Generating EXPRESS Data Models from SBVR

Imran Sarwar Bajwa¹, *, Nadeem Sarwar¹, and M. Asif Naeem²

¹Department of Computer Sciences & Information Technology, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan
²School of Computer and Mathematical Sciences, Auckland University of Technology, Auckland, 1142, New Zealand

Abstract: EXPRESS is a standard graphical documentation for data models. It is a helpful buddy to the EXPRESS dialect for showing element and sort definitions, connections and cardinality. This graphical documentation underpins a subset of the EXPRESS dialect. One of the preferences of utilizing EXPRESS over EXPRESS is that the structure of an information model can be exhibited in a more reasonable way. In software engineering the graphical representation of the software structure is very necessary because understanding of coding is very difficult. Semantic Business Vocabulary Rules (SBVR) specifications many authors have done model to model transformation. SBVR2EXPRESS is also possible. In this study we propose an EXPRESS data model using Natural language. That model manually implements on two case studies and generate the EXPRESS diagram using SBVR rules. That diagram full fills the all software requirement of the Software engineers. Model representation is very helpful to develop large scale of systems like, Aerospace, medical science and other industries where the representation of system working is very important. This work is very helpful to the Data manger and IT managers to represent their organization structure.

Keywords: EXPRESS, SBVR, natural language rules

1. INTRODUCTION

The field of exploration and real issues of learning space are exhibited, highlights the tended to research issue, and portrays the examination inspirations and the significant exploration targets. In Data Modeling language Different types of Data Modeling languages have been available, There are two types of Modeling Languages these are textual and graphical. In Textual modeling language is use standard keywords associated by natural language expression and phrases or variables make computer explainable terms. Diagram technique with named symbols are use in Graphical Modeling language to represent concepts and lines that join the arbitrary signs and show the connection between them and various elements represent by using graphical notations. The EXPRESS [2] is an example of a textual modeling language and graphical modeling language. The EXPRESS-G and EXPRESS (International Standard Organization 10303-11) are international standard nonspecific purpose data modeling language [3, 4].

2. RELATED WORK

Ma and Wang [3] focused on the components of EXPRESS representing the data base modeling with possibility distribution and FUZZY sets are extended and integrating such EXPRESS FUZZY data base models. The EXPRESS FUZZY models are mapped to object oriented FUZZY data base by using formal techniques is being presented. The data access standard interface requirements are investigated for functioning the EXPRESS defined information in database [3]. Similar to this work, Zhao and Liu [4] developed a technique which is based on ontology for interoperable meaning and logic of model knowledge. Two languages namely OWL and SWRL for web semantics are used to develop information model of product. They disused...
EXPRESS language as traditional language they presented information model of EXPRESS oriented product using representation technique. They introduce the importance of representing technique is configured from EXPRESS to SWRL/OWL [4, 5]. Another similar work was [6] implementation of an EXPRESS model on a database vault, questioned upon and controlled. They introduce a product advancement stage which aids creators to make EXPRESS models, to produce a proportional database composition and to control this mapping. Likewise they can make an introductory EXPRESS model utilizing EXPRESS-G. EXPRESS code is then created and further altering embraced utilizing a particular EXPRESS proofreader. They can likewise imagine the legacy chain of importance of the complete pattern. The EXPRESS pattern is changed over to an article situated database framework and got to by means of a STEP Data Access Interface (SDAI). They portray every segment of the advancement stage furthermore inspect the conceivable improvements. The ISO STEP particulars plan to give a viable means by which item data can be imparted, traded in the middle of utilizations and ventures. EXPRESS is a modeling language inside the STEP particulars and it is utilized to portray item information [6].

Kahn et al. [11] describe a structure for controlling EXPRESS models and the objective is to hold the STEP idea of the direct mapping of a data model to a usage, yet to do as such in a manner that empowers elective execution techniques to be received. In this structure, called STEPWISE, permits the client to point out controls and model changes keeping in mind the end goal to change over models from one structure into an alternate [11]. Similarly, Sukys et al. [12] presented transformation framework to transform questions in structured language SBVR to SPARQL queries for ontologies defined in Web Ontology Language (OWL) 2.0 and supplemented with semantic web rules SWRL. This transformation depends on OWL 2 ontology related with corresponding SBVR vocabulary and rules. They considers a family of transformations and metamodels required for relating ontologies, rules, SPARQL queries and real business data supported by computerized information systems, as well as establishes requirements for harmonizing the coexistence and preserving semantics of these different representations [12].

A process for planning change was presented as a medium for capturing structural and behavioral qualities of a model change [19], that supports layouts which, when instantiatted, naturally create proportional formal specification with investigation capacities. They demonstrated with a little illustration, UML Class to Relational Database change, and verification utilizing Alloy, and the improvement of model changes is normally an specially appointed action in MDE. Changes are designing items and can be created in a decently composed manner, in the same way as other programming items, and model change advancement methodology can deliver changes communicated in various styles, change designs can be utilized to strengthen such distinctive properties to be developed.

López-Ortega [14] introduced a system, which is utilized to turn upward the meaning of element and relations of the operation information and to make the acclimating Data Manipulation Language (DML) articulations execute and afterward the mapping tenet of EXPRESS information [14] style to social database, including fundamental information sort, object information sort and legacy, is depicted. The mapping method bantered about can be utilized as a part of all STEP information constructions, and STEP is an item demonstrating approach that considers all the peculiarities of an item, including geometry and hierarchical information [16]. Another technique was developed that was based on ontology for interoperable meaning and logic of model knowledge. Two languages namely OWL and SWRL for web semantics are used to develop information model of product. They disused EXPRESS language as traditional language they presented information model of EXPRESS oriented product using representation technique. They introduce the importance of representing technique is configured from EXPRESS to SWRL/OWL [17]. A methodology was presented for changing business rules (BR) made in regular tongue (Natural Language) [18]. All the related work discussed in this section highlights that there is currently no approach available to generate EXPRESS models from natural language or SBVR text.
3. PRELIMINARIES

3.1 EXPRESS Data Modeling Language

A standard language for modeling material extracted from software products is called EXPRESS and the exchange type of model STEP (10303) standard in ISO is formalized for EXPRESS and the evolved standard is 10303-11. The problem domain representation into EXPRESS model is achieved through different schemas which are used for grouping elements having similar semantic and persistence. The different data types including plain types e.g. String or integer containers e.g. LIST or SET, enumerations and data types defined by users are assisted by EXPRESS language entity types is a key type in EXPRESS language to define the elements and constraints [2].

Information Modeling is a system used to portray and analyze data necessities anticipated that would backing the business structures inside the degree of looking at information systems in affiliations and the technique of data showing incorporates capable data modelers working almost with business accomplices, As well as potential customers of the information structure.

3.2 Semantic Business Vocabulary Rules

SBVR is an openly accessible determination from the Object Management Group [1] expected the premise for a prescribed and documented characteristic dialect revelatory depiction of a complex substance, for example, a commercial, SBVR is prearranged to validate difficult agreeability standards, for example, operational principles for an undertaking, security strategy, standard consistence, or administrative agreeability rules.

3.3 Model Transformation

A model change, in model-driven designing is an automatable method for guaranteeing that a group of models is reliable, in an exact sense which the product architect can characterize and the point of utilizing a model change is to spare exertion and lessen lapses via computerizing the building and adjustment of models where conceivable. A model change may be composed in a broadly useful programming dialect; however specific model change dialects are additionally accessible.

4. USED APPROACH

4.1 SBVR to EXPRESS Transformation Framework

In this framework, we extract the SBVR specifications then find the SBVR elements and describe them individually then implement them on EXPRESS metamodel elements which are finding from EXPRESS metamodel using transformation rules as shown in Fig. 1. Here, the researcher compares both metamodel elements and describe them individually and extract them.

4.2 Semantic Analysis of SBVR Specifications

To distinguish the SBVR vocabulary, semantic part naming is performed and Semantic part marking or topical part marking is a shared methodology utilized as a part of shallow semantic parsing, and the SBVR components, for example, thing

![Fig. 1. SBVR to EXPRESS transformation framework.](image-url)
idea, singular idea, article sort, verb ideas are distinguished from the SBVR data.

4.3 Extracting SBVR Elements
By using the following conservative of mappings are used to extract of SBVR elements is as following:

4.3.1. Extracting Noun Concept
Different methods are used to describe the noun and noun phrases in English language. In old-fashioned syntax, nouns are imparted to be words that mention to individuals, things, places, or immaterial concepts and while up-to-date morphology discover this classification to be challenging because it depend on generic nouns such as objects to exactly define what a noun is, in our community thoughtful of what nouns are complies to the old-fashioned meaning but In SBVR, the mutual noun phrase (co-actors, actors, recipients, thematic things,) or nouns, are plotted as the object types e.g. employee, tools, chair, etc.

4.3.2. Extracting Individual Concepts
Here, each single type of individual concepts has exactly one unique element, which is the instance of the proper nouns (recipients, co-actors, actors and thematic objects) and distinct concept, is plotted to the distinct ideas. In the distinct noun concept ‘Bahawalpur’ is a one instance and it is a specific state in the Punjab of Pakistan. The General concepts, Individual noun concepts, Verb concepts, noun concept and concept type are also defined individual concepts.

4.3.3. Extracting Fact Types
In SBVR fact types (kinds of facts, such as “Employee works for Department”) are explain obligation the subordinate and exploit verbs are characterized as Association, Property, Property association, Partitive verb ideas to building a Fact Types. Examples of fact-based representation of a database schema, the information structure implied

---

**Fig. 2.** Output EXPRESS model.
4.3.4. Extracting Characteristics

In SBVR, trademark which is key to comprehension an idea. An idea has an inferred trademark just on the off chance that it takes after by sensible ramifications from some blend of consolidations of attributes by ideas and/or structural standards that the trademark is constantly ascribed to every example of the idea. Furthermore clarify parallel verb idea, unitary thing idea and individual thing idea. Every example of every trademark sort is a trademark. The augmentation of the trademark sort “shading” incorporates the attributes ‘thing is blue’, ‘thing is red’, ‘thing is green’ and so forth. The core package contains all of the generally required modeling elements of EXPRESS, along with some basic metamodel artifacts, and it is the foundation on which all of the other packages are built.

The Core Package is the minimal implementation of the EXPRESS metamodel, and the Core package contains all of the generally required modeling elements of EXPRESS, such as scopes and naming concepts, schemas, data types, entities, attributes, and relationships along with domain constraints. The Core package also includes the abstract classes Expression and Instance, which serve as linking points for detailed models contained in other packages.

4.1.5 Transforming SBVR to EXPRESS

Semantic Business Vocabulary Rules are transformed to the EXPRESS and present the rules of transformation. The all elements of the EXPRESS model is mapped with the elements of SBVR by using rules of transformation and the above section is already discussed the extraction of SBVR specifications which are used in this transformation. The following section describes the transformation method of mapping SBVR elements with element of EXPRESS data models:

4.3.6. Mapping Verb Concepts to Methods

In this section we transform or map verb concepts (action verbs) of the SBVR are related to a noun concept are transformed and Verb concepts map to fact types, each fact type being a set of possible ground facts that can be formulated based on the verb concept and that use reference schemes to identify, for each fact, each thing that fills each role.

4.3.7. Mapping Noun Concept to EXPRESS Class

In the Semantic Business Vocabulary Rule, entirely the Noun Concepts are transformed to EXPRESS classes in an EXPRESS Core Package model. The idea is the importance of a thing or thing expression.

4.3.8. Mapping Individual Concepts to Objects

In this phase transform all individual concepts and its sub elements of Semantic Business Vocabulary Rule to the objects in an EXPRESS Core model and all the general, individual, noun concepts, verb concepts and concept types are mapped to object and with its elements. Individual noun concepts logically map to singleton types of individuals. Each single type of individual has exactly one element, which is the instance of the individual noun concept.

4.3.9 Mapping Definitional Rule to Global Rule

A Schema Element denoting a collection of Named Rules for the interaction of the Extents of one or more Entity Types and It corresponds to the RULE declaration in EXPRESS; Every Global Rule is also an Algorithm Scope and may define Common Elements and Variables.

4.3.10. Mapping Characteristics to Attribute

The concept of SBVR element characteristics is mapped with the element of EXPRESS element Attribute. The characteristic “driver is of age” by this meaning: “the age of the driver is at least the EU-Rent Minimum Driving Age.” In this element dualistic verb concept, unitary noun concept and individual noun concept are also sub elements are mapped with the sub elements of Attribute like. Entity Definition, Inverse attribute, explicit attribute, Derived attribute and Base type.
5. IMPLEMENTATION

Our aim was to define computerized transformations from natural language to an EXPRESS model. In this thesis, we focused on model to model transformation from a SBVR specification to an EXPRESS model and proposed text to model transformations. We made the assumptions that the input SBVR specification is consistent and complete. Our approach was divided into two steps:

- Firstly, extracting SBVR elements. SBVR elements are extracted using NL2SBVR tool [9]; and
- Defining applied Business Rules transformations to the control flow of the identified elements of EXPRESS and top21otentiallyrefinethem.

6. CASE STUDIES

The following case study will help us to generate the EXPRESS data model by using SBVR rules and system requirements:

“A library issues loan items to student. Each student is known as a member and is issued a membership card that shows a unique member number. It is necessary that the membership number and other details on a student must be kept such as a name, address, and date of birth. The library is made up of a number of subject sections. Each section is denoted by a classification mark. A loan item is identified by a bar-code. There are exactly two types of loan items, language tapes, and books. A language tape has a title-language, and level. A book has a title, and author(s). It is a possibility that each item can be borrowed, reserved or renewed to extend a current loan. When an item is issued the student’s membership-number is scanned via a barcode reader or entered manually. If the membership is valid and the number of items on loan less than 8, the book barcode is read, either via the barcode reader or entered manually. If the item can be issued (e.g. not reserved) the item is stamped and then issued. The library must support the facility for an item to be searched and for a daily update of records.”

6.1 SBVR Vocabulary Generation

The SBVR specification (output of NL2SBVR tool) was given as input to the SBVR2EXPRESS tool that is an Eclipse plugin implemented in java as a proof of concept. The following text is shown in SBVR Structured English notation [1] that represents each SBVR vocabulary type in predefined different colours. The SBVR specification after extracting SBVR vocabulary is as follows:

“A library issues loan items to each student. Each student is known as a member and is issued a membership card that shows a unique member number. It is necessary that the membership number and other details on a student must be kept such as a name, address, and date of birth. The library is made up of a number of subject sections. Each section is denoted by a classification mark. A loan item is identified by a bar-code. There are exactly two types of loan items, language tapes, and books. A language tape has a title-language, and level. A book has a title, and author(s). It is a possibility that each student may borrow up to at most 8 items. It is a possibility that each item can be borrowed, reserved or renewed to extend a current loan. When an item is issued the student’s membership-number is scanned via a barcode reader or entered manually. If the membership is valid and the number of items on loan at most 8, the book’s bar-code is read, either via the barcode reader or entered manually. It is possibility that if the item can be issued the item is stamped and then issued. It is necessary that the library must support the facility for an item to be searched and for a daily update of records.”

Afterwards, the extracted SBVR vocabulary was mapped to the EXPRESS elements. Following information was extracted in OO analysis phase:

6.2 SBVR Based Software Requirements

For this situation study Library information Management System is an (invented) books library, utilized as a contextual analysis as a part of the SBVR determination. The business prerequisites for Library information Management System incorporate the accompanying:

- Authors wrote different books
- Books have different types
- Different catalogue has record of books
• Each book has account in library
• Each student has account in library
• Librarian searches the records of students and books
• Librarian manages the records of students and books

6.3 SBVR Rules Generation

SBVR Rule: It is necessary that each student has exactly one account.

SBVR Rule: It is necessary that each catalogue includes exactly one book.

SBVR Rule: It is necessary that the Librarian manage each account.

SBVR Rule: It is necessary that each Patron uses exactly all accounts.

SBVR Rule: It is obligatory that the Book issuance of a book is at most 2 months.

6.4 Generated EXPREES Diagram

EXPRESS generated Library information Management System diagram is shown in Figure 2.

7. EXPERIMENTS & RESULTS

Our assessment strategy is taking into account three things and technique that are portrayed. For the evaluation purpose, we have used three metrics: Precision, Recall and F-Measure.

8. RESULTS AND DISCUSSION

There were seven sentences in the used case study problem. The largest sentence was composed of 58 words and the smallest sentence contained 18 words. The average length of all sentences is 26. The major reason to select this case study was to test our tool with the complex examples. The correct, incorrect, and missing SBVR elements are shown in Table 3.

Results of each SBVR element describe in above Table separately. According to our evaluation methodology, Table 4 shows sample elements are 54 in which 48 are True 04 are False and 02 are error SBVR elements. The above table describes the Recall and precision of examples for generating EXPRESS data model. In Table 4, the average recall for SBVR software requirements are premeditated 96.42% while average precision is calculated 93.10%. The normal F-value is figured 84.90% that is encouraging for introductory analysis. We can’t hope to measure up our outcomes to whatever other apparatus as no other device is accessible that can produce EXPRESS information model using SBVR.

<table>
<thead>
<tr>
<th>Elements of SBVR Metamodel</th>
<th>mapped to</th>
<th>Elements of EXPRESS Metamodel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual Concept</td>
<td>→</td>
<td>Object</td>
</tr>
<tr>
<td>Noun Concept</td>
<td>→</td>
<td>Class</td>
</tr>
<tr>
<td>Verb Concept</td>
<td>→</td>
<td>Methods</td>
</tr>
<tr>
<td>Business Rule</td>
<td>→</td>
<td>Schema</td>
</tr>
<tr>
<td>Definitional Rule</td>
<td>→</td>
<td>Global Rules</td>
</tr>
<tr>
<td>Characteristic</td>
<td>→</td>
<td>Attribute</td>
</tr>
<tr>
<td>Elementary Concepts</td>
<td>→</td>
<td>Base Type</td>
</tr>
<tr>
<td>Constraints</td>
<td>→</td>
<td>Constraints</td>
</tr>
<tr>
<td>PartitiveFactType</td>
<td>→</td>
<td>Generalization</td>
</tr>
<tr>
<td>Associative FactType</td>
<td>→</td>
<td>Association</td>
</tr>
<tr>
<td>CategorizationFactType</td>
<td>→</td>
<td>Aggregation</td>
</tr>
<tr>
<td>Quantification</td>
<td>→</td>
<td>Cardinalities</td>
</tr>
</tbody>
</table>
Table 2. List of Generated SBVR vocabulary.

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun Concept</td>
<td>08</td>
<td>Student, library, subject section, section, loan items, language tapes,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>books, car movement, receiving branch, geographical movement type.</td>
</tr>
<tr>
<td>Verb Concepts</td>
<td>06</td>
<td>Membership, bar code, records, reader, item, records, issued, receiving</td>
</tr>
<tr>
<td></td>
<td></td>
<td>branch</td>
</tr>
<tr>
<td>Individual Concepts</td>
<td>08</td>
<td>Borrowed, issued, manually, Author, title, searched, car group,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>movement-id, rental car, movement-id</td>
</tr>
<tr>
<td>Characteristics</td>
<td>03</td>
<td>Renewed, served, extend, searched, stamped</td>
</tr>
<tr>
<td>Quantifications</td>
<td>05</td>
<td>Unique, scanned, valid, daily, reader, in-country car movement</td>
</tr>
<tr>
<td>Constraints</td>
<td>08</td>
<td>Manually, facility, read, title language, membership, tapes, borrowed,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>one-way car, movement</td>
</tr>
<tr>
<td>Associative Fact Types</td>
<td>06</td>
<td>Section, department, class, catalogue, sending branch.</td>
</tr>
<tr>
<td>Partitive Fact Types</td>
<td>04</td>
<td>Possibility, manually, round-trip car movement</td>
</tr>
<tr>
<td>Categorization Fact Types</td>
<td>06</td>
<td>At most, stamped, record, international car movement</td>
</tr>
</tbody>
</table>

Table 3. Recall and Precision of examples for EXPRESS model from case study.

<table>
<thead>
<tr>
<th>Type/Metrics</th>
<th>N\textsubscript{example}</th>
<th>N\textsubscript{True}</th>
<th>N\textsubscript{False}</th>
<th>N\textsubscript{error}</th>
<th>Rec%</th>
<th>Prec%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software Requirements</td>
<td>54</td>
<td>48</td>
<td>04</td>
<td>02</td>
<td>96.42</td>
<td>93.10</td>
</tr>
</tbody>
</table>

Table 4. Evaluation results.

<table>
<thead>
<tr>
<th>Input</th>
<th>N\textsubscript{example}</th>
<th>N\textsubscript{True}</th>
<th>N\textsubscript{False}</th>
<th>N\textsubscript{error}</th>
<th>Rec%</th>
<th>Prec%</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>48</td>
<td>40</td>
<td>6</td>
<td>2</td>
<td>84.32</td>
<td>87.28</td>
<td>80.76</td>
</tr>
<tr>
<td>Sample 2</td>
<td>50</td>
<td>45</td>
<td>4</td>
<td>1</td>
<td>83.56</td>
<td>84.41</td>
<td>83.43</td>
</tr>
<tr>
<td>Sample 3</td>
<td>45</td>
<td>35</td>
<td>5</td>
<td>5</td>
<td>84.28</td>
<td>86.50</td>
<td>82.66</td>
</tr>
<tr>
<td>Sample 4</td>
<td>54</td>
<td>48</td>
<td>4</td>
<td>2</td>
<td>85.43</td>
<td>85.84</td>
<td>84.68</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>84.26</td>
<td>85.98</td>
<td>84.90</td>
</tr>
</tbody>
</table>

9. CONCLUSIONS AND FUTURE WORK

This paper deals with the generation of EXPRESS models from SBVR details to EXPRESS Core Package components. This generation begins with the SBVR parsing and by utilizing ordinary NLP methodologies and second issue of change of SBVR metamodel components to EXPRESS metamodel components was tended to by utilizing model change innovation. SiTra library was utilized with the end goal of the model change. Moreover, the mechanized article arranged investigation of SBVR particulars of programming necessities was finished. The outcomes demonstrate that the displayed methodology is a superior approach when contrasted with the other accessible methodologies. Our tool SBVR2EXPRESS gives a higher exactness when contrasted with other accessible NL-based tools, what’s more better precision,
SBVR has likewise empowered to concentrate OO data, for example, affiliation variety, collections, speculations and occurrences as other NL-based apparatuses can’t process and concentrate this data.

The possible future development of the extracted SBVR elements and transformation to EXPRESS elements is mapping to its graphical representation. In this paper, all work that we have planned were not completed, we transform only a few elements of both metamodels. Mapping of other elements are remaining and we will do it further in future. After the complete transformation it will further helpful and use full in the field of Data modelling and representation of data in graphically with notations. Our focus on the exact and complete model transformation between EXPRESS and SBVR

10. REFERENCES