Emerging Sciences, Lahore, Pakistan.

conversion from one system to another may not be completely correct. Moreover, different notations are not compatible with each other. ER model is taught in all Computer Science (CS) universities as it is a core topic in CS curricula (IS, 2010). Due to different notations, the selection of a particular notation to teach ER model becomes difficult. As a result, the choice of ER model notation varies across universities. Users usually construct the semantic data model mainly based on the personal preferences like ease of use, convention and technical hindrances instead of considering the aspect of provision of all the constructs necessary to develop a semantic data model (Helen C. Purchase, 2004**).** Another factor that can influence the selection of ERD notation is to find the support of text book or CASE tool available for any selected notation.

As there are a numerous number of notations currently in use around the world, a survey has been conducted to find the most popular notations being taught in top universities of the world. For data gathering, we select top 100 universities according to the 2011 QS world university ranking (QS World University Rankings, 2011)**.** For the selection of books and notations followed by these universities, we go through the database course outlines followed by these universities. On the basis of frequency of usage, we put together all the notations found in our survey, in tabular form. The coverage of these notations in text books and CASE Tools is presented in graphical form. We have selected nine notations as presented in **Table1** being taught in top universities.

Table 1 USAGE OF NOTATIONS BY UNIVERSITIES

NOTATIONS		OCCURRENCES
1.	Korth & Silberschatz (Silberschatz, 2010)	47
2.	Elmasri & Navathe (R. Elmasri, 2011)	25
3.	Chen (Peter-Pin-Shan, 1976), (Chen", 1983)	4
4.	Bachman (Bachman, 1992)	4
5.	Batini, Ceri , and Navathe (C. Batini, 1992)	2
6.	Information Engineering (KnowledgeWare, 1991), (IEF, 1990), (Martin, 1990)	2
7.	Teorey (Teorey, 1990), (Teorey, 1999)	1
8.	Oracle's CASE*METHOD (Barker, 1990)	1
9.	IDEF1X Information Mode (T.Bruce, 1992)	1
10.	Others	13

On the basis of this survey we have identified the support of these notations available in any text book and/or CASE tool **Fig 1**. An in-depth analysis of the detailed constructs supported by each notation has been performed with the perspective of purposing a notation for ERD containing maximum constructs support. This proposed notation has support to 12 top level constructs which are further broken down into 29 detailed level constructs. **Table 3** shows both level of constructs and their support in 9 selected notations. The construct support has been characterized into three levels X = No Support, P=Partial Support and BLANK= Full Support.

The research work is geared toward selecting optimal notations for ER diagram from the selected set by scrutinizing the notations for ER constructs support and understandability. For the selection of symbols for this proposed ER notation, we have reused the existing popular/common notation with extension. For the proposed ERD notation, we choose a notation from the selected set having minimum limitations in providing the support to the ER modeling constructs.

Our main objective to propose a notation having support to all constructs. To form the selected notation into optimal ERD notation, it must be extended with the missing construct. We have also proposed our own symbol set for the constructs those are not supported by the selected notation. Using this optimal notation, any real world scenario can be modeled without any limitation.

The rest of the research paper has been arranged in this way: Section II contains survey details to select the popular notations and to identify the support of text books or CASE tools for these notations. In section III, to propose a notation for ERD containing maximum construct support, a meticulous analysis of the detailed constructs supported by each notation has been performed. In section IV, for the selection of symbols, select a notation from the selected set with minimum limitations and then extend this selected notation with the missing symbols. Section V briefly discusses the related work. Section VI concludes the paper and also discusses the future work.

2. Popular erd notations

2.1. Usage of ERD notations by universities

The survey is conducted to find the support of different models in top universities of the world. We gathered data according to the university ranking, the books followed in those universities and the notations supported in those books. Then we put together all the notations' data on the basis of highest usage and displayed in tabular form.

The QS World University Rankings 2011 (QS World University Rankings, 2011) is among the most trusted world university rankings. Here one can find world's leading overall universities, the best universities by subject rank area, and the best universities as voted by employers. The data has been collected from different universities' websites and through emails.

Table1 shows the usage of notations by the universities in tabular form.

2.2. Usage of ERD Notations by Books and CASE Tools

The Figure 1 below shows the usage of ERD notations by books and CASE tools. The index for the list of ERD notations is mentioned in Table 2.

Figure 1 USAGE OF ERD NOTATIONS BY BOOKS & CASE TOOLS

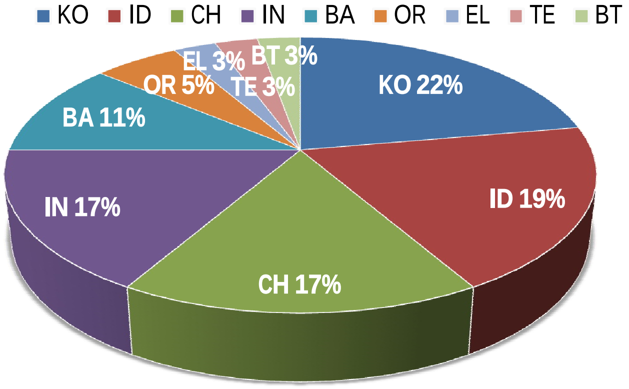


Table 2 Index for ERD Notations

CH	CHEN
TE	TEOREY
EL	ELMASRI & NAVATHE
BA	BACHMAN
KO	KORTH & SILBERSCHATZ
ID	IDEF1X
BT	BATINI, CERI, and NAVATHE
OR	ORACLE'S CASE METHOD
IN	INFORMATION ENGINEERING

This survey is an attempt to find the most widely taught notations in the universities for conducting the course of Database Systems. We collected the data in the form of a table and formulated results in graphical and tabular form. Another analysis derived from this survey identifies the books and CASE tools those are using these notations. The result of the survey shows that the most extensively used notations are Korth & Silberschatz notation and Elmasri & Navathe notation. These results are helpful in proposing the optimal ERD notation keeping in the view the limitations of notations under consideration.

3. PROPOSED OPTIMAL ERD NOTATION - CONSTRUCTS

On the basis of survey conducted in section 2, we have listed down the constructs supported by each ERD notation thus we are proposing a notation for ERD containing maximum constructs support. This proposed notation has support to 12 top level constructs which are further broken down into 29 detailed level constructs.

As the proposed ERD notation supports maximum number of constructs, almost all the real world problems can be modeled using this notation. This section provides a brief description of the constructs supported by our proposed ERD notation.

1) Entity Type:

Entity is an object that is represented in the database. Entities which have same attributes are grouped into an entity type (Peter-Pin-Shan, 1976)**.**

Strong entity: is an entity that must have self identifier. Its existence may or may not depend on another entity.

Weak entity: is an entity that does not have self identifier. Its identifier always includes the identifier of its parent entity and some attributes of weak entity as a partial key. Existence of weak entity always depends on its parent entity.

2) Attribute:

Attribute is a characteristic of an entity or a relationship. Following are the different types of attributes (Peter-Pin-Shan, 1976).

Single-valued attribute: is an attribute that consists of a single atomic value. For example, birth date attribute of employee

Composite attribute: is one that comprises of different components that make up the attribute. For example, name attribute of employee comprises of three components i.e. first name, middle initial, and last name.

Multi-valued attribute: is an attribute that can have more than one value for a particular entity. For example, an employee may have multiple contact numbers.

Complex attribute: is an attribute that is hybrid of composite and multivalued attribute. For example, an employee may have multiple addresses and each address is a composite of house no, street no, area, city, and country.

Derived attribute: is calculated from other attributes or relations. For example, calculating annual salary attribute of employee by multiplying monthly salary by 12.

3) Key:

Key is an attribute that uniquely identifies an entity (Peter-Pin-Shan, 1976). For example, employee number.

Composite Key: a key that is composed of more than one attributes. For example, vehicle number attribute of employee comprises of city code and serial number.

Partial key: attributes of weak entity that are part of its identifier. For example, dependent name attribute of dependent weak entity type without employee no attribute of employee parent entity type.

Foreign Key: a set of referring attributes of an entity type to another set of referenced attributes of same or other entity type is called a foreign key (Bachman, 1992)**,** (T.Bruce, 1992).

4) Relationship Type:

Relationship is an association between two or more distinct entities. Following are the two main types of relationship (Peter-Pin-Shan, 1976).

Identifying Relationship: is a relationship between owner entity and weak entity, where the existence of the weak entity depends on the owner entity.

Non-Identifying Relationship: is a regular relationship between two or more independent entities.

5) Relationship Degree:

Relationship degree refers to the number of entity types involved in the relationship (Peter-Pin-Shan, 1976).

Degree one (unary/recursive) relationship: have one entity type
Degree two (binary) relationship: have two entity types
Degree n (n-ary) relationship: have n number of entity types

6) Aggregation:

Allows relationship between aggregate entity set and other entity set. For example, employee works on a specific project using multiple tools i.e. works on relationship set among employee and project entities is an aggregate entity set, which have ‘using’ relationship with tools entity set (D.Smith, 1977)**,** (Silberschatz, 2010).

7) Structural Constraints:

Structural constraints are used to specify the limit on entities to participate in the relationship set. There are two main types of structural constraints (Peter-Pin-Shan, 1976).

Cardinality ratios constraints: are used to specify the maximum limit on entities to participate in the relationship set. The possible cardinality ratios are 1:1, 1:N, N:1, and M:N.

Participation constraints: are used to specify the minimum limit on entities to participate in the relationship set. The possible participation constraints are total and partial.

Total participation constraint specifies that each entity of an entity set must participate in the relationship set.

Partial participation constraint specifies that an entity of an entity set may or may not participate in the relationship set.

8) Subclasses, Super classes, and Inheritance:

An entity type may have some attributes which are common to all of its entities but there may be some attributes which are specific to some of its entities. An entity type that has common attributes is called superclass and an entity type that has only specific attributes is called subclass (D.Smith, 1977)**,** (Silberschatz, 2010).

Entity set of a subclass must be a subset of entity set of its superclass. For example, if C is the superclass and S is one of its subclass then, S ⊆ C. An entity in the subclass represents the same corresponding real world entity in the superclass.

Entity of subclass inherits all the attributes and relationships of superclass.

9) Generalization/specialization:

Generalization/specialization are used to specify the superclass/subclass association among entity types (D.Smith, 1977)**,** (Silberschatz, 2010).

Specialization is the top-down process of creating a set of subclasses of a superclass.

Generalization is the bottom-up process of creating a generalized superclass from several classes.

10) Constraints on Specialization and Generalization:

Predicate-defined subclass: Condition determines subclass members.

User-defined subclass: No condition determines subclass members; users determine subclass members when they add (R. Elmasri, 2011).

Attribute defined specialization: when all the subclasses of a specialization are predicate-defined.

Two constraints that can employ to specialization/generalization are disjointness and completeness.

Disjointness constraints: Disjoint employ that an entity in the superclass can participate in at most one of its subclasses of a specialization.

Overlapping (not disjoint) employ that an entity in the superclass can participate in more than one of its subclasses of a specialization.

Completeness constraints: Total employ that every entity in the superclass must participate in at least one of its subclasses of a specialization.

Partial employ that an entity in the superclass may or may not participate in any of its subclasses of a specialization.

11) Shared Subclass:

A shared subclass is a subclass which has multiple superclasses but it has only one superclass within each distinct superclass/subclass relationship (R. Elmasri, 2011).

For example, if S is a shared subclass and C_n are of its superclasses then

$$S \subseteq C_1 \cap C2 \cap \ldots \cap C_n$$

12) Category Subclass:

A category subclass is subclass which has multiple superclasses within distinct superclass/subclass relationship (R. Elmasri, 2011).

For example, if S is a category subclass and C_n are of its superclasses then

$$S \subseteq C_1 \cup C2 \cup \ldots \cup C_n$$

4. PROPOSED OPTIMAL ERD NOTATION : SYMBOLS

The research work is geared toward selecting optimal notation for ER diagram from the selected set by scrutinizing the notations for ER constructs support and understandability. For the selection of symbols for this proposed ER notation following are two ways: **1-** One way is to introduce entirely new symbols. But this approach is not highly recommended as untried symbols will not easily be adopted by the users. Hence there is a chance of failure of this notation. **2-** The other way is to reuse the existing popular/common notation with possible extension. This approach will be more appropriate as majority of the symbols are known to the users.

Using the second approach of selection of symbols, we have selected a notation from the selected set of notations having minimum limitations in providing the support to the ER modeling constructs.

In Table 3 we have listed down the limitations in the selected most popular ERD notations. The index for the list of ERD notations is mentioned in Table 2. The level of support/limitations can be represented as: No Support → X, Partial Support → P and Full support → BLANK

As per the survey and deep analysis of all the most popular notations we found that Elmasri's notation poses minimum limitations. These are two limitations in providing support to foreign key representation and aggregation.

We will be using the (min, max) notation for optimal ERD notation. The reason is that UML which is a standard also uses this method instead of the separate cardinality and participation representations. Moreover, the separate representation is suitable for binary relationships but for n-ary relationships where n>2, (min, max) notation is more articulate whereas separate representation becomes confusing.

Table 3- Construct Limitations in ERD Notations

CONSTRUCTS		CH	TE	EL	BA	KO	ID	BT	OR	IN
Entity Type	**Strong**									
	Weak							P	P	
Attribute	**Single**		X							X
	Composite		X		X		P		X	X
	Multi Valued		X		X		P		X	X
	Derived		X		X		P	X	X	X
	Complex		X		X	X	P	X	X	X
Key	**Primary**		X							X
	Composite		X						X	X
	Partial		X				X		X	X
	Foreign	X	X	X		X		X	X	X
Relationship Degree	**Unary**									
	Binary									
	N-ary				X	P	X	P	X	X
Relationship Type	**Identifying**									
	Non-Identifying									
Cardinality Constraints	**1:1**									
	1:N									
	M:N									
Participation Constraints	**Total**									
	Partial									
Generalization/ Specialization										
Generalization/ Specialization Constraints	**Disjoint**	X	X		X					
	Overlapping	X	X		X				X	
	Total	X	X		X					
	Partial	X	X		X				X	
Shared Subclass		X	X		X	X	X	X	X	X
Category Subclass		X	X		X	X	X	X	X	X
Aggregation		X	X	X	X		X	X	X	X

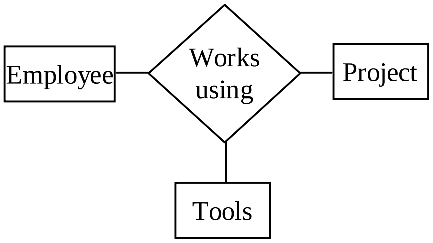
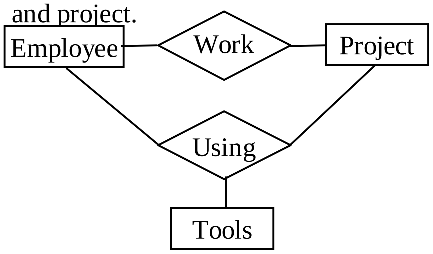
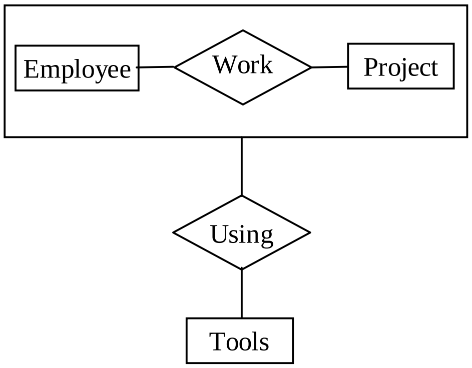
The Figure 2 below depicts the need of adding the missing constructs into our proposed solution. In figure 2(a) below shows the requirement for adding the support for aggregation construct. Without it certain requirement can't be represented accurately and precisely. The figure 2(b) below shows the requirement to represent the structural constraints using min and max notation. The figure 2(c) below shows the requirement for adding the support for foreign key and ERD level. Without it, complex rules are required for derivation of foreign key from cardinality ratio and participation constraints.

Figure 2-Proposed Solution for ERD Limitations

2 (a)

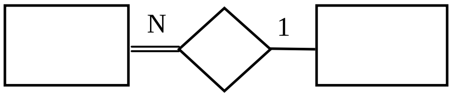
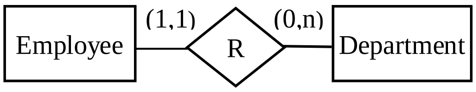
Requirement: Aggregation

Example: an employee works on a specific project using multiple tools.

EXISTING MODEL	PROPOSED MODEL
Semantically not correct, since it requires each project to use tools.  Redundant relationship between tools and project. 	**Consider the relationship set works and the entity sets employee and projects as a higher level entity set – an aggregate entity set. Make relationship between this aggregate entity set and tools entity**. 

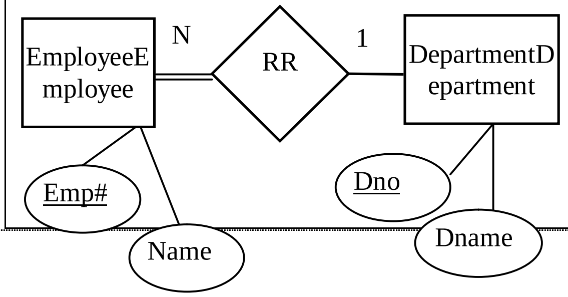
2(b)

Requirement: Representation of structural constraints (cardinality ratios & participation constraints)

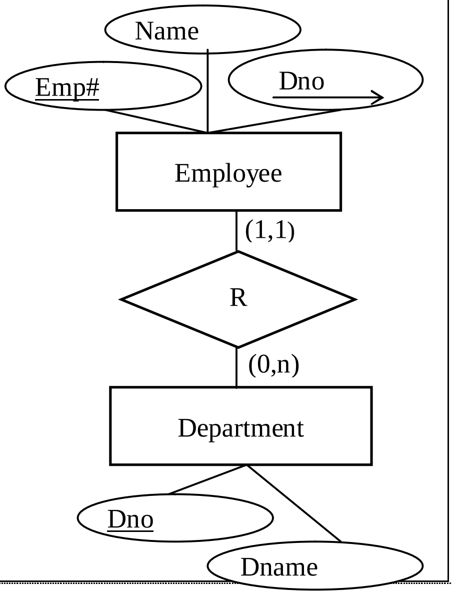
EXISTING MODEL	PROPOSED MODEL
General specification of structural constraints. 	Precise specification of structural constraints. In case of no constraint, use min=0 and max=n as a default. 

2(c)

<table>
<tr><th colspan="2">Requirement: Foreign Key</th></tr>
<tr><th>EXISTING MODEL</th><th>PROPOSED MODEL</th></tr>
<tr><td>
When mapping, foreign key is derived from cardinality ratios and participation constraints using the following complex mapping rules [15]:
(1:N): include the key of the entity with cardinality ratio 1 in the entity with cardinality ration N as a foreign key.
(M:N): includes the keys of the participating entity types in relationship type relation as foreign keys.
(1:1): - in case of one side is partial and other side is total, (1:N) rule is apply by considering total side as N side.
- in case of total participation on both sides, merge both side entity types in one entity type.
-in case of partial participation on both sides, (M:N) rule is apply.



</td><td>
Direct representation of foreign key at ERD level, avoids application of complex mapping rules to derive foreign key.



</td></tr>
</table>

5. related work

One of the major parts of database design is Data Modeling. Data modeling deals with the structure, organization, and effective use of data and the information they represent (D.C Tsichritzis, 1982). Such semantic modeling of the data has been helped by data models such as entity relationship data model (Peter-Pin-Shan, 1976) which models the data requirements of an enterprise as set of entities and relationships. Although there are numerous sources of literature that discuss the specifications, constructs and notations of various entity relationship diagrams, only a few number of papers performed comparisons for a set of notations. However, these sources do not focus on the detailed level constructs. The selection criteria for the set of notations is purely qualitative approach as authors choose notations of their own choice rather than following some systematic way like survey. Also, most of the papers discuss support of different notations for various constructs, whereas our research focused on highlighting the limitations in providing support to various constructs. This has been achieved with the help of comparative study and same like UML, we have proposed an optimized ERD notation that will provide maximum support to know ERD constructs.

(II Yeol Song, 1995) employs some part of the methodology as this research paper provides comparison using different notations to determine constructs representation in various ERD notations. However, this paper does not discuss, in depth, the detailed comparison for ERDs. Only seven top level constructs are discussed. The detailed constructs like attribute type (multi valued attribute, derived attribute, composite attribute, complex attribute), category type etc. are not discussed in this paper. Further no optimal model has been proposed that will support all the detailed level constructs.

(Helen C. Purchase, 2004) discusses few comparisons identified via figures for only two notations, which is much less than those handled in this research paper. Furthermore, it has provided efficiency and cost analysis that is not required for our research.

Another source (Peter-Pin-Shan, 1976) has discussed in detail how models differ in terms of their interpretation of attributes and relationship concepts. It has discussed a framework based on the following factors:

Models allow n-ary and binary relationships.

Models allow attributes of relationships, attributes for entities or no attributes at all.

Additionally, more research papers and books have been consulted for getting the complete construct support for the notations under consideration. Across the sources under consideration, there is no availability of information published regarding the comparison of these ERDs and also no optimal standard for ERDs has been suggested.

6. Conclusion and future work

There are several diagrammatic notations for ERD that have been constructed and are being used in published text, educational materials and diagrammatic modeling systems. Due to the lot of variations in the different notations, there are a number of issues to be faced like maintenance, reusability and compatibility hence selection of ERD notation is very difficult. The selection is usually based on the personal preferences like ease of use, convention and technical hindrances instead of considering the aspect of provision of all the constructs necessary to develop semantic data model (Helen C. Purchase, 2004).

The main objective of this research is to propose an ERD notation having minimum limitations and have support to maximum number of detailed level constructs. To find a comprehensive list of constructs containing 12 top level and 29 detailed level constructs a study/survey is being conducted. Next to find the symbols for this optimal notation, we chose the symbols of the popular notation having minimum limitations in providing support to the optimal ERD constructs. We propose our own symbol set for the constructs not supported by the selected notation. Using this optimal notation we can precisely and accurately model a real world scenario without any limitation.

A CASE Tool can be built that will support this proposed ERD notation. Furthermore a framework can also be built to provide cross transformation with ERD notations. This will also take care of the gaps and limitations among these notations at the time of transformation through the suggested optimal ERD notation. In addition to this an automated tool can also be developed to transform these notations with ease to users/ practitioners.

7. REFERENCES

1. Andrea De Lucia, C. G. (2010). An experimental comparison of ER and UML class. *Empir Software Eng*, 15:455–492.

2. Bachman. (1992). "Bachman Analyst, Bachman Information Systems Incorporated".

3. Barker, R. (1990). *"CASE*METHODTM: Entity Relationship Modeling".* New York: Addison-Wesley Publishing Company New York.

4. C. Batini, S. C. (1992). *"Conceptual Database Design: an Entity- Relationship Approach".* Benjamin/Cummings Publishing, Redwood City, CA.

5. Chen", ". P.-S. (1983). “A Preliminary framework for Entity-Relationship Models”.

6. D.C Tsichritzis, F. H. (1982). *"Data Models".* Prentic-Hall.

7. D.Smith, J. (1977). "Database Abstractions : Aggregation and Generlization". *1(1) : 105 - 133.*

8. Date, C. (2004). *"An Introduction to Database Systems", 8th edition.* Addison-Wesley.

9. H. S. Thompson, D. B. (2001). *"XML Schema Part 1 : Structures".* W3C Recommendation.

10. Helen C. Purchase, R. W. (2004). "Comprehension of diagram syntax: an empirical study of Entity Relationship notations". *International Journal of Human-Computer Studies* , pp187-203.

11. IEF. (1990). *"Technology Overview".* Texas Instrument.

12. II Yeol Song, M. E. (1995). "A Comparative Analysis of Entity-Relationship Diagrams". *Journal of Computer and Software Engineering, Vol. 3, No.4* , pp. 427-459.

13. IS. (2010). *" Curriculum Guidelines for Undergraduate Degree Programs in Information Systems".* ACM/AIS Joint Task Force.

14. KnowledgeWare. (1991). "ADW Case Tool Seminar".

15. Martin, J. (1990). *" Information Engineering: Planning & Analysis, Book II"* . Englewood Cliffs, NJ: Prentice-Hall.

16. Peter-Pin-Shan, C. (1976). "The entity-relationship model-toward a unified view of data". *ACM Transactions on Database Systems, 1, 1,* , pp. 9-36.

17. QS World University Rankings, b. S. (2011). *"Computer Science & Information Systems Rankings.".*

18. R. Elmasri, S. N. (2011). *"Fundamentals of Database Systems" 6th edition.* Addison- Wesley.

19. Silberschatz, H. K. (2010). *"Database System Concepts", 6th edition.* New York, N.Y: McGraw-Hill.

20. T. Bray, J. P.-M. (1998). *"Extensible Markup Language (XML) 1.0".* W3C Recomdations.

21. T.Bruce. (1992). *Designing Quality Databases with IDEF1X Information Models.* New York, New York: Dorset House Publishing.

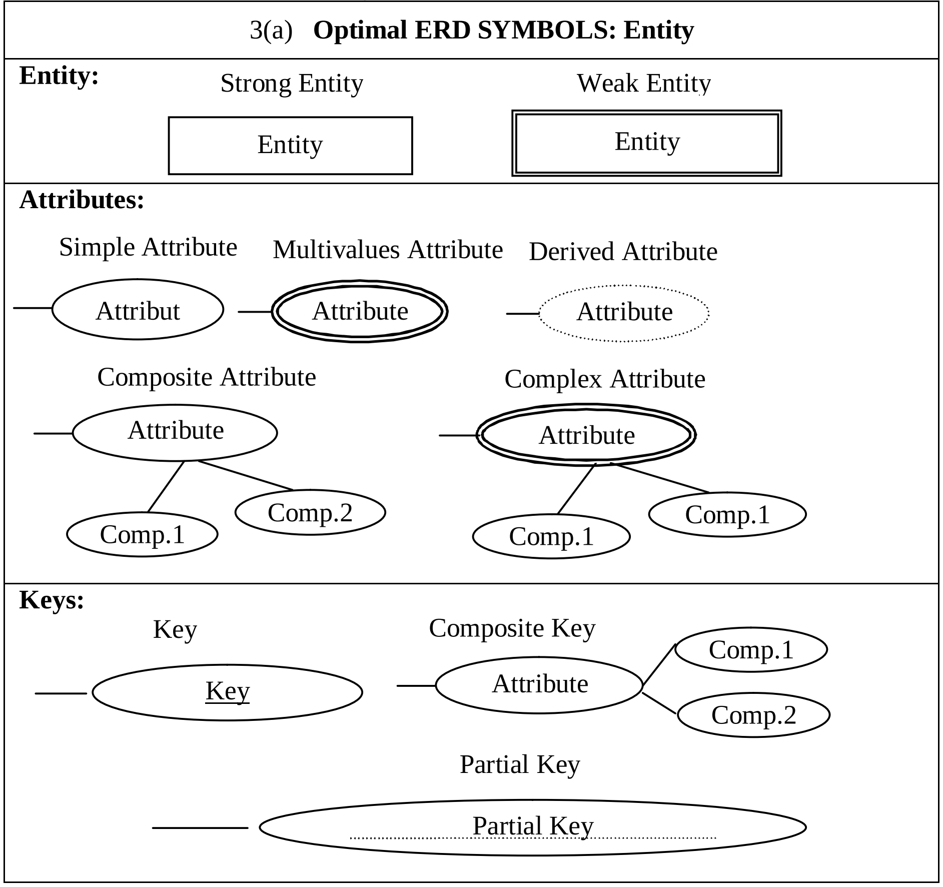
22. Teorey, T. J. (1999). *"Database Modeling and Design:The Fundamental Principles", 3rd edition.* San Francisco, CA: Morgan Kauffmann.

23. Teorey, T. J. (1990). *"Database Modelling and Design: The Entity-Relationship Approach".* San Mateo, CA: Morgan Kauffmann.

Appendix

The figure 3 below shows the ERD constructs and symbols of Elmasri ERD notations which can be subdivided into three groups: The figure 3(a) related to Entity Type, Attribute Type and Keys. The figure 3(b) represents the group of ERD constructs and symbols related to Relation Types, Relationship Degree. The figure 3(c) represents the group of ERD Constructs and Symbols related to Generalization/Specialization, Shared Subclass and Category Subclass.

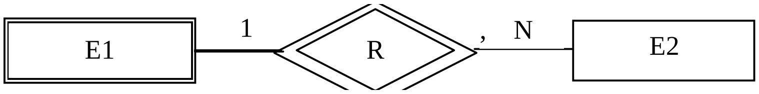
Figure 3- Elmasri ERD Constructs and Symbols



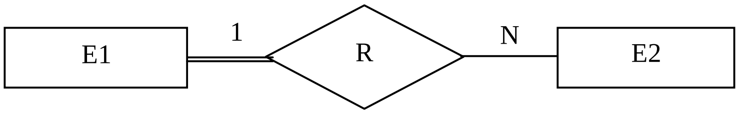
3(b) **Optimal ERD SYMBOLS: Relationship**

Relationship Types and Structural Constraints using (min, max) notation:

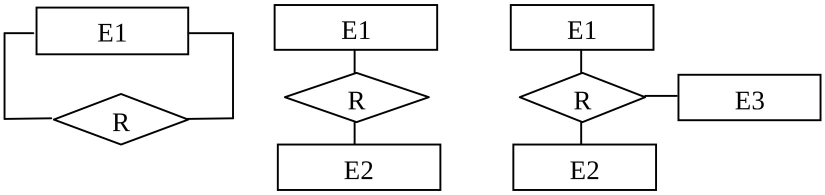
Identifying Relationship



Non-Identifying Relationship



Relationship Degree:


(1) Recursive Relationship

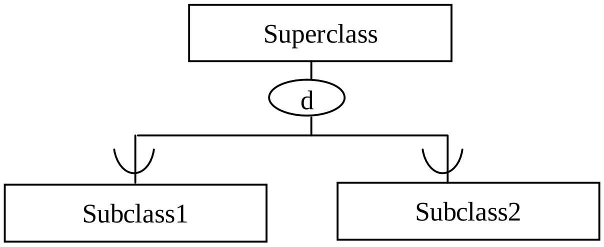
(2) Binary Relationship

(3) n-ary Relationship

3(c) **Optimal ERD SYMBOLS: Extended ER**

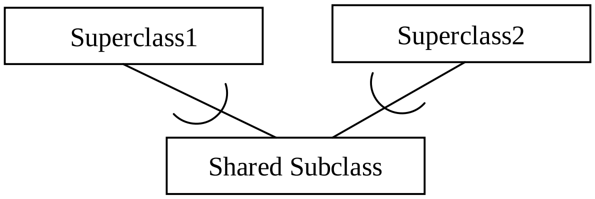
Generalization/Specialization:

Use ‘d’ for disjoin, ‘o’ for overlapping, single line for partial participation and double line for total.



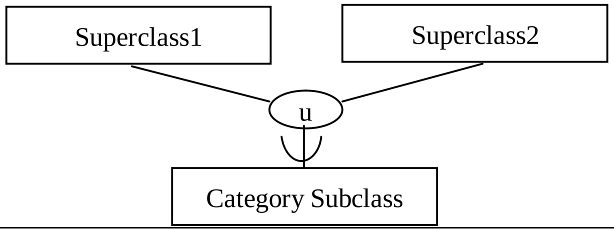
Shared Subclass:

It is a subset of intersection of all superclasses of same entity type.



Category Subclass:

It is a subset of union of all superclasses of same/different entity types.


Superclass1
Superclass2
u
Category Subclass